Systematic uncertainties in parton distribution functions from lattice QCD simulations at the physical point

Constantia Alexandrou,[1,2] Krzysztof Cichy,[3] Martha Constantinou,[4] Kyriakos Hadjiyiannakou,[2] Karl Jansen,[5] Aurora Scapellato,[1,6] and Fernanda Steffens[7]

[1]*Department of Physics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus*
[2]*Computation-based Science and Technology Research Center, The Cyprus Institute, 20 Kavafi Str., Nicosia 2121, Cyprus*
[3]*Faculty of Physics, Adam Mickiewicz University, Umultowska 85, 61-614 Poznań, Poland*
[4]*Temple University, 1925 N. 12th Street, Philadelphia, Pennsylvania 19122-1801, USA*
[5]*NIC, DESY, Platanenallee 6, D-15738 Zeuthen, Germany*
[6]*University of Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany*
[7]*Institut für Strahlen- und Kernphysik, Rheinische Friedrich-Wilhelms-Universität Bonn, Nussallee 14-16, 53115 Bonn*

(Received 16 February 2019; published 17 June 2019)

We present a detailed study of the helicity-dependent and helicity-independent collinear parton distribution functions (PDFs) of the nucleon, using the quasi-PDF approach. The lattice QCD computation is performed employing twisted mass fermions with a physical value of the light quark mass. We give a systematic and in-depth account of the salient features entering in the evaluation of quasi-PDFs and their relation to the light-cone PDFs. In particular, we give details for the computation of the matrix elements, including the study of the various sources of systematic uncertainties, such as excited-states contamination. In addition, we discuss the nonperturbative renormalization scheme used here and its systematics, effects of truncating the Fourier transform and different matching prescriptions. Finally, we show improved results for the PDFs and discuss future directions, challenges and prospects for evaluating precisely PDFs from lattice QCD with fully quantified uncertainties.

DOI: 10.1103/PhysRevD.99.114504

I. INTRODUCTION

Quantum chromodynamics (QCD), being the fundamental theory of the strong interactions, describes the interactions among quarks and gluons both in the perturbative regime, as well as in the nonperturbative regime with the emergence of complex structures like hadrons. As the latter requires a nonperturbative approach, the use of a discretized Euclidean spacetime to numerically solve QCD has been proven to be extremely successful in the computation of physical properties of hadrons, such as their masses and decay constants. The rich and diverse structure of hadrons is described in terms of observables like various kinds of form factors, revealing different structural aspects, such as the average distribution of electric charge or parton (quark and gluon) momentum and spin inside the hadron. Even more information is contained in distribution functions, such as parton distribution functions (PDFs), generalized parton distributions and transverse-momentum-dependent PDFs, which decompose the aggregate observables into probability densities for partons with specific longitudinal momentum, transverse momentum and momentum transfer within a scattering event. The simplest of these are the PDFs, which have been studied intensively and continuously in experimental facilities over the last few decades, most notably for the proton, and also provide input in collider experiments. It is well established that the main source of information on PDFs are global QCD analyses, providing accurate results due to theoretical advances and new data emerging from accelerators, covering different kinematical regions (see e.g., Refs. [1–12]). Such analyses rely on the factorization framework, in which scattering cross sections, generally obtaining contributions from all energy scales, are written as convolutions of perturbatively computable coefficient functions describing the high-energy scales, and the low-energy PDFs. Despite the tremendous progress in phenomenological parametrizations of PDFs, the procedures are not without ambiguities [7], as there are still kinematical regions that are not easily accessible experimentally, e.g., the large Bjorken-x region. Uncertainties on the PDFs in the large-x range propagate to smaller-x regions when QCD evolution is applied.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. 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[3].

A calculation of PDFs from first principles is a valuable addition to the global fitting analyses and of crucial importance for the deeper understanding of the inner structure of hadrons. Apart from providing a fundamental framework for the study of these quantities, it can also serve as input for experimental analyses in collision experiments. The nonperturbative nature of PDFs makes lattice QCD an ideal *ab initio* formulation to determine them, utilizing large-scale simulations. A novel method to extract parton distribution functions from lattice QCD was proposed five years ago by Ji [13]. It is based on considering matrix elements probing purely spatial correlations, making them accessible in Euclidean lattice QCD.

The first studies on quasi-PDFs within lattice QCD [14–16] appeared soon after their proposal by Ji. These papers and the follow-up ones [17,18] focused mainly on the feasibility of the approach where only the bare nucleon matrix elements of boosted nucleons are calculated for nonsinglet operators. The first computations that utilized a proper procedure for renormalization and matching to $\overline{\text{MS}}$ light-cone PDFs and using simulations with physical pion masses were published in Ref. [19] for the unpolarized isovector PDFs, $u(x) - d(x)$, and helicity PDF, $\Delta u(x) - \Delta d(x)$ and in Ref. [20] for the isovector transversity PDF, $h_1^u(x) - h_1^d(x)$. These works were published in shorter letters, where many of the technical and methodological details of the calculations needed to be left out.

It is the purpose of this paper to fill this gap and provide the above-mentioned details in a comprehensive way, give the present understanding and control of systematics effects appearing in the computations and to present the status of the calculations of PDFs for our twisted mass setup. In particular, we extend the work of Refs. [19,20] and give a systematic and in-depth account of the salient features entering in the evaluation of quasi-PDFs and their relation to the light-cone PDFs. We explain in detail the computation of the matrix elements including the study of the various sources of systematic uncertainties, such as excited-states contamination. An improved renormalization procedure is presented using three ensembles for the nonperturbative extraction of the renormalization functions, allowing to take the chiral limit, as well as an investigation of other sources of systematic uncertainties related to the renormalization, such as the finite volume and scale dependence. In addition, different prescriptions for applying the Fourier transform and the matching prescription are employed and critically compared.

On the physics side, in order to compare with existing phenomenological estimates, we convert our results in the RI′ scheme into the $\overline{\text{MS}}$ scheme following the procedure given in Ref. [21]. We then apply appropriate matching to relate quasi-PDFs in the $\overline{\text{MS}}$ scheme to light-cone PDFs extracted from global fits. With the detailed analysis of systematic effects, as described in this work, having a physical value of the pion mass and the improved renormalization and matching prescriptions, we are finally in a position to indeed make contact with the phenomenological analyses of PDFs and we present final results of our lattice PDF calculations in Sec. VI below. In addition to our work, results on the quasidistribution approach were reported in Refs. [22–32] that include, besides the nucleon PDFs, exploratory studies of pion distribution amplitudes and PDFs as well as of gluon PDFs. For complete references and an overview of the theoretical and numerical developments, see Ref. [33].

On more general grounds, it is very important to realize that quasi-PDFs and light-cone PDFs have been shown to share the same infrared physics [34–37], which is the fundamental observation that allows one to relate both quantities using perturbation theory, provided that the hadron is moving with a large, although necessarily finite, momentum in a chosen spatial direction. It has also been proven that quasi-PDFs can be extracted from lattice QCD in Euclidean spacetime [36] and that they do not suffer from power-divergent mixings with lower-dimensional operators [38–40]. A factorization formula makes it possible to extract the PDFs from the quasi-PDFs, an operation called matching [19,20,29,31,34,35,37,41–45]. In general, this procedure is based on a newly developed large-momentum effective theory [46], and it is renormalizable to all orders in perturbation theory [47–50]. Other approaches for a direct computation of the x dependence of PDFs include the hadronic tensor [51–53], fictitious heavy quark [54], auxiliary light quark [55], good lattice cross sections [35,37] (closely related to the auxiliary light quark method), "OPE without OPE" [56] and pseudo-PDFs [57–60], where the latter can be seen as a generalization of PDFs off the light cone. These alternative approaches have been explored in lattice QCD, and recent results can be found in Refs. [40,56,61–68]. The new formulation and its successful implementation within lattice QCD has also led to a wider interest in phenomenological studies using models and toy theories of QCD [57,58,69–76]. A detailed overview of the current status of lattice QCD calculations of PDFs and other partonic distributions can be found in the recent reviews of Refs. [33,77].

The remainder of the paper is organized as follows. In Sec. II, we provide the general theoretical aspects, lattice QCD action and parameters. We discuss the computation of the bare matrix elements, along with the lattice techniques necessary to attain high statistical accuracy. In Sec. III, we present our results on bare matrix elements, together with various analysis methods that are required to control excited-states contamination. Section IV describes in detail the nonperturbative renormalization program, which investigates pion mass dependence, volume effects and renormalization scale dependence. The matching procedure is presented in Sec. V and we demonstrate that a modified $\overline{\text{MS}}$ scheme is preferable. Different prescriptions are presented and we compare the effect on the final PDFs. Section V also

includes an investigation of systematic uncertainties resulting from the Fourier transform. In Sec. VI, we present our final results with a discussion on the systematic uncertainties, while in Sec. VII we conclude.

II. METHODOLOGY

A. From PDF to quasi-PDF

The original definition of the quark momentum distribution within a hadron can be derived from the operator product expansion (OPE) of hadronic deep-inelastic scattering and is given by

$$q(x)=\int_{-\infty}^{+\infty} d\xi^- e^{-ixP^+\xi^-}\langle N|\bar{\psi}(\xi^-)\Gamma W(\xi^-,0)\psi(0)|N\rangle, \quad (1)$$

where x is the momentum fraction of the quark, $|N\rangle$ is the hadron state at rest (assumed to be the nucleon), the momentum $P^+=(P^0+P^3)/\sqrt{2}=M/\sqrt{2}$, $\xi^-=(\xi^0-\xi^3)/\sqrt{2}$, $W(\xi^-,0)$ is the Wilson line connecting ξ^- and 0, and Γ is a Dirac matrix whose structure defines the momentum distributions of quarks having spin parallel or perpendicular to the nucleon momentum. The factorization scale is implicit. Equation (1) is light-cone dominated, as it receives contributions only in the region where $\xi^2=t^2-\vec{r}^2=0$. Consequently, a direct implementation of this equation is impossible on the lattice, where only a single point can satisfy the Euclidean light-cone condition $\xi^2=t^2+\vec{r}^2=0$. The alternative approach proposed by Ji overcomes this difficulty by reconstructing light-cone distributions from correlation functions where the bilinear fermionic fields, ψ and $\bar{\psi}$, are separated by a purely spatial distance z and the nucleon is boosted with a finite momentum. Quasi-PDFs are defined as

$$\tilde{q}(x,P_3)=\int_{-z_{\max}}^{+z_{\max}}\frac{dz}{4\pi}e^{-ixP_3z}\langle N|\bar{\psi}(0,z)\Gamma W(z,0)\psi(0,0)|N\rangle, \quad (2)$$

where $|N\rangle$ represents a nucleon, which is boosted in the z direction with momentum $P=(P_0,0,0,P_3)$. On the lattice, quasi-PDFs are thus computed using matrix elements of nonlocal operators containing a straight Wilson line of finite length z, that varies from 0 to some maximum value, $z_{\max}$. This aspect will be further discussed in the following sections. The standard light-cone PDFs can be recovered from quasi-PDFs using the large momentum effective theory [46], developed specifically for these operators. Eventually, quasi-PDFs are matched to the physical PDFs through the factorization

$$q(x,\mu)=\int_{-\infty}^{\infty}\frac{d\xi}{|\xi|}C\left(\xi,\frac{\mu}{P_3}\right)\tilde{q}\left(\frac{x}{\xi},\mu,P_3\right)+O\left(\frac{m_N^2}{P_3^2},\frac{\Lambda_{\rm QCD}^2}{P_3^2}\right), \quad (3)$$

where $q(x,\mu)$ is the light-cone PDF at the scale μ and we follow the usual choice of common factorization and renormalization scales, $\mu_F=\mu_R\equiv\mu$. C is the matching kernel, which can be computed perturbatively and has so far been evaluated to the one-loop level. For sufficiently large momenta, ideally much larger than the nucleon mass, m_N, and $\Lambda_{\rm QCD}$, higher-order corrections in α_s are very small and the extraction of quark distributions is fully robust. Presently, such a situation cannot be accomplished and it is part of this paper to investigate whether practically reachable momenta are sufficient.

B. Lattice gauge ensemble

The analysis is performed using a gauge ensemble of two dynamical degenerate light quarks ($N_f=2$), generated by the ETM[1] Collaboration [78]. The light quark mass has been tuned to reproduce the physical value of the pion mass, henceforth referred to as the *physical point*. The lattice volume is $48^3\times 96$, the lattice spacing $a=0.0938(2)(3)$ fm [79], the spatial lattice extent $L\approx 4.5$ fm and $m_\pi L=2.98$. Based on a model prediction of excited-states effects in certain hadron structure quantities [80], the dependence on $m_\pi L$ is mild (see, e.g., Fig. 3 of Ref. [81]). The complete list of parameters of this ensemble is reported in Table I.

The gauge configurations were generated with the Iwasaki-improved gauge action [82]. In the fermionic sector, the twisted mass fermion action [83,84] with a clover term [85] was employed. The fermion action is given by

$$S_F[\chi,\bar{\chi},U]=\sum_x\bar{\chi}(x)\left(D_W[U]+i\mu_l\gamma^5\tau^3+\frac{i}{4}c_{\rm SW}\sigma^{\mu\nu}\mathcal{F}^{\mu\nu}[U]\right)\times\chi(x), \quad (4)$$

where D_W is the Wilson-Dirac operator, μ_l is the bare twisted mass for the light quarks, $\tau^3=\mathrm{diag}(1,-1)$ is the third Pauli matrix in flavor space and the last term includes the field-strength tensor $\mathcal{F}^{\mu\nu}[U]$ weighted by $c_{\rm SW}$, known as the Sheikoleslami-Wohlert (clover) coefficient. In the gauge ensemble used in this work, the clover parameter is taken to be $c_{\rm SW}=1.57551$ [86]. In Eq. (4) $\chi(x)=(u,d)^T$ denotes the light quark doublet in the "twisted basis" at maximal twist. The fermion fields in the "physical basis," denoted by $\psi(x)$, are recovered by the following chiral rotation:

$$\psi(x)\equiv e^{i\frac{\pi}{2}\gamma^5\tau^3}\chi(x), \qquad \bar{\psi}(x)\equiv\bar{\chi}(x)e^{i\frac{\pi}{2}\gamma^5\tau^3}, \quad (5)$$

[1]Effective as of this year, the European Twisted Mass Collaboration has officially changed its name to the Extended Twisted Mass Collaboration, as now it also comprises members from non-European institutions. Along with the name change, there is also a new logo.

TABLE I. Parameters of the ensemble used in this work. The nucleon mass (m_N), the pion mass (m_π) and the lattice spacing (a) have been determined in Ref. [79].

$\beta=2.10$, $c_{\rm SW}=1.57751$, $a=0.0938(3)(2)$ fm, $r_0/a=5.32(5)$	
$48^3\times 96$, $L\approx 4.5$ fm	$a\mu = 0.0009$
	$m_\pi = 0.1304(4)$ GeV
	$m_\pi L = 2.98$
	$m_N = 0.932(4)$ GeV

where $\alpha = \pi/2$ at maximal twist. In the next sections of this paper, the interpolating fields and the nucleon matrix elements have to be understood with quark fields in the physical basis, unless otherwise specified.

The introduction of the twisted mass term in the lattice action has a series of advantages in hadron structure calculations, such as excluding zero eigenvalues from the spectrum of the Wilson-Dirac operator allowing a speed-up in the numerical simulations. Moreover, it also simplifies renormalization properties of operators [84,87–89] and at maximal twist it provides an automatic $O(a)$ improvement. However, the isospin symmetry breaking of the twisted mass formulation could lead to instabilities when simulations are carried out at light quark masses close to the physical value. The isospin breaking, which manifests itself in the neutral pion being lighter than the charged one, is reduced with the addition of the clover term in the fermion action [78].

C. Nucleon bare matrix elements

In this work, we evaluate the unpolarized, helicity and transversity quasi-PDFs using three values for the nucleon momentum, that is, $P_3 = 6\pi/L$, $P_3 = 8\pi/L$ and $P_3 = 10\pi/L$, corresponding in physical units to 0.83, 1.11 and 1.38 GeV, respectively. The matrix element of interest is given by

$$h_\Gamma(P_3, z) = \langle N|\bar\psi(0,z)\Gamma W(z,0)\psi(0,0)|N\rangle, \qquad (6)$$

for a straight Wilson line, W, with varying length from $z = 0$, corresponding to the standard ultralocal operators, up to half of the spatial extension, $L/2$. Γ is the Dirac structure leading to different types of PDFs, as discussed below. The matrix elements of Eq. (6) are extracted from a ratio of two- and three-point functions, averaged over the gauge field ensemble. The two-point and three-point functions are given by

$$C^{\rm 2pt}(\mathbf{P}, t, 0) = \mathcal{P}_{\alpha\beta}\sum_{\mathbf{x}} e^{-i\mathbf{P}\cdot\mathbf{x}}\langle 0|N_\alpha(\mathbf{x}, t)\bar N_\beta(\mathbf{0}, 0)|0\rangle, \qquad (7)$$

$$C^{\rm 3pt}(\mathbf{P}, z; t_s, \tau, 0) = \tilde{\mathcal{P}}_{\alpha\beta}\sum_{\mathbf{x},\mathbf{y}} e^{-i\mathbf{P}\cdot\mathbf{x}}\langle 0|N_\alpha(\mathbf{x}, t_s) \times \mathcal{O}(\mathbf{y}, \tau; z)\bar N_\beta(\mathbf{0}, 0)|0\rangle, \qquad (8)$$

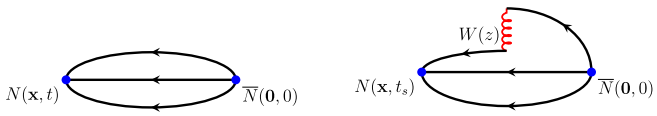

FIG. 1. Schematic representation of the nucleon two- (left) and three-point (right) functions. The creation point of the nucleon is at $(\mathbf{0}, 0)$, while the annihilation is at $(\mathbf{x}, t_s)$ and $(\mathbf{x}, t_s)$ for the two- and three-point functions, respectively. The solid lines represent the quark propagators and the curly line denotes the Wilson line of length z.

where $N_\alpha(x)$ is the proton interpolating field, t (t_s) is the time separation of the sink relative to the source in the two-point (three-point) function,[2] τ is the insertion time of the $\mathcal{O}$ operator and $\mathcal{P}_{\alpha\beta}$, $\tilde{\mathcal{P}}_{\alpha\beta}$ are the parity projectors for the two- and three-point functions, respectively. For the two-point functions, we use the plus and minus parity projectors $\mathcal{P}^\pm = \frac{1\pm\gamma^0}{2}$ and average over the forward and backward correlators. The projector $\tilde{\mathcal{P}}_{\alpha\beta}$ depends on the operator under study. The operator $\mathcal{O}(\mathbf{y}, \tau; z)$ has the following form:

$$\mathcal{O}(\mathbf{y}, \tau; \mathbf{z}) = \bar\psi(\mathbf{y} + \mathbf{z}, \tau)\Gamma W(\mathbf{y} + z, \mathbf{y})\psi(\mathbf{y}, \tau), \qquad (9)$$

where the Wilson line is aligned in the direction of the nucleon momentum and ψ is the fermion doublet of flavors (u, d). We compute the isovector combination $u - d$, inserting a τ_3 in flavor space. With this choice, the disconnected diagrams cancel and the total contribution is obtained from the connected diagram shown in Fig. 1.

As mentioned in Sec. II A, the choice for the Dirac structure in the operator of Eq. (1) gives access to different quark distribution functions. In fact, to extract a given PDF, different Γ matrices may be used, such as γ^3 (parallel to the Wilson line direction) or γ^0 (temporal direction) for the unpolarized PDF. However, some choices are preferable (e.g., γ^0) as they avoid finite mixing under renormalization, which was found to be allowed in lattice regularization [21]. In the absence of mixing, the operators renormalize multiplicatively, and one avoids systematic uncertainties related to the elimination of mixing. In this work, we compute matrix elements of the operators with

(1) $\Gamma = \gamma^3$, γ^0 for the unpolarized distribution, $\tilde q(x) = \tilde q(x)_\uparrow + \tilde q_\downarrow(x)$,
(2) $\Gamma = \gamma^5\gamma^3$ for the helicity distribution, $\Delta\tilde q(x) = \tilde q_\uparrow(x) - \tilde q_\downarrow(x)$, and
(3) $\Gamma = \sigma^{3j}$ for the transversity distribution, $\delta\tilde q(x) = \tilde q_\perp(x) + \tilde q_\top(x)$.

[2]We use different symbols to emphasize that the three-point function is computed only for selected values of the source-sink separation, $t_s/a = 8, 9, 10, 12$ in this work, while the two-point function is evaluated for the full range of time separations to constrain more precisely the gap between the ground state and the first excited state in two-state fits; see Sec. III C.

$q_\uparrow$ ($q_\downarrow$) and $q_\perp$ ($q_\top$) indicate quarks with helicity aligned (antialigned) with that of a longitudinally and transversely polarized proton, respectively. The index j entering the matrix element of the transversity distribution is purely spatial and denotes the direction of the quark spin, which is orthogonal to the proton momentum. Each choice of an operator requires an appropriate parity projector $\tilde{\mathcal{P}}_{\alpha\beta}$, that is, $\frac{1+\gamma^0}{2}$ for the unpolarized, $i\gamma^3\gamma^5\frac{1+\gamma^0}{2}$ for the helicity, and $i\gamma^k\frac{1+\gamma^0}{2}$ for the transversity PDF. The desired matrix elements of Eq. (6) are finally extracted from fitting the ratio of three-point and two-point functions to isolate the ground state. One choice is a constant fit as a function of the time insertion of the operator over which the ratio is time independent (plateau region). Other choices are two-state fits that take into account the first excited-states contributions, as discussed in Sec. III with a comparison among them. Thus, the desired matrix element in the plateau fit is given by

$$h_\Gamma(p,z) \overset{0\ll\tau\ll t_s}{=} \mathcal{K}\frac{\langle C^{3\mathrm{pt}}(p,z;t_s,\tau,0)\rangle_G}{\langle C^{2\mathrm{pt}}(p,t_s,0)\rangle_G}, \tag{10}$$

where $\mathcal{K}$ is a kinematic factor equal to $\mathcal{K}=iE/p$ (with $E=\sqrt{m_N^2+p^2}$) for the vector current γ^3, whereas $\mathcal{K}=1$ for all other choices of the Dirac structure used. A gauge average ($\langle\cdots\rangle_G$) is performed on the two- and three-point functions prior to taking the ratio.

D. Lattice techniques

To evaluate the correlation functions of Eqs. (7) and (8), a state with the same quantum numbers as the nucleon is created at the source and annihilated at the sink. In our calculation, we employ point sources generated with the proton interpolating field $N_\alpha(x)=\epsilon^{abc}u_\alpha^a(x)(d^{b^T}(x)\mathcal{C}\gamma^5u^c(x))$, where $\mathcal{C}=\gamma^0\gamma^2$ is the charge-conjugation matrix. In general, the overlap of the states generated with such interpolating fields with the desired ground state is improved by employing smearing. Gaussian smearing [90,91] of quark fields at the source and the sink, used in combination with APE smearing on the gauge field, is a very effective approach to improve ground-state dominance. For nucleon states at rest, previous studies [92] performed on the gauge ensemble employed in this work, have extracted the optimal parameters for both Gaussian and APE smearing, tuned to approximately reproduce a nucleon state with an rms radius of about 0.5 fm and maximize the overlap with the ground state. These parameters are $(N_G,\alpha_G)=(50,4)$ for Gaussian smearing and $(N_{\mathrm{APE}},\alpha_{\mathrm{APE}}=0.5)$ for APE smearing.

However, in the computation of PDFs, we need to optimize the overlap to a boosted nucleon state while keeping the statistical noise minimal. We, thus, modify the standard Gaussian smearing by using momentum-smeared interpolating fields [93]. This has proven extremely advantageous when the nucleon is boosted with a large momentum, as is the case for quasi-PDFs. The momentum smearing modifies the standard Gaussian smearing function by including a complex phase factor that affects only the gauge links along the direction of the boost, namely

$$\mathcal{S}_{\mathrm{mom}}=\frac{1}{1+6\alpha_G}\left(\psi(x)+\alpha_G\sum_j U_j(x)e^{i\xi\mathbf{P}\cdot\hat{\mathbf{j}}}\psi(x+\hat{j})\right), \tag{11}$$

where U_j denotes a gauge link in the j direction. The value of the free momentum smearing parameter ξ depends, in general, on the parameters of the gauge ensemble and on the nucleon momentum. The tuning of ξ is necessary in order to improve the overlap of our interpolator with the boosted proton. We, thus, have optimized the parameter ξ for each value of the momentum employed by minimizing the statistical errors of the nucleon two-point functions. In Fig. 2, we demonstrate the effect of the momentum smearing, by plotting the scaling of the error of the two-point correlator and the effective energy. The results shown have been extracted using 100 measurements for the nucleon boost $P_3=8\pi/L$. For comparison, we include also the results using the standard Gaussian smearing ($\xi=0$). As can be seen, the errors in the correlation functions reduce dramatically as the value of ξ increases and convergence is observed in the range $\xi\in[0.6\text{–}0.75]$. Thus, any value of ξ in this window of values leads to a similar signal-to-noise ratio. The tuning procedure for the other two momenta, $P_3=6\pi/L$ and $P_3=10\pi/L$, leads to the same conclusions and we fix $\xi=0.6$ throughout this work.

The three-point functions are computed using the sequential method [94], which has the advantage of summing automatically the sink spatial volume to produce the sequential propagator for all insertion points. Compared to the stochastic method [95], the sequential method does not introduce stochastic noise, but has the drawback that new inversions of the Dirac matrix have to be carried out for different source-sink separations and nucleon momenta. For details on the comparison between the stochastic method (with pure Gaussian smearing) and the sequential method (with and without momentum smearing), we refer to Refs. [15,18]. In addition, the sequential method is preferable when the momentum smearing technique on the quark fields is employed.

To increase the number of measurements of the three-point functions, we use multiple source positions (N_{src}) for each configuration. To confirm that the data extracted from different source positions on the same configuration are statistically independent, we study the scaling of the statistical error with the increase of N_{src}. In Fig. 3 (left), we show the error on the matrix elements for $z=0$,

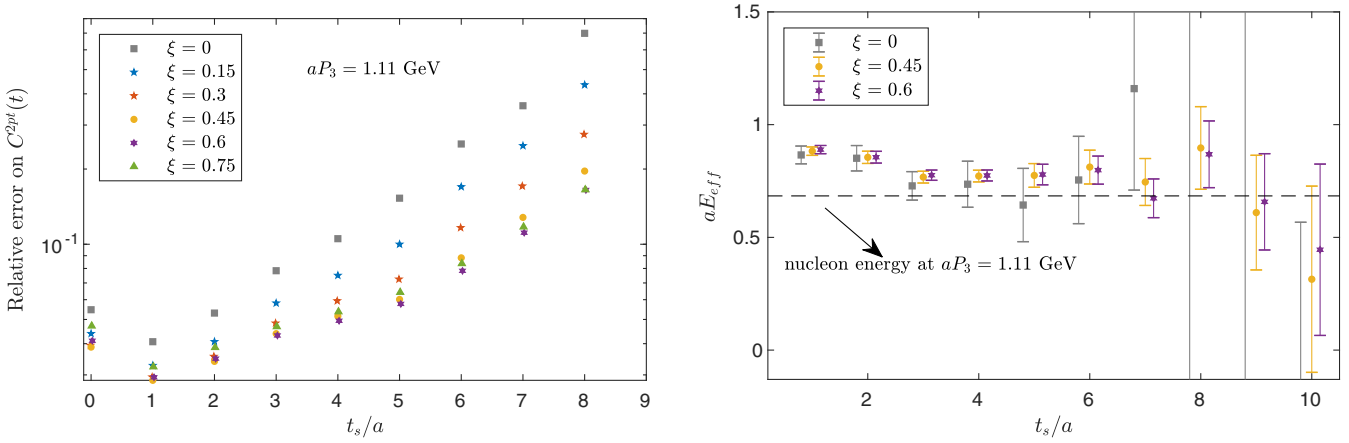

FIG. 2. Left: Relative error of the nucleon two-point correlator as a function of the Euclidean time t/a, for $P_3 = 8\pi/L$ and different values of the momentum smearing parameter ξ. The value $\xi = 0$ (black points) corresponds to the standard Gaussian smearing. Right: Effective energy of the nucleon with boost $P_3 = 8\pi/L$ for different values of ξ.

obtained from the analysis of a sample of gauge configurations for a nucleon boost of $6\pi/L$. For comparison purposes, the absolute error is normalized to the one obtained using only one source position per configuration. As can be seen, the errors for the three choices of Γ structure decrease approximately as $1/\sqrt{N_{\rm src}}$, which is the scaling expected if the measurements are not correlated.

To further decrease the statistical uncertainties, we boost the nucleon along different spatial directions and orientations, that is, $\pm x$, $\pm y$ and $\pm z$, with the Wilson line taken always along the axis in which the spatial component of the momentum is nonzero. The correlation functions obtained from these $N_{\rm dir} = 6$ possible directions lead to the same physical results due to the spatial rotational symmetry on the lattice, and therefore can be averaged within each configuration. The statistical error reduction is close to the ideal behavior $1/\sqrt{N_{\rm dir}}$, as shown in the right panel of Fig. 3.

