Valence parton distribution function of pion from fine lattice

Taku Izubuchi,[1,2] Luchang Jin,[2,3] Christos Kallidonis,[4] Nikhil Karthik,[1] Swagato Mukherjee,[1]
Peter Petreczky,[1] Charles Shugert,[1,4] and Sergey Syritsyn[2,4]

[1]*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2]*RIKEN-BNL Research Center, Brookhaven National Lab, Upton, New York 11973, USA*
[3]*Physics Department, University of Connecticut, Storrs, Connecticut 06269-3046, USA*
[4]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*

We present a lattice QCD study of the valence parton distribution inside the pion within the framework of Large Momentum Effective Theory. We use a mixed action approach with 1-HYP smeared valence Wilson clover quarks on $2 + 1$ flavor HISQ sea with the valence quark mass tuned to 300 MeV pion mass. We use $48^3 \times 64$ lattice at a fine lattice spacing $a = 0.06$ fm for this computation. We renormalize the quasi parton distribution function matrix element in the nonperturbative regularization independent momentum subtraction (RI-MOM) scheme. As a byproduct, we test the validity of a 1-loop matching procedure by comparing the RI-MOM renormalized quasi parton distribution function matrix element with off-shell quark external states as computed in the continuum 1-loop perturbation theory with the lattice results at $a = 0.04$ and 0.06 fm. By applying the RI-MOM to $\overline{\text{MS}}$ one-loop matching, implemented through a fit to phenomenologically motivated parton distribution functions, we obtain the valence parton distribution function of pion.

I. INTRODUCTION

QCD factorization allows us to calculate the cross section of hard hadronic processes in terms of the convolution of partonic cross section and parton distribution functions [1]. Parton distribution for a hadron can be defined using hadronic matrix elements of appropriately chosen gauge invariant operators separated along the light cone. For example, the quark parton distribution function (PDF) of a hadron H can be defined in terms of an operator bilocal in quark field ψ as [1,2]

$$f(x) = \frac{1}{4\pi} \int d\xi^- e^{ixP^+\xi^-} \langle H(P)|\bar{\psi}(\xi^-)\gamma^+ W(\xi^-,0)\psi(0)|H(P)\rangle, \quad (1)$$

where $W(\xi^-,0) = \mathcal{P}e^{ig\int_0^{\xi^-} d\xi^- A^+}$ is the path-ordered straight Wilson line on the light cone, and the light-cone coordinates $\xi^\pm = (t \pm z)/\sqrt{2}$. A straightforward first principle calculation of PDF is not possible because lattice QCD is formulated in the Euclidean space-time, and, thus, it cannot access quantities defined on the light cone. To circumvent this problem, it has recently been proposed to

calculate the quasi parton distribution function (qPDF), $\tilde{q}(x, P_z)$, defined in terms of matrix elements of equal time, but spatially separated, quark bilinears [3] evaluated in a hadron state boosted to a large momentum P_z:

$$\tilde{q}(x,P_z) = \frac{1}{4\pi} \int dz e^{-ixP^z z} \langle H(P_z)|\bar{\psi}(z)\Gamma W(z,0)\psi(0)|H(P_z)\rangle, \quad (2)$$

where Γ is either γ_z or γ_t for the unpolarized parton distribution addressed in this paper. Here, $W(z,0)$ is a straight spatial Wilson line joining the quark and antiquark. For sufficiently boosted hadrons, one can use the Large Momentum Effective Theory (LaMET) [4] to relate the qPDF to PDF through a convolution with a matching kernel C as

$$\tilde{q}(x,\mu_L,P_z) = \int_{-1}^{+1} \frac{dy}{|y|} C\left(\frac{x}{y}, \frac{yP_z}{\mu}, \frac{\mu_L}{yP_z}\right) f(y,\mu). \quad (3)$$

Here μ_L and μ are the renormalization scales of the schemes in which the qPDF and PDF are defined. For the latter, $\overline{\text{MS}}$ scheme is used and μ is referred to as the factorization scale. The matching kernel is perturbative and hence universal for all the hadrons. Therefore, it is calculated using quark external states in a chosen gauge. Such calculations at 1-loop order have been performed using the cutoff scheme [5] as well as in the $\overline{\text{MS}}$ scheme [6–8]. There are also

related approaches to calculate the PDF from the lattice that use similar logic but differ in details, like the pseudo-PDF approach proposed in Refs. [9,10] and the use of good lattice cross sections [11,12]. The latter includes the current-current correlators [13].

Using LaMET and related approaches, various attempts have been made to calculate the unpolarized and polarized isovector quark distribution of the nucleon [7,14–17]. The first studies of the valence quark distribution for the pion have also been presented [13,18]. One important issue in the calculation of the PDF from the lattice is the renormalization and matching. As indicated above, the PDF and qPDF are usually defined in different renormalization schemes. The qPDF, which is calculated on the lattice, needs a nonperturbative renormalization scheme because of the self-energy divergence of the Wilson line [19], and this is usually implemented using the regularization independent momentum subtraction (RI-MOM) scheme [20] defined using external off-shell quark states accessible on the lattice. Then, one has to match the qPDF in this lattice renormalization scheme to the PDF in the $\overline{\text{MS}}$ scheme through Eq. (3). This is achieved through the convolution using the matching kernel between the RI-MOM and $\overline{\text{MS}}$ schemes that is perturbatively calculated in the continuum theory using dimensional regularization [8]. One could also define the qPDF operator in the $\overline{\text{MS}}$ scheme and then perform the matching between PDF and qPDF [6,21]. The current status of this field, including the comparison with the phenomenological PDF and the issue of renormalization, is reviewed in Refs. [22–24].

In principle, Eq. (3) offers a way to calculate PDF from the lattice, but it is unclear as to what extent this is actually feasible given the various assumptions that go along with the equation implicitly. For example, at any finite hadron momentum P_z, the Eq. (3) suffers from $\mathcal{O}(\Lambda_{\text{QCD}}^2/(x^2 P_z^2))$ higher twist corrections. This is closely related to the assumption that the perturbative calculation, currently truncated at 1-loop order, is able to capture the renormalization as well as the matching of the qPDF matrix element over a range of quark-antiquark separations, z—to be in the perturbative regime, one would expect z to be smaller than or about $\mathcal{O}(1)$ fm. It is also important to ensure that $aP_z < 1$ to make sure we are not overcome with lattice artifacts [25]. Therefore, a closer look at this new methodology is warranted and is actively being studied [7,26]. The aim of this paper is to explore these issues further by using finer lattices than what are being used in the qPDF literature, and use the pion as a case study. The smaller mass of the pion makes it easier to achieve a large boost, the numerical calculations are expected to be less expensive, and it also helps suppress the target mass correction by ensuring $m_\pi \ll P_z$. We focus on the valence PDF of the pion since it can be accessed using the isotriplet $u - d$ PDF, and, thereby, avoid mixing with the gluon sector. In our study, we will use the renormalization and matching strategy

outlined in [7,20]. The pion valence PDF has been determined through leading-order and next-to-leading-order analyses of the experimental data [27–35], but it is much less constrained than the nucleon PDF, and, therefore, the lattice calculations may have more impact in this case, especially in constraining the $x \to 1$ limit which is not yet well established.

The paper is organized as follows. In Sec. II, we discuss our lattices setup. In Sec. III, we present the calculations of the two-point function of the boosted pion and check how reliable the extractions of the ground state and the first excited state are. In Sec. IV, we present our results for the pion three-point function that defines the qPDF. Here, we also discuss the problem of excited state contamination. In Sec. V, we discuss the nonperturbative renormalization as well as the validity of 1-loop matching. Our results on the renormalized pion qPDF and the matching to PDF are presented in Sec. VI. Some technical aspects of the calculations are discussed in the Appendices. Preliminary results on this work have been reported in conference proceedings [36–38].