Despite the use of the momentum and APE smearing, the exponential decrease of the signal-to-noise ratio persists as the nucleon momentum and source-sink separation increase. To increase statistics at a reduced cost, we employ, together with the momentum and APE smearing, the covariant approximation averaging (CAA) [96] that belongs to the class of truncated solver methods with a bias correction. For each configuration, $N_{\rm LP}$ low-precision (LP) inversions of the Dirac matrix are carried out from a set of random source positions $\{X_{\rm src}\}$ and the bias from the measurements is removed using a small number $N_{\rm HP}$ of high-precision (HP) inversions. Denoting with $C_{\rm LP}$ and $C_{\rm HP}$ the correlation functions produced with low and high precision of the solver, respectively, the improved

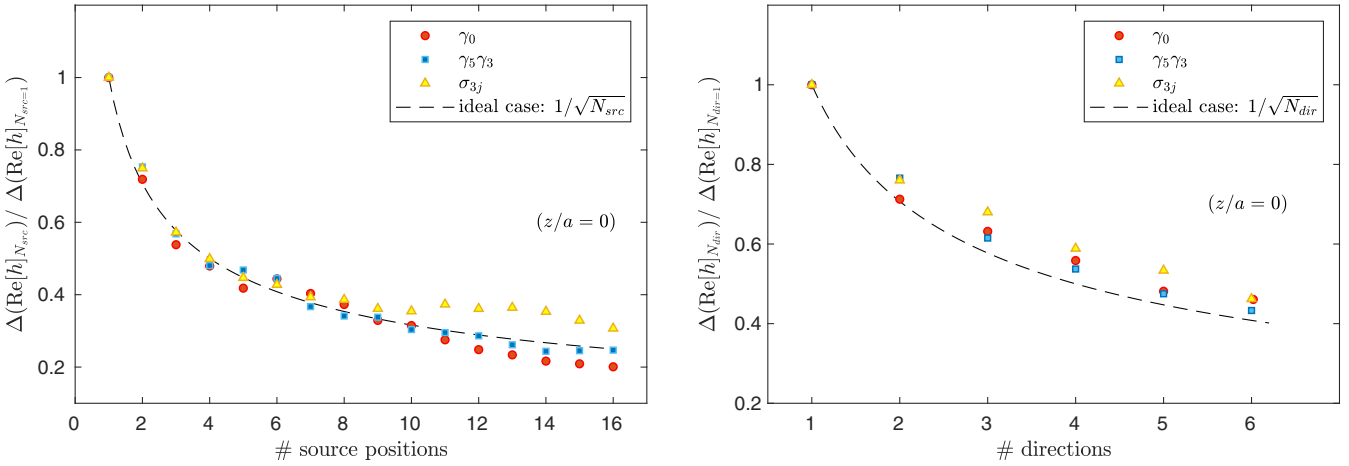

FIG. 3. Scaling of the statistical errors of the matrix elements, for the operator with $z = 0$, varying the number of source positions on each configuration (left) and averaging over the directions of the nucleon boost (right). The absolute errors for the unpolarized, helicity and transversity matrix elements (red circles, blue squares, and yellow triangles, respectively) are normalized to the one obtained from only one point source or one boost direction. The results are compared with the ideal scaling (dashed line in both figures) expected for uncorrelated measurements.

correlation functions C^{imp} for each configuration are defined by

$$C^{\text{imp}} = \frac{1}{N_{\text{LP}}}\sum_{i=1}^{N_{\text{LP}}} C_{i,LP} + \frac{1}{N_{\text{HP}}}\sum_{i=1}^{N_{\text{HP}}}(C_{i,HP} - C_{i,LP}), \quad (12)$$

where, in the second sum, C_{LP} and C_{HP} are computed at the same source position, as otherwise the bias cannot be corrected. The error for a given observable scales with the ratio $N_{\text{HP}}/N_{\text{LP}}$ as

$$\text{error} \propto \sqrt{2(1-r_c) + \frac{N_{\text{HP}}}{N_{\text{LP}}}}, \quad (13)$$

where r_c is the correlation coefficient among nucleon correlators computed at high and low precision. A compromise is needed to have $r_c \simeq 1$, while keeping the inversions as fast as possible and $N_{\text{LP}} \gg N_{\text{HP}}$. Thus, a tuning of the precision of the solver has to be carried out. To invert the Dirac operator, we use the adaptive multigrid solver with twisted mass fermion support [97] and require the residual to be $r_{\text{HP}} = 10^{-10}$ for HP inversions. After testing different values of the residual for LP inversions, we find that the stopping criterion

$$r_{\text{LP}} \equiv \frac{|\text{residue}|_{\text{LP}}}{|\text{source}|} = 2 \times 10^{-3} \quad (14)$$

guarantees a correlation coefficient $r_c \geq 0.999$ with a considerable speed-up in the inversion time. Moreover, taking 15 HP inversions as the reference setup, a comparison of the HP and CAA estimates for the two- and three-point correlators verified that the bias introduced from LP inversions is negligible compared to the gauge noise of our measurements when one HP inversion for each configuration is performed. Thus, to extract the nucleon matrix elements for quasi-PDFs at momenta $8\pi/L$ and $10\pi/L$, we use the CAA setup with $(N_{\text{HP}}, N_{\text{LP}}) = (1, 16)$.

The quasi-PDFs are computed at different source-sink time separations t_s in order to investigate excited-states effects. In particular, we use $t_s/a = 8$, 9, 10, 12 for the unpolarized and $t_s/a = 8$, 10, 12 for the helicity and transversity cases. For a detailed discussion on the excited-states effects, we refer to Sec. III C. In the remaining part of this section and unless otherwise stated, we focus on the results extracted from $t_s = 12a$, which is the one where excited states are found to be suppressed for all three PDFs when the statistical precision is around 10%. The number of configurations and the total statistics for each operator and momentum are reported in Table II. In the number of total measurements, we include a factor of 6, coming from the average over correlators computed from the boost aligned along the $\pm x$, $\pm y$, $\pm z$ directions, and a factor of 16 (15) from the source positions for $P_3 = 6\pi/L$ ($P_3 = [8, 10]\pi/L$). This translates into an additional factor of 96 for $P_3 = 6\pi/L$ and 90 for $P_3 = 8, 10\pi/L$, where the CAA is employed. Moreover, we note that only for the transversity PDF at $P_3 = 6\pi/L$, we have averaged over the matrix elements computed for the two possible choices of Dirac matrices (for example, σ^{31} and σ^{32} for a boost in the z direction). This contributes an additional factor of 2 to the total number of measurements at this momentum, as well as to the computational cost (due to the different parity projectors needed for σ^{31} and σ^{32}).

Although large nucleon momenta are needed to approach light-cone PDFs, high values of the momentum on a Euclidean lattice may lead to substantial cutoff effects if the condition $P_3 \ll 1/a$ is not satisfied. One possible check of cutoff effects is via the computation of the dispersion relation. To this end, we compute the nucleon energies for the momenta $\{0, 2, 4, 6, 8, 10\}\pi/L$ using the momentum smearing method, and check whether the relativistic dispersion relation $E^2 = m_N^2c^4 + P_3^2c^2$ is satisfied for our results. As can be seen in Fig. 4, no deviations from the continuum energy-momentum relation are observed for the values employed in this work (up to 1.38 GeV), which is well below the inverse lattice spacing ($10\pi/L \approx 0.65/a$). Moreover, by performing a two-parameter fit to the lattice data, we obtain $c^2 = 1.00(3)$ for the squared speed of light and $a^2m_N^2c^4 = 0.207(4)$ for the nucleon mass in lattice units, which is compatible with the value

TABLE II. The statistics of our calculation at a source-sink separation of $t_s = 12a$, for each current insertion and each momentum. N_{conf} is the number of gauge configurations, N_{HP} (N_{LP}) is the number of high- (low-)precision measurements, and N_{meas} is the total number of measurements. A factor of 6 is included in the total measurements due to the averaging of data with the Wilson line and momentum boost in the $\pm x$, $\pm y$, $\pm z$ directions. An additional factor of 2 compared to all other cases is included for the transversity at $P_3 = \frac{6\pi}{L}$ to take into account the averages over the two possible Dirac structures, as explained in the text.

$P_3 = \frac{6\pi}{L}$					$P_3 = \frac{8\pi}{L}$				$P_3 = \frac{10\pi}{L}$				
Ins.	N_{conf}	N_{HP}	N_{meas}	Ins.	N_{conf}	N_{HP}	N_{LP}	N_{meas}	Ins.	N_{conf}	N_{HP}	N_{LP}	N_{meas}
γ^3	100	16	9600	γ^3	425	1	16	38 250	γ^3	811	1	16	72 990
γ^0	50	16	4800	γ^0	425	1	16	38 250	γ^0	811	1	16	72 990
$\gamma^5\gamma^3$	65	16	6240	$\gamma^5\gamma^3$	425	1	16	38 250	$\gamma^5\gamma^3$	811	1	16	72 990
σ^{3j}	50	16	9600	σ^{3j}	425	1	16	38 250	σ^{3j}	811	1	16	72 990

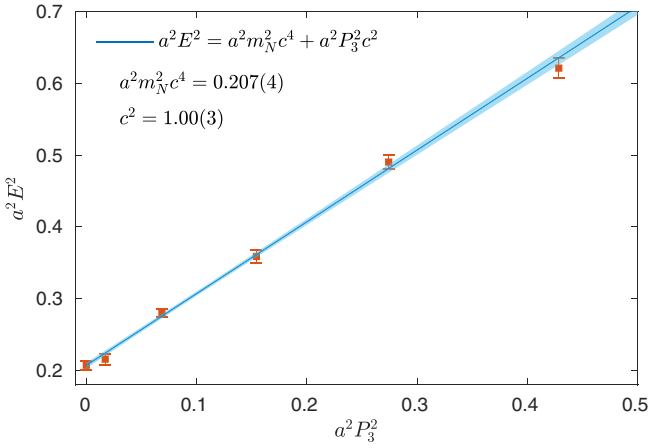

FIG. 4. Nucleon energy as a function of the momentum in lattice units. From the fit of the continuum dispersion relation, $E^2 = m_N^2c^4 + P_3^2c^2$, to the lattice data (red squares), we find a value of the squared speed of light equal to $c^2 = 1.00(3)$, and the nucleon mass in lattice units $a^2m_N^2c^4 = 0.207(4)$, which is also in agreement with the estimate given in Ref. [79].

$am_N = 0.4436(11)$ extracted at zero momentum [79]. Whether this finding of only small cutoff effects in the dispersion relation also holds in the case of the PDFs we are interested in, can eventually only be answered when results at several lattice spacings are available.

III. TECHNIQUES FOR THE EVALUATION OF BARE QUASI-PDF NUCLEON MATRIX ELEMENTS IN LATTICE QCD

In this section, we discuss crucial aspects related to the lattice QCD computation of bare quasi-PDF nucleon matrix elements, such as dependence on the number of stout smearing iterations used in the operator, the choice of the Dirac structure and identification of excited-states contamination.

A. Stout smearing

We apply three-dimensional stout smearing to the gauge links of the Wilson line of the operator, following the prescription of Ref. [98]. This reduces the power divergence that is present in the matrix elements of fermion operators with Wilson lines (see, e.g., Refs. [99,100], as well as Ref. [33] for a review of recent investigations of the power divergence). The application of smearing was especially crucial in the first calculations of quasi-PDFs [14,16–18] when the renormalization procedure was not yet developed. Even though the complete renormalization procedure was developed recently [101], a few iterations of stout smearing are still useful for noise reduction in the renormalized matrix elements. In Fig. 5, we show examples of the effect of stout smearing for the case of the bare matrix elements of the unpolarized (insertion γ^0), helicity and transversity quasi-PDFs for a nucleon momentum of $P_3 = 8\pi/L$ without stout smearing, using 0, 5, 10 and 15 stout steps. As expected, the stout smearing modifies the matrix elements, increasing the values of the real and imaginary parts at each z. We find convergence of the matrix elements after a few iterations of stout smearing and, thus, in what follows we discuss in detail results obtained with 5 stout steps. The renormalized matrix elements are expected to no longer show any dependence on the smearing. We discuss the details of the renormalization in Sec. IV.

B. Choice of the Dirac structure

Although the natural choice to extract the unpolarized PDF would be $\gamma^+ = (\gamma^0 + \gamma^3)/\sqrt{2}$, a recent study [21] has revealed that the matrix element with γ^3 exhibits mixing with the twist-3 scalar operator (later confirmed by symmetry properties [23]) and the computation of a mixing renormalization matrix is required [101]. However, the twisted mass formulation has the advantage that the mixing is between the vector and pseudoscalar operators. The latter vanishes in the continuum limit and only contributes as a discretization effect that increases the gauge noise. Consequently, it may be neglected in the first approximation, since eventually one is interested in the results for $a \to 0$. In this work, we compute both matrix elements, but use the data for γ^0 to extract our final results for the unpolarized quasi-PDF. Even though the mixing is a purely lattice artifact for twisted mass fermions, we expect increased noise contamination for the case of γ^3, which is clearly visible in the lattice data, as shown in Fig. 6. This behavior is momentum independent and we compare the matrix elements for the two γ structures at a momentum of $8\pi/L$. The statistical errors for γ^3 are twice as large compared to the ones for γ^0 with the same number of measurements. We, thus, conclude that the insertions γ^3 or γ^+ are not optimal for extracting the unpolarized PDFs within a lattice QCD calculation.

We compute the matrix elements for the three values of the nucleon momentum given in Table II with the associated number of measurements. The momentum dependence of the resulting matrix elements is shown in Fig. 7 for the three operators considered in this work. We find that with increasing momentum, the matrix elements decay to zero faster and, for the highest momentum employed, both the real and imaginary parts are compatible with zero for $z \geq 11a \simeq 1.03$ fm.

C. Excited states contamination

Identification of the nucleon state is crucial in order to extract the correct nucleon matrix elements from lattice QCD measurements. This requires a careful analysis of excited states. An additional challenge is the need to boost the nucleon with a relatively large momentum, something that it is not needed in e.g., studies of nucleon charges and

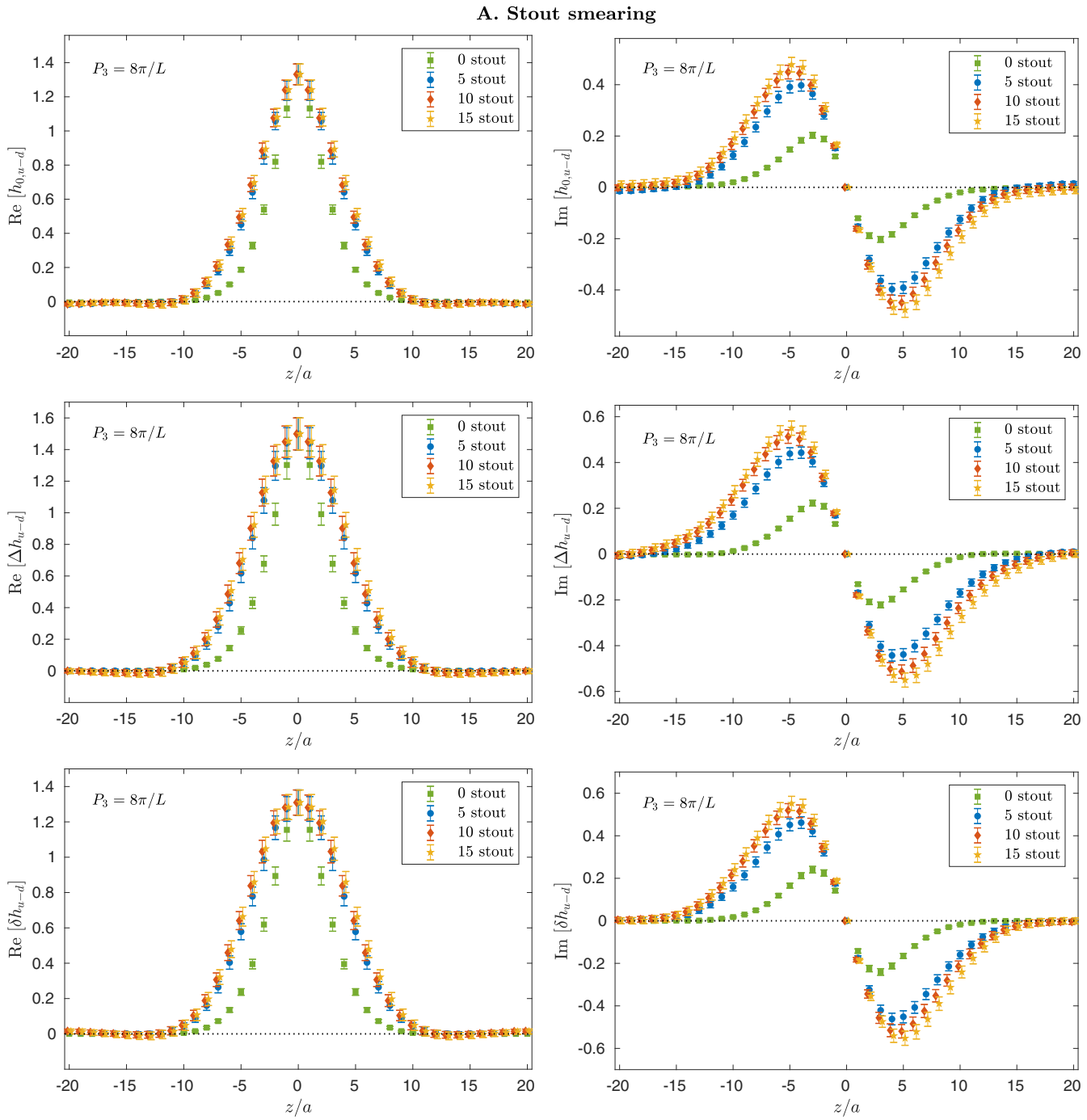

FIG. 5. Real and imaginary parts of the bare matrix elements for 0, 5, 10, and 15 steps of stout smearing at a nucleon momentum of $P_3 = 8\pi/L$. From top to bottom: The bare nucleon matrix element for the unpolarized, helicity and transversity operators.

form factors. Requiring a large nucleon momentum in combination with using simulations at the physical point leads to a more severe contamination due to the excited states, since the spectrum is denser. Therefore, a thorough assessment of the excited states is even more essential for the reliable extraction of PDFs. We extend the study of excited states first presented in Ref. [102] in order to eliminate, as much as possible, one of the major systematic uncertainties in the evaluation of nucleon matrix elements from which the PDFs are extracted within lattice QCD. In what follows, we review the analysis methods that we employ to isolate the ground state. As will be discussed in this subsection, the excited-states contributions are milder for the matrix elements of the unpolarized operator as compared to the matrix elements of the helicity and transversity operators.

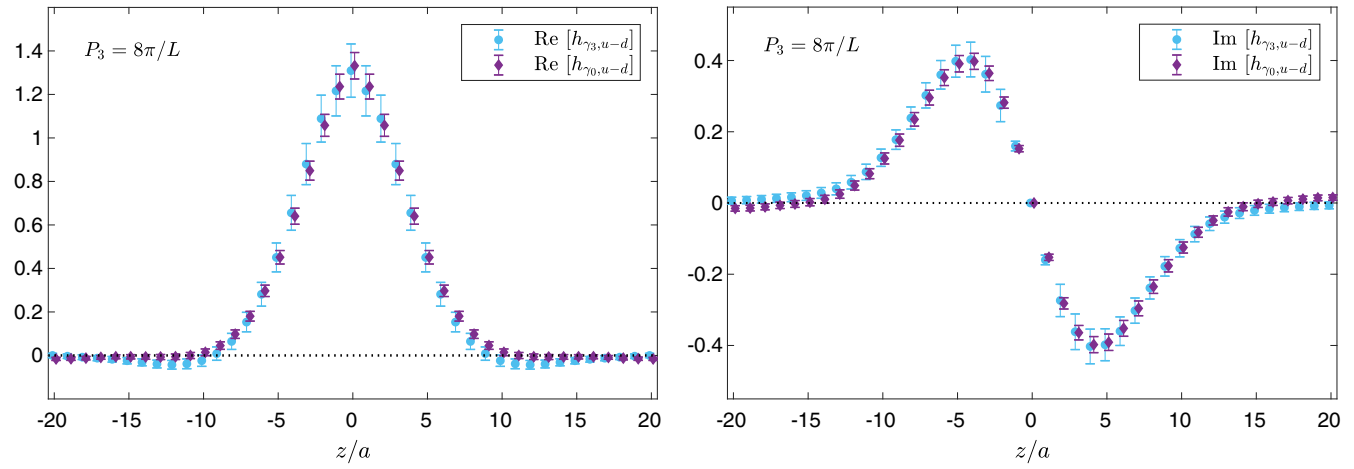

FIG. 6. Comparison of the spatial γ^3 (cyan circles) and temporal γ^0 (purple diamonds) unpolarized bare matrix elements for 5 stout steps and momentum $8\pi/L$.

To extract the ground-state contribution in the correlation functions, we employ three analysis methods, briefly described below.

(1) *Plateau method (single-state fit)*: In this method, one looks for a region where the ratio of Eq. (10) becomes independent of the insertion time, τ, which indicates a (possibly partial) suppression of excited states. The ratio in this region is fitted to a constant value, yielding the matrix element of the ground state. Indeed, inserting into the two- and three-point functions two sets of complete eigenstates of the QCD Hamiltonian ($|n(P_3)\rangle$ and $|n'(P_3)\rangle$, where we indicated that the states are momentum dependent, but we keep this dependence implicit below to simplify the notation) with quantum numbers of the nucleon, the ratio of Eq. (10) can be written as[3]

$$\frac{C^{3\mathrm{pt}}(P_3,z;t_s,\tau)}{C^{2\mathrm{pt}}(P_3;t_s)} = \frac{\sum_{n,n'}\langle N|n'\rangle\langle n|N\rangle\langle n'|\mathcal{O}|n\rangle e^{-E_{n'}(t_s-\tau)}e^{-E_n\tau}}{\sum_n|\langle N|n\rangle|^2 e^{-E_n t_s}}, \tag{15}$$

where $|N\rangle$ is the nucleon state, $E_{n,n'}$ are energies of the eigenstates $|n\rangle$, $|n'\rangle$ (where $|0\rangle$ is the ground state, $|1\rangle$ is the first excited state, etc.) and the time slice of the source is set to zero. Isolating in the ratio the contribution of the ground state and expanding the sum up to the first excited state, Eq. (15) reads

$$\frac{C^{3\mathrm{pt}}(P_3,z;t_s,\tau)}{C^{2\mathrm{pt}}(P_3;t_s)} = \frac{\langle 0|\mathcal{O}|0\rangle + f_{10}\langle 1|\mathcal{O}|0\rangle e^{-\Delta E(t_s-\tau)} + f_{10}^{\dagger}\langle 0|\mathcal{O}|1\rangle e^{-\Delta E\tau} + |f_{10}|^2\langle 1|\mathcal{O}|1\rangle e^{-\Delta E t_s} + \ldots}{1 + |f_{10}|^2 e^{-\Delta E t_s} + \ldots}, \tag{16}$$

where we introduced the constant $f_{10} = \langle N|1\rangle/\langle N|0\rangle$ and $\Delta E = E_1 - E_0$. As can be seen, the excited-states contributions fall off exponentially and for $\Delta E t_s \gg 0$, $\Delta E\tau \gg 0$ and $\Delta E(t_s-\tau) \gg 0$, the first time-independent term dominates, yielding the matrix element of the nucleon state, $\langle 0|\mathcal{O}|0\rangle$. Thus, fitting the ratio within the plateau region to a constant value yields the desired nucleon matrix element. Since statistical errors grow exponentially with t_s, the challenge is to identify the smallest value of t_s that suppresses the contributions of excited states in the ratio to a level negligible as compared to the statistical accuracy.

(2) *Summation method*: This method was introduced in Ref. [103] and entails the sum of the ratio of Eq. (16) over the insertion time τ, excluding the time slices of the source and the sink. The sum is a geometric series in the terms involving the excited-states contributions and yields the expression

$$\mathcal{R}(P_3,z;t_s) \equiv \sum_{\tau=a}^{t_s-a}\frac{C^{3\mathrm{pt}}(P_3,z;t_s,\tau)}{C^{2\mathrm{pt}}(P_3;t_s)} = C + \mathcal{M}t_s + \mathcal{O}(e^{-\Delta E t_s}), \tag{17}$$

where C is a constant and the desired matrix element $\mathcal{M} = \langle 0|\mathcal{O}|0\rangle$ can be extracted from a linear two-parameter fit. As a consequence, this procedure

[3]All correlation functions in this section should be understood as averaged over the gauge field configurations ensemble.

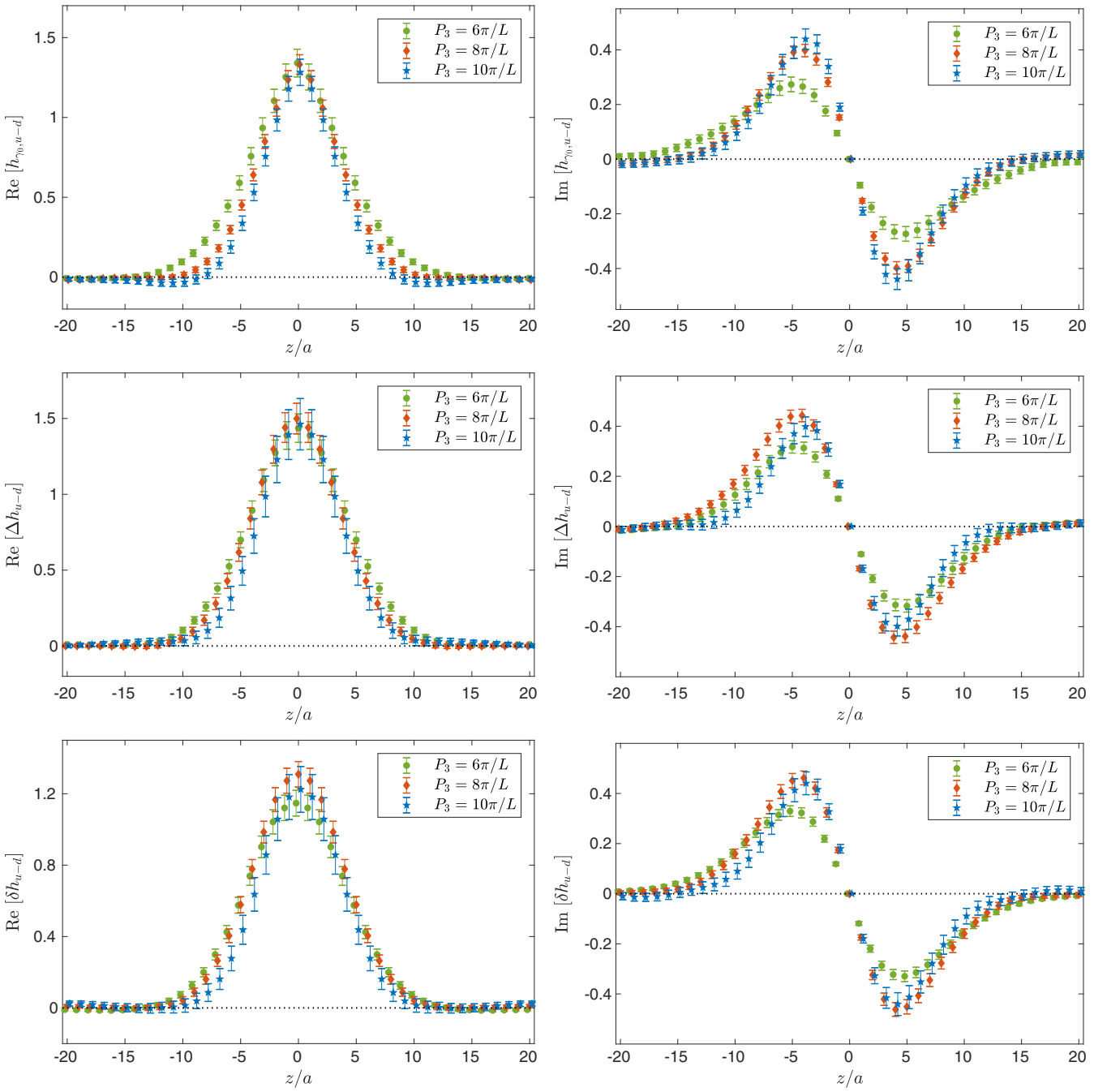

FIG. 7. Bare matrix elements for momenta $6\pi/L$ (green circles), $8\pi/L$ (red diamonds), and $10\pi/L$ (blue stars), extracted at a source-sink separation of $t_s = 12a \simeq 1.13$ fm. From top to bottom: Matrix elements of the unpolarized, helicity and transversity operators.

yields, in general, results with larger uncertainties as compared to the plateau method, but has the advantage of suppressing the excited-states contamination by a faster-decaying factor $e^{-\Delta E t_s}$, as compared to the leading exponential factors in Eq. (16).

(3) *Two-state fits*: In this method, one retains the first excited-states contributions in the two-point and three-point correlation functions. We use two types of fits: (a) a fit to the two-point correlator followed by a fit to the three-point function, and (b) a simultaneous fit to both the two- and three-point correlators.

We refer to the type (a) fit as a sequential fit and parametrize the nucleon two-point function with momentum P_3 as

$$C^{\text{2pt}}(P_3;t) = |A_0|^2 e^{-E_0 t}(1 + |f_{10}|^2 e^{-\Delta E t}), \quad (18)$$

with the fit parameters being the ground-state amplitude, $|A_0|^2 = |\langle N|0\rangle|^2$, the ground-state energy E_0, $|f_{10}|^2$ and ΔE. The three-point correlator can be written as

$$\begin{aligned} C^{3\mathrm{pt}}(P_3, z; t_s, \tau) = |A_0|^2 e^{-E_0 t_s}(&\langle 0|\mathcal{O}|0\rangle \\ &+ f_{10}\langle 1|\mathcal{O}|0\rangle e^{-\Delta E(t_s-\tau)} \\ &+ f_{10}^{\dagger}\langle 0|\mathcal{O}|1\rangle e^{-\Delta E\tau} \\ &+ |f_{10}|^2\langle 1|\mathcal{O}|1\rangle e^{-\Delta E t_s}). \end{aligned} \tag{19}$$

The amplitudes $|A_0|^2$ and $|f_{10}|^2$ and energies E_0 and ΔE are determined from first fitting the two-point function in Eq. (18) and used as inputs into the three-point function, which is fitted separately for the real and imaginary parts. The fit parameters are $\mathrm{Re}\langle 0|\mathcal{O}|0\rangle$, $\mathrm{Re}\tilde{f}_{10} \equiv \mathrm{Re} f_{10}\langle 1|\mathcal{O}|0\rangle$, $\mathrm{Re}\tilde{f}_{01} \equiv \mathrm{Re} f_{10}^{\dagger}\langle 0|\mathcal{O}|1\rangle$, $\mathrm{Re}\langle 1|\mathcal{O}|1\rangle$ and $\mathrm{Im}\langle 0|\mathcal{O}|0\rangle$, $\mathrm{Im}\tilde{f}_{10}$, $\mathrm{Im}\tilde{f}_{01}$, $\mathrm{Im}\langle 1|\mathcal{O}|1\rangle$, respectively. Through this fitting procedure that entails four fit parameters for the two-point function fit and an additional four parameters for the three-point function, we extract the desired matrix element, $\langle 0|\mathcal{O}|0\rangle$.

The call the type (b) fit simultaneous and it is a combined fit to Eqs. (18) and (19), using all eight parameters. We expect that both fit types will lead to very similar results. Proper analysis requires the evaluation of two- and three-point correlation functions on exactly the same gauge field configurations and for the same set of source positions. Only in such a case can the correlations among the fit parameters be probed in a fully consistent manner. In practice, this is a severe problem for the observables we are interested in, which is due to the exponential increase of the noise with increasing source-sink time separation. When the time separation grows by one lattice spacing, the necessary statistics to suppress noise to the same level as before increases by a factor of 2–3 in our data. Thus, small source-sink time separations yield a precise signal already for $\mathcal{O}(10^3)$ measurements, while correlators at $t_s = 12a$ require $\mathcal{O}(10^5)$ measurements to reach a similar precision. For a given amount of computational resources, one, therefore, has to make a choice between using the same statistics for all t_s values and thus obtaining very precise data at small values of t_s and considerably less precise ones for larger values of t_s, or using different statistics and achieving similar precision for all values of t_s. The severe drawback of the former choice is that the two-state fits are dominated by the precise results at small t_s that can lead to biased results for the ground-state matrix element. This bias becomes more severe as the boost increases due to the denser spectrum prohibiting the robust identification of the effects of excited states.

1. Two-point correlator

Before we compare different analysis methods for the extraction of quasi-PDF matrix elements, we discuss the two-state fit of the two-point correlator at our largest momentum, using Eq. (18). We obtain the boosted nucleon ground-state energy of $aE_0^{\text{2-state}} = 0.773(18)$, which is in line with a single-state fit, leading to a value $aE_0^{\text{1-state}} = 0.788(9)$. These values are illustrated in the left panel of Fig. 8. As we showed in Fig. 4, the obtained ground-state energy satisfies the continuum dispersion relation. The energy difference between the ground state and the first excited state is $a\Delta E = 0.36(7)$, which in physical units is 0.76(15) GeV. This can be compared with the expectation from the approximation of noninteracting stable hadrons in a box, which leads, at our volume and the physical pion mass ($m_\pi L \approx 3$), to the following splittings (for $N\pi\pi$, $N\pi$, $N\pi\pi\pi\pi$, $N\pi\pi\pi$) with respect to the ground state (N): ≈0.27, 0.36, 0.54, and 0.63 GeV, respectively. The above values hold for a nucleon at rest, whereas the spectrum becomes denser with increasing nucleon momentum. In the interacting case, the spectrum obviously changes, but its general features are unchanged; for a comprehensive discussion on the spectrum of excitations, we refer to the recent review of Ref. [81].