II. LATTICE SETUP

We performed the calculations of the pion two-point and three-point functions needed to obtain the qPDF using the Wilson-Clover action for valence quarks on 1-HYP smeared gauge configurations [39] and the highly improved staggered quark (HISQ) action [40] in the sea. We used the $2 + 1$ flavor gauge configurations corresponding to lattice size $48^3 \times 64$ and the lattice spacing of $a = 0.06$ fm generated by the HotQCD collaboration [41]. In addition to this ensemble, we also used 64^4 highly improved staggered quark lattices [41] with the lattice spacing $a = 0.04$ fm for the study of the nonperturbative renormalization (NPR). In both the ensembles, the sea quark mass was tuned to a pion mass of 160 MeV. A similar setup was used by the PNDME collaboration, albeit for $2 + 1 + 1$ flavor MILC configurations (c.f., Ref. [7]). For the valence quark masses, we used the values $am = -0.0388$ (i.e., $\kappa = 0.12623$) for the $a = 0.06$ fm ensemble and $am = -0.033$ (i.e., $\kappa = 0.12604$) for the $a = 0.04$ fm ensemble, which are tuned such that the pion mass, m_π, is 300 MeV. We did not see any exceptional configurations for these valence quark masses in our calculations.

We used higher statistics at smaller quark-antiquark separations z than at larger ones; to be exact, we used 216, 100, and 48 gauge configurations for $|z|/a \in [0, 8]$, $(8, 16]$, and $(16, 24]$, respectively. We further improved the statistics by using the All-Mode Averaging (AMA) [42] technique in the computations of the two- and three-point functions, with 32 sloppy calculations to one exact solve for each configuration. For the exact and sloppy inversions, we used the stopping criterion of 10^{-10} and 10^{-4}, respectively. In our study, we will consider the valence quark distribution, which in turn is related to the isovector $u - d$

quark distribution in the pion, and thus we do not compute the quark line disconnected diagrams.

For a reliable extraction of qPDF, a good overlap of the source operator with the pion state is necessary so as to project out the ground state at as small source-sink separation as possible. The quark sources with a Gaussian profile, typically implemented through a gauge covariant Wuppertal smearing [43], are used for this purpose when the hadron is at rest. However, for the fast moving hadrons that are required in the qPDF framework, the use of the Gaussian sources is no longer sufficient, and this necessitates the usage of the boosted Gaussian sources [44] instead. Since we are interested in the calculation of the pion two- and three-point function at several values of the pion momenta and several source-sink separations, we found it more practical to implement the Gaussian sources by using the Coulomb gauge instead of implementing the Wuppertal smearing. We found the optimal size of the Gaussian profile to be about 0.3 fm, which roughly corresponds to 90 steps of Wuppertal smearing. We checked that in terms of the signal-to-noise ratio, the Wuppertal and Coulomb-gauge Gaussian sources are similar (see Appendix E). In the next section, we discuss the boosted sources in detail and the energy levels of the boosted pion. In Appendix A, we have explained the construction of boosted sources in detail.

Out of 216 gauge configurations used in our calculations, 24 gauge configurations were analyzed using the graphics processing unit (GPU) cluster in Brookhaven National Lab to calculate two-point and three-point correlation functions. These calculations were performed entirely on GPU using the QUDA suite [45–47], including the inversion of the fermion operator with a multigrid algorithm, communication between GPU devices to perform covariant shifts, and the necessary spin-color matrix multiplications. In QUDA, the communications between GPUs on the same node are implemented through message passing interface or as direct peer-to-peer communications between the GPU devices. We have found that on rare occasion the QUDA peer-to-peer communications did not finish by the time the computations started. These rare glitches happened randomly. We checked, however, that these glitches did not affect our results noticeably compared to other errors.

III. TWO-POINT FUNCTION OF THE BOOSTED PION

We calculated the two-point functions of the positively charged pion ($\pi^+ = \bar{d}u$),

$$C_{2\text{pt}}^{ss'}(t, P_z) = \langle [\pi_s^+(t, \vec{P})][\pi_{s'}^+(0, \vec{P})]^\dagger \rangle, \qquad (4)$$

for a spatial pion momentum $\vec{P} = (0, 0, P_z)$ which is nonzero only along the z-direction, using the pion source and sink $\pi_s^+(0, \vec{P})$ and $\pi_{s'}^+(t, \vec{P})$, respectively. The values of momenta in lattice units are $aP_z = \pm 2\pi n_z/48$ for n_z ranging from 0 to 5, which in physical units correspond

to $P_z = 0$, 0.43, 0.86, 1.29, 1.72, and 2.15 GeV, respectively. We always used the Coulomb gauge Gaussian smeared-source ($s = S$), and either a smeared-sink ($s' = S$) or point-sink ($s' = P$). In the rest of the paper, we will refer to the smeared-source and smeared-sink setup to be SS, and we will refer to the smeared-source point-sink set-up as SP.

For the lowest two momenta, we used the usual Gaussian sources. To improve the signal for the higher momenta, we followed Ref. [44] and used boosted sources in which the valence quarks are boosted to a momentum $k_z = \zeta P_z$, with ζ being a tunable parameter. Naively, one might expect that the optimal choice would be $\zeta = 0.5$. However, we found that the optimal choice of ζ for the pion in terms of the signal-to-noise ratio is between 0.6–0.75. For $P_z = 0.86$ GeV the signal-to-noise ratio is not very sensitive to the value of ζ. These findings are in agreement with Ref. [44]. We discuss the optimization of boosted sources further in Appendix E. Since we need to create a source for each value of ζ, we used $k_z = 2(2\pi/48)$ for $n_z = 2$, 3 and $k_z = 3(2\pi/48)$ for $n_z = 4$, 5, corresponding to the choices of the parameter $\zeta = 1, 2/3, 3/4$, and $3/5$ for $n_z = 2$, 3, 4, and 5, respectively. We have shown the corresponding effective masses for the SS two-point functions in Fig. 1. By using the boosted smeared sources, one can see that a reasonable signal for the two-point correlation function can be obtained up to source-sink separations $t = 12a$ for all momenta except for the highest momentum $P_z = 2.15$ GeV. Simply from the data points in Fig. 1, we see that the effective masses approach a plateau
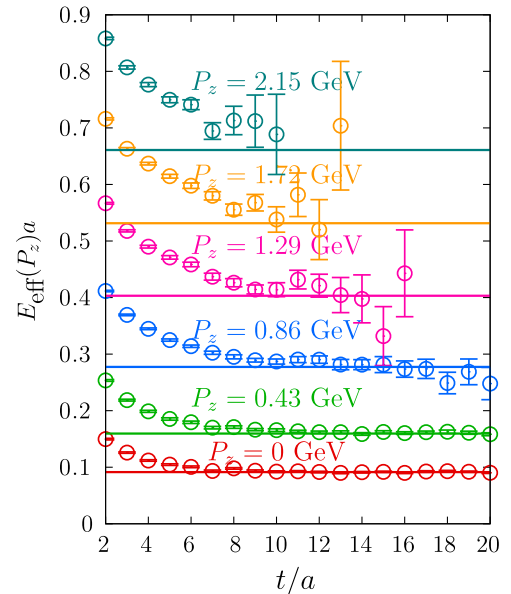

FIG. 1. The effective masses E_{eff} from the pion two-point functions with the boosted Coulomb gauge Gaussian source and sink for different momenta as a function of the source-sink separation t. The horizontal lines are the energy levels from the continuum dispersion relation with $m_\pi = 300$ MeV.

corresponding to the continuum dispersion relation $E_\pi(P_z) = \sqrt{P_z^2 + m_\pi^2}$, shown as the horizontal lines. The effective mass approaches the plateau region at larger source-sink separations when the momentum is increased, as one would expect from the shrinking gap between the ground and excited states as the pion is boosted.