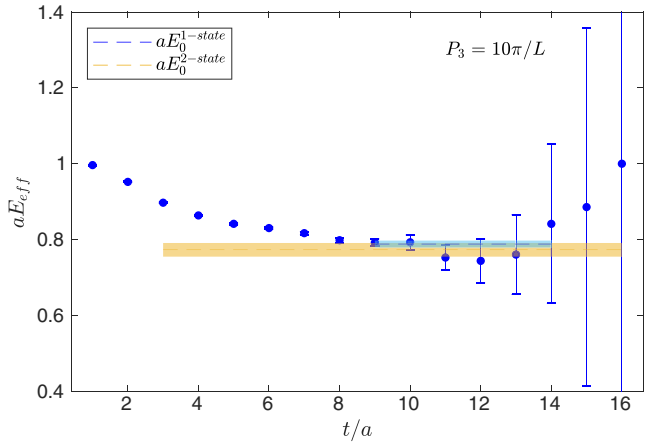

FIG. 8. Effective mass together with the extracted values of the ground-state energy from one-state (blue band) and two-state fits (orange band). The length of the bands indicates the fit range. The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

The energy extracted for the first excited state does not match the energy expected for the two- and three-particle states lying below. Furthermore, it is known that a two-state fit will model, in one exponential, all the states above the ground state that have an overlap with the interpolating field. Thus, one expects a bias in the determination of the first excited state. Consequently, the conclusions from two-state fits need to be checked against the plateau values or three-state fits, if the accuracy of the data allows for this. If consistency is found, this indicates ground-state dominance up to the achieved statistical precision.

2. Matrix elements of the unpolarized PDF

We first present the analysis of excited states in the nucleon matrix elements for the unpolarized operator using

TABLE III. Numbers of measurements of the matrix element for the unpolarized operator (γ^0) at each source-sink separation and $P_3 = 10\pi/L$.

$P_3 = 10\pi/L$	$t_s = 8a$	$t_s = 9a$	$t_s = 10a$	$t_s = 12a$
N_{conf}	48	98	100	811
N_{meas}	4320	8820	9000	72 990

the γ^0 structure and the largest momentum, $P_3 = 10\pi/L$, where excited-states effects are expected to be most severe. We use four values of the source-sink time separation, namely $t_s/a = 8$, 9, 10, 12 or in physical units $t_s \simeq 0.75$, 0.84, 0.94, 1.13 fm. We opt to increase the number of measurements as we increase t_s, to have statistical errors that are approximately the same for consequent values of t_s. This enables us to perform a reliable analysis at each value of t_s. We use the same CAA setup for all separations, as explained above. In Table III, we collect the statistics for each value of t_s.

The plateau method is employed to analyze the data from each t_s value and the results are shown in Fig. 9. For $t_s = 8a$, the real part of the matrix element shifts towards smaller values at each z/a compared to larger t_s, while the effect in the imaginary part is less prominent. Within statistical uncertainties, a convergence between the results at $t_s = 10a$ and $t_s = 12a$ is observed in both the real and imaginary parts. Comparing the data for the four t_s values, we observe signs of a nonmonotonic behavior that affects the real and imaginary parts differently, depending on the value of z. This can introduce a complicated effect in the determination of the PDF. Ideally, one must compute the matrix elements for several large enough values of t_s with equally small statistical errors and demonstrate convergence to one value. However, the exponential increase of the noise-to-signal ratio seen in Fig. 2 and the need for new sets of sequential inversions for each t_s, require computational resources beyond what is currently available, placing limitations on the maximum value of t_s. Nevertheless, if consistency can be found between the converged plateau values of the single-state fits and other methods, ground-state dominance can be reliably established.

We apply the summation method for two sets of t_s values, namely (a) $t_s = 8a$, $9a$, $10a$, $12a$ and (b) $t_s = 9a$, $10a$, $12a$, with the results shown in Fig. 10. The first set yields results that, although consistent with those from the second fit, are systematically higher. Furthermore, for large values of z/a, the slope obtained using the (b) set is significantly larger than that from the (a) set and $\chi^2/\text{d.o.f.} > 2$ for $z/a \approx 10$–15, indicating that the (a) fits do not provide a good description of the data. This is a consequence of the ratio at $t_s = 8a$ being considerably below the ones at other source-sink separations. Figure 11 shows examples of summation fits for our data. In the right panel, only the (b) fit yields an acceptable $\chi^2/\text{d.o.f.}$ of around 0.16, while the full fit to all t_s values has $\chi^2/\text{d.o.f.} \approx 2.4$. Thus, the fitting ansatz of Eq. (17) does not provide a correct description of the data and the large-z values in Fig. 10 for $t_s = [8a$–$12a]$ are not reliable. The effect is visible also for small z/a (see the left panel of Fig. 11 for $z/a = 0$ fits) and may result in overestimating the real part of the matrix elements. At $z/a = 0$, the extracted value of the matrix element is 1.21(16) for the (a) set and 1.02(24) for the (b) set, after applying in both cases the renormalization factor $Z_V = 0.7565$ [104]. We note that the matrix element should be equal to 1 upon renormalization, which is satisfied only by the set (b), while the value from the summation method that includes the smallest t_s in the fit is considerably larger. All of the above points lead to the conclusion that using too small source-sink separations gives incorrect results. We will thus omit $t_s = 8a$ in our summation method analysis, and take fits (b) as our final estimates from the summation method. All of these fits, both in the real and imaginary parts, have $\chi^2/\text{d.o.f.} \lesssim 1$, indicating that the exponential contributions

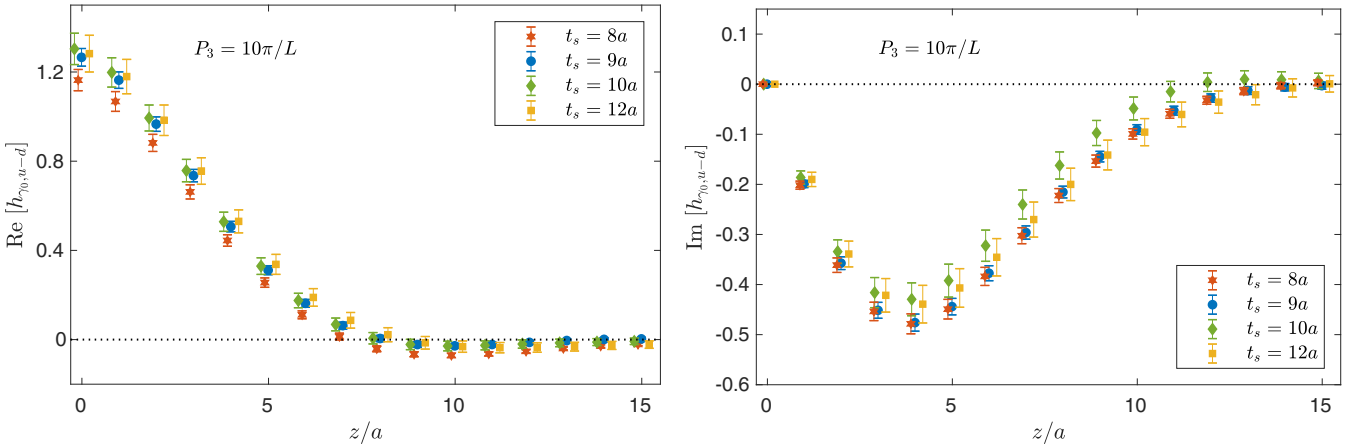

FIG. 9. Real (left) and imaginary (right) parts of the matrix element for the unpolarized PDF, at a source-sink separation of $8a$ (red stars), $9a$ (blue circles), $10a$ (green diamonds) and $12a$ (yellow squares). The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

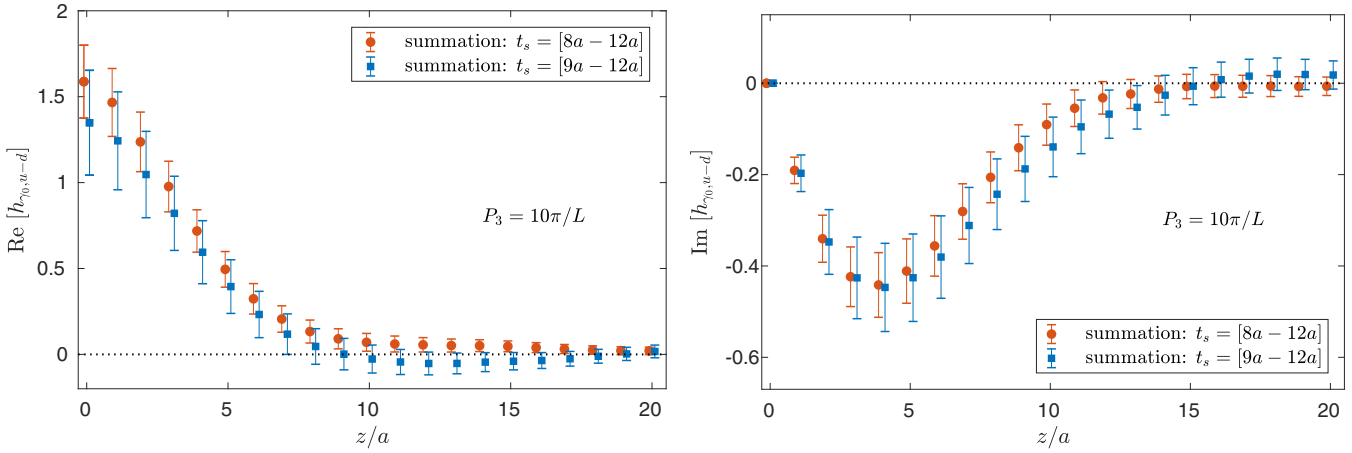

FIG. 10. Real (left) and imaginary (right) parts of the matrix element for the unpolarized PDF from the summation method, using source-sink separations $8a$–$12a$ (red circles) and $9a$–$12a$ (blue squares). The nucleon is boosted with momentum $10\pi/L \simeq 1.38$ GeV.

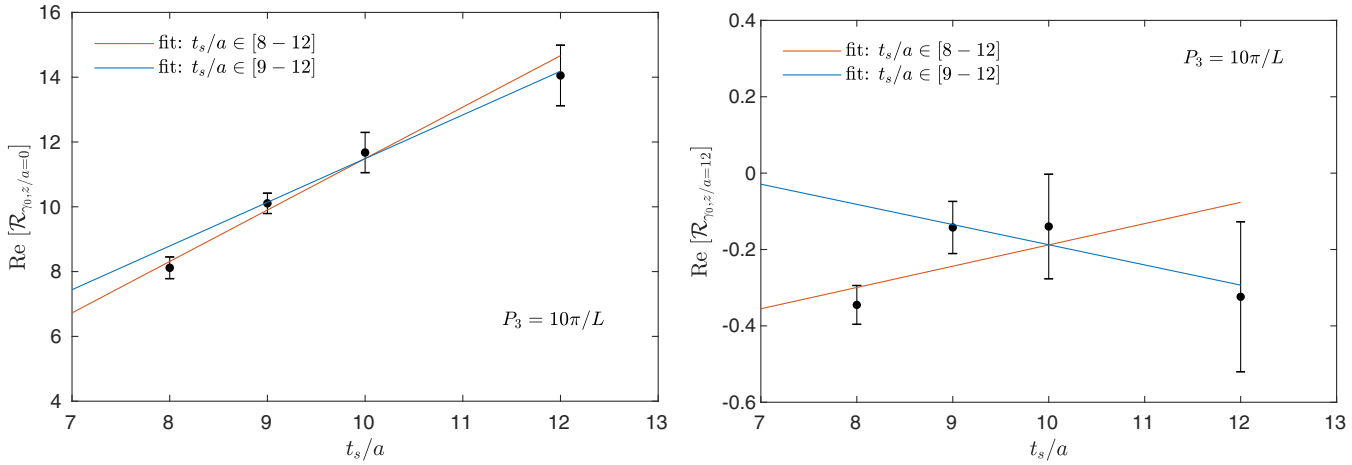

FIG. 11. Summation fits for the real part of the matrix element for the unpolarized PDF, for $z/a = 0$ (left) and $z/a = 12$ (right). In both cases, the (a) fits to all source-sink separations and the (b) fits to the three largest separations are shown. The nucleon boost is $10\pi/L \simeq 1.38$ GeV.

from higher excited states in Eq. (17) are sufficiently suppressed, yielding a good estimate of the ground-state matrix element $\mathcal{M}$. However, the precision of the summation method estimates is much worse than the one from the plateau or two-state fits (see below).

We present now the two-state fits and compare results from the sequential and simultaneous two-state fits. In addition, we check the robustness of these fits by using three or four values of t_s. The fits using only two t_s values lead to considerably larger errors (with compatible central values) and hence, we do not show them. The different choices of fits with three t_s values lead to very similar results in terms of the central values and their errors. Thus, in Fig. 12, we present results from one choice of three source-sink separations, namely $t_s = 8a, 9a, 10a$, and from the fit to all $t_s = 8a, 9a, 10a, 12a$. We conclude that the simultaneous and sequential fits lead to statistically identical results for both sets of t_s values. Including the results for the largest t_s yields consistent fits, but with errors that are slightly smaller (up to 10–15%) for most z values. This is visible particularly for the simultaneous fits at all z and for sequential fits at small z values in the real part. Note that the central values are essentially the same, suggesting that excited states are sufficiently suppressed.

In Fig. 13, we collect the results of the analyses using the three aforementioned procedures, for all z/a values, and in Fig. 14 we present part of the same data for two selected values of the Wilson line length, $z/a = 5$ and $z/a = 10$, for better visibility. Excited states effects are clearly visible for $t_s = 8a$, especially in the real part when using the plateau method. For $t_s = 9a$ and $10a$, we observe a small tension in the imaginary part: the plateau values are consistently lower for $t_s = 9a$ and consistently higher for $t_s = 10a$, with respect to the two-state fits. We note that, in particular, there

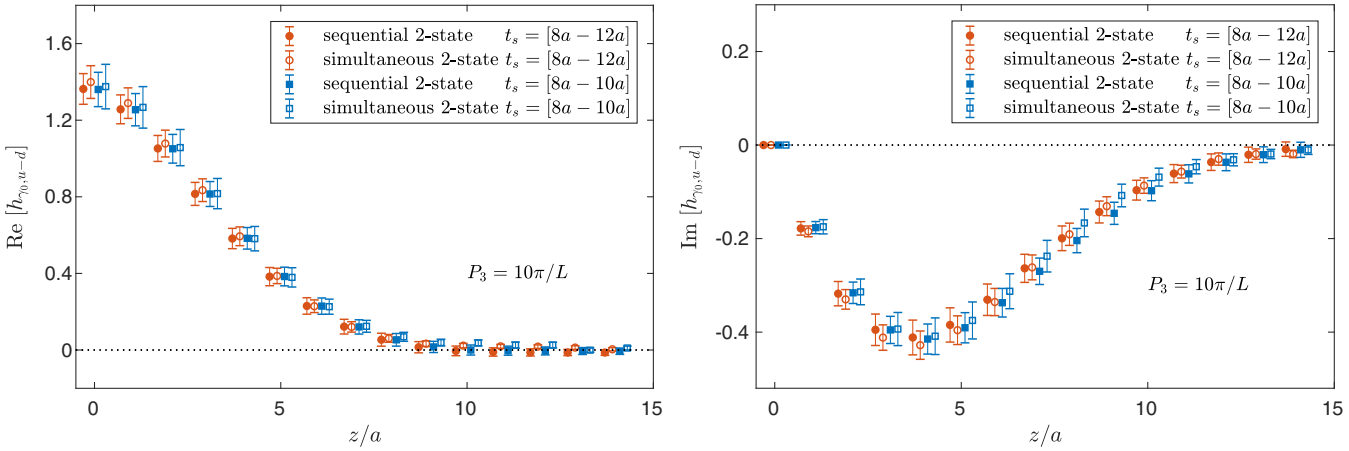

FIG. 12. Real (left) and imaginary (right) parts of the matrix element for the unpolarized PDF from two-state fits, using source-sink separations $8a$–$12a$ (circles) and $8a$–$10a$ (squares). The sequential fit (filled symbols) is also compared to the simultaneous fit (open symbols). The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

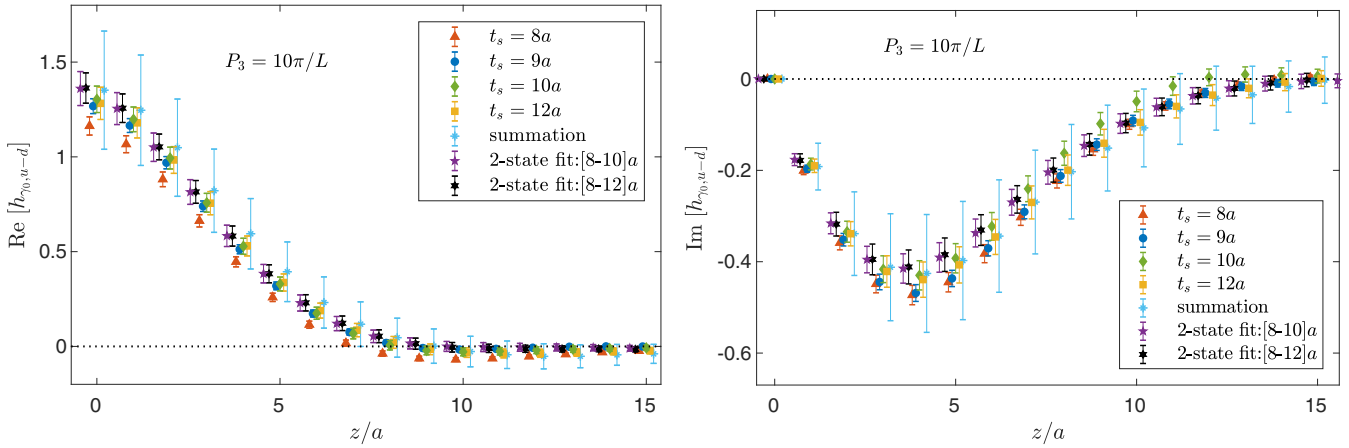

FIG. 13. Real (left) and imaginary (right) parts of the matrix element for the unpolarized PDF from the plateau method (points labeled with appropriate t_s), the summation method (using $t_s \geq 9a$) and sequential two-state fits (to $t_s = 8a$, $9a$, $10a$ and to all t_s). The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

is tension between the two-state fit using $t_s = 8a$–$10a$ and the plateau values for $t_s = 10a$, at $z/a \approx 10$, which fails to satisfy our criterion of consistency between two-state fits and plateau fits. Hence, data at source-sink separations below $12a$ are likely still contaminated by excited states, although statistical fluctuation as the source of the trend cannot be excluded. We find that the results extracted from the plateau method using $t_s = 10a$ and $t_s = 12a$ are in agreement with each other and the ones from $t_s = 12a$ are compatible with those extracted using the all-t_s two-state fits for all values of z/a and for both the real and imaginary parts. The errors in the summation method are too large to draw any meaningful conclusions. Given that this investigation is carried out for the largest value of the momentum, we can take as our final value for the nucleon matrix elements of the unpolarized operator the data at $t_s = 12a$ for all the momentum values. We note that increasing the nucleon boost would need a similarly thorough reanalysis of the effects of the excited states. Likewise, increased statistical precision could reveal excited-states contamination even at this nucleon momentum. Thus, we emphasize that the attained conclusion about ground-state dominance is valid only within the present statistical uncertainties of $\mathcal{O}(10)\%$.

3. Matrix elements of the helicity and transversity PDFs

In this subsection, we discuss the excited-states effects for the two other Dirac structures used in our work—axial and tensor—associated with the helicity and transversity PDFs, respectively. We perform a similar analysis as for the

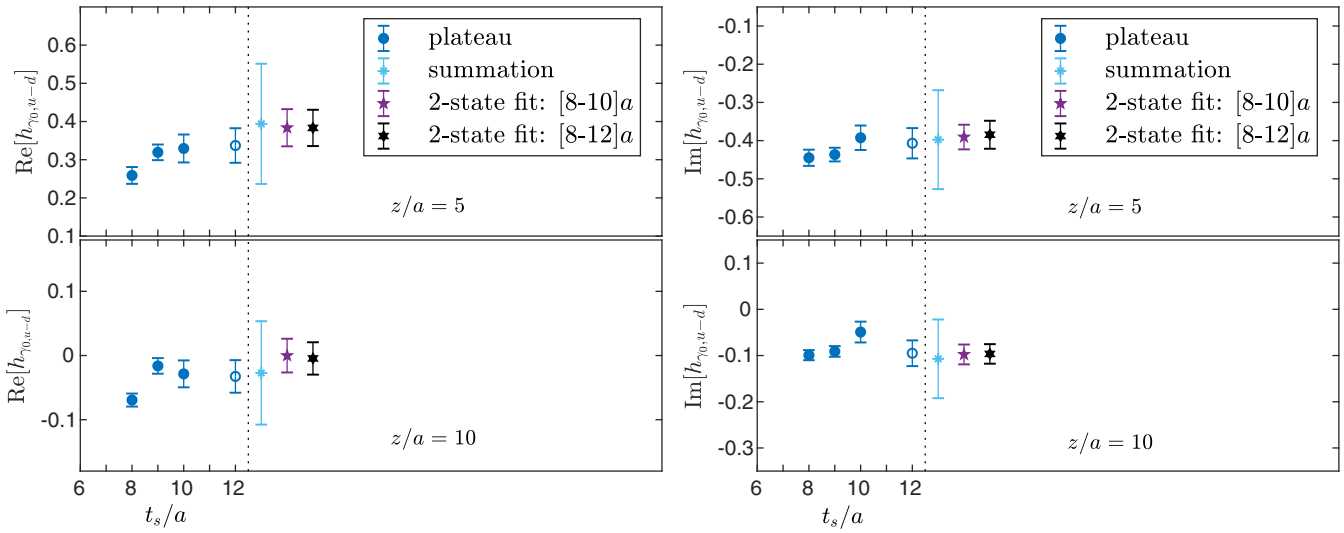

FIG. 14. Real (left) and imaginary (right) parts of the matrix element for the unpolarized PDF at fixed $z/a = 5$ (upper) and $z/a = 10$ (lower). We compare results from the plateau method, the summation method (using $t_s \geq 9a$) and sequential two-state fits (to $t_s = 8a$, $9a$, $10a$ and to all t_s). The final value that we use for extracting PDFs is indicated with an open symbol. The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

unpolarized case at the largest momentum, $P_3 = 10\pi/L$, and again use four values of the source-sink separation, namely $t_s/a = 8$, 9, 10, 12, or in physical units $t_s \simeq \{0.75, 0.84, 0.94, 1.13\}$ fm. The numbers of measurements are listed in Table IV. The methodology of our investigation is the same as for the unpolarized case, i.e., for each t_s value, we perform single-state fits within a plateau region, and we combine data at different t_s using the summation method approach, as well as two-state sequential and simultaneous fits.

In Figs. 15 and 16, we plot the full z dependence of bare matrix elements for helicity and transversity PDFs, respectively. Figure 17 displays a zoom-in for two selected values of z/a, i.e., $z/a = 5$ and $z/a = 10$. In this case, the fits using the summation method have good quality ($\chi^2/\text{d.o.f.} \lesssim 1$) even when including $t_s = 8a$ and hence, we use all source-sink separations for this approach. For the real part, the two-state fits and the summation method yield results compatible with plateau fits for $t_s = 12a$. The only observed tensions in the real part are rather small, possibly statistical fluctuations, and are visible for the plateau fits using $t_s/a = 8$, 9 versus the summation method, for intermediate and large z/a in the helicity case. However, for the imaginary part, we consistently observe a rather strong dependence of the plateau values extracted when using $t_s = 12a$ as compared to those from $t_s/a = 8$, 9, 10, with 2σ to 3σ tensions. The latter are also incompatible with the results extracted using two-state fits to all t_s and the summation method, both of which are consistent with the plateau results for $t_s = 12a$. Moreover, the two-state fits using $t_s/a = 8$, 9, 10 are incompatible with plateau fits for $t_s = 10a$ (see right panel of Fig. 15), which again violates our criterion for consistency, necessitating the use of $t_s = 12a$. This behavior reinforces the conclusion reached in the case of the bare matrix elements for the unpolarized quasi-PDF, namely that the tensions observed between results from source-sink separations below 1 fm are indeed manifestations of excited-states contamination and ground-state dominance is achieved, to around 10% statistical accuracy, only at $t_s = 12a$. Hence, to extract PDFs, we take the plateau results for $t_s = 12a$ as our preferred ones, since they are more precisely determined as compared to both the results extracted using the two-state and the summation approaches, but show full consistency with them. The more severe excited-states effects observed in the cases of helicity and transversity are in accordance with observations connected to the extraction of the nucleon axial and tensor charges, where excited-states contamination is more significant as compared to the case of the vector operator. Thus, studies that aim at a higher precision in the determination of quasi-PDF at increasing values of the nucleon boost would need large statistics to account for the increased errors resulting from having a nucleon state with large boost, but also for investigating excited states.

TABLE IV. Statistics used in the study of excited states for the polarized cases ($\gamma^5\gamma^3$: helicity; σ^{3j}: transversity) at $P_3 = 10\pi/L$.

$P_3 = 10\pi/L$	$t_s = 8a$	$t_s = 9a$	$t_s = 10a$	$t_s = 12a$
N_{conf}	36	50	88	811
N_{meas}	3240	4500	7920	72 990

4. Remarks and computational costs

This concludes our investigation of excited-states effects. We emphasize that the spectrum of nucleon excitations is

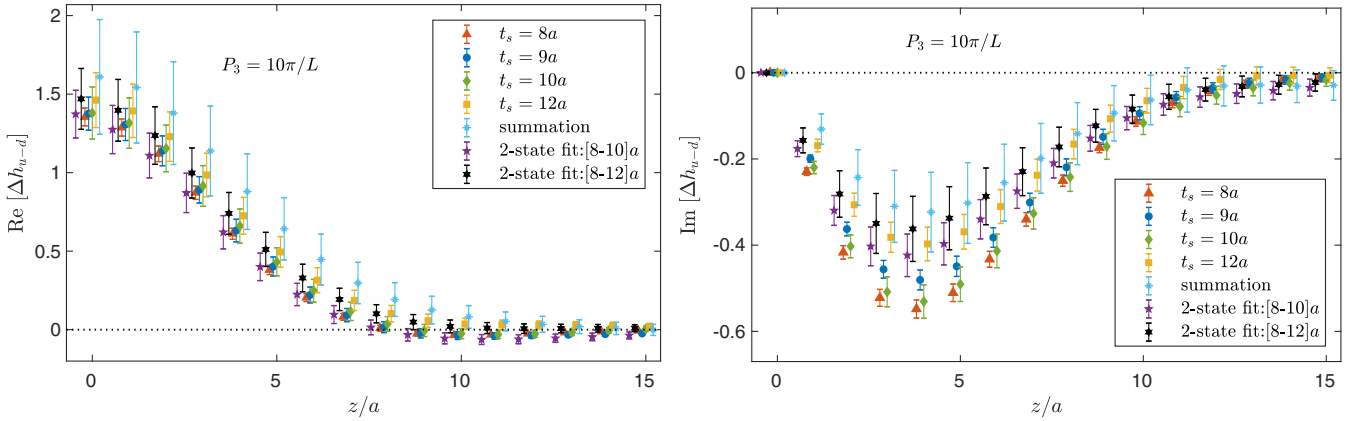

FIG. 15. Real (left) and imaginary (right) parts of the matrix element for the helicity PDF from the plateau method (points labeled with appropriate t_s), the summation method (using all t_s) and sequential two-state fits (to $t_s = 8a$, $9a$, $10a$ and to all t_s). The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

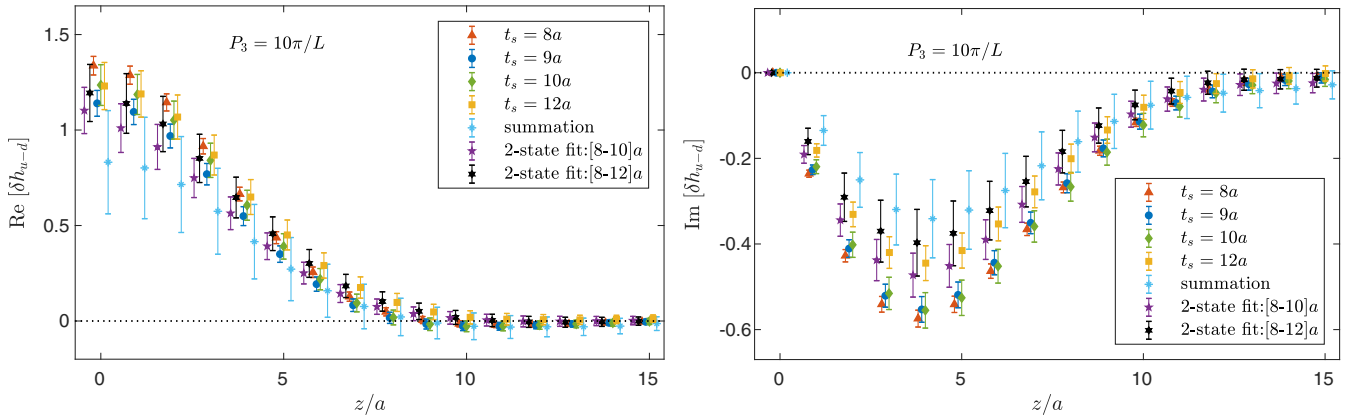

FIG. 16. Real (left) and imaginary (right) parts of the matrix element for the transversity PDF from the plateau method (points labeled with appropriate t_s), the summation method (using all t_s) and sequential two-state fits (to $t_s = 8a$, $9a$, $10a$ and to all t_s). The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

rich, particularly for a boosted nucleon with quarks with physical masses. Thus, one method of extracting bare matrix elements can be misleading, as the fitted energy gap between the ground state and the explicitly modeled first excited state suggests that there are tens of excited states. In such a situation, excited states need to be suppressed by going to large enough source-sink separations and robust statements can be made only when they are based on compatible results between different methods—in our case the plateau fits at the largest source-sink separation of around 1.1 fm, the two-state fits and the summation method. We note that the largest $t_s = 12a$ is crucially needed to establish this compatibility. Having only $t_s = 8a$, $9a$, $10a$ in the two-state fits and comparing to plateau fits at $t_s = 10a$, one would conclude that there is significant tension between the former and the latter. This is best illustrated in the imaginary part of the bare matrix elements for the helicity quasi-PDF; see the right panels of Fig. 15.

Finally, we would like to make some remarks on the computational resources that go well beyond what one needs for typical hadron structure calculations. The quasi-PDFs will reproduce the light-cone PDFs in the limit of large boosts. How large the boost should be, needs further investigation. Within the lattice QCD formulation, as already explained, one cannot increase the momentum of the nucleon to arbitrarily large values. The reasons are as follows:

(1) As the momentum increases, the signal-to-noise ratio rapidly deteriorates, despite the utilization of special methods, such as smearing techniques that reduce the noise. We find that to increase the momentum from $6\pi/L \approx 0.83$ to $8\pi/L \approx 1.11$ GeV, we need to increase the statistics by a factor of around 3.7 (for unpolarized PDFs) or around 6 (for helicity and transversity PDFs) and to increase to $10\pi/L \approx 1.38$ GeV by an additional factor of 3.7 or 6, respectively, in order to keep the statistical error approximately the same at $t_s = 12a$.