As we will discuss next, we used source-sink separations $t = 8a$, $10a$, and $12a$ for the computation of three-point functions. Therefore, we needed to analyze the excited state contribution to the SS and SP correlators to perform the infinite source-sink extrapolations. For this, we performed multiexponential fits on the SS and SP pion two-point functions in the interval $t \in [t_{min}, t_{max}]$ in order to extract the energy levels. For fixed t_{min}, we varied t_{max} and checked the sensitivity of the result to t_{max}. Then, we repeated the procedure for different values of t_{min}. We found that we were able to reliably extract the ground state $E_\pi(P_z)$ as well as the first excited state $E_1(P_z)$ using the four-parameter two-state fits to the SP correlator instead of using the SS correlator. This could be due to the fact that the contributions from the high-lying energy levels are smaller in the SP

correlator compared to the SS correlator stemming from the possible cancellations between the positive as well as the negative amplitudes that are allowed in the SP correlator. In the top-left and bottom-left panels of Fig. 2, we have shown the systematics of the two-state fits to the SP correlator. In the top-left and the bottom-left figures, we have shown the dependence of the best fit values of E_π (blue circles) and E_1 (black circles) as a function of t_{min} used in the fits. For a given t_{min}, the data points from two values of t_{max} have been clubbed together for $P_z = 0$ and 1.29 GeV, respectively, and it demonstrates that there is no dependence on t_{max}. The ground state is seen to compare well with the expectation from the dispersion relation shown by the black solid lines. The red band shows the values of E_1 chosen as the best estimate of the first excited state.

On the top-right and bottom-right panels of Fig. 2, we show similar plots for the first excited state E_1 as estimated using the SS correlator. The statistical errors of the excited state energy E_1 in the simple two-state fits (magenta diamond) quickly grow large with increasing t_{min}, and, thus, these fits turned out to be of limited use. Therefore, we performed constrained two- and three-state exponential fits for the SS correlator with the ground state energy fixed to $E_\pi = \sqrt{P_z^2 + m_\pi^2}$ with $m_\pi = 300$ MeV, and imposing a prior on E_1 using its best estimate from the SP correlator— that is, we added the term $(E_1 - E_{1,\text{prior}})^2 / \sigma_{\text{prior}}^2$ to the χ^2 with $E_{1,\text{prior}}$ and σ_{prior} being the mean and error of E_1, respectively, as determined from the SP correlator. The t_{min} dependence of the resulting E_1 from the constrained two-state fit (black circles) and constrained three-state fit (blue triangles) are shown in the top-right and bottom-right panels. The two-state fits of the SS two-point correlator largely overestimate the energy of the first excited state for small t_{min} whether or not priors are used, and there is a significant dependence on t_{min}. One should use $t_{min} \geq 6a$ to obtain reliable results for the first excited state from the SS correlator. The three-state fits with priors on E_π and E_1 give energies of excited states that are the same within errors for the SS two-point correlators and show almost no t_{min} dependence. In summary, we determined the lowest three energy levels using the SP correlator and then determined the corresponding amplitudes $|A_n| = |\langle 0|\pi_S^+(0, \vec{P})|E_n, P_z\rangle|$ of these excited states in the SS correlator through a constrained fit analysis.

In Fig. 3, we show the three energy levels obtained from the different fits discussed above, as a function of P_z. For $P_z = 0$, we compare our result with the energy levels that would correspond to the pion resonances $\pi(1300)$ and $\pi(1800)$ from the particle data group (PDG) [48]. In order to account for the 300 MeV pion mass, we shifted the PDG values by 0.161 GeV as an approximation and these are shown as the two arrows in Fig. 3. Our estimate of the first excited state energy agrees with this shifted mass of $\pi(1300)$. We also show the expected P_z-dependence of $E_1(P_z)$ assuming a particlelike dispersion, and this
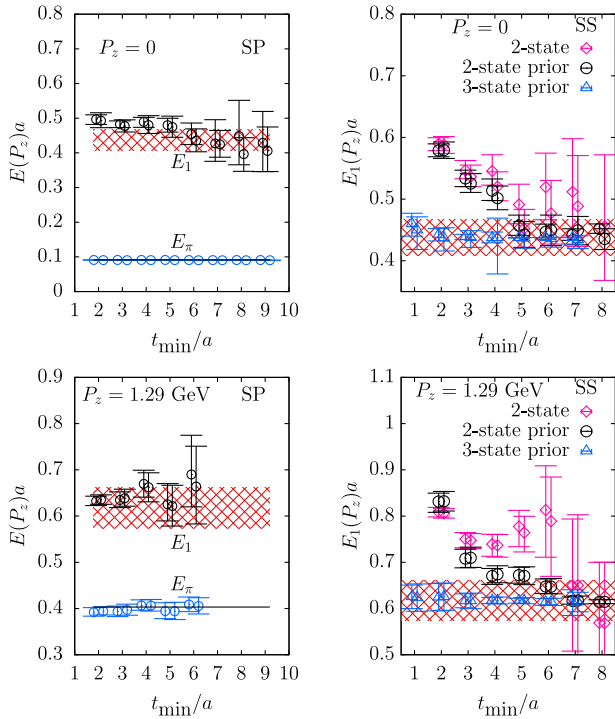

FIG. 2. The systematical dependence of the ground state E_π and the first excited state E_1 on the fit range $[t_{min}, t_{max}]$ is shown. In the left panels, we show such a dependence for the pion SP correlator at two different P_z. For each t_{min}, data from $t_{max} = 24a$ and $32a$ are shown. The black solid line is the value of E_π expected from the continuum dispersion relation. The red patterned band is our best estimate of E_1 using the SP correlator. In the right panels, the fit systematics of E_1 for the SS correlator is shown. The red band is the prior used for E_1 from the SP correlator (same as the one in the left panels). The different symbols are the various fit strategies.
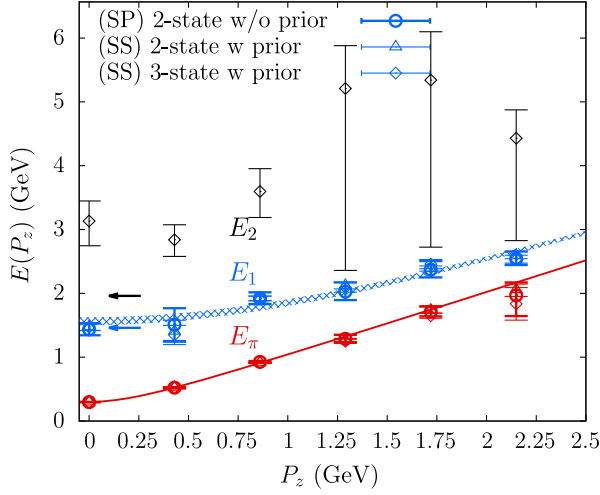
FIG. 3.   The energies of the ground state and the first two excited states as functions of P_z. The red, blue, and black symbols correspond to E_π, E_1, and E_2, respectively. For each color, the different symbols correspond to different fitting methods (2-state and 3-state fit with or without prior on the ground state) and the types of source-sink (SP or SS). The lines show the expected dispersion relations for the pion and its first excited state. The arrows are the PDG values of $\pi(1300)$ and $\pi(1800)$ which are shifted to account for $m_\pi = 300$ MeV.

describes the actual data very well. The energy of the second excited state is much larger than expected, meaning that the third state effectively parametrizes several higher lying states. As one can see from the figure the energy gap between E_π and E_1 shrinks with increasing P_z as expected. The results on the excited state energies will be important for the analysis of the pion three-point function discussed in the next section.