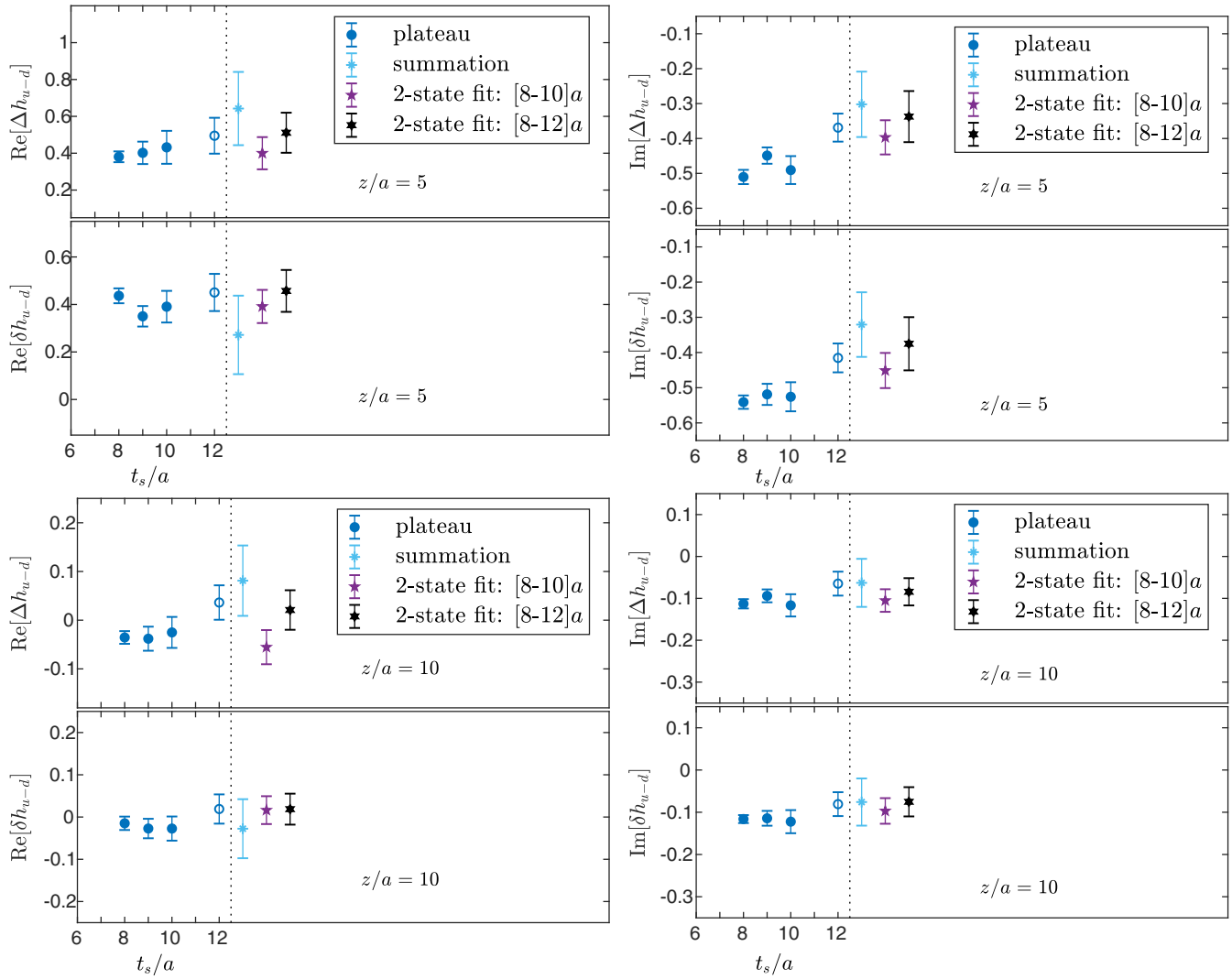

FIG. 17. Real (left) and imaginary (right) parts of the matrix element for the helicity and transversity PDFs at fixed $z/a = 5$ (upper) and $z/a = 10$ (lower). We compare results from the plateau method, the summation method (using all t_s) and sequential two-state fits (to $t_s = 8a, 9a, 10a$ and to all t_s). The final value that we use for extracting PDFs is indicated with an open symbol. The nucleon is boosted with $10\pi/L \simeq 1.38$ GeV.

(2) A careful study of excited states must be carried out and becomes increasingly more difficult as the momentum increases, because of the denser spectrum. As the source-sink separation increases, the statistical errors grow exponentially, e.g., for $t_s = 12a$, we needed a factor of about 10–12 (similar for all Dirac structures) more statistics for the same error as for $t_s = 10a$ (at the largest boost). It is imperative to have large enough source-sink separations for at least three values of t_s with comparable errors to perform a reliable analysis of excited-states effects and extract the ground-state matrix element.

Therefore, in order to reach a nucleon momentum of e.g., 2 GeV at $t_s = 12a$, for unpolarized PDFs, we estimate that one would need $\mathcal{O}(100)$ million core hours (Mch) on a typical supercomputer, as compared to $\mathcal{O}(5)$ Mch at $P_3 \approx 1.38$ GeV studied in this work and for a momentum of 3 GeV, we would need $\mathcal{O}(10\,000)$ Mch. For the polarized PDFs, the projected estimate for $P_3 \approx 3$ GeV is $\mathcal{O}(10^6)$ Mch. In addition, one may have to increase the source-sink separation to account for the increased excited-states contamination. If $t_s = 14a$ is used, then the computational resources for a boost of 3 GeV would be $\mathcal{O}(10^5)/\mathcal{O}(10^7)$ million core hours for the unpolarized/polarized case, which is prohibitively expensive, given currently available computers. These requirements may be alleviated with the possible development of better algorithms, enhancing the signal-to-noise ratio at large nucleon boosts and large source-sink separations. Another way to milden the need for huge computational resources is the derivation of two-loop matching and conversion factors, which is foreseen in the near future. In this way,

quasi-PDFs may be robustly connected to light-cone PDFs already at lower momenta.

We emphasize that the appropriate strategy is to not increase the nucleon boost uncontrollably, and to rely on precise data only at low source-sink separations. Failing to keep the statistical errors approximately constant as one increases t_s may introduce uncontrolled systematic errors, since the fits will be determined mostly by the more accurate data at small t_s. As we have shown, it is essential that all source-sink separations have approximately equal errors for a reliable extraction of the matrix elements and this is the criterion we adopt. Unlike other studies [26,28,31], we do not rely solely on two-state fits of data from source-sink separations with widely varied errors, which are thus dominated by the precise data at the small values of t_s. Such an approach can be uncontrolled and lead to a systematic bias in the final results.

IV. RENORMALIZATION

Renormalization is needed in order to relate the bare lattice QCD matrix elements to physical results. This is achieved by removing ultraviolet divergences, as well as finite dependence on the lattice action.[4] In the case of quasi-PDFs, one needs to also eliminate additional divergences arising due to the utilization of operators with a finite Wilson line. Renormalization of Wilson loops was addressed a long time ago using dimensional regularization (DR) for smooth contours [99], as well as for contours containing singular points [100]. Based on arguments valid to all orders in perturbation theory, it was demonstrated that smooth Wilson loops in DR are finite functions of the renormalized coupling, while the presence of cusps and self-intersections introduces logarithmically divergent multiplicative renormalization factors, referred to as Z factors. More importantly, it was shown that other regularization schemes are expected to lead to further Z factors, which are power-law divergent with respect to the dimensionful ultraviolet cutoff. This also appears in the lattice formulation, where a divergence arises as a function of the lattice spacing, that increases exponentially with the length of the Wilson line as $\sim e^{z/a}$. Such a divergence must be removed prior to the extrapolation to the continuum limit. We describe in this section our renormalization program that includes the removal of the exponential divergence. A recent work on the perturbative renormalization of Wilson-line fermion operators of the type given in Eq. (9) has identified the mixing pattern among nonlocal straight Wilson-line operators, and led to the development of an appropriate renormalization prescription for both the multiplicative renormalization and the mixing coefficient [21]. The nonperturbative renormalization program that we developed is a generalization of the RI′ scheme [105], which is appropriate for operators that contain a Wilson line [101].[5] The Z factors are extracted by imposing the following conditions:

$$\frac{Z_{\mathcal{O}}^{\mathrm{RI}'}(z,\mu_0,m_\pi)}{Z_q^{\mathrm{RI}'}(\mu_0,m_\pi)}\frac{1}{12}\mathrm{Tr}[\mathcal{V}(z,p,m_\pi)(\mathcal{V}^{\mathrm{Born}}(z,p))^{-1}]\Big|_{p^2=\mu_0^2}=1, \tag{20}$$

$$Z_q^{\mathrm{RI}'}(\mu_0,m_\pi)\frac{1}{12}\mathrm{Tr}[(S(p,m_\pi))^{-1}S^{\mathrm{Born}}(p)]\Big|_{p^2=\mu_0^2}=1. \tag{21}$$

We use the general notation $Z_{\mathcal{O}}$ and consider $\mathcal{O}=V_0, A, T$ corresponding to the unpolarized, helicity and transversity operators, respectively. $Z_{\mathcal{O}}$ and Z_q are the renormalization functions of the operator and the quark field, respectively. Both $Z_{\mathcal{O}}$ and Z_q are scheme and scale dependent, and are expected to have some dependence on the pion mass. Also, $Z_{\mathcal{O}}$ is a function of the length of the Wilson line, z. $\mathcal{V}$ is the amputated vertex function of the operator and S is the fermion propagator, while $\mathcal{V}^{\mathrm{Born}}$ and S^{Born} are the corresponding tree-level values. Note that this condition is applied independently for each value of z. The RI′ renormalization scale, μ_0, is chosen to be democratic in the spatial directions, that is $a\mu_0=\frac{2\pi}{L_s}(n_t+\frac{\pi}{2},n,n,n)$, which minimizes the ratio $P4\equiv\sum_i\mu_i^4/(\sum_i\mu_i^2)^2$ and suppresses discretization effects [104,107].

For the calculation of renormalization functions, we employ the momentum source method [104,108] that has the advantage of yielding results of high statistical accuracy. This method requires new inversions for each momentum used, but significant reduction in the gauge noise is observed, which by far outweighs the additional computational cost. Data are produced using the three $N_f=2$ ensembles given in Table V that have a different value of the pion mass. The twisted mass parameters of the light quarks in the sea and valence sectors have been set equal (isospin limit and unitary setup). Even though the ensemble simulated with the smallest value of the pion mass has a larger volume, the Z factors are obtained at the same values of $(a\mu_0)^2$ and $P4$. In the current work, we improve our previous analysis on the Z factors of one-derivative operators [101] by performing the following:

(1) A chiral extrapolation using the three ensembles at different values of the pion mass.
(2) A fit on the chiral data to eliminate residual dependence on the RI′ scale. We use several values of the RI′ scale that cover the range $(a\mu_0)^2\in[1\text{–}4]$. An extensive study on the choice of the renormalization scale and the corresponding systematic uncertainties can be found in Ref. [101].

[4]Elimination of residual dependence on the lattice formulation requires continuum extrapolation.

[5]For alternative approaches, using, e.g., an auxiliary field method [106], see the recent review of Ref. [33].

TABLE V. Parameters of the ensembles used to compute the Z factors extracted in this work. For details, see Ref. [79].

$\beta = 2.10$,	$c_{\rm SW} = 1.57751$,	$a = 0.0938(3)(2)$ fm
$48^3 \times 96$	$a\mu = 0.0009$	$m_\pi = 130$ MeV
$24^3 \times 48$	$a\mu = 0.003$	$m_\pi = 235$ MeV
$24^3 \times 48$	$a\mu = 0.006$	$m_\pi = 340$ MeV

A. Pion mass dependence

The pion mass dependence for the Z factors for operators with a finite Wilson line has never been explored, and it is expected that small values of z will have a weak dependence on m_π, as observed in the case of local operators computed within the same setup. However, it is not known how the Z factors will behave when the length of the Wilson line is large. To obtain the Z factors in the chiral limit, we fit the data from the three ensembles at each value of z. This fit is applied to the real and imaginary parts independently. The data for $Z_{\mathcal{O}}^{\rm RI'}$ are expected to have a quadratic dependence on the pion mass (equivalently linear with respect to the twisted mass parameter) and are fitted using

$$Z_{\mathcal{O}}^{\rm RI'}(z,\mu_0,m_\pi) = Z_{\mathcal{O},0}^{\rm RI'}(z,\mu_0) + m_\pi^2 Z_{\mathcal{O},1}^{\rm RI'}(z,\mu_0). \quad (22)$$

The chirally extrapolated value is given by the fit parameter $Z_{\mathcal{O},0}^{\rm RI'}(z,\mu_0)$. Since the fitted data are obtained on different ensembles, we use the super-jackknife method (see, e.g., Ref. [109]) to correctly calculate the statistical error. This method is applicable to both correlated and uncorrelated data.

In Fig. 18, we show the pion mass dependence of the real and imaginary parts of the Z factors for all three operators. For a clear presentation, we focus on $z/a = 1$, 5 and we plot against m_π^2 for the scale $(a\mu_0)^2 = 2.5$. We find that the dependence is almost constant in m_π^2 and the chiral fit yields a slope consistent with zero for both the real and imaginary parts for small values of z (see, e.g., $z/a = 1$). As z increases, a nonzero slope is observed, with the dependence being linear, as expected. The dashed line corresponds to the chiral fit of Eq. (22), and the open symbols are the extrapolated values $Z_{\mathcal{O},0}^{\rm RI'}(z,\mu_0)$.

B. Volume effects

Another source of systematic uncertainty entering the determination of matrix elements is due to the finite lattice extent. Finite-volume effects are expected to be suppressed as $\exp(-m_\pi L)$ and based on empirical studies, $m_\pi L \geq 4$ is considered sufficient in most applications. However, most lattice QCD studies deal with matrix elements of local operators, and as discussed in Ref. [110], the length of the Wilson line entering the operator may enhance finite-volume effects. To date, this systematic uncertainty has not been investigated due to the absence of lattice QCD

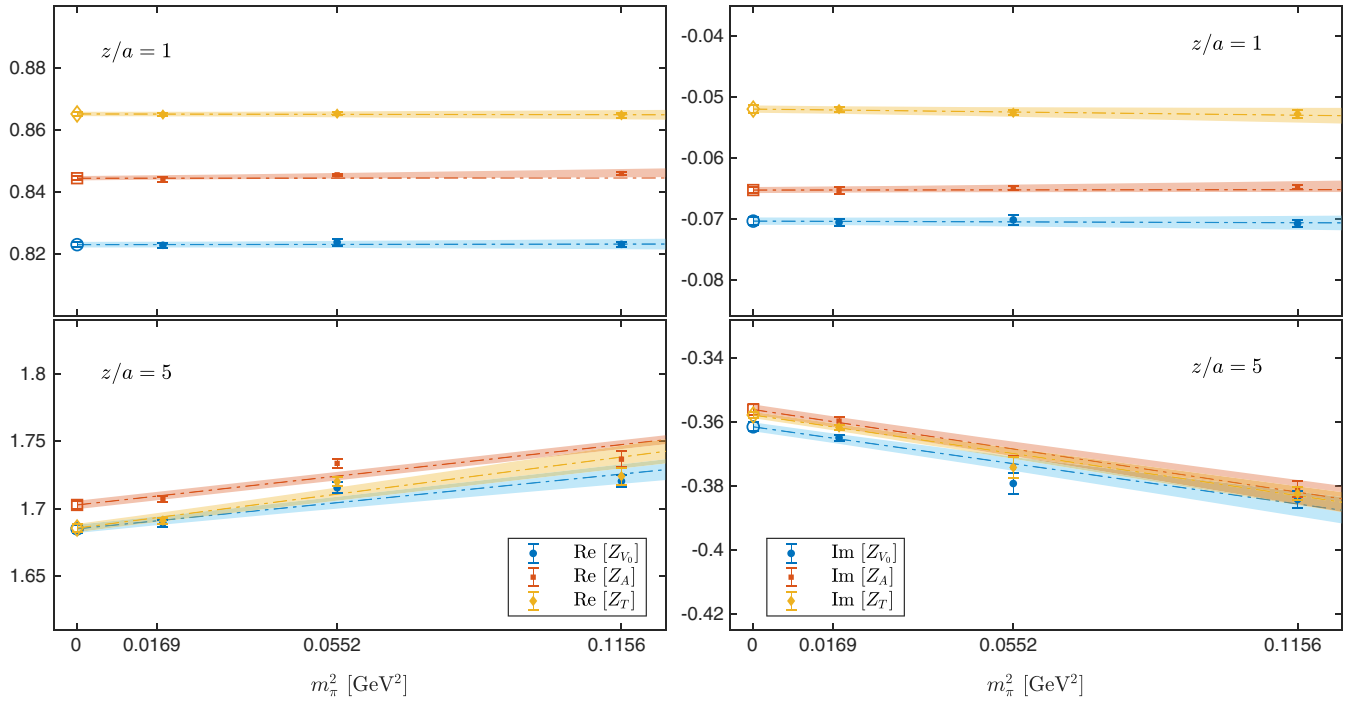

FIG. 18. Real (left) and imaginary (right) parts of the Z factor $Z_{\mathcal{O}}^{\rm RI'}(z,\mu_0,m_\pi)$ for the unpolarized (blue circles), helicity (red squares) and transversity (orange diamonds) PDFs as a function of m_π^2 of the ensemble used. The scale is $(a\mu_0)^2 = 2.5$. The upper and lower panels show the results for $z/a = 1$ and $z/a = 5$, respectively. The dashed line corresponds to the chiral fit of Eq. (22), and the open symbols are the extrapolated values, $Z_{\mathcal{O},0}^{\rm RI'}(z,\mu_0)$.

computations using ensembles with different volumes, while keeping the rest of the parameters constant. Such a study requires significant computational resources and is one of our future goals. Volume effects in the renormalization functions can be easily examined due to the reduced computational resources compared to the matrix elements. Therefore, we compute the Z factors using the $48^3 \times 96$ ensemble of Table V and a $64^3 \times 128$ ensemble with the same pion mass. In Fig. 19, we plot the ratios

$$\mathcal{R}_{\mathcal{O}}(z) \equiv \frac{\mathrm{Re}[Z^{\mathrm{RI}'}_{\mathcal{O},64}(z,\mu_0,m_\pi)]}{\mathrm{Re}[Z^{\mathrm{RI}'}_{\mathcal{O},48}(z,\mu_0,m_\pi)]},$$
$$\mathcal{I}_{\mathcal{O}}(z) \equiv \frac{\mathrm{Im}[Z^{\mathrm{RI}'}_{\mathcal{O},64}(z,\mu_0,m_\pi)]}{\mathrm{Im}[Z^{\mathrm{RI}'}_{\mathcal{O},48}\mathcal{O}(z,\mu_0,m_\pi)]}, \tag{23}$$

as a function of the length of the Wilson line, for all three types of operators. The ratios are defined at the pion mass $m_\pi = 130$ MeV and a renormalization scale $(a\mu_0)^2 = 2.5$. The additional subscripts 64 and 48 indicate the lattice volume. We find that both the real and imaginary parts do not show a statistically significant dependence on the volume, as the ratios take maximum values of 1.02 and 1.03, respectively, for z/a up to 15, which is well within the range of interest. In addition, we find an almost linear increase of $\mathcal{R}_{\mathcal{O}}(z)$, while $\mathcal{I}_{\mathcal{O}}(z)$ has an oscillatory dependence on the volume.

Hence, we conclude that volume effects in the Z factors are small and have little effect on the renormalized matrix elements. This partly results from the fact that the bare matrix elements go to zero in the large-z region. In determining the final values for the Z factors, we do not include the $64^3 \times 128$ ensemble, because the ratio $P4$ is large for most of the scales $(a\mu_0)^2 \in [1\text{–}4]$ compared to the smaller volumes, leading to contamination from finite-a effects, as seen in Ref. [104]. However, for $(a\mu_0)^2 = 2.5$ which was used in the comparison, $P4$ is the same as the one obtained from the smaller volumes, which allows one to isolate the volume effects. Finite-a effects are in fact under investigation for this class of nonlocal operators in lattice perturbation theory [111].

C. Conversion to the standard and modified $\overline{\mathrm{MS}}$ schemes

In order to compare renormalized lattice QCD matrix elements with phenomenological results extracted from global analyses, the Z factors must be converted to the same scheme and evolved to the same scale as those used in the phenomenological analyses. Traditionally, the chosen scheme is the $\overline{\mathrm{MS}}$ scheme and the scale $\bar{\mu}$ is typically set to 2 GeV. The appropriate conversion factors for the nonlocal operators with a straight Wilson line are taken from Ref. [21], where a calculation was carried out to one-loop level in perturbation theory, using dimensional regularization. Technical complications related to the nonlocality of the operators under study make it very hard to extend such a calculation to higher loops, as done for local operators, which are usually known to three and four loops. As a consequence, it is expected that the Z factors will have a residual dependence on the initial RI′ scale μ_0. Thus, one typically computes the Z factors at several values of the RI′ scale, as set by Eq. (20), and then uses an appropriate conversion to bring each $Z^{\mathrm{RI}'}_{\mathcal{O},0}(z,\mu_0)$ to $Z^{\overline{\mathrm{MS}}}_{\mathcal{O},0}(z, 2\ \mathrm{GeV})$. The remaining dependence on μ_0 can be studied with a linear fit in $(a\mu_0)^2$, which is the leading order of a more complicated scale dependence. The final values are obtained by taking $(a\mu_0)^2 \to 0$ using

$$Z^{\overline{\mathrm{MS}}}_{\mathcal{O}}(z,\bar{\mu},\mu_0) = \mathcal{Z}^{\overline{\mathrm{MS}}}_{\mathcal{O},0}(z,\bar{\mu}) + (a\mu_0)^2 \mathcal{Z}^{\overline{\mathrm{MS}}}_{\mathcal{O},1}(z,\bar{\mu}). \tag{24}$$

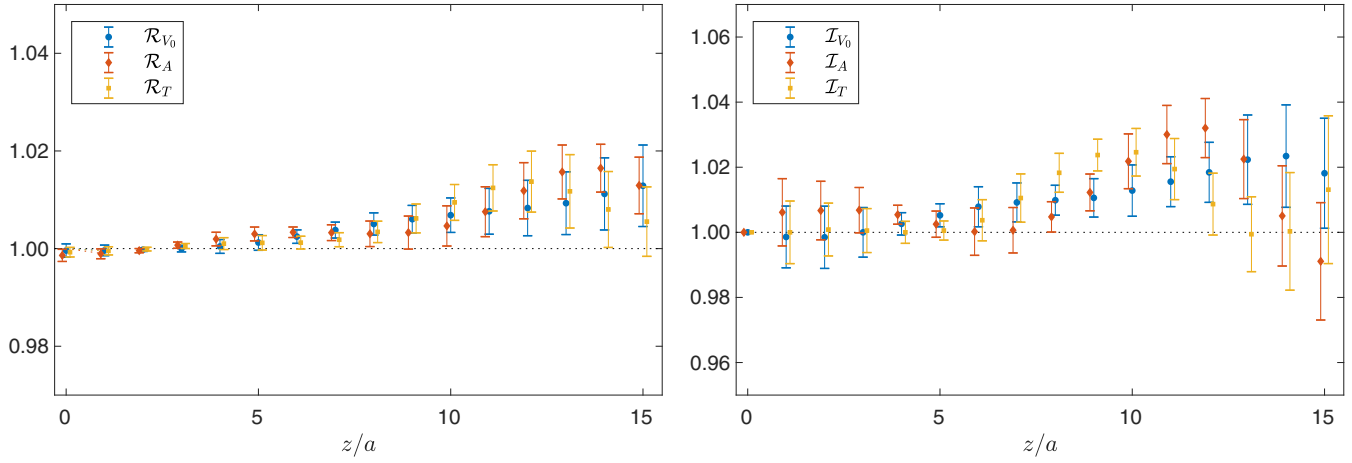

FIG. 19. Plots of $\mathcal{R}_{\mathcal{O}}(z)$ and $\mathcal{I}_{\mathcal{O}}(z)$ to study volume effects between the ensembles $48^3 \times 96$ and $64^3 \times 128$, for the unpolarized (circles), helicity (diamonds) and transversity (squares) operators as a function of the length of the Wilson line, z/a. The RI′ scale is $(a\mu_0)^2 = 2.5$ and the pion mass is 130 MeV.

For $Z_{\mathcal{O}}^{\overline{\rm MS}}(z,\bar{\mu},\mu_0)$, we use the chirally extrapolated Z factors converted to $\overline{\rm MS}$ at the scale $\bar{\mu}=2$ GeV. For simplicity, in the notation we dropped the subscript "0" appearing in the fit of Eq. (22). The desired quantity is the fit parameter $\mathcal{Z}_{\mathcal{O},0}^{\overline{\rm MS}}(z,\bar{\mu})$.

Our renormalization program entails a new element, namely the use of a modified $\overline{\rm MS}$ scheme (M$\overline{\rm MS}$). The development of such a scheme was motivated by the fact that the existing matching formulas do not satisfy particle number conservation (e.g. Ref. [45]). The matching using the M$\overline{\rm MS}$ scheme was already presented in our recent work [19,20]. Here we complement the previous analyses by giving the appropriate conversion of the Z factors to the M$\overline{\rm MS}$ scheme, instead of $\overline{\rm MS}$. We find that the resulting modification is numerically very small, but moves the final values of the PDFs towards the phenomenological ones; see Sec. V A for details of the M$\overline{\rm MS}$ scheme and Sec. V B for numerical effects when applied to our data. In a nutshell, an additional conversion factor is needed to bring $\mathcal{Z}_{\mathcal{O},0}^{\overline{\rm MS}}(z,\bar{\mu})$ to $\mathcal{Z}_{\mathcal{O},0}^{\rm M\overline{MS}}(z,\bar{\mu})$ via

$$\mathcal{Z}_{\mathcal{O},0}^{\rm M\overline{MS}}(z,\bar{\mu})=\mathcal{Z}_{\mathcal{O},0}^{\overline{\rm MS}}(z,\bar{\mu})\mathcal{C}^{\overline{\rm MS},\rm M\overline{MS}}, \tag{25}$$

which has been computed perturbatively in dimensional regularization to the one-loop level and is presented in the next section. The expression for the conversion $\mathcal{C}_{\mathcal{O}}^{\overline{\rm MS},\rm M\overline{MS}}$ is different for each operator under study and its general form is given by

$$\mathcal{C}_{\mathcal{O}}^{\overline{\rm MS},\rm M\overline{MS}}=1+\frac{C_F g^2}{16\pi^2}\left[e_{\mathcal{O}}^{(1)}+e_{\mathcal{O}}^{(2)}\ln\left(\frac{\bar{\mu}^2}{4\mu_F^2}\right)+e_{\mathcal{O}}^{(3)}\left(\frac{i\pi|\mu_F z|}{2\mu_F z}-\ln(|\mu_F z|)-\mathrm{Ci}(\mu_F z)-i\mathrm{Si}(\mu_F z)+\ln(\mu_F z)\right)\right.$$
$$\left.+e_{\mathcal{O}}^{(4)}(e^{i\mu_F z}(2\mathrm{Ei}(-i\mu_F z)+i\pi\mathrm{sgn}(\mu_F z)-\ln(-i\mu_F z)+\ln(i\mu_F z)))\right], \tag{26}$$

where μ_F is the factorization scale that is taken equal to the $\overline{\rm MS}$ scale, that is, $\mu_F=\bar{\mu}=2$ GeV. The expression also contains the special functions Ci (cosine integral), Si (sine integral) and Ei (exponential integral), as well as the sign function (sgn). The coefficients $e_{\mathcal{O}}^{(i)}$ depend on the operator and their numerical values are given in Table VI.

In Fig. 20, we show $\mathcal{Z}_{\mathcal{O},0}^{\rm M\overline{MS}}(z,\bar{\mu})$ after multiplying $\mathcal{Z}_{\mathcal{O},0}^{\overline{\rm MS}}(z,\bar{\mu})$ with the conversion factor given in Eq. (25). The data are shown against $(a\mu_0)^2$ to demonstrate the dependence of the Z factors on the initial RI′ scale. As done for the previous figures, we choose representative values of the length of the Wilson line, namely $z/a=1$ and $z/a=5$. We find that the imaginary part has a stronger dependence on the μ_0 value. Note, however, that the imaginary part is an order of magnitude smaller than the real part of the matrix element and thus this dependence is still suppressed in the total matrix element, especially in the region $(a\mu_0)^2\in[1\text{–}2]$, which is an indication of nonperturbative effects. Such behavior is also observed in local operators for $(a\mu_0)^2<1$. This tendency seems to affect the region $(a\mu_0)^2<2$ for nonlocal operators with $z/a\geq 5$. Since we are using perturbative expressions for the conversion to the M$\overline{\rm MS}$ scheme, it is important to choose a region where perturbation theory is valid, and thus, we choose the values obtained from $(a\mu_0)^2\in[2\text{–}4]$ for both the real and imaginary parts. There is a systematic uncertainty attached to the Z factor due to the choice of the fit range, and here we used the ranges $(a\mu_0)^2\in[1\text{–}3]$, [1–4], [2–3], [2–4] to estimate the systematic uncertainty. Even though the real and imaginary parts of the Z factors do have mild dependence on the range, the final values are consistent. Therefore, we do not give any systematic uncertainty from the $(a\mu_0)^2\to 0$ extrapolation. In the same figure, we also show the improvement of the nonperturbative results when subtracting $\mathcal{O}(g^2a^\infty)$ effects computed perturbatively in Ref. [104] on the same ensembles. They are obtained by replacing Eq. (20) with

$$\frac{Z_{\mathcal{O}}^{\rm RI'}(z,\mu_0,m_\pi)}{Z_q^{\rm RI'}(\mu_0,m_\pi)-A_q^\infty(\mu_0)}\times\frac{1}{12}\mathrm{Tr}[\mathcal{V}(z,p,m_\pi)(\mathcal{V}^{\rm Born}(z,p))^{-1}]\Big|_{p^2=\mu_0^2}=1, \tag{27}$$

where the denominator has been modified by subtracting the $\mathcal{O}(g^2a^\infty)$ artifacts in Z_q, denoted by $A_q^\infty(\mu_0)$. As expected, the differences between the subtracted results of Eq. (27) and the unsubtracted results of Eq. (20) are small, and mostly affect the real part of the Z factors. In addition, subtraction of the artifacts in Z_q is not sufficient to

TABLE VI. Numerical values for the coefficients $e_{\mathcal{O}}^{(i)}$ appearing in the conversion function from $\overline{\rm MS}$ to M$\overline{\rm MS}$ in Eq. (26).