IV. EXTRACTION OF THE BARE QUASI-PDF MATRIX ELEMENTS FROM THE THREE-POINT FUNCTIONS

The next step is the calculation of the bare qPDF matrix element

$$h_\Gamma^B(z, P_z) = \langle E_\pi, P_z | \mathcal{O}_\Gamma(z; \tau) | E_\pi, P_z \rangle, \quad (5)$$

where the bilocal $u - d$ qPDF operator in a time-slice τ involving a quark and an antiquark separated along the z-direction by $\mathcal{L} = (0, 0, 0, z)$ is given by

$$\mathcal{O}_\Gamma(z; \tau) = \sum_{\vec{x}} (\bar{u}_x W_{x,x+\mathcal{L}} \Gamma u_{x+\mathcal{L}} - \bar{d}_x W_{x,x+\mathcal{L}} \Gamma d_{x+\mathcal{L}}), \text{ where}$$

$$\Gamma = \gamma_t, \gamma_z, 1; \quad W_{x,x+\mathcal{L}} = \prod_{x'=x}^{x+\mathcal{L}} U_3(x'), \quad (6)$$

and it is made gauge-invariant by the Wilson line $W_{x,x+\mathcal{L}}$. The Dirac γ matrices in the qPDF operator are in the

Minkowskian convention. The state $|E_\pi, P_z\rangle$ denotes the on-shell ground state pion with momentum P_z. In addition to the natural choices of $\Gamma = \gamma_t$ and γ_z that approach γ^+ in the light-cone limit, we also considered $\Gamma = 1$. This choice of Γ is needed because under renormalization, $\mathcal{O}_{\gamma_z}(z)$ mixes with $\mathcal{O}_1(z)$ [6]. We applied one-level of HYP smearing to the links entering $W_{x,x+\mathcal{L}}$ in order to reduce the noise. Since the qPDF calculation involves values of $z \sim \mathcal{O}(a)$, we checked that there is no significant difference between the renormalized matrix elements using the smeared and unsmeared Wilson line. To obtain the bare matrix element $h_\Gamma^B(P, z)$, we computed the three-point function at different source-sink separations t and operator insertion point τ,

$$C_{3pt}^{SS}(t, \tau; z, P_z) = \langle [\pi_S^+(t, \vec{P})] \mathcal{O}_\Gamma(z; \tau) [\pi_S^+(0, \vec{P})]^\dagger \rangle, \quad (7)$$

and constructed the ratio of the three-point function to two-point function,

$$R(t, \tau; z, P_z, \Gamma) = \frac{C_{3pt}^{SS}(t, \tau; z, P_z)}{C_{2pt}^{SS}(t, P_z)}. \quad (8)$$

The reader can refer to Appendix B for a detailed description of the construction of three-point functions. The two-point function is always real when the source and sink are of the same type. The three-point function for the $u - d$ qPDF operator $\mathcal{O}_\Gamma(z)$ in a pion external state is real at all z for $\Gamma = \gamma_z, \gamma_t$ and purely imaginary for $\Gamma = 1$ (refer Appendix C). Inserting a complete set of states in the above equation,

$$R(t, \tau; z, P_z, \Gamma)$$
$$= \frac{\sum_{n,n'} A_n A_{n'}^* \langle E_n, P | \mathcal{O}_\Gamma(z) | E_{n'}, P \rangle e^{-(E_{n'} - E_n)\tau - E_n t}}{\sum_m |A_m|^2 e^{-E_m t}}, \quad (9)$$

with $E_{n+1} \geq E_n$, and $E_0 = E_\pi$. It is easy to see that in the infinite t limit, $R(t, \tau; z, P_z, \Gamma)$ is equal to $h_\Gamma^B(z, P)$. The above equation holds for the infinite time extent. For a finite time extent, the effects of the periodic boundary condition should be taken into account. This turns out to be important for $P_z = 0$, while for nonzero P_z the effect is negligible as discussed in Appendix G. In practice, one truncates the sums in Eq. (9) at some value n, and then obtains $h_\Gamma^B(z, P_z)$ by fitting the t and the τ dependence of $R(t, \tau; z, P_z, \Gamma)$ using $\langle E_n, P | \mathcal{O}_\Gamma(z) | E_{n'}, P \rangle$ as fit parameters. In the fits, the values of A_n and E_n were held fixed at values determined from the two-state fit analysis on the SS correlators. In what follows, we will refer to this method of fitting using n-state ansatz to the data between $\tau/a > \tau_o$ and $\tau/a < t/a - \tau_o$ as $\texttt{Fit}(n, \tau_o)$. In this method, the excited states are suppressed by $\exp(-(E_n - E_\pi)t/2)$. For $z = 0$, it is easy to see that $R(t, \tau; z, P_z, \Gamma)$ is symmetric in τ around the midpoint $\tau - t/2$. For $z \neq 0$ and $P_z \neq 0$, we only have the following relation (see Appendix D):