$\mathcal{O}$	$e_{\mathcal{O}}^{(1)}$	$e_{\mathcal{O}}^{(2)}$	$e_{\mathcal{O}}^{(3)}$	$e_{\mathcal{O}}^{(4)}$
V	−5	−3	+3	−3/2
A	−7	−3	+3	−3/2
T	−4	−4	+4	−4/2

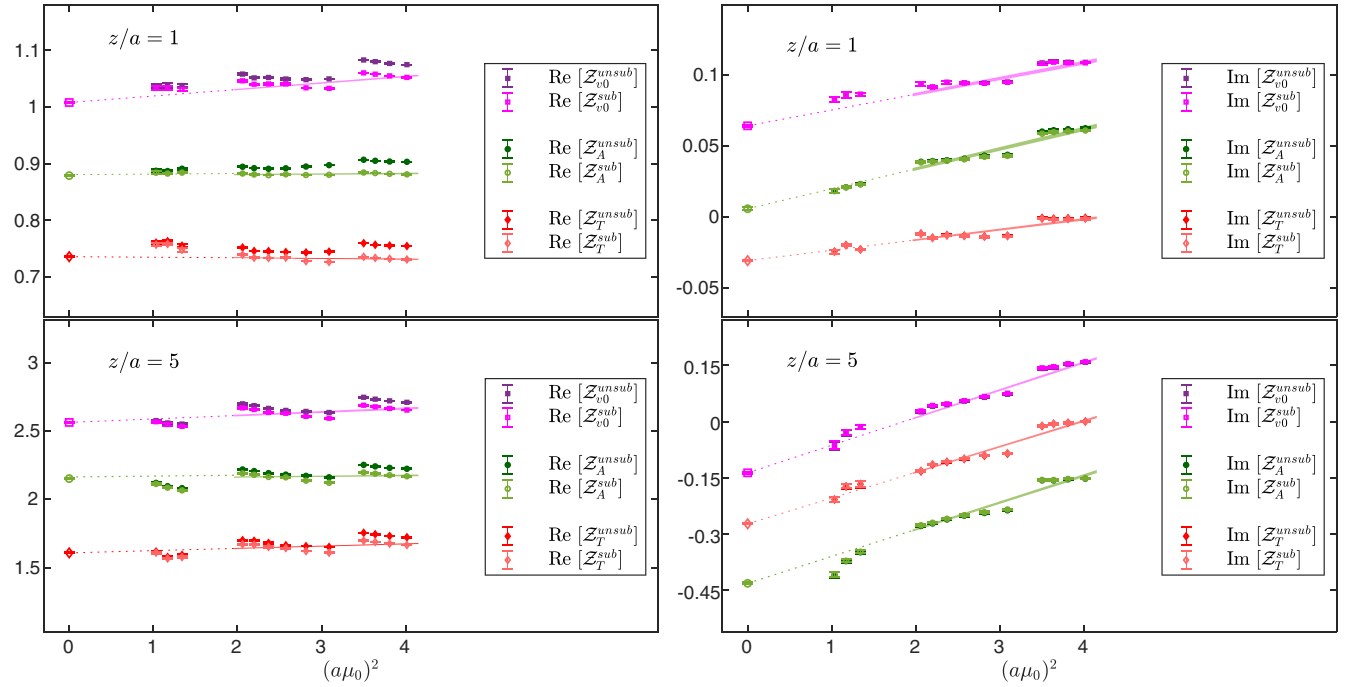

FIG. 20. Real (left) and imaginary (right) parts of $Z_{\mathcal{O}}^{\overline{\text{MMS}}}(z, \bar{\mu}, \mu_0)$ for the unpolarized (squares), helicity (circles) and transversity (diamonds) operators as a function of the initial RI′ scale. The upper (lower) panel corresponds to $z/a = 1$ ($z/a = 5$). The dashed line corresponds to the fit of Eq. (24), and the open symbols are the extrapolated values $\mathcal{Z}_{\mathcal{O},0}^{\overline{\text{MMS}}}$. We show results for all Z factors with Z_{V_0} and Z_T shifted vertically for clarity of the presentation. The darker colored filled symbols correspond to the purely nonperturbative estimates, while the corresponding lighter colored open symbols have been partly improved using the perturbative results of the finite lattice spacing effects to $\mathcal{O}(g^2 a^\infty)$ computed in Ref. [104] on the same ensembles.

eliminate the dependence on $a\mu_0$. For the latter to be achieved, subtraction of the artifacts in $\mathcal{V}(z, p, m_\pi)$ is crucial, as has been discussed in Ref. [104] for the local operators. Such a computation for the nonlocal operators is not only more technically complicated, but also requires the addition of stout smearing in the links of the operator resulting in lengthy expressions. This calculation is under way and will be presented in a separate publication [111].

The final values for the Z factors are given in the $\overline{\text{MMS}}$ scheme upon a linear fit in $(a\mu_0)^2 \to 0$. In the region $(a\mu_0)^2 \in [2\text{–}4]$ [Eq. (24)]. They are shown in Fig. 21 for the vector operator. As can be seen, the statistical errors of the Z factors are much smaller as compared to those of the

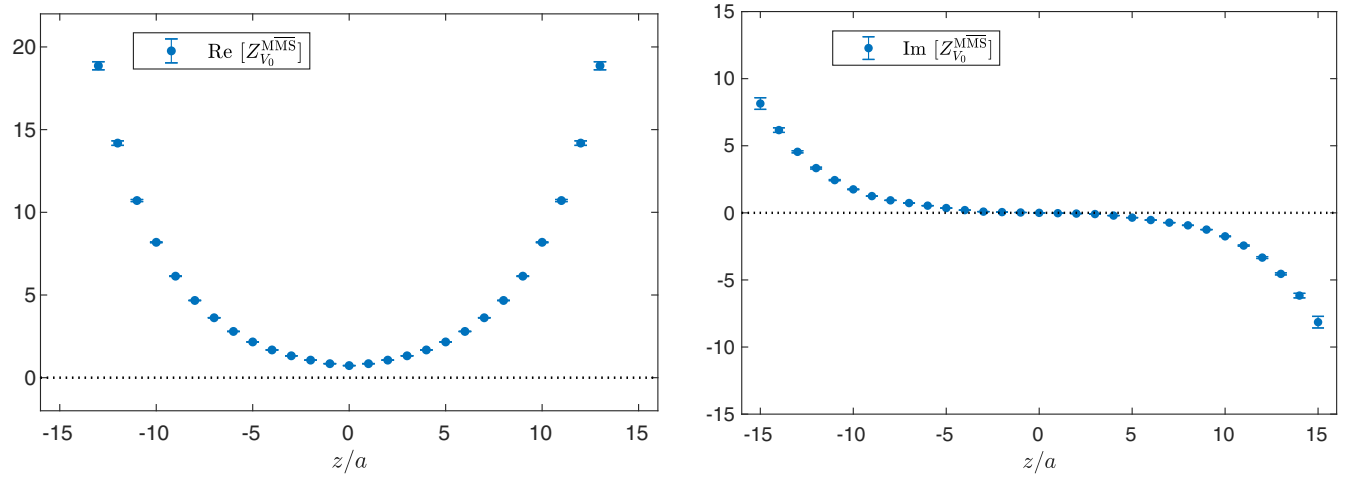

FIG. 21. Final values of the Z factors in the $\overline{\text{MMS}}$ scheme for the vector operator upon chiral and $(a\mu_0)^2$ extrapolation. The real and imaginary parts are shown in the left and right panels, respectively.

matrix elements, because of the momentum source method used in the calculation of the vertex functions. The real part increases significantly with z, which is due to the power-law divergence of the Wilson line. Note that for $z/a > 10$, the bare matrix elements decay to zero, and thus, the renormalized matrix elements, given by the complex multiplication

$$h_\Gamma^{\overline{\rm MS}}(P_3, z, \bar{\mu}) = h_\Gamma^{\rm bare}(P_3, z) \cdot \mathcal{Z}_{\Gamma,0}^{\overline{\rm MS}}(z, \bar{\mu}), \quad (28)$$

are also zero but with large statistical uncertainty originating from the large values of the real part of the Z factors.

Both the matrix elements and the renormalization functions of a given operator share similar properties with respect to z (symmetric real part and antisymmetric imaginary part), with the difference that the Z factor decreases when a smearing technique is applied to the Wilson line, while the matrix element increases. Of course, the dependence in the smearing is nonmonotonic due to the complex nature of the matrix elements and Z factors, but the stout smearing dependence is expected to cancel out in the renormalized matrix elements. This is demonstrated in Fig. 22, where we compare the renormalized matrix elements for the helicity operator extracted using 0, 5, 10 and 15 stout smearing steps, in the $\overline{\rm MS}$ scheme at 2 GeV and for a momentum of $\frac{6\pi}{L}$. This confirms that the elimination of the power divergences is correctly realized via the renormalization program, yielding compatible results for the renormalized matrix elements for different stout iterations. We find that this holds also for the renormalized matrix elements of the unpolarized and transversity operators and therefore any stout step may be used without changing the final physical result. It is worth mentioning that the agreement is more striking upon the $(a\mu_0)^2 \to 0$ extrapolation of the Z factors, as the central values of the data from different stout steps are very close to each other. Nevertheless, a similar comparison using a specific value of $(a\mu_0)^2$ shows that the agreement still holds within statistical uncertainties.

V. MATCHING TO LIGHT-CONE PDFs

A. Derivation of the matching formulas

In this section, we discuss, in detail, the matching procedure that relates the quasi-PDFs, renormalized in some scheme, to light-cone PDFs in the same or another renormalization scheme, typically chosen to be the $\overline{\rm MS}$ scheme. We derive new matching formulas that relate $\overline{\rm MS}$-renormalized quasi-PDFs to $\overline{\rm MS}$-renormalized light-cone PDFs and conserve the particle number. To satisfy this condition, we introduce a modification of the $\overline{\rm MS}$ scheme, i.e., the M$\overline{\rm MS}$ scheme, which was already partially discussed in the previous section, as it is a procedure applied also on the conversion factors.

Quasi-PDFs can be obtained as a Fourier transform (FT) of renormalized matrix elements, $h_\Gamma(P_3, z)$,

$$\tilde{q}(x, P_3) = \frac{2P_3}{4\pi} \sum_{z=-z_{\rm max}}^{z_{\rm max}} e^{-ixP_3 z} h_\Gamma(P_3, z). \quad (29)$$

To relate the quasi-PDFs $\tilde{q}(x, P_3)$ to light-cone PDFs, one relies on a perturbative matching procedure [34,43–46]. To one-loop order, and in the Feynman gauge, one needs to compute self-energy corrections, which include the usual self-energy plus the virtual "sail" and "tadpole" diagrams, and the vertex corrections, with the corresponding real "sail" and "tadpole" diagrams (see Fig. 23). We use operators with four Dirac structures, namely γ^0 and γ^3 for the unpolarized distribution, $\gamma^3\gamma^5$ for the helicity, and $\gamma^3\gamma^j$, $j = 1, 2$ for the transversity case.

Since the extraction of the matching formula follows a similar process for all four operators, we use the γ^0 structure

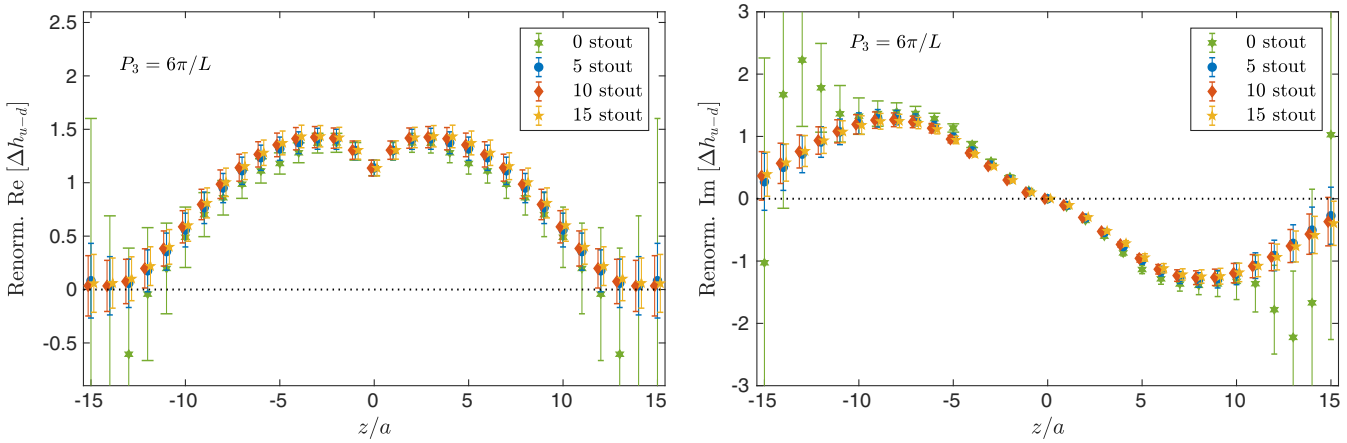

FIG. 22. Real (left) and imaginary (right) parts of renormalized matrix elements for the axial operator (helicity PDF) for a momentum of $\frac{6\pi}{L}$, as a function of the length of the Wilson line. Green stars/blue circles/red diamonds/orange stars correspond to 0/5/10/15 iterations of stout smearing.

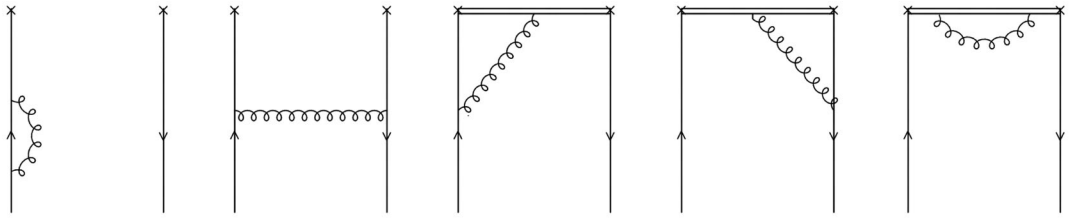

FIG. 23. One-loop diagrams entering the perturbative calculation for the matching between quasi-PDFs and light-cone PDFs.

as an example and we only present the final results for the other cases. As already mentioned, we take the nucleon momentum in the third direction, $P = (P^0, 0, 0, P^3)$. It is assumed that, before gluon emission, the quark momentum p is collinear to the nucleon momentum, i.e. $p = (p^0, 0, 0, p^3)$. It also obeys the Dirac equation $\not{p}u(p) = 0$, and carries a fraction of the momentum y of the parent hadron. After gluon emission, the quark has momentum $\xi p_3 = \xi y P_3$. When the bare quark distributions are dressed by taking the $p_3 \to \infty$ limit, we have the usual quark distributions in the infinite-momentum frame or, equivalently, on the light cone. Defining $x \equiv \xi y$, the nonsinglet quark distribution at one loop is given by

$$q_{\gamma^0}(x,\Lambda) = q_{\text{bare}}(x) + \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y}\left(\Pi_{\gamma^0}(\Lambda)\delta\left(1-\frac{x}{y}\right) + \Gamma_{\gamma^0}\left(\frac{x}{y},\Lambda\right)\right) q_{\text{bare}}(y), \tag{30}$$

where Π_{γ^0} denotes the self-energy corrections, Γ_{γ^0} represents the vertex corrections, and Λ stands for the IR and UV regulators in any given scheme. Using DR to regulate both the IR and UV divergences, the resulting one-loop correction for $0 < \xi < 1$ in the $\overline{\text{MS}}$ scheme is

$$\Gamma^{\overline{\text{MS}}}_{\gamma^0}\left(\xi, \frac{\bar\mu}{\mu_F}\right) = \frac{1+\xi^2}{1-\xi}\left(-\frac{1}{\epsilon_{\text{IR}}} + \ln\left(\frac{\bar\mu^2}{\mu_F^2}\right)\right), \tag{31}$$

where the pole from the UV divergence has already been subtracted, and $\bar\mu$ is the corresponding renormalization scale in the $\overline{\text{MS}}$ scheme. The pole from the soft IR, $1/\epsilon_{\text{IR}}$, can be absorbed into the bare distribution, at the factorization scale μ_F, but it is explicitly written in Eq. (31), as it must cancel a similar pole arising in the one-loop correction to the quasi-PDF. The fact that the poles in $1/\epsilon_{\text{IR}}$ are the same for the light-cone PDF and the quasi-PDF, is a crucial observation that allows one to match them using perturbation theory. There are remaining IR divergences, which are located at $\xi = 1$ and have their origin in the emission of soft gluons. However, they cancel between the vertex and self-energy one-loop corrections, which are related by $\Pi_{\gamma^0}(\Lambda) = -\int_0^1 d\xi \Gamma_{\gamma^0}(\xi,\Lambda)$.

To obtain the quasi-PDF, the quark is dressed with finite p_3. To simplify notation, we write the one-loop correction for positive x only, that is

$$\tilde q_{\gamma^0}(x,P_3,\Lambda) = q_{\text{bare}}(x) + \frac{\alpha_s}{2\pi}\int_0^1 \frac{dy}{y}\left(\tilde\Pi_{\gamma^0}(p_3,\Lambda)\delta\left(1-\frac{x}{y}\right) + \tilde\Gamma_{\gamma^0}\left(\frac{x}{y},p_3,\Lambda\right)\right) q_{\text{bare}}(y). \tag{32}$$

Because $\tilde\Gamma_{\gamma^0} \neq 0$ for $\xi > 1$, the lower limit of integration in Eq. (32) goes to zero, and the quasi-PDF has support for $x > 1$. Using DR to perform the integrals, the one-loop correction to the vertex is given by

$$\tilde\Gamma_{\gamma^0}\left(\xi,\frac{p_3}{\mu_F}\right) = \begin{cases} \frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi}{\xi-1}\right)+1, & \xi > 1, \\ \frac{1+\xi^2}{1-\xi}\left(-\frac{1}{\epsilon_{\text{IR}}}+\ln\left(\frac{4\xi(1-\xi)(p_3)^2}{\mu_F^2}\right)\right) - \frac{2\xi}{1-\xi} + \xi, & 0 < \xi < 1, \\ \frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi-1}{\xi}\right)-1, & \xi < 0. \end{cases} \tag{33}$$

Note that the vertex corrections have no poles when DR is used, but they are nonzero outside the physical region $0 < \xi < 1$, a behavior that is different from the one-loop correction to the light-cone PDF. Thus, in addition to the IR divergences at $\xi = 1$, there are UV divergences associated with the $\xi \to \pm\infty$ limits when Eq. (33) is integrated. So far, these UV divergences have not been subtracted, and this is reflected by the Λ dependence in the functional form of $\tilde\Pi_{\gamma^0}(\Lambda, p_3/\mu_F)$. DR is again used to isolate the poles, in which case $d\xi \to d\xi(\xi p_3)^{d-1}(\bar\mu^2 e^{\gamma_E}/4\pi)^{1/2-d/2}$. For the region $\xi > 1$, the self-energy is

$$\tilde\Pi_{\gamma^0}(\epsilon,\xi>1) = \int_1^\infty d\xi \xi^{d-1}(p_3/\bar\mu)^{d-1}(e^{\gamma_E}/4\pi)^{1/2-d/2} \times\left(\frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi-1}{\xi}\right)-1\right), \tag{34}$$

where $d = 1–2\epsilon$, with $\epsilon > 0$. Using the plus prescription for the collinear divergence at $\xi = 1$, Eq. (34) can be written as

$$\tilde{\Pi}_{\gamma^0}(\epsilon, \xi > 1) = \frac{7}{4} - \frac{3}{4}\ln\left(\frac{p_3^2}{\bar{\mu}^2}\right) + \frac{3}{4}\left(\frac{1}{\epsilon} + \gamma_E - \ln(4\pi)\right). \tag{35}$$

Similar computations can be done for the other regions, and the renormalized one-loop self-energy in the $\overline{\text{MS}}$ scheme for the quasi-PDF is given by

$$\tilde{\Pi}_{\gamma^0}^{\overline{\text{MS}}}\left(\xi, \frac{p_3}{\mu_F}, \frac{\mu_F}{\bar{\mu}}\right) = \frac{5}{2} + \frac{3}{2}\ln\left(\frac{\mu_F^2}{4\bar{\mu}^2}\right) - \int d\xi \begin{cases} \frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi}{\xi-1}\right) + 1 - \frac{3}{2\xi}, & \xi > 1, \\ \frac{1+\xi^2}{1-\xi}\left(-\frac{1}{\epsilon_{\text{IR}}} + \ln\left(\frac{4\xi(1-\xi)(p_3)^2}{\mu_F^2}\right)\right) - \frac{2\xi}{1-\xi} + \xi, & 0 < \xi < 1, \\ \frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi-1}{\xi}\right) - 1 - \frac{3}{2(1-\xi)}, & \xi < 0, \end{cases} \tag{36}$$

with the corresponding renormalization function, $\tilde{Z}_{\Pi}^{\overline{\text{MS}}} = 1 + \frac{\alpha_s}{2\pi} C_F \frac{3}{2}\left(\frac{1}{\epsilon} + \gamma_E - \ln(4\pi)\right)$. Upon integration of Eq. (33) and employing the $\overline{\text{MS}}$ scheme, the Ward identity of the resulting local current is naturally respected.

For the computation of the x dependence of the distributions, however, the convolution involving the vertex correction prevents the aforementioned cancellation from occurring, and one needs to impose a prescription to ensure conservation of the quark number for the nonsinglet distributions. From the equations for the quark distributions, Eq. (30), and quasidistributions, Eq. (32), and including the negative x regions we have to one loop

$$q_{\gamma^0}(x, \bar{\mu}) = \int_{-\infty}^{+\infty} \frac{dy}{|y|} C_{\gamma^0}^{\overline{\text{MS}}}\left(\frac{x}{y}, \frac{\bar{\mu}}{p_3}, \frac{\bar{\mu}}{\mu_F}\right) \tilde{q}_{\gamma^0}(y, P_3, \bar{\mu}), \tag{37}$$

where $C_{\gamma^0}^{\overline{\text{MS}}}(\xi) = \delta(1-\xi) - \alpha_s C_F[(\tilde{\Pi}_{\gamma^0}^{\overline{\text{MS}}} - \Pi_{\gamma^0}^{\overline{\text{MS}}})\delta(1-\xi) + \tilde{\Gamma}_{\gamma^0}(\xi) - \Gamma_{\gamma^0}^{\overline{\text{MS}}}(\xi)]/2\pi$. The dependence on p_3, $\bar{\mu}$, and μ_F is implicit in $C^{\overline{\text{MS}}}(\xi)$. $\tilde{\Gamma}(\xi)$ is the bare function, hence the absence of an $\overline{\text{MS}}$ superscript. The poles of $1/\epsilon_{\text{IR}}$, on the other hand, automatically cancel in $C_{\gamma^0}^{\overline{\text{MS}}}$, which is explicitly written as

$$C_{\gamma^0}^{\overline{\text{MS}}}\left(\xi, \frac{\bar{\mu}}{p_3}, \frac{\bar{\mu}}{\mu_F}\right) = \delta(1-\xi) + \frac{\alpha_s C_F}{2\pi} \begin{cases} \left(\frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi}{\xi-1}\right) + 1 + \frac{3}{2\xi}\right)_{+(1)} - \frac{3}{2\xi}, & \xi > 1 \\ \left(\frac{1+\xi^2}{1-\xi}\left[\ln\left(\frac{p_3^2}{\bar{\mu}^2}\right) + \ln(4\xi(1-\xi))\right] - \frac{\xi(1+\xi)}{1-\xi}\right)_{+(1)}, & 0 < \xi < 1 \\ \left(-\frac{1+\xi^2}{1-\xi}\ln\left(\frac{-\xi}{1-\xi}\right) - 1 + \frac{3}{2(1-\xi)}\right)_{+(1)} - \frac{3}{2(1-\xi)}, & \xi < 0 \end{cases} + \frac{\alpha_s C_F}{2\pi}\delta(1-\xi)\left(\frac{3}{2}\ln\left(\frac{\mu_F^2}{4\bar{\mu}^2}\right) + \frac{5}{2}\right). \tag{38}$$

For the rest of this paper, we will employ the usual choice of renormalization and factorization scales, $\mu_F = \bar{\mu}$. Equation (38) has an imbalance when integrated, because $\tilde{\Gamma}_{\gamma^0}$ outside the physical region picks up a logarithmic divergence in the momentum fraction as $\xi \to \pm\infty$, while this divergence has already been removed from the self-energy part in Eq. (36). This implies an anomalous nonconservation of the quark number, even if the classical current is conserved. In Ref. [45], $C_{\gamma^0}^{\overline{\text{MS}}}$ has also been computed and a plus prescription at infinity was used to cancel the UV divergences between the vertex and self-energy corrections. This explains the source of the difference between our result in Eq. (38) and that in Eq. (68) of Ref. [45]. Namely, in our case, the term proportional to the Dirac δ function depends on the ratio of the factorization and renormalization scales, while in the case of Ref. [45], it depends on the quark momentum p_3. Both prescriptions, however, do not conserve quark number.

To treat the unbalanced divergence, we introduce a modified $\overline{\text{MS}}$ scheme ($\text{M}\overline{\text{MS}}$), which has already been discussed in the previous section and used in our previous work [19,20]. In this scheme, an extra subtraction is made outside the physical region of the unintegrated vertex corrections, which, in practice, renormalizes the ξ dependence for $\xi > 1$ and $\xi < 0$ and removes the potential divergences,

$$\tilde{Z}^{\overline{\rm MMS}}_{\Gamma_{\gamma^0}}(\xi) = 1 - \frac{\alpha_s}{2\pi} C_F \frac{3}{2}\left(-\frac{1}{\xi}\theta(\xi-1) - \frac{1}{1-\xi}\theta(-\xi)\right) - \frac{\alpha_s C_F}{2\pi}\delta(1-\xi)\left(\frac{3}{2}\ln\left(\frac{1}{4}\right)+\frac{5}{2}\right). \quad (39)$$

One can write Eq. (39) in z space, noticing that ξ outside the physical region is now renormalized, at the scale $\bar{\mu}$. The inverse Fourier transform is then from ξ to $z\bar{\mu}$, and the extra subtraction in z space is written as

$$\begin{aligned} Z^{\overline{\rm MMS}}_{\Gamma_{\gamma^0}}(z\bar{\mu}) = 1 &- \frac{\alpha_s}{2\pi} C_F\left(\frac{3}{2}\ln\left(\frac{1}{4}\right)+\frac{5}{2}\right) \\ &+ \frac{3}{2}\frac{\alpha_s}{2\pi} C_F\left(i\pi\frac{|z\bar{\mu}|}{2z\bar{\mu}} - \mathrm{Ci}(z\bar{\mu}) + \ln(z\bar{\mu}) - \ln(|z\bar{\mu}|) - i\mathrm{Si}(z\bar{\mu})\right) \\ &- \frac{3}{2}\frac{\alpha_s}{2\pi} C_F e^{iz\bar{\mu}}\left(\frac{2\mathrm{Ei}(-iz\bar{\mu}) - \ln(-iz\bar{\mu}) + \ln(iz\bar{\mu}) + i\pi\mathrm{sgn}(z\bar{\mu})}{2}\right). \end{aligned} \quad (40)$$

For consistency, Eq. (40) has been also applied to the renormalization functions to bring them to the $\overline{\rm MMS}$ scheme, as described in Sec. IV C. Thus, these renormalization functions are obtained as follows: the Z factors calculated in the RI′ scheme are converted to the $\overline{\rm MS}$ scheme according to the perturbative formulas of Ref. [21] and then multiplied by the factor given in Eq. (26) to convert them to the $\overline{\rm MMS}$ scheme. An important self-consistency check is that the expression of Eq. (40) must cancel the $z \to 0$ divergence in $\ln(z^2)$ present in the $\overline{\rm MS}$ scheme [21,47]. Indeed, in the limit $z \to 0$, one has

$$\begin{aligned} Z^{\overline{\rm MMS}}_{\Gamma_{\gamma^0}}(z \to 0) &= 1 - \frac{\alpha_s C_F}{2\pi}\left(\frac{3}{2}\ln\left(\frac{\bar{\mu}^2 z^2 e^{2\gamma_E}}{4}\right)+\frac{5}{2}\right) \\ &= Z^{\rm ratio}_{\Gamma_{\gamma^0}}(z\bar{\mu}), \end{aligned} \quad (41)$$

i.e., the Z factor of the vertex corrections at $z = 0$ is the same in our scheme and in the "ratio" scheme introduced in Ref. [45] and both cancel the divergence. The latter scheme was proposed as an alternative to our solution of the current conservation problem when using the pure $\overline{\rm MS}$ expression of Eq. (68) of Ref. [45]. The "ratio" scheme is another modification of the $\overline{\rm MS}$ scheme that also needs an additional conversion factor in order to bring the renormalization functions in the $\overline{\rm MS}$ scheme into the "ratio" scheme. The difference with respect to our solution is that the form of the matching kernel [or of $Z^{\rm ratio}_{\Gamma_{\gamma^0}}(z\bar{\mu}) = \int \frac{d\xi}{2\pi} e^{i\xi z\bar{\mu}} \tilde{Z}^{\overline{\rm MMS}}_{\Gamma_{\gamma^0}}(\xi)$ after doing the Fourier transform to z space] implies that the ξ dependence of the matching equation in the physical region is also renormalized, while our approach does not modify this region. Thus, the "ratio" scheme can induce potentially large modifications to the matched PDF, as we show below numerically.

We apply $\tilde{Z}^{\overline{\rm MMS}}_{\Gamma_{\gamma^0}}(\xi)$ to Eq. (38) to obtain a matching kernel that keeps the norm of the nonsinglet distributions unchanged. We also include the results for the γ^3 and $\gamma^3\gamma^5$ Dirac structures. Because quarks are taken to be massless, the one-loop corrections are the same for the γ^3 and $\gamma^3\gamma^5$ cases. Compared to the γ^0 case, a shift of $+2(1-\xi)$ inside the plus prescription in the physical region of Eq. (38) is needed, and also, the factor of $5/2$ in the last line of Eq. (38) becomes $7/2$. The matching equation is then written as

$$q^{\overline{\rm MS}}(x,\bar{\mu}) = \int_{-\infty}^{+\infty}\frac{dy}{|y|} C^{\overline{\rm MMS}}\left(\frac{x}{y},\frac{\bar{\mu}}{p_3}\right)\tilde{q}^{\overline{\rm MMS}}(y,P_3,\bar{\mu}), \quad (42)$$

with the matching kernels, for the different Dirac structures, given by

$$C^{\overline{\rm MMS}}_{\gamma^0,\gamma^3,\gamma^3\gamma^5}\left(\xi,\frac{\bar{\mu}}{p_3}\right) = \delta(1-\xi) + \frac{\alpha_s C_F}{2\pi}\begin{cases} \left(\frac{1+\xi^2}{1-\xi}\ln\left(\frac{\xi}{\xi-1}\right)+1+\frac{3}{2\xi}\right)_{+(1)}, & \xi > 1, \\ \left(\frac{1+\xi^2}{1-\xi}\left[\ln\left(\frac{p_3^2}{\bar{\mu}^2}\right)+\ln(4\xi(1-\xi))\right] - \frac{\xi(1+\xi)}{1-\xi} + 2\iota(1-\xi)\right)_{+(1)}, & 0 < \xi < 1, \\ \left(-\frac{1+\xi^2}{1-\xi}\ln\left(\frac{-\xi}{1-\xi}\right)-1+\frac{3}{2(1-\xi)}\right)_{+(1)}, & \xi < 0, \end{cases} \quad (43)$$

where $\iota = 0$ for γ^0, and $\iota = 1$ for γ^3 and $\gamma^3\gamma^5$. We note that an alternative procedure to get $\overline{\rm MS}$-renormalized light-cone PDFs is to directly match to them from RI-renormalized quasi-PDFs. The relevant formulas for this were derived in Ref. [44] for the operators with Dirac structures γ^3 and $\gamma^3\gamma^5$ and in Ref. [29] for γ^0. Such a procedure is equivalent, to

one-loop order, to the one described above, but with different higher-order effects. This choice is investigated numerically below, by comparing the direct matching from a variant of the RI scheme to the $\overline{\text{MS}}$ scheme with a two-step procedure that first brings the renormalization functions to the M$\overline{\text{MS}}$ scheme and then performs the M$\overline{\text{MS}}$-to-$\overline{\text{MS}}$ matching.

For the transversity case, the computation (albeit simpler) is similar because the usual one-loop vertex correction is zero in the Feynman gauge. The quark self-energy is not zero, but the unbalanced terms cancel each other in the matching equation, since they are the same for $\Pi_{\gamma^3\gamma^j}$ and $\tilde{\Pi}_{\gamma^3\gamma^j}$. Thus, $Z^{\text{M}\overline{\text{MS}}}_{\Gamma_{\gamma^3\gamma^j}}(z\bar{\mu})$ is given by the same functional form as Eq. (40), with the replacement of the numerical factors $3/2$ and $5/2$ by 2. $C^{\text{M}\overline{\text{MS}}}_{\gamma^3\gamma^j}$ is given by

$$C^{\text{M}\overline{\text{MS}}}_{\gamma^3\gamma^j}\left(\xi,\frac{\bar{\mu}}{p_3}\right)=\delta(1-\xi)+\frac{\alpha_s}{2\pi}C_F\begin{cases}\left[\frac{2\xi}{1-\xi}\ln\left(\frac{\xi}{\xi-1}\right)+\frac{2}{\xi}\right]_{+(1)}, & \xi>1,\\ \left[\frac{2\xi}{1-\xi}\left(\ln\left(\frac{p_3^2}{\bar{\mu}^2}\right)+\ln(4\xi(1-\xi))\right)-\frac{2\xi}{1-\xi}\right]_{+(1)}, & 0<\xi<1,\\ \left[-\frac{2\xi}{1-\xi}\ln\left(\frac{\xi}{\xi-1}\right)+\frac{2}{1-\xi}\right]_{+(1)}, & \xi<0.\end{cases}\tag{44}$$

The matching kernel for the transversity has been calculated previously, with a hard cutoff and a quark mass to regularize the UV and the IR divergences, respectively, but without renormalization [34]. More recently, following the idea of Ref. [44] to match directly from RI-renormalized quasi-PDFs to light-cone PDFs in the $\overline{\text{MS}}$ scheme, the matching for the transversity was also derived in Ref. [31].

B. Comparison of matching in the $\overline{\text{MS}}$, M$\overline{\text{MS}}$ and ratio schemes

Our choice for the matching to extract the light-cone PDFs from quasi-PDFs is to use a minimal modification of the $\overline{\text{MS}}$ scheme that ensures current conservation. Thus, the matching formula is applied on a quasi-PDF renormalized in a modified $\overline{\text{MS}}$ scheme, the M$\overline{\text{MS}}$ scheme, to yield the corresponding light-cone PDF in the pure $\overline{\text{MS}}$ scheme. The difference between the $\overline{\text{MS}}$ and M$\overline{\text{MS}}$ schemes was also taken into account at the level of the renormalization functions in Sec. IV. Note that in our previous work [19,20], we did not modify the conversion and, thus, our quasi-PDFs were renormalized in the standard $\overline{\text{MS}}$ scheme. The formulas for the conversion modification, derived in the current paper, were not available at the time of our previous work and we argued that the numerical effect of the modification is subleading with respect to other uncertainties (as the modification in the $\overline{\text{MS}}$ is minimal and only in the unphysical regions). Moreover, it disappears in the infinite-momentum limit. Having now derived the relevant formulas to convert renormalization functions from the $\overline{\text{MS}}$ scheme to M$\overline{\text{MS}}$, we test this and validate that it is indeed a small effect.

In Fig. 24, we compare matched PDFs obtained from quasi-PDFs renormalized in the $\overline{\text{MS}}$ scheme, where the additional conversion of Eq. (26) is not included, and in the modified M$\overline{\text{MS}}$ scheme. As anticipated, the numerical effect is very small, particularly in the unpolarized case. Nevertheless, it is important to take the conversion modification into account to have a self-consistent procedure. We note that the modification brings the matched PDFs slightly towards phenomenological extractions, which is reassuring.

Another possibility to obtain a matching formula with current conservation is to use the procedure proposed in Ref. [45], the so-called "ratio" scheme, already mentioned in Sec. V A. It consists in a different way of changing the $\overline{\text{MS}}$ scheme to achieve current conservation, including a modification of the physical region. Thus, the effect on the matched PDFs is expected to be larger numerically than when using the M$\overline{\text{MS}}$ scheme. We show the comparison of the matching from our previous procedure and from the "ratio" scheme in Fig. 25. Indeed, the "ratio" scheme has a noticeable effect on the matched PDF, which is particularly large in the small-x region. Theoretically, both the "ratio" and M$\overline{\text{MS}}$ schemes are valid modifications of the $\overline{\text{MS}}$ scheme. Obviously, they are equivalent at the one-loop level, but they have significantly different higher-order effects. Hence, they can both be used to match quasi-PDFs to light-cone PDFs and the difference between them can serve as an estimate of truncation effects at one loop. For the remainder of the paper, we use the M$\overline{\text{MS}}$ scheme, as the one for which the higher-order effects are likely to be smaller, since this scheme is a milder modification of the $\overline{\text{MS}}$ scheme, i.e., one that does not affect the physical region in the matching.

C. Comparison of $\overline{\text{MS}} \to \overline{\text{MS}}$ and RI $\to \overline{\text{MS}}$ matching

As we discussed above, our renormalization and matching procedure proceeds in two steps. First, the renormalization functions computed in the RI′ scheme are converted

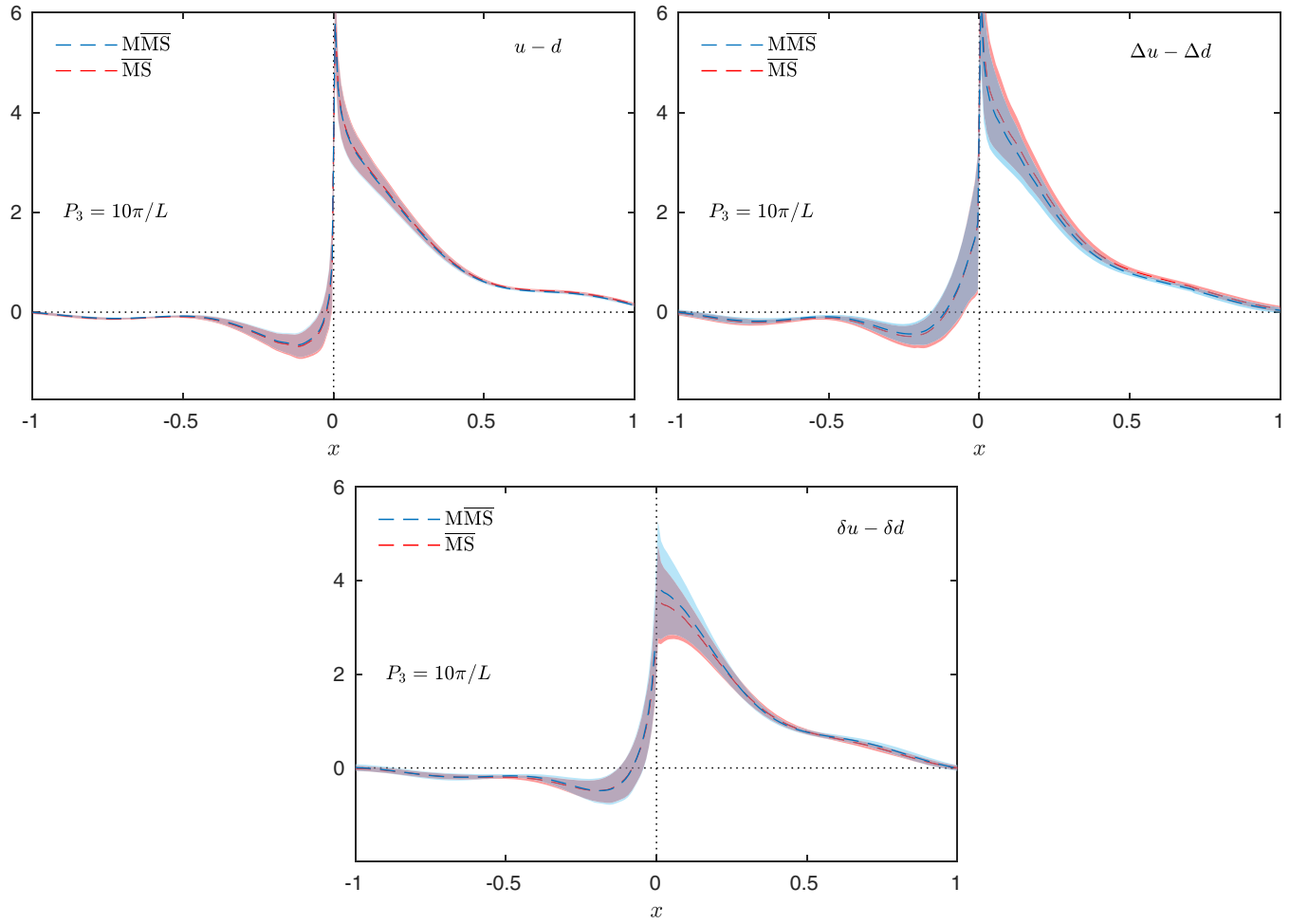

FIG. 24. Comparison of matched PDFs from quasi-PDFs renormalized in the $\overline{\text{MS}}$ (red band) scheme and in the M$\overline{\text{MS}}$ (cyan band) scheme, for the unpolarized (top left panel), helicity (top right panel) and transversity (bottom panel) cases and for a nucleon momentum of $P_3 = 10\pi/L$.

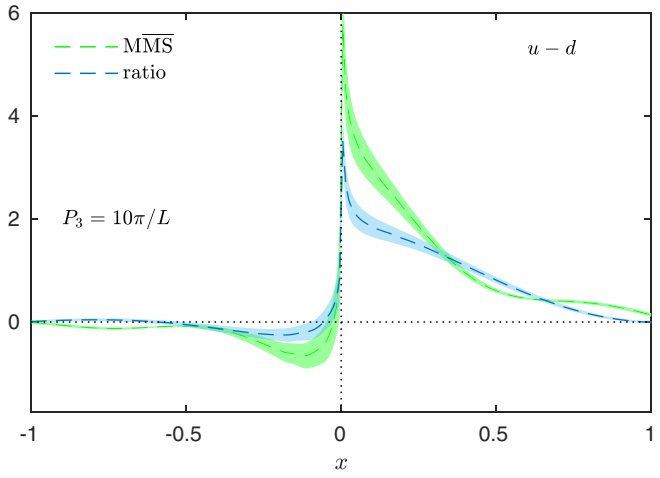

FIG. 25. Comparison of matched PDFs from quasi-PDFs renormalized in the M$\overline{\text{MS}}$ scheme (green band) and in the "ratio" scheme (cyan band) for the unpolarized case. The nucleon momentum is $P_3 = 10\pi/L$.

to the M$\overline{\text{MS}}$ scheme and evolved to a chosen scale, $\bar{\mu} = 2$ GeV, using one-loop perturbative formulas derived in Ref. [21], and including the additional conversion from the $\overline{\text{MS}}$ scheme to the M$\overline{\text{MS}}$ scheme, derived in this work. The resulting M$\overline{\text{MS}}$ Z factors are used to renormalize the bare matrix elements, which are then Fourier transformed from z space to x space, yielding the renormalized quasi-PDF in the M$\overline{\text{MS}}$ scheme. In the second step, a matching procedure is applied that brings the M$\overline{\text{MS}}$-renormalized quasi-PDF to the light-cone PDF in the $\overline{\text{MS}}$ scheme, at the same renormalization scale $\bar{\mu}$. We will refer to this procedure as a two-step procedure.

In Ref. [44], a one-step procedure was proposed. It consists in applying the renormalization functions computed in the RI scheme to the bare matrix elements and taking the Fourier transform of RI-renormalized matrix elements to obtain a quasi-PDF renormalized in the RI scheme. The matching procedure then serves to simultaneously bring the quasi-PDFs to the light-cone PDFs and to

convert from the RI scheme to the $\overline{\text{MS}}$ scheme, evolving them from the given RI scale to the reference scale of 2 GeV. For this procedure, we use the formulas derived in Ref. [29] and consider the operator with the γ_0 Dirac structure. We choose the variant with the $\not{p}$ projection that differs from the projection that we used in our variant of the RI′ scheme. We refer to this choice of the projection as $\text{RI}_{\not{p}}$.

The comparison of the one-step and two-step procedures is presented in Fig. 26. In the left panel, we show our quasi-PDFs, renormalized in the $\text{RI}_{\not{p}}$ scheme and in the $\text{M}\overline{\text{MS}}$ scheme. The corresponding renormalization functions are applied to the same lattice data for bare matrix elements. The results of applying the one-step and two-step procedures are given in the right panel. We observe that the results of both procedures are fully compatible for small positive and negative x. For large positive x, the two-step procedure gives a PDF that goes to zero more slowly, while the PDF from the one-step procedure crosses zero at $x \approx 0.6$ and remains negative until $x \approx 1$ (reaching a minimum of around -0.16 at $x \approx 0.8$). In the large negative-x region, the situation is analogous: the one-step (two-step) procedure leads to the PDF approaching zero from below (above).

Both the one-step and two-step procedures use a one-loop computation in continuum perturbation theory. Obviously, both contain higher-order contributions that cannot be quantified unless the two-loop formulas are available. Hence, neither of them can be considered to be the preferred one. It might happen that one of the procedures evinces smaller higher-loop effects without any theoretical arguments to support this, but it is not possible to say which one. The comparison between them is, thus, useful, as it reveals systematic effects due to higher-order terms, which may differ for different x regions. The present study suggests that the large-$|x|$ regions suffer more from such higher-order effects. As Fig. 26 indicates, there may be significant two-loop effects in the conversion/evolution/matching procedures. Thus, a two-loop computation is indeed necessary. The two-step procedure allows for separating the conversion, the evolution and the matching and hence, a two-loop computation even for one of these, which is simpler than the full computation of the one-step procedure, can provide some insight.

D. Truncation of the Fourier transform

We turn now to the investigation of systematic effects related to the FT that is applied on the renormalized matrix elements in z space, giving the quasi-PDFs in x space. In principle, the FT integrates over all Wilson line lengths from zero to infinity, whereas on the lattice we have available for a finite number of discrete lengths z/a. Hence, one needs to decide about the maximum value of z/a taken in the discretized FT integral. Ideally, it should be a value for which both the real part and the imaginary part of the matrix elements have decayed to zero. For unrenormalized matrix elements, the choice of such a value poses no problem. As can be seen in Fig. 27, the bare matrix elements are zero for any $|z/a| \gtrsim 15$. However, renormalized matrix elements involve multiplication of the bare ones by complex Z factors that mix bare real and imaginary parts, according to

$$\text{Re}[h^{\text{ren}}] = \text{Re}[Z]\text{Re}[h^{\text{bare}}] - \text{Im}[Z]\text{Im}[h^{\text{bare}}], \quad (45)$$

$$\text{Im}[h^{\text{ren}}] = \text{Re}[Z]\text{Im}[h^{\text{bare}}] + \text{Im}[Z]\text{Re}[h^{\text{bare}}]. \quad (46)$$

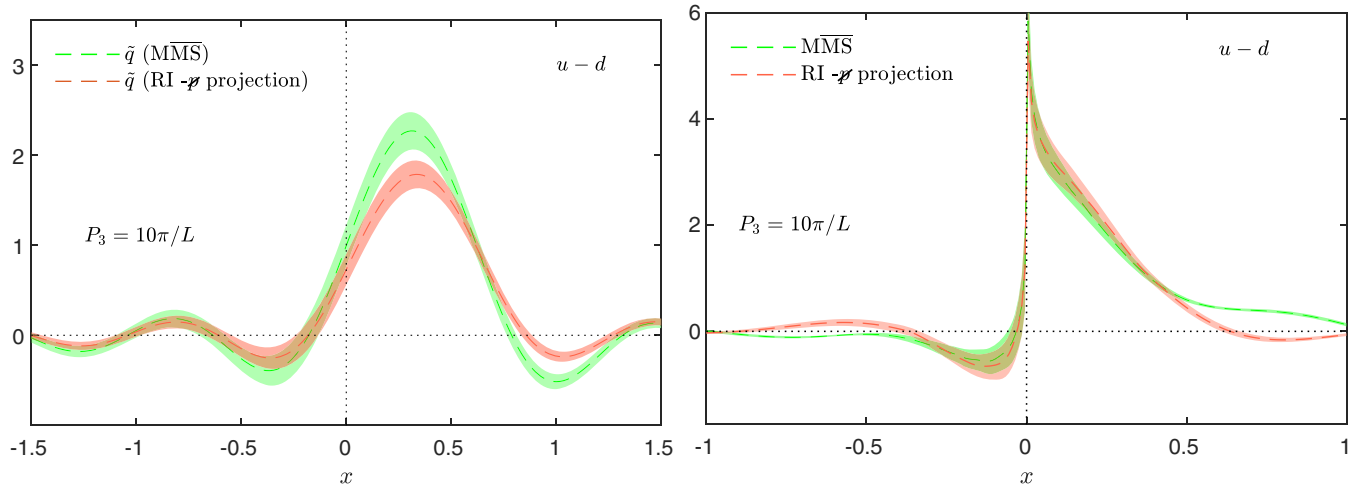

FIG. 26. Left panel: Comparison of renormalized quasi-PDFs in the RI scheme with $\not{p}$ projection (red band) and the $\text{M}\overline{\text{MS}}$ scheme (green band). Right panel: Matched PDFs obtained from the same bare matrix elements, but with different scheme conversion, evolution and matching procedures: either one-step (RI $\to$ $\overline{\text{MS}}$) or two-step (RI′ $\to$ $\text{M}\overline{\text{MS}}$ $\to$ $\overline{\text{MS}}$) (see text for details). This is for the unpolarized case, with a nucleon momentum of $P_3 = 10\pi/L$.

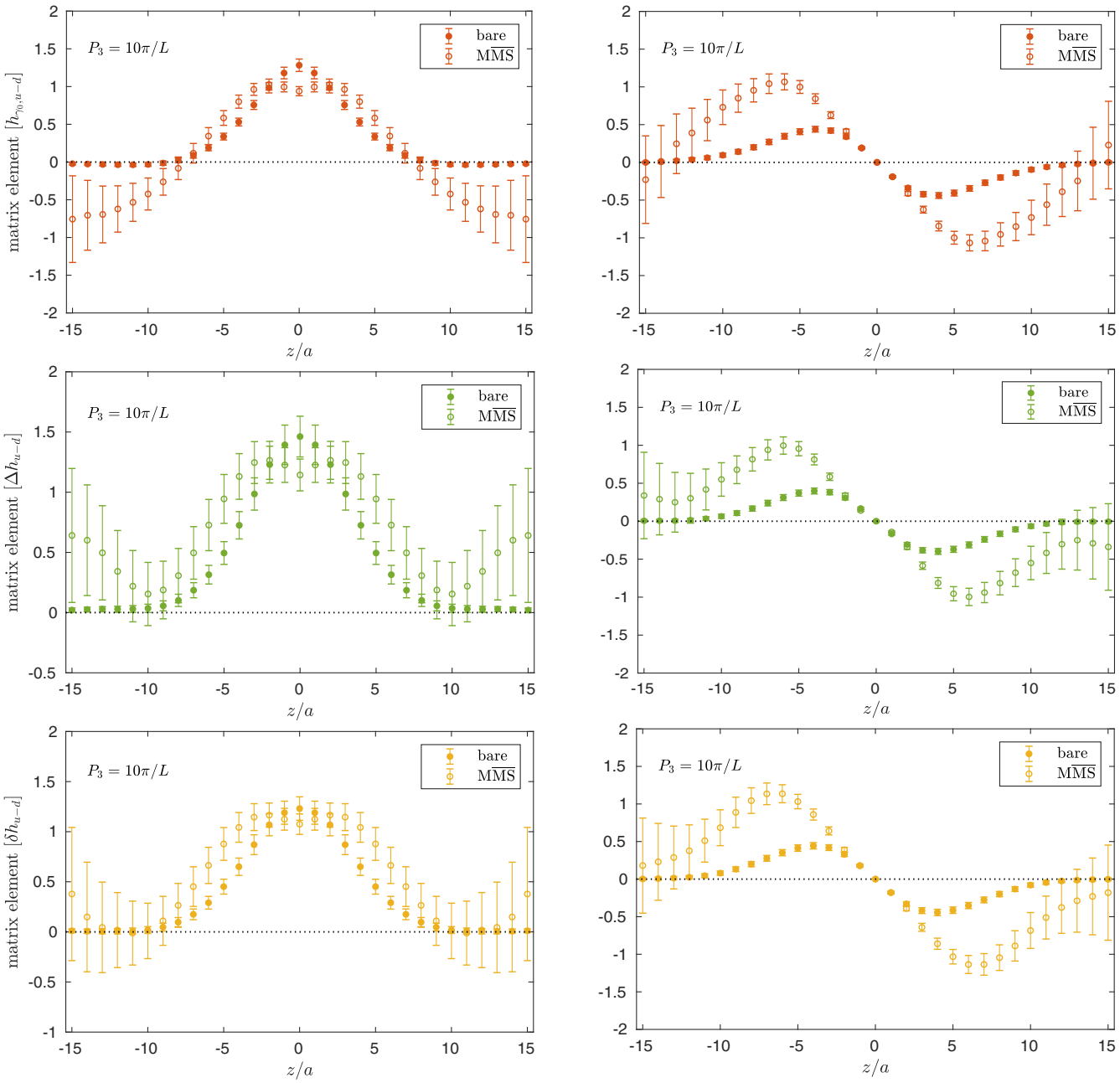

FIG. 27. Real (left) and imaginary (right) parts of the bare (filled symbols) and $\overline{\text{MMS}}$-renormalized (open symbols) matrix elements for the unpolarized (upper row), helicity (middle row) and transversity PDFs (lower row) (for a nucleon momentum of $P_3 = 10\pi/L$).

As a result, the real part of the renormalized matrix elements can be nonzero even if the real part of the bare matrix elements has already decayed to zero. This stems from an unphysical truncation effect in the perturbative conversion between the intermediate RI′ scheme Z factors and the $\overline{\text{MS}}$ ones. Analogously, the imaginary part of the renormalized matrix elements gets contaminated by $\text{Im}[Z]$ which multiplies the real part of the bare matrix elements and leads to nonphysical contributions. Moreover, the large values of Z factors at large Wilson line lengths amplify the bare matrix elements and even if the latter are compatible with zero, this amplification introduces very large noise to the data, which propagates through the FT and matching to the final PDFs. These effects call for a truncation of the renormalized matrix elements at some justified value of z/a, which we call z_{max}/a. In Fig. 28, we illustrate the z_{max}/a dependence of the resulting final PDFs for all operators and for our largest momentum. The smallest (largest) value of z_{max}/a is chosen according to where the real (imaginary) part of the renormalized matrix elements is compatible with zero.

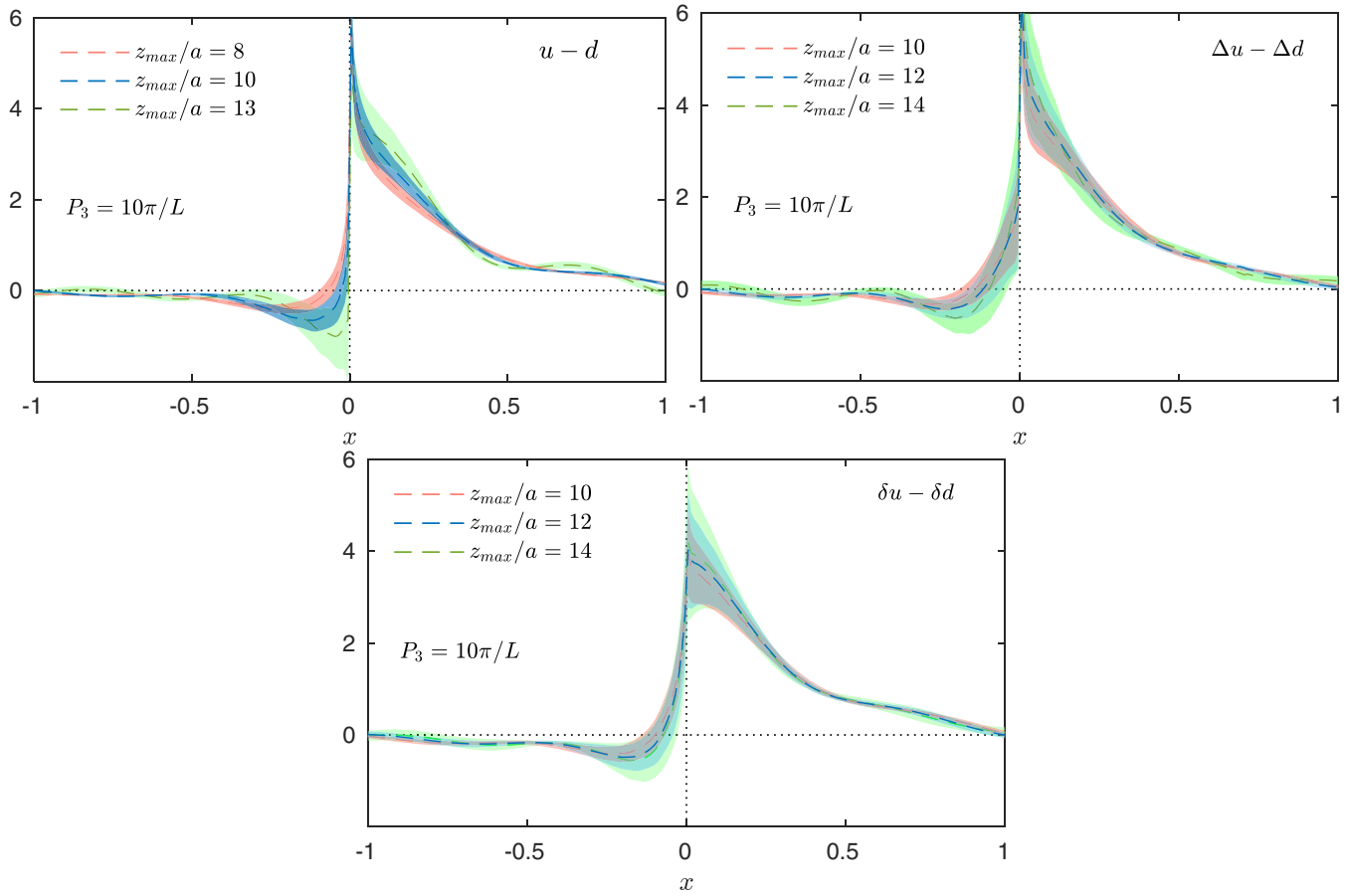

FIG. 28. Matched unpolarized PDF (upper right), helicity PDF (lower left) and transversity PDF (lower right) for different values of the cutoff $z_{\max}/a = 10$ (red), 12 (blue), 14 (green). Our final choice is always the middle value of $z_{\max}/a$ from the ones shown. The nucleon momentum is $P_3 = 10\pi/L$ for all plots.

For the transversity case (lower plot of Fig. 28), we observe very little dependence on the choice of $z_{\max}/a$ (apart from increased statistical noise for larger values of $z_{\max}/a$) and we choose $z_{\max}/a = 12$ as the value for our final plots. In this case, both the real and imaginary parts of the renormalized matrix elements are compatible with zero. Hence, the unphysical effects of truncating the Fourier transform at finite $z_{\max}/a$ are minimal. For the unpolarized case, a value of $z_{\max}/a$ where $\mathrm{Re}[h^{\mathrm{ren}}]$ and $\mathrm{Im}[h^{\mathrm{ren}}]$ are both zero does not exist. Hence, we observe that the variation of $z_{\max}/a$ leads to a visible effect in the upper left plot of Fig. 28. However, matched PDFs from all considered values of the cutoff are compatible with one another. In the end, we choose the middle value, $z_{\max}/a = 10$, as the final one. In this way, the unphysical effect of truncation is split between the real and imaginary parts. For the former, the matched PDF gets a contribution from unphysical negative values of $\mathrm{Re}[h^{\mathrm{ren}}]$ (resulting from the imaginary part of Z factors being nonzero due to the truncation of the perturbative conversion between renormalization schemes). In turn, for the latter, part of the contribution from the imaginary part of the renormalized matrix elements is missing. In the helicity case, $\mathrm{Re}[h^{\mathrm{ren}}]$ decays to zero at a similar value of z/a as for the other cases and, contrary to the unpolarized matrix elements, does not go below zero. However, $|\mathrm{Im}[h^{\mathrm{ren}}]|$, reaches a minimum around $z/a \approx 12$–13 and then increases again, to become compatible with zero at $z/a \approx 17$–18, within huge errors. For our preferred value of the cutoff, we choose $z_{\max}/a = 12$, where the real part is compatible with zero and the absolute value of the imaginary part is locally minimal. Varying $z_{\max}/a$ (upper right plot of Fig. 28) leads to changes in the matched helicity PDF that are statistically insignificant.

Apart from the nonphysical effects appearing when the renormalized matrix elements are not compatible with zero at $z_{\max}/a$, further nonphysical effects are introduced by the FT if the decay of h^{ren} is not fast enough. Namely, the periodicity of the Fourier transform leads to unphysical oscillations in the quasi-PDFs and finally in the matched PDFs. This effect is visible in all of our final distributions as a distorted approach of PDFs to zero for x between 0.5 and 1, as a mild oscillatory behavior for large negative x, and as an unphysical minimum in the antiquark part at $-x \approx 0.1$–0.2. We note that these symptoms

are interconnected, e.g., vector current conservation implies that a too large integral of the quark part, caused by the FT truncation effects, must lead to a too low value of the integral in the antiquark part, signaled by the negative value of $\bar{d}-\bar{u}$ for low negative x. Decomposing the FT into cosine and sine transforms, one can observe that the quasi-PDF is bound to be negative if the matrix elements entering the transforms do not decay to zero fast enough. The rate of decay of these matrix elements depends on the nucleon boost, as was demonstrated in Fig. 7. For larger momenta, the decay becomes faster, i.e., zero is reached at a smaller value of z/a, for both the real and imaginary parts. Thus, the unphysical oscillations that we observe in the data clearly depend on the nucleon boost. However, as argued in a study performed in parallel to our work [112], the issue is more general and profound. They showed that the problem of reconstructing a distribution from a naively discretized and truncated Fourier transform, i.e., with only a few available values of the matrix elements (corresponding to integer values of z/a) extending to a finite Wilson line length (up to some $z_{\max}/a$), is mathematically ill defined. They advocated the use of advanced reconstruction techniques that can improve the situation considerably. Among such techniques that were tested in that paper are the Backus-Gilbert method, a neural-network reconstruction and Bayesian methods. Applying them to our current data is beyond the scope of this work, but will be pursued by us in the future.

We also remark that the FT problem is aggravated in the process of renormalization. In Fig. 29, we display the results for the unpolarized quasi-PDF for $z_{\max}=10a$, comparing the bare, RI′ and $\overline{\text{MS}}$ renormalized cases. As can be seen, the oscillations do not emerge when Fourier transforming bare matrix elements that decay to zero faster. The amplification of the matrix elements by the Z factors is responsible for their relatively slow decay, which enhances the reconstruction problem. At this stage, one needs to remember that the Z factors, as calculated now, are subject to two kinds of nonphysical effects: lattice artifacts in their nonperturbative evaluation on a hypercubic lattice in the RI′ scheme, and truncation effects in the perturbative conversion to the $\overline{\text{MS}}$ scheme and evolution to the reference scale of 2 GeV. After curing these effects, by computing the higher-order conversion formula and by subtracting lattice artifacts computed in lattice perturbation theory [111], we expect that the oscillatory behavior may be considerably less prominent, but we do not expect that the final PDFs will be devoid of it. This is supported also by the behavior observed for the quasi-PDF renormalized using the RI′ scheme, which has no conversion truncation effects, where we find that the oscillation is reduced with respect to the $\overline{\text{MS}}$-renormalized quasi-PDF, as seen in Fig. 29. However, even in the RI′ scheme, the bare matrix elements are still amplified by the Z factors, whose real part

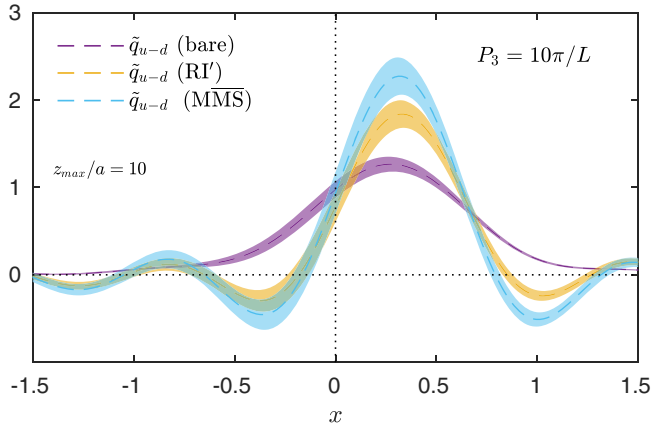

FIG. 29. The unpolarized quasi-PDFs for $P_3=10\pi/L$. We show the bare matrix element (purple), RI′ (yellow) and $\overline{\text{MS}}$ (cyan) quasi-PDFs for $z_{\max}=10a$.

is exponentially increasing due to the presence of the power divergence related to the Wilson line.

An approach that may address nonphysical oscillations was proposed in Ref. [24]. The idea is to rewrite the FT that yields the quasi-PDF, using integration by parts as follows:

$$\tilde{q}(x)=h(z)\frac{e^{ixzP_3}}{2\pi ix}\Big|_{-z_{\max}}^{z_{\max}}-\int_{-z_{\max}}^{z_{\max}}\frac{dz}{2\pi}\frac{e^{ixzP_3}}{ix}h'(z), \qquad (47)$$

where $h'(z)$ is the derivative of the matrix elements with respect to the Wilson line length z. We note that the difference between the standard FT and the "derivative" method vanishes if $h(|z|\geq z_{\max})=0$. The integration-by-parts step is exact; however, the proposed approximation is to neglect the surface term, even if it is nonvanishing. The latter happens when the matrix elements have not decayed to zero at $z=z_{\max}$. The resulting FT, given by the second term of Eq. (47), thus, involves the derivatives of renormalized matrix elements, which decay to zero faster than the matrix elements themselves. As argued above, a fast enough decay of the transformed matrix elements is important to avoid oscillations. In this so-called "derivative" method, the oscillations are effectively transferred to the neglected surface term. We note, however, that this procedure is dangerous, due to the $1/x$ factor of the surface term, which leads to an uncontrolled effect for small values of x, where also the exponential of the surface term has a large real part. This aspect is illustrated in Fig. 30 for all three types of PDFs. The explicit $1/x$ factor amplifies the final PDF for small $|x|$, leading to its huge values, particularly in the small positive-x region. This is very visible in the case of the unpolarized PDF, where the results are statistically the most precise. In the polarized cases, in the small-$|x|$ regions, one observes that the statistical errors are significantly enhanced by the "derivative" method. All three final PDFs evince divergences to $-\infty$ as $x\to 0$ in the antiquark part, in contradiction with the behavior observed

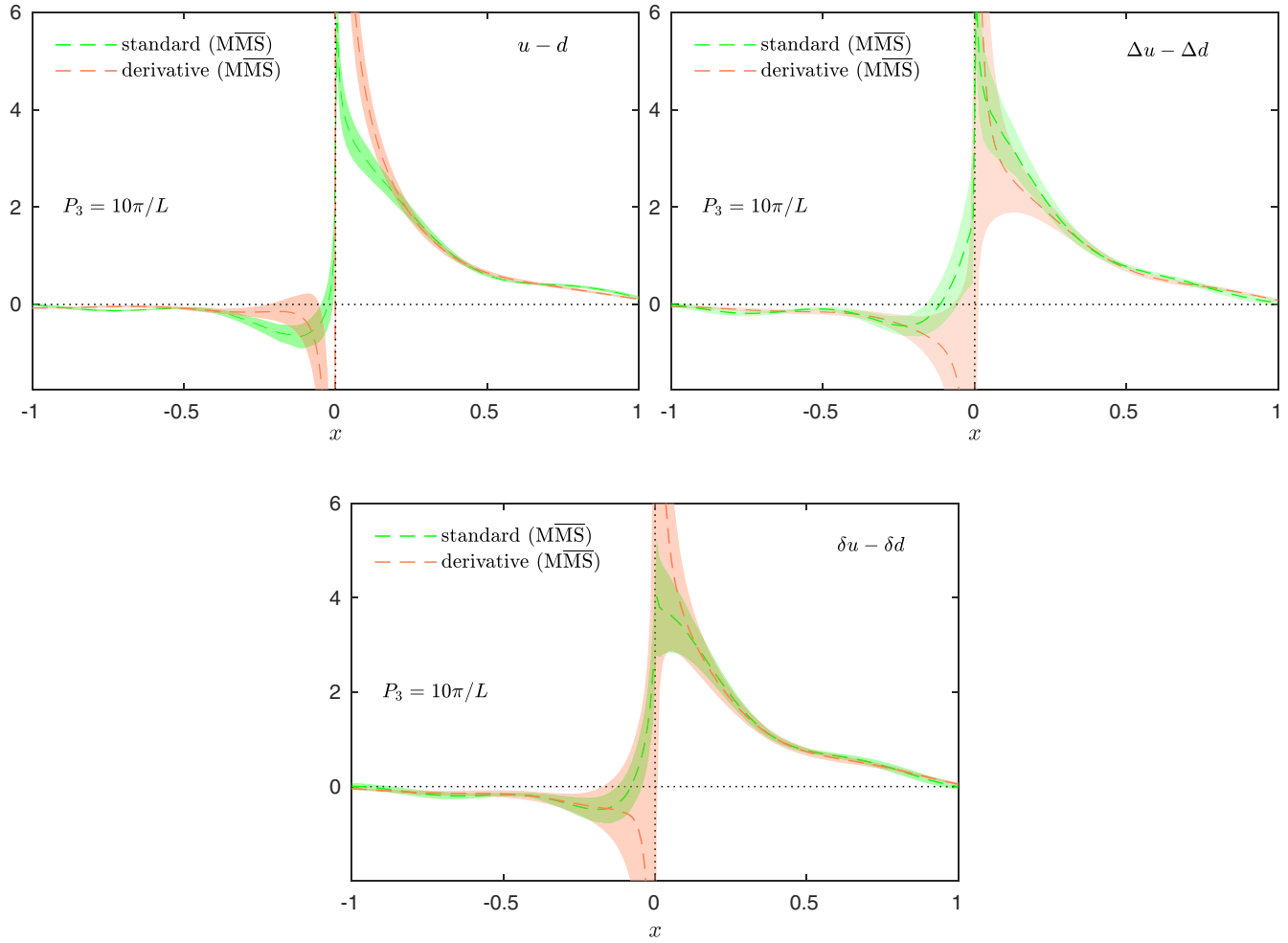

FIG. 30. Comparison of matched PDFs obtained from the same renormalized matrix elements, but with different Fourier transform definitions: the standard one (green band) or the "derivative" one (red band). The matching procedure is the same for both and proceeds via the M$\overline{\text{MS}}$ scheme. We show the unpolarized (upper left), helicity (upper right) and transversity cases (lower). The nucleon momentum is $P_3 = 10\pi/L$.

in phenomenological PDF extractions. Furthermore, the oscillatory behavior is very mildly affected, undermining the very aim of using the "derivative" procedure.

The "derivative" method was also analyzed by the authors of Ref. [112], who showed that this method does not alter the ill-conditioned behavior encountered for a naive truncated FT limited to $\mathcal{O}(10)$ values of z/a. Thus, it does not offer a genuine solution to the FT issue. We are led to conclude that the problem does not disappear, since the Fourier transform that defines the quasi-PDF is still truncated and part of the truncation encoded in the surface term may be hiding physical information relevant for reconstructing the full PDF. In other words, for a fully reliable extraction of PDFs, the surface term has to go to zero when applying actual lattice QCD data, and not by setting it to zero by hand. Another disadvantage of using the "derivative" method is that it introduces additional discretization effects, which need to be controlled. For all the above reasons, we opt not to use the "derivative" method and treat the residual oscillations as an indication that advanced reconstruction methods need to be used instead of the naive approaches. Moreover, this aspect may be improved with a nucleon boost that is increased in a controlled manner (by bare matrix elements that decay to zero faster) and by refinements of the renormalization procedure. All of these are important directions for further study and we plan to explore these ideas in the future.

VI. FINAL RESULTS AND DISCUSSION

In this section, we present our final results for the collinear nonsinglet quark PDFs, whose determination was guided by the extensive investigation of several systematic effects presented in the previous sections. Although it is not possible to quantify all of the systematics, this investigation clearly points to the steps that one must follow in the future in order to arrive at precise extractions of PDFs from lattice QCD, with systematic uncertainties

under control. We will be discussing these steps together with the presentation of the final results.

In Fig. 31, we show the quasi-PDFs, the matched PDFs and the final PDFs that take into account the nucleon mass corrections (NMCs) for the largest value of the momentum. The nucleon mass corrections are performed according to the formulas of Ref. [17]. They are $\mathcal{O}(m_N^2/P_3^2)$ corrections to the PDF resulting from the finite mass of the nucleon and they can be derived in a closed form to all orders. We note that the one-loop matching that connects the quasi-PDFs to the light-cone PDFs is still relatively large for our largest nucleon boost. Increasing the nucleon boost while excited-states contamination is fully under control is one aspect that has to be addressed in future computations. Another is the two-loop perturbative conversion between renormalization schemes and evolution to the chosen reference scale. This will further reduce the dependence on the initial RI′ scale, and may be used to quantify truncation effects in the conversion factor. NMCs, on the other hand, are very small and negligible in most regions of x and thus not expected to contribute to large systematic effects.

In Fig. 32 we show the dependence of the final PDFs on the nucleon boost. For the unpolarized PDF, we observe a strong effect when increasing the momentum from $6\pi/L$ to $8\pi/L$, which seems to converge after that yielding values that are compatible with those obtained for a nucleon boost of $10\pi/L$ over almost the whole range of x. In addition, the oscillatory behavior (at the level of both quasi-PDFs and matched PDFs) becomes milder as the momentum increases. In particular, there is an additional minimum of the final PDFs for the smallest considered momentum, at $x \approx 0.5$ in the unpolarized case (see the upper left plot of Fig. 32), which disappears with momentum increase. As we discussed above, this behavior results from the fact that for larger boosts, both bare and renormalized matrix elements become consistent with zero at smaller Wilson line lengths z; see Fig. 7. For example, the real (imaginary) part of the bare matrix elements corresponding to the unpolarized case becomes compatible with zero for $z/a = 8$ (13) at $P_3 = 10\pi/L$ and for $z/a = 13$ (18) for $P_3 = 6\pi/L$, the latter aggravating the truncation problem, as discussed in the previous section. Moreover, the reconstruction of the

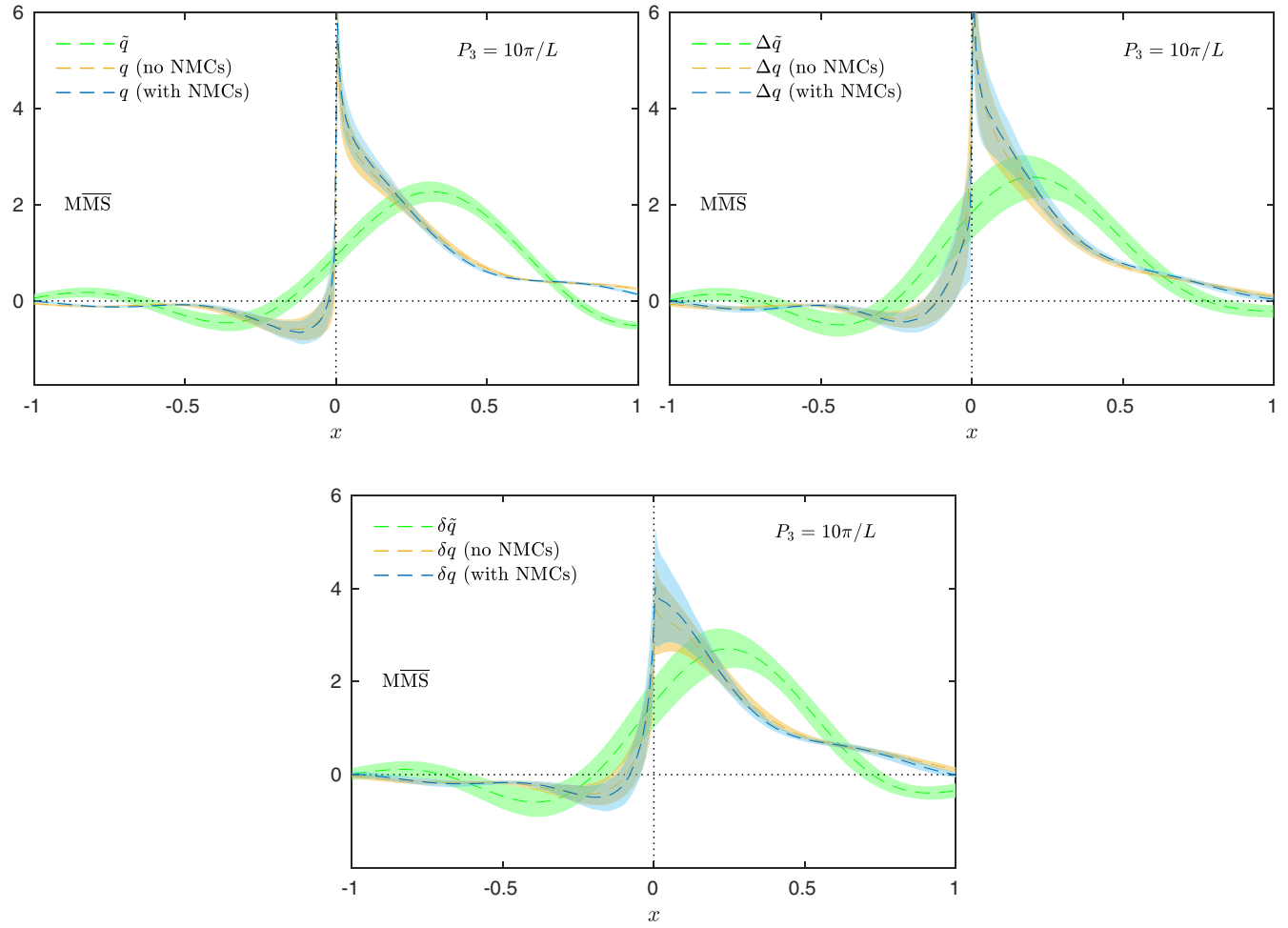

FIG. 31. Effect of matching and nucleon mass corrections in the unpolarized PDF (upper left), helicity PDF (upper right) and transversity PDF (lower). The nucleon momentum is $P_3 = 10\pi/L$ for all plots.

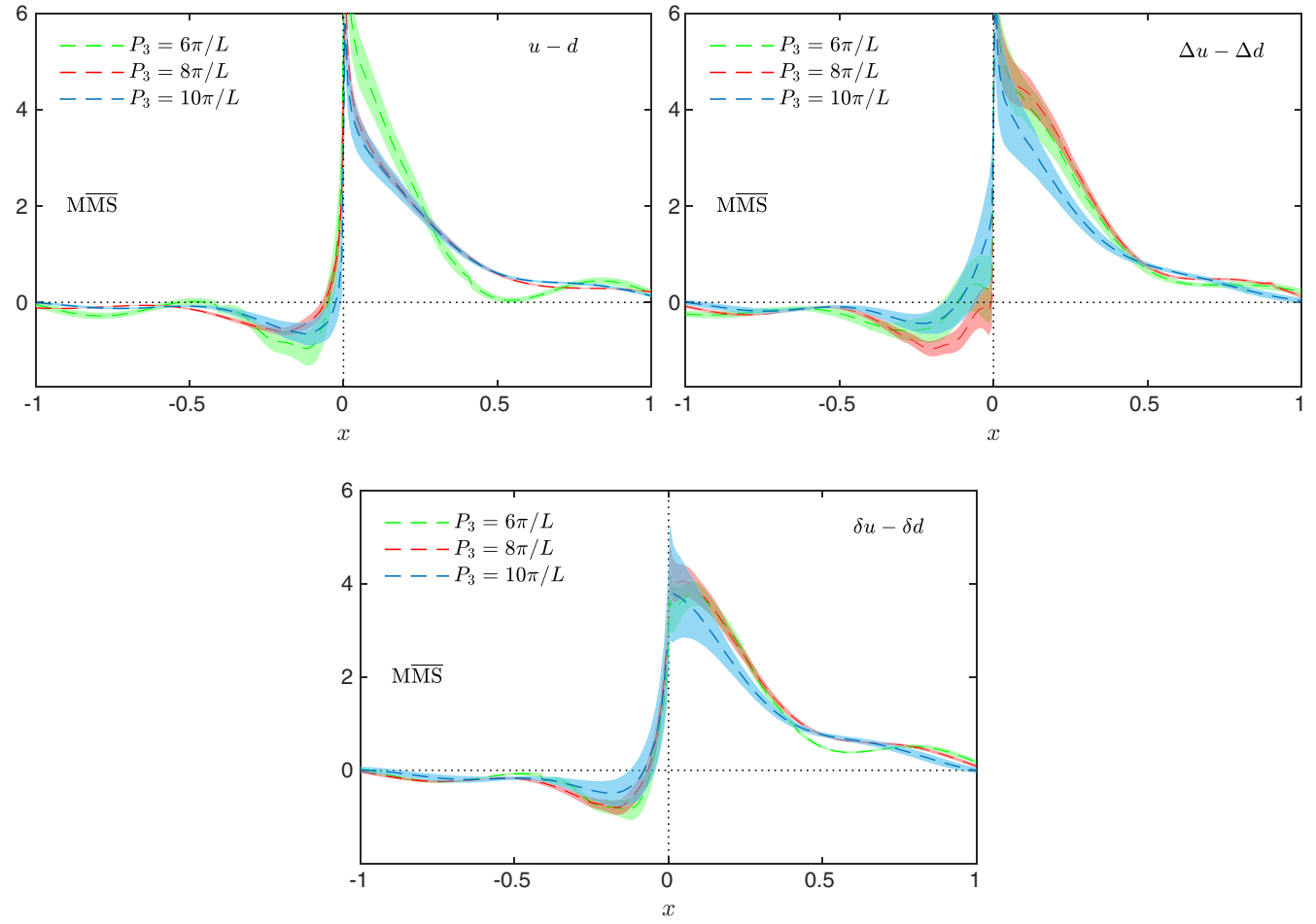

FIG. 32. Nucleon boost dependence for the unpolarized PDF (upper left), helicity PDF (upper right) and transversity PDF (lower), using $P_3 = 6\pi/L$ (green curve), $P_3 = 8\pi/L$ (red curve) and $P_3 = 10\pi/L$ (blue curve).

x-dependent distributions is more problematic with renormalized data, as again argued in Sec. V D. Also, the decay of the matched PDF to zero at $x = \pm 1$ is best achieved at the largest momentum. For the helicity PDF, on the other hand, increasing the momentum from $6\pi/L$ to $8\pi/L$ is a minor effect, i.e., the final PDFs are compatible with each other for both of these momenta, over almost the whole range of x. However, when increasing the momentum from $8\pi/L$ to $10\pi/L$, a significant shift is observed, especially for small values of x. Thus, it is likely that higher-twist effects are more pronounced for the helicity PDF. The oscillatory behavior is also milder in this case as the momentum increases, and the final PDF is compatible with zero at $x = 1$, which is not the case for the two smaller momenta. For the transversity PDF, the momentum dependence is the mildest and results using momenta of $8\pi/L$ and $10\pi/L$ are compatible over a wide range of x values. Similarly to the unpolarized and helicity PDFs, the unphysical oscillations are significantly reduced and the transversity PDF vanishes at $x = 1$ for a momentum of $10\pi/L$.

In Fig. 33, we analyze the momentum dependence of the final matched PDFs more closely. In the unpolarized case, we observe compatibility between the data for the largest two momenta for all values of x illustrated in the upper panel. This is in accordance with theoretical expectations: the higher-twist effects (HTEs) are suppressed by $\mathcal{O}(\Lambda^2_{\text{QCD}}/P_3^2)$, which amounts to $\mathcal{O}(5\%)$, $\mathcal{O}(7\%)$ and $\mathcal{O}(13\%)$, for $P_3 = 10\pi/L$, $P_3 = 8\pi/L$ and $P_3 = 6\pi/L$, respectively. With our statistical precision, around 10% for most ranges of x, HTEs are expected to be hidden within statistical uncertainties when considering the two largest momenta. Obviously, the prefactor of $\mathcal{O}(\Lambda^2_{\text{QCD}}/P_3^2)$ terms is unknown and can also depend on x, but the lattice data seem to favor $\mathcal{O}(1)$ values (or smaller) in the unpolarized case. For the helicity PDF (middle panel of Fig. 33), we observe larger deviations between $P_3 = 10\pi/L$ and $P_3 = 8\pi/L$, particularly for intermediate values of x, i.e., $x \approx 0.3$. The central value for the largest momentum is approximately 25% lower than the one for $P_3 = 8\pi/L$. This can indicate enhanced HTEs [a prefactor of $\mathcal{O}(\Lambda^2_{\text{QCD}}/P_3^2)$ larger than one] or a statistical fluctuation.

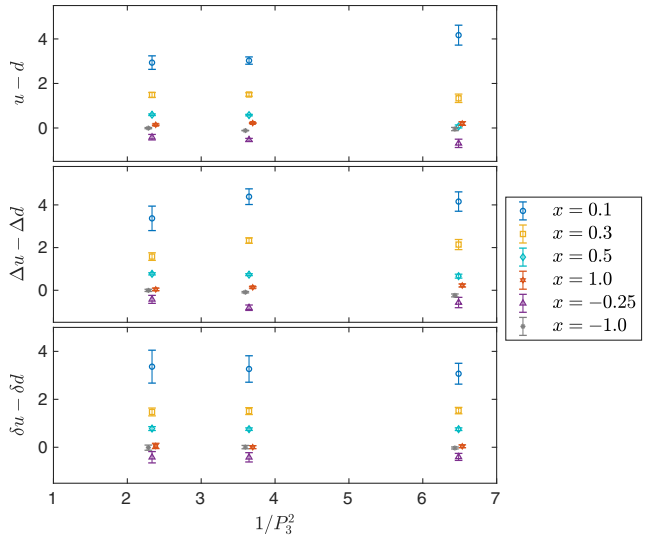

FIG. 33. Dependence of the unpolarized PDF (upper), helicity PDF (middle) and transversity PDF (lower) on the nucleon momentum for six selected values of x.

The latter interpretation may be more favored, given the fact that the results are compatible for the other illustrated values of x. Moreover, HTEs are typically expected to be enhanced for low and high x, as concluded e.g., in Ref. [113]. Even though the results of that paper are not directly applicable to our setup, since the renormalization is performed by forming suitable ratios that have canceled divergences, this generic feature of HTE enhancement is very likely to hold. Nevertheless, the behavior that we obtained for the helicity PDF at intermediate x needs to be checked in future studies. In the transversity case (lower panel of Fig. 33), we observe compatible results for all x values and for all three momenta. This suggests, again, rather small HTEs. All of this corroborates the expectation that HTEs are not the dominant systematic effect in the extracted PDFs, i.e., they are, most likely, hidden in the $\mathcal{O}(10\%)$ statistical uncertainties if the nucleon boost is 1.1–1.4 GeV.

We compare our final PDFs at the largest momentum with the phenomenologically extracted ones in Fig. 34. For the

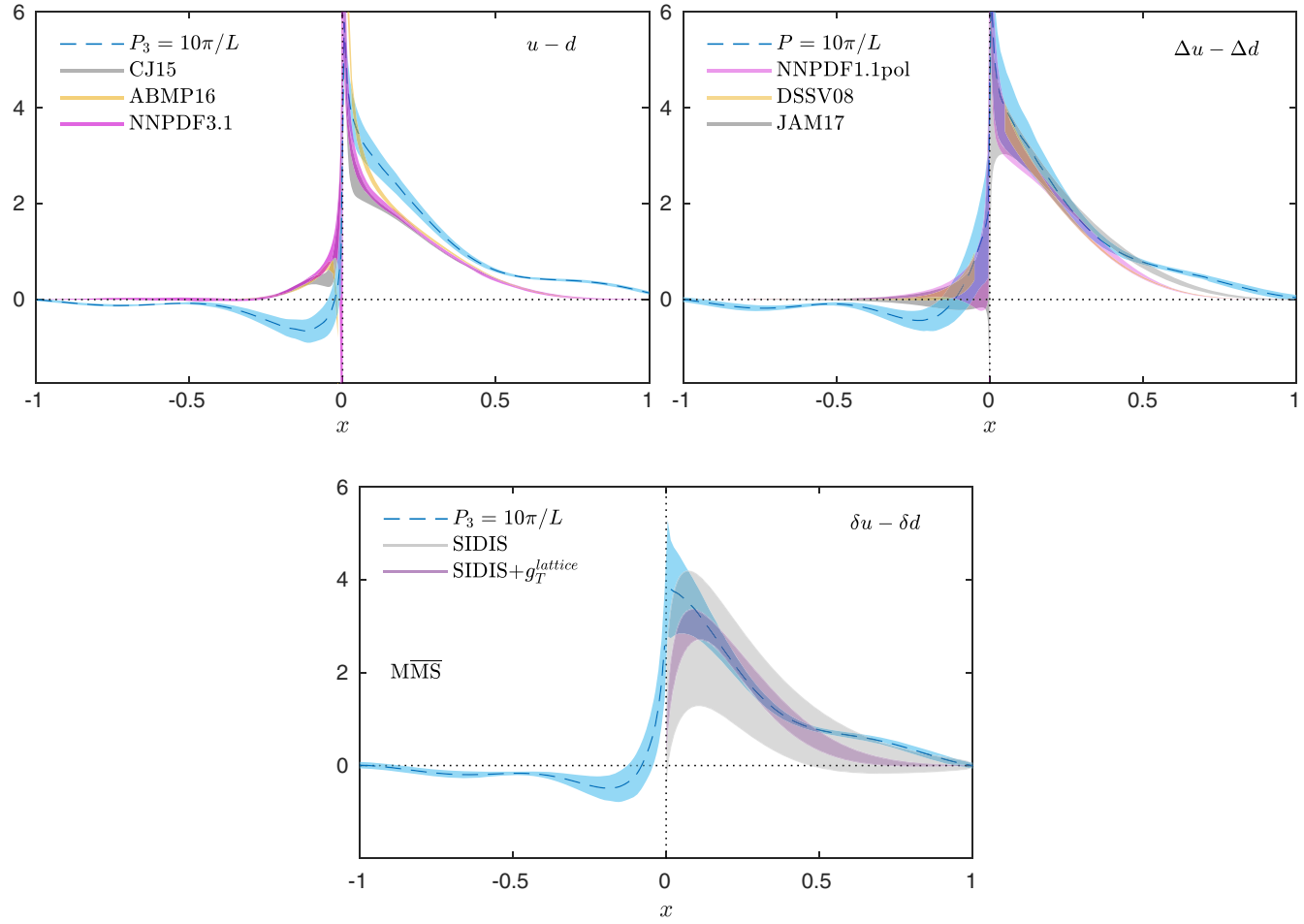

FIG. 34. Final results for the unpolarized PDF (upper left), helicity PDF (upper right) and transversity PDF (lower), using the largest momentum $P_3=10\pi/L$ (blue curve). The global fits of Refs. [114–116] (unpolarized), Refs. [117–119] (helicity), and Refs. [120] (transversity) are shown for qualitative comparison.

phenomenological results, we use the ones from CJ15 [114], ABMP16 [115] and NNPDF3.1 [116] for the unpolarized PDF, the ones from DSSV08 [117], NNPDF1.1pol [118] and JAM17 [119] for the helicity PDF and two parametrizations for the transversity PDF: one extracted from experimental semi-inclusive deep-inelastic scattering data and one where the tensor charge computed in lattice QCD was used as input in the phenomenological analysis [120]. We stress that the comparison with phenomenological PDFs is intended to be qualitative, since we only include statistical uncertainties. The unpolarized PDF has a similar slope to the phenomenological PDFs, but lies above them in the positive-x region. Increasing the momentum should bring the PDF closer to the phenomenological values. In Fig. 32, we choose not to show the phenomenological curves for clarity of the plots, but looking at the position of these curves in Fig. 34, it is clear that a larger boost indeed moves the lattice-extracted PDFs in the direction of the phenomenological PDFs. Our results for the helicity PDF are compatible with phenomenological ones for $x \lesssim 0.4$–0.5. While this may indicate faster convergence of the quasi-PDF, it can also be the result of different systematic effects, with some of them possibly canceling out. Likewise, our result for the transversity PDF is in agreement with both the phenomenologically extracted one, as well as the one using the lattice determination of the tensor charge as input, for $x \lesssim 0.4$–0.5. An interesting feature is that the precision of our results is better than in both phenomenological determinations. Thus, lattice QCD holds the prospect of significantly impacting our knowledge of the transversity PDF, in particular because we expect that in the next generation of lattice QCD computations of PDFs a better control of the systematic effects will be achieved.

The errors shown in the above figures are statistical, and significant effort is needed to properly quantify the systematic uncertainties present in the various steps of the analysis. Currently, these constitute the dominant source of uncertainty, and addressing them will allow one to draw final conclusions from the lattice results. Based on the current studies, it is not possible to quantify and hierarchize them. Typical systematics related to the lattice calculation include discretization effects, finite-volume effects and the role of the pion mass value. The pion mass of the ensemble used here is already at its physical value and hence, there is no need for a chiral extrapolation. The latter would be an important source of uncertainty, as the fit function is not known. We remark that, in practice, the effect of a nonphysical pion mass can be large, as we demonstrated in Ref. [19], comparing the physical pion mass PDF with one at $m_\pi \approx 375$ MeV.

The parameters of the ensembles are expected to satisfy certain criteria for the range of values of the pion mass, the volume and the lattice spacing, to study the following uncertainties.

(1) *Cutoff effects*: A reliable control of cutoff effects requires at least three values of the lattice spacing smaller than 0.1 fm. Normally, cutoff effects are found to be relatively small in lattice hadron structure calculations. In the quasi-PDF computation, the nucleon is boosted to momenta for which P_3 becomes significant in comparison to the inverse lattice spacing and this may lead to increased cutoff effects. We note that for our largest momentum, we have $aP_3 = 0.65$ which is below the lattice cutoff (unlike Refs. [26,28,31] where the employed nucleon momenta are significantly above the lattice cutoff), and the continuum dispersion relation is still satisfied, as shown in Fig. 4. Still, it is unclear to what extent the good quality of the dispersion relation translates into discretization effects of the matrix elements considered here.

(2) *Finite-volume effects*: Similarly to discretization effects, finite-volume effects are also usually found to be rather small in hadron structure observables. The situation with quasi-PDFs is likely to be somewhat more complicated, since we use operators with Wilson lines of significant length. The volume behavior of such extended operators was considered by Briceño *et al.* [110] within a model using current-current correlators in a scalar theory. Despite the fact that the model is not directly applicable to our investigation, it does provide a warning that the suppression of finite-volume effects for matrix elements of spatially extended operators may change from the standard $\exp(-m_\pi L)$ to $(L^m/|L-z|^n)\exp(-m_N(L-z))$, where m and n are undetermined exponents, which may become dominant for large z. Thus, finite-volume effects may turn out to be a significant source of systematics and their investigation will be crucial in the future.

(3) *Systematic uncertainties in the determination of the renormalization functions*: Uncertainties also arise in the computation of renormalization functions due to the breaking of rotational invariance. We have partly improved our work by subtracting lattice artifacts in the renormalization factor of the quark field, computed in lattice perturbation theory (see Sec. IV). A similar subtraction for the Z factors of local operators was very successful [104,107,121], and we intend to perform such a subtraction in the future [111]. We note that the pion mass dependence and finite-volume effects in the renormalization functions studied here were found to be insignificant.

(i) *Uncertainties specific to the quasi-PDF approach*: The most significant uncertainty of this type is the truncation of the Fourier transform and the finite number of evaluations of the matrix elements, when going from the z space of the latter to the x space of partonic distributions. This issue can be alleviated at larger nucleon boosts, which implies a faster decay of matrix elements to zero and, thus, mildened

oscillations. However, according to the findings of the recent Ref. [112], the problem needs a different treatment, using advanced reconstruction techniques. Other *ad hoc* prescriptions, such as the "derivative" method, do not actually remove the source of the oscillations. In this technique, it is unclear how to handle the small-$|x|$ region, since the ignored surface term contains an explicit $1/x$ factor. In addition, the need to discretize the derivative of the matrix elements introduces additional cutoff effects. The effects from the finite momentum are likely to be within the order of magnitude of their theoretical expectation, $\mathcal{O}(\Lambda^2_{\rm QCD}/P_3^2)$, which amounts to around 5% at our largest momentum. Such higher-twist effects can be suppressed by going to larger boosts or alternatively, by explicitly computed and subtracted. However, the former is difficult, as the signal is exponentially decaying at larger momenta and the larger momentum also significantly increases excited-states contributions, particularly when using simulations at the physical pion mass. As we argued, good control over excited states is necessary for extracting reliable physical results. Thus, a detailed study is required to maintain ground-state dominance, following the approach explained in detail in this work. In this paper, we showed that excited states are suppressed below $\sim$10% if the source-sink separation is 1.1 fm at $P_3 \approx 1.4$ GeV, as established with a detailed analysis based on three methods. We note that at the physical pion mass and at small source-sink separations, the contamination comes from tens of excited states, and thus, a one-state or two-state fit alone is not reliable. Compatibility between one-state and two-state fits provides a strong argument that ground-state dominance has been achieved. Another possibly significant source of systematic errors are the perturbative truncation effects in the conversion to the $\overline{\rm MS}$ scheme, evolution to the reference scale and matching, all of which are currently known to the one-loop level. We observed that the effect of the latter is significant at our largest momentum and hence two-loop effects are likely to be substantial. Therefore, a two-loop computation would lead to better connection to light-cone PDFs with smaller values of the nucleon momentum.

VII. CONCLUSIONS

In this work we presented a detailed investigation of the formalism employed to extract x-dependent PDFs within lattice QCD. The analysis was carried out using one ensemble of $N_f = 2$ twisted mass fermions with the quark mass fixed to its physical value. Results on the collinear unpolarized, helicity and transversity PDFs were obtained for the isovector flavor combination. The quasi-PDF method used here relies on a computation of equal-time correlators involving boosted nucleons, with the momentum increased to large enough values, so that the large momentum effective theory is applicable. In this work, we investigated three values of the momentum, reaching a maximum of 1.38 GeV. All necessary components to obtain light-cone PDFs have been considered, namely, the calculation of the bare nucleon matrix elements, renormalization, matching to light-cone PDFs and subtraction of finite nucleon mass corrections. The work presented here is an extension of our earlier work [19,20], and includes details on technical and theoretical aspects, as well as improvements in various aspects of the calculation and examination of systematic effects. In particular, we provided the lattice techniques used in the evaluation of the matrix elements, such as momentum smearing and methods to eliminate (within statistical accuracy) excited-states contributions. Theoretical developments associated with the nonperturbative renormalization were presented, where a chiral extrapolation and a modification of the conversion to the M$\overline{\rm MS}$ scheme have been employed. The latter is a variant of the $\overline{\rm MS}$ scheme that ensures particle number conservation in the matching procedure. Issues related to the Fourier transform as well as different matching prescriptions were also explored.

Particular attention was given to discussing the role of systematic effects in the various steps of the analysis. Typical systematics related to the lattice calculation that are also common in other hadron structure calculations include discretization effects and finite-volume effects. These can be addressed and eliminated by simulations using additional gauge field configuration ensembles. At present, these are part of the unquantified uncertainties, that need to be addressed in future studies. Another typical systematic enters lattice data if the pion mass of the ensemble is at an unphysical value. In this work, we used simulations at the physical point and chiral extrapolation was not needed. Analyses using nonphysical values of the quark mass would introduce an important source of uncertainty, as the fit function to extrapolate the results to the physical point is not known.

Apart from the aforementioned systematics that are common in other hadron structure investigations, there are additional ones that are specific to the quasi-PDF approach. As discussed in the previous sections, significant effects are caused by the truncated and discretized Fourier transform. We will explore whether the reconstruction techniques advocated in Ref. [112] will indeed cure this problem. Also, increasing the nucleon boost can lead to further systematic uncertainties. High values of the momenta reduce the appearance of oscillations, but at the same time, they increase the number of contributing excited states. Systematics related to the truncation of the conversion factor and the matching formula to the one-loop level are non-negligible. In particular, having a matching formula to two loops may lead to better convergence to

light-cone PDFs at smaller nucleon momenta. Hence, a two-loop computation is strongly desired and would help to establish a better connection to light-cone PDFs with smaller values of the nucleon momentum.

Despite these uncertainties, this work demonstrates tremendous progress in the determination of PDFs from the quasidistribution approach [33]. Lattice QCD results confirm the feasibility of extracting PDFs from first-principle calculations. The success in the quasi-PDF approach for the nucleon has also resulted in studies of other hadrons and alternative approaches, as summarized in Ref. [33]. The theoretical and technical aspects are now well understood for the nucleon studies and addressing the systematic uncertainties is the next step that will be enabled by using the methodology developed, in combination with the availability of increased computational resources. In fact, our preliminary results for an $N_f = 2 + 1 + 1$ twisted mass ensemble with physical values of the quark masses, a lattice spacing of 0.082 fm and a larger volume of $64^3 \times 128$ have been shown in Ref. [122] and demonstrate the future direction and progress of such computations. Furthermore, the production of another ensemble at a finer lattice spacing is already underway.

ACKNOWLEDGMENTS

We would like to thank all members of ETMC for their constant and pleasant collaboration. We also thank Y. Zhao for useful discussions. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 642069 (HPC-LEAP). K. C. was supported by National Science Centre (Poland) Grant SONATA BIS No. 2016/22/E/ST2/00013. F. S. was funded by DFG Project No. 392578569. M. C. acknowledges financial support by the U.S. Department of Energy, Office of Nuclear Physics, within the framework of the TMD Topical Collaboration, as well as, by the National Science Foundation under Grant No. PHY-1714407. A. S. is partly supported by the project COMPLEMENTARY/0916/0015 funded by the Cyprus Research Promotion Foundation. This research used computational resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility supported under Contract No. DE-AC05-00OR22725. Additional computational resources were provided by John-von-Neumann-Institut für Computing (NIC) for the JURECA supercomputer (allocation ECY00), the Prometheus supercomputer at the Academic Computing Centre Cyfronet AGH in Cracow (grant ID *quasipdfs*), the Eagle supercomputer at the Poznan Supercomputing and Networking Center (Grant No. 346), and the Okeanos supercomputer at the Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw (Grant IDs gb70-17, ga71-22).

[1] E. Perez and E. Rizvi, The quark and gluon structure of the proton, Rep. Prog. Phys. **76**, 046201 (2013).
[2] A. De Roeck and R. S. Thorne, Structure functions, Prog. Part. Nucl. Phys. **66**, 727 (2011).
[3] S. Alekhin *et al.*, The PDF4LHC working group interim report, arXiv:1101.0536.
[4] R. D. Ball *et al.*, Parton distribution benchmarking with LHC data, J. High Energy Phys. 04 (2013) 125.
[5] C. A. Aidala, S. D. Bass, D. Hasch, and G. K. Mallot, The spin structure of the nucleon, Rev. Mod. Phys. **85**, 655 (2013).
[6] S. Forte and G. Watt, Progress in the determination of the partonic structure of the proton, Annu. Rev. Nucl. Part. Sci. **63**, 291 (2013).
[7] P. Jimenez-Delgado, W. Melnitchouk, and J. F. Owens, Parton momentum and helicity distributions in the nucleon, J. Phys. G **40**, 093102 (2013).
[8] J. Rojo *et al.*, The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II, J. Phys. G **42**, 103103 (2015).
[9] J. Butterworth *et al.*, PDF4LHC recommendations for LHC Run II, J. Phys. G **43**, 023001 (2016).
[10] A. Accardi *et al.*, A critical appraisal and evaluation of modern PDFs, Eur. Phys. J. C **76**, 471 (2016).
[11] J. Gao, L. Harland-Lang, and J. Rojo, The structure of the proton in the LHC precision era, Phys. Rep. **742**, 1 (2018).
[12] H.-W. Lin *et al.*, Parton distributions and lattice QCD calculations: A community white paper, Prog. Part. Nucl. Phys. **100**, 107 (2018).
[13] X. Ji, Parton Physics on a Euclidean Lattice, Phys. Rev. Lett. **110**, 262002 (2013).
[14] H.-W. Lin, J.-W. Chen, S. D. Cohen, and X. Ji, Flavor structure of the nucleon sea from lattice QCD, Phys. Rev. D **91**, 054510 (2015).
[15] C. Alexandrou, K. Cichy, V. Drach, E. Garcia-Ramos, K. Hadjiyiannakou, K. Jansen, F. Steffens, and C. Wiese, First results with twisted mass fermions towards the computation of parton distribution functions on the lattice, Proc. Sci., LATTICE2014 (**2014**) 135 [arXiv:1411.0891].
[16] C. Alexandrou, K. Cichy, V. Drach, E. Garcia-Ramos, K. Hadjiyiannakou, K. Jansen, F. Steffens, and C. Wiese, Lattice calculation of parton distributions, Phys. Rev. D **92**, 014502 (2015).
[17] J.-W. Chen, S. D. Cohen, X. Ji, H.-W. Lin, and J.-H. Zhang, Nucleon helicity and transversity parton

distributions from lattice QCD, Nucl. Phys. **B911**, 246 (2016).
[18] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, F. Steffens, and C. Wiese, Updated lattice results for parton distributions, Phys. Rev. D **96**, 014513 (2017).
[19] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato, and F. Steffens, Light-Cone Parton Distribution Functions from Lattice QCD, Phys. Rev. Lett. **121**, 112001 (2018).
[20] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato, and F. Steffens, Transversity parton distribution functions from lattice QCD, Phys. Rev. D **98**, 091503 (2018).
[21] M. Constantinou and H. Panagopoulos, Perturbative renormalization of quasi-parton distribution functions, Phys. Rev. D **96**, 054506 (2017).
[22] J.-H. Zhang, J.-W. Chen, X. Ji, L. Jin, and H.-W. Lin, Pion distribution amplitude from lattice QCD, Phys. Rev. D **95**, 094514 (2017).
[23] J.-W. Chen, T. Ishikawa, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Parton distribution function with nonperturbative renormalization from lattice QCD, Phys. Rev. D **97**, 014505 (2018).
[24] H.-W. Lin, J.-W. Chen, T. Ishikawa, and J.-H. Zhang (LP3 Collaboration), Improved parton distribution functions at the physical pion mass, Phys. Rev. D **98**, 054504 (2018).
[25] J.-W. Chen, L. Jin, H.-W. Lin, A. Schäfer, P. Sun, Y.-B. Yang, J.-H. Zhang, R. Zhang, and Y. Zhao, Kaon distribution amplitude from lattice QCD and the flavor SU(3) symmetry, Nucl. Phys. **B939**, 429 (2019).
[26] J.-W. Chen, L. Jin, H.-W. Lin, Y.-S. Liu, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Lattice calculation of parton distribution function from LaMET at physical pion mass with large nucleon momentum, arXiv:1803.04393.
[27] J.-W. Chen, L. Jin, H.-W. Lin, Y.-S. Liu, A. Schäfer, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, First direct lattice-QCD calculation of the x-dependence of the pion parton distribution function, arXiv:1804.01483.
[28] H.-W. Lin, J.-W. Chen, L. Jin, Y.-S. Liu, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Proton Isovector Helicity Distribution on the Lattice at Physical Pion Mass, Phys. Rev. Lett. **121**, 242003 (2018).
[29] Y.-S. Liu, J.-W. Chen, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Unpolarized quark distribution from lattice QCD: A systematic analysis of renormalization and matching, arXiv:1807.06566.
[30] Z.-Y. Fan, Y.-B. Yang, A. Anthony, H.-W. Lin, and K.-F. Liu, Gluon Quasi-Parton-Distribution Functions from Lattice QCD, Phys. Rev. Lett. **121**, 242001 (2018).
[31] Y.-S. Liu, J.-W. Chen, L. Jin, R. Li, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Nucleon transversity distribution at the physical pion mass from lattice QCD, arXiv:1810.05043.
[32] P. Petreczky, T. Izubuchi, L. Jin, C. Kallidonis, N. Karthik, S. Mukherjee, C. Shugert, and S. Syritsyn, Pion structure from Lattice QCD, arXiv:1812.04334.
[33] K. Cichy and M. Constantinou, A guide to light-cone PDFs from Lattice QCD: An overview of approaches, techniques and results, Adv. High Energy Phys. **2019**, 3036904 (2019).
[34] X. Xiong, X. Ji, J.-H. Zhang, and Y. Zhao, One-loop matching for parton distributions: Nonsinglet case, Phys. Rev. D **90**, 014051 (2014).
[35] Y.-Q. Ma and J.-W. Qiu, Extracting parton distribution functions from lattice QCD calculations, Phys. Rev. D **98**, 074021 (2018).
[36] R. A. Briceño, M. T. Hansen, and C. J. Monahan, Role of the Euclidean signature in lattice calculations of quasidistributions and other nonlocal matrix elements, Phys. Rev. D **96**, 014502 (2017).
[37] Y.-Q. Ma and J.-W. Qiu, Exploring Partonic Structure of Hadrons using ab initio Lattice QCD Calculations, Phys. Rev. Lett. **120**, 022003 (2018).
[38] X. Ji, J.-H. Zhang, and Y. Zhao, More on large-momentum effective theory approach to parton physics, Nucl. Phys. **B924**, 366 (2017).
[39] A. V. Radyushkin, Structure of parton quasi-distributions and their moments, Phys. Lett. B **788**, 380 (2019).
[40] J. Karpie, K. Orginos, and S. Zafeiropoulos, Moments of Ioffe time parton distribution functions from nonlocal matrix elements, J. High Energy Phys. 11 (2018) 178.
[41] X. Ji, A. Schäfer, X. Xiong, and J.-H. Zhang, One-loop matching for generalized parton distributions, Phys. Rev. D **92**, 014039 (2015).
[42] X. Xiong and J.-H. Zhang, One-loop matching for transversity generalized parton distribution, Phys. Rev. D **92**, 054037 (2015).
[43] W. Wang, S. Zhao, and R. Zhu, Gluon quasidistribution function at one loop, Eur. Phys. J. C **78**, 147 (2018).
[44] I. W. Stewart and Y. Zhao, Matching the quasi parton distribution in a momentum subtraction scheme, Phys. Rev. D **97**, 054512 (2018).
[45] T. Izubuchi, X. Ji, L. Jin, I. W. Stewart, and Y. Zhao, Factorization theorem relating Euclidean and light-cone parton distributions, Phys. Rev. D **98**, 056004 (2018).
[46] X. Ji, Parton physics from large-momentum effective field theory, Sci. Chin. Phys. Mech. Astron. **57**, 1407 (2014).
[47] T. Ishikawa, Y.-Q. Ma, J.-W. Qiu, and S. Yoshida, Renormalizability of quasiparton distribution functions, Phys. Rev. D **96**, 094019 (2017).
[48] X. Ji, J.-H. Zhang, and Y. Zhao, Renormalization in Large Momentum Effective Theory of Parton Physics, Phys. Rev. Lett. **120**, 112001 (2018).
[49] J.-H. Zhang, X. Ji, A. Schäfer, W. Wang, and S. Zhao, Accessing Gluon Parton Distributions in Large Momentum Effective Theory, Phys. Rev. Lett. **122**, 142001 (2019).
[50] Z.-Y. Li, Y.-Q. Ma, and J.-W. Qiu, Multiplicative Renormalizability of Operators defining Quasiparton Distributions, Phys. Rev. Lett. **122**, 062002 (2019).
[51] K.-F. Liu and S.-J. Dong, Origin of Difference between Anti-d and Anti-u Partons in the Nucleon, Phys. Rev. Lett. **72**, 1790 (1994).
[52] K. F. Liu, S. J. Dong, T. Draper, D. Leinweber, J. H. Sloan, W. Wilcox, and R. M. Woloshyn, Valence QCD: Connecting QCD to the quark model, Phys. Rev. D **59**, 112001 (1999).
[53] K.-F. Liu, Parton degrees of freedom from the path integral formalism, Phys. Rev. D **62**, 074501 (2000).

[54] W. Detmold and C. J. D. Lin, Deep-inelastic scattering and the operator product expansion in lattice QCD, Phys. Rev. D **73**, 014501 (2006).

[55] V. Braun and D. Mueller, Exclusive processes in position space and the pion distribution amplitude, Eur. Phys. J. C **55**, 349 (2008).

[56] A. J. Chambers, R. Horsley, Y. Nakamura, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, K. Somfleth, R. D. Young, and J. M. Zanotti, Nucleon Structure Functions from Operator Product Expansion on the Lattice, Phys. Rev. Lett. **118**, 242001 (2017).

[57] A. Radyushkin, Nonperturbative evolution of parton quasi-distributions, Phys. Lett. B **767**, 314 (2017).

[58] A. Radyushkin, Target mass effects in parton quasi-distributions, Phys. Lett. B **770**, 514 (2017).

[59] A. V. Radyushkin, Quasi-parton distribution functions, momentum distributions, and pseudo-parton distribution functions, Phys. Rev. D **96**, 034025 (2017).

[60] A. V. Radyushkin, Quark pseudodistributions at short distances, Phys. Lett. B **781**, 433 (2018).

[61] J. Liang, K.-F. Liu, and Y.-B. Yang, Lattice calculation of hadronic tensor of the nucleon, EPJ Web Conf. **175**, 14014 (2018).

[62] W. Detmold, I. Kanamori, C. J. D. Lin, S. Mondal, and Y. Zhao, Moments of pion distribution amplitude using operator product expansion on the lattice, arXiv: 1810.12194.

[63] G. S. Bali *et al.*, Pion distribution amplitude from Euclidean correlation functions, Eur. Phys. J. C **78**, 217 (2018).

[64] G. S. Bali, V. M. Braun, B. Gläßle, M. Göckeler, M. Gruber, F. Hutzler, P. Korcyl, A. Schäfer, P. Wein, and J.-H. Zhang, Pion distribution amplitude from Euclidean correlation functions: Exploring universality and higher twist effects, Phys. Rev. D **98**, 094507 (2018).

[65] K. Orginos, A. Radyushkin, J. Karpie, and S. Zafeiropoulos, Lattice QCD exploration of parton pseudo-distribution functions, Phys. Rev. D **96**, 094503 (2017).

[66] J. Karpie, K. Orginos, A. Radyushkin, and S. Zafeiropoulos, Parton distribution functions on the lattice and in the continuum, EPJ Web Conf. **175**, 06032 (2018).

[67] A. Radyushkin, One-loop evolution of parton pseudo-distribution functions on the lattice, Phys. Rev. D **98**, 014019 (2018).

[68] R. S. Sufian, J. Karpie, C. Egerer, K. Orginos, J.-W. Qiu, and D. G. Richards, Pion valence quark distribution from matrix element calculated in lattice QCD, Phys. Rev. D **99**, 074507 (2019).

[69] L. Gamberg, Z.-B. Kang, I. Vitev, and H. Xing, Quasi-parton distribution functions: A study in the diquark spectator model, Phys. Lett. B **743**, 112 (2015).

[70] I. Vitev, L. Gamberg, Z. Kang, and H. Xing, A study of quasi-parton distribution functions in the diquark spectator model, Proc. Sci., QCDEV2015 (**2015**) 045 [arXiv:1511.05242].

[71] Y. Jia and X. Xiong, Quasidistribution amplitude of heavy quarkonia, Phys. Rev. D **94**, 094005 (2016).

[72] Y. Jia, S. Liang, X. Xiong, and R. Yu, Partonic quasidistributions in two-dimensional QCD, Phys. Rev. D **98**, 054011 (2018).

[73] A. Bacchetta, M. Radici, B. Pasquini, and X. Xiong, Reconstructing parton densities at large fractional momenta, Phys. Rev. D **95**, 014036 (2017).

[74] W. Broniowski and E. Ruiz Arriola, Nonperturbative partonic quasidistributions of the pion from chiral quark models, Phys. Lett. B **773**, 385 (2017).

[75] W. Broniowski and E. Ruiz Arriola, Partonic quasidistributions of the proton and pion from transverse-momentum distributions, Phys. Rev. D **97**, 034031 (2018).

[76] S. Bhattacharya, C. Cocuzza, and A. Metz, Generalized quasi parton distributions in a diquark spectator model, Phys. Lett. B **788**, 453 (2019).

[77] C. Monahan, Recent developments in x-dependent structure calculations, Proc. Sci., LATTICE2018 (**2018**) 018 [arXiv:1811.00678].

[78] A. Abdel-Rehim *et al.* (ETM Collaboration), First physics results at the physical pion mass from $N_f = 2$ Wilson twisted mass fermions at maximal twist, Phys. Rev. D **95**, 094515 (2017).

[79] C. Alexandrou and C. Kallidonis, Low-lying baryon masses using $N_f = 2$ twisted mass clover-improved fermions directly at the physical pion mass, Phys. Rev. D **96**, 034511 (2017).

[80] M. T. Hansen and H. B. Meyer, On the effect of excited states in lattice calculations of the nucleon axial charge, Nucl. Phys. **B923**, 558 (2017).

[81] J. Green, Systematics in nucleon matrix element calculations, Proc. Sci., LATTICE2018 (**2018**) 016 [arXiv: 1812.10574].

[82] Y. Iwasaki, Renormalization group analysis of lattice theories and improved lattice action. II. Four-dimensional non-Abelian SU(N) gauge model, arXiv:1111.7054.

[83] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz (Alpha Collaboration), Lattice QCD with a chirally twisted mass term, J. High Energy Phys. 08 (2001) 058.

[84] R. Frezzotti and G. C. Rossi, Chirally improving Wilson fermions. 1. O(a) improvement, J. High Energy Phys. 08 (2004) 007.

[85] B. Sheikholeslami and R. Wohlert, Improved continuum limit lattice action for QCD with Wilson fermions, Nucl. Phys. **B259**, 572 (1985).

[86] S. Aoki *et al.* (JLQCD, CP-PACS Collaboration), Nonperturbative O(a) improvement of the Wilson quark action with the RG-improved gauge action using the Schrodinger functional method, Phys. Rev. D **73**, 034501 (2006).

[87] K. Jansen, C. McNeile, C. Michael, K. Nagai, M. Papinutto, J. Pickavance, A. Shindler, C. Urbach, and I. Wetzorke (XLF Collaboration), Flavor breaking effects of Wilson twisted mass fermions, Phys. Lett. B **624**, 334 (2005).

[88] F. Farchioni, R. Frezzotti, K. Jansen, I. Montvay, G. C. Rossi, E. Scholz, A. Shindler, N. Ukita, C. Urbach, and I. Wetzorke, Twisted mass quarks and the phase structure of lattice QCD, Eur. Phys. J. C **39**, 421 (2005).

[89] F. Farchioni *et al.*, Numerical simulations with two flavors of twisted-mass Wilson quarks and DBW2 gauge action, Eur. Phys. J. C **47**, 453 (2006).

[90] S. Gusken, A Study of smearing techniques for hadron correlation functions, Nucl. Phys. B, Proc. Suppl. **17**, 361 (1990).

[91] C. Alexandrou, S. Gusken, F. Jegerlehner, K. Schilling, and R. Sommer, The static approximation of heavy—light quark systems: A systematic lattice study, Nucl. Phys. **B414**, 815 (1994).
[92] A. Abdel-Rehim *et al.*, Nucleon and pion structure with lattice QCD simulations at physical value of the pion mass, Phys. Rev. D **92**, 114513 (2015).
[93] G. S. Bali, B. Lang, B. U. Musch, and A. Schäfer, Novel quark smearing for hadrons with high momenta in lattice QCD, Phys. Rev. D **93**, 094515 (2016).
[94] G. Martinelli and C. T. Sachrajda, A lattice study of nucleon structure, Nucl. Phys. **B316**, 355 (1989).
[95] C. Alexandrou, S. Dinter, V. Drach, K. Jansen, K. Hadjiyiannakou, and D. B. Renner (ETM Collaboration), A stochastic method for computing hadronic matrix elements, Eur. Phys. J. C **74**, 2692 (2014).
[96] T. Blum, T. Izubuchi, and E. Shintani, New class of variance-reduction techniques using lattice symmetries, Phys. Rev. D **88**, 094503 (2013).
[97] C. Alexandrou, S. Bacchio, J. Finkenrath, A. Frommer, K. Kahl, and M. Rottmann, Adaptive aggregation-based domain decomposition multigrid for twisted mass fermions, Phys. Rev. D **94**, 114509 (2016).
[98] C. Morningstar and M. J. Peardon, Analytic smearing of SU(3) link variables in lattice QCD, Phys. Rev. D **69**, 054501 (2004).
[99] V. S. Dotsenko and S. N. Vergeles, Renormalizability of phase factors in the nonabelian gauge theory, Nucl. Phys. **B169**, 527 (1980).
[100] R. A. Brandt, F. Neri, and M.-a. Sato, Renormalization of loop functions for all loops, Phys. Rev. D **24**, 879 (1981).
[101] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, H. Panagopoulos, and F. Steffens, A complete non-perturbative renormalization prescription for quasi-PDFs, Nucl. Phys. **B923**, 394 (2017).
[102] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, A. Scapellato, and F. Steffens, Light-cone PDFs from lattice QCD, Proc. Sci., LATTICE2018 (**2018**) 094 [arXiv:1811.01588].
[103] L. Maiani, G. Martinelli, M. L. Paciello, and B. Taglienti, Scalar densities and Baryon mass differences in lattice QCD with Wilson fermions, Nucl. Phys. **B293**, 420 (1987).
[104] C. Alexandrou, M. Constantinou, and H. Panagopoulos (ETM Collaboration), Renormalization functions for Nf = 2 and Nf = 4 twisted mass fermions, Phys. Rev. D **95**, 034505 (2017).
[105] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, A general method for nonperturbative renormalization of lattice operators, Nucl. Phys. **B445**, 81 (1995).
[106] J. Green, K. Jansen, and F. Steffens, Nonperturbative Renormalization of Nonlocal Quark Bilinears for Parton Quasidistribution Functions on the Lattice Using an Auxiliary Field, Phys. Rev. Lett. **121**, 022004 (2018).
[107] M. Constantinou *et al.* (ETM Collaboration), Non-perturbative renormalization of quark bilinear operators with $N_f = 2$ (tmQCD) Wilson fermions and the tree-level improved gauge action, J. High Energy Phys. 08 (2010) 068.
[108] M. Göckeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, P. E. L. Rakow, A. Schäfer, G. Schierholz, and A. Schiller, Nonperturbative renormalization of composite operators in lattice QCD, Nucl. Phys. **B544**, 699 (1999).
[109] A. Ali Khan *et al.* (CP-PACS Collaboration), Light hadron spectroscopy with two flavors of dynamical quarks on the lattice, Phys. Rev. D **65**, 054505 (2002); Eratum, Phys. Rev. D **67**, 059901(E) (2003).
[110] R. A. Briceño, J. V. Guerrero, M. T. Hansen, and C. J. Monahan, Finite-volume effects due to spatially nonlocal operators, Phys. Rev. D **98**, 014511 (2018).
[111] M. Constantinou and H. Panagopoulos, Improvement of the renormalization functions of non-local operators with straight Wilson lines (in preparation).
[112] J. Karpie, K. Orginos, A. Rothkopf, and S. Zafeiropoulos, Reconstructing parton distribution functions from Ioffe time data: from Bayesian methods to Neural Networks, J. High Energy Phys. 04 (**2019**) 057.
[113] V. M. Braun, A. Vladimirov, and J.-H. Zhang, Power corrections and renormalons in parton quasi-distributions, Phys. Rev. D **99**, 014013 (2019).
[114] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato, Constraints on large-x parton distributions from new weak boson production and deep-inelastic scattering data, Phys. Rev. D **93**, 114017 (2016).
[115] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte, Parton distribution functions, α_s, and heavy-quark masses for LHC Run II, Phys. Rev. D **96**, 014011 (2017).
[116] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017).
[117] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, Extraction of spin-dependent parton densities and their uncertainties, Phys. Rev. D **80**, 034030 (2009).
[118] E. R. Nocera, R. D. Ball, S. Forte, G. Ridolfi, and J. Rojo (NNPDF Collaboration), A first unbiased global determination of polarized PDFs and their uncertainties, Nucl. Phys. **B887**, 276 (2014).
[119] J. J. Ethier, N. Sato, and W. Melnitchouk, First Simultaneous Extraction of Spin-Dependent Parton Distributions and Fragmentation Functions from a Global QCD Analysis, Phys. Rev. Lett. **119**, 132001 (2017).
[120] H.-W. Lin, W. Melnitchouk, A. Prokudin, N. Sato, and H. Shows, First Monte Carlo Global Analysis of Nucleon Transversity with Lattice QCD Constraints, Phys. Rev. Lett. **120**, 152502 (2018).
[121] M. Constantinou, R. Horsley, H. Panagopoulos, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, and J. M. Zanotti, Renormalization of local quark-bilinear operators for $N_f = 3$ flavors of stout link nonperturbative clover fermions, Phys. Rev. D **91**, 014502 (2015).
[122] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, A. Scapellato, and F. Steffens, Quasi-PDFs from Twisted mass fermions at the physical point, Proc. Sci., LATTICE2018 (**2018**) 095 [arXiv:1811.01589].