$$\langle E_n, P_z | \mathcal{O}_\Gamma(z) | E_{n'}, P_z \rangle \rangle^* = \phi_\Gamma \langle E_{n'}, -P_z | \mathcal{O}_\Gamma(z) | E_n, -P_z \rangle \rangle, \tag{10}$$

where $\phi_\Gamma = 1$ for $\Gamma = \gamma_t$, 1 and $\Phi_\Gamma = -1$ for $\Gamma = \gamma_z$. Thus, generically the matrix elements $\langle E_n, P_z | \mathcal{O}_\Gamma(z) | E_{n'}, P_z \rangle \rangle$ and $\langle E_{n'}, P_z | \mathcal{O}_\Gamma(z) | E_n, P_z \rangle \rangle$ are independent and the number of fit parameters is thereby increased. Based on the above relation, we constructed appropriate averages using both the positive and negative values of momenta to increase the statistics. However, in practice the gain was marginal.

We demonstrate the extraction of the matrix element using the fit method in Fig. 4, where the ratios $R(t, \tau; z, P_z, \Gamma)$ are shown for $\Gamma = \gamma_t$ and γ_z qPDFs for the $P_z = 1.29$ GeV pion. The results on $R(t, \tau; z, P_z, \Gamma)$ and the extraction of matrix elements for the other values of P_z are given in Appendix G. In the figure, we show the data at $t/a = 8$, 10, and 12 along with the result of Fit(2,2). Using the fit, the results for the $t \to \infty$ extrapolations are shown with the horizontal bands. We also performed the three-state fit of $R(t, \tau; z, P_z, \Gamma)$ and the picture looks similar. In this case, the data points at all $\tau - t/2$ could be described by the fit. The $t \to \infty$ extrapolations from the three-state fit gave results consistent with the two-state ones, albeit with larger errors. A closer look at Fig. 4 (and also from Fig. 21) reveals that the excited state contribution

is larger for $\Gamma = \gamma_z$ than for $\Gamma = \gamma_t$. Furthermore, the excited state contribution grows with increasing z. The nonsymmetric nature of $R(t, \tau; z, P_z, \Gamma)$ for $z \neq 0$ is also apparent in the figure. We expect that $h_\Gamma(z = 0, P) = 1$ for $\Gamma = \gamma_t$ because of the charge conservation once a proper renormalization is implemented and the continuum limit is taken. Our extrapolation procedure gives a result for the bare matrix element which is larger than one at all the values of momenta as can be seen in Fig. 4, as well as from Fig. 22 in Appendix G, where, in addition, one can also see that $h_{\gamma_t}^B(z = 0, P_z)$ is independent of P_z. Thus, any deviation of $h_{\gamma_t}^B(z = 0, P_z)$ away from unity should be taken care of by the renormalization. We will see in Sec. V that this is indeed the case.

Alternatively we can use the summation method [49] to obtain h_Γ^B. Here one sums over all τ/a minus a certain number of end points τ_o

$$R_{\text{sum}}(t; z, \Gamma) = \sum_{\tau/a=\tau_o}^{t/a-\tau_o} R(t, \tau; z, P_z, \Gamma). \tag{11}$$

We will refer to this method as $\texttt{sum}(\texttt{sum}(\tau_o))$. For large t, one would find a linear behavior in t of R_{sum} as

$$R_{\text{sum}}(t; z, \Gamma) \simeq (t - 2\tau_o)h_\Gamma^B(z, P_z) + \text{const} + \mathcal{O}(e^{-(E_1-E_\pi)t}). \tag{12}$$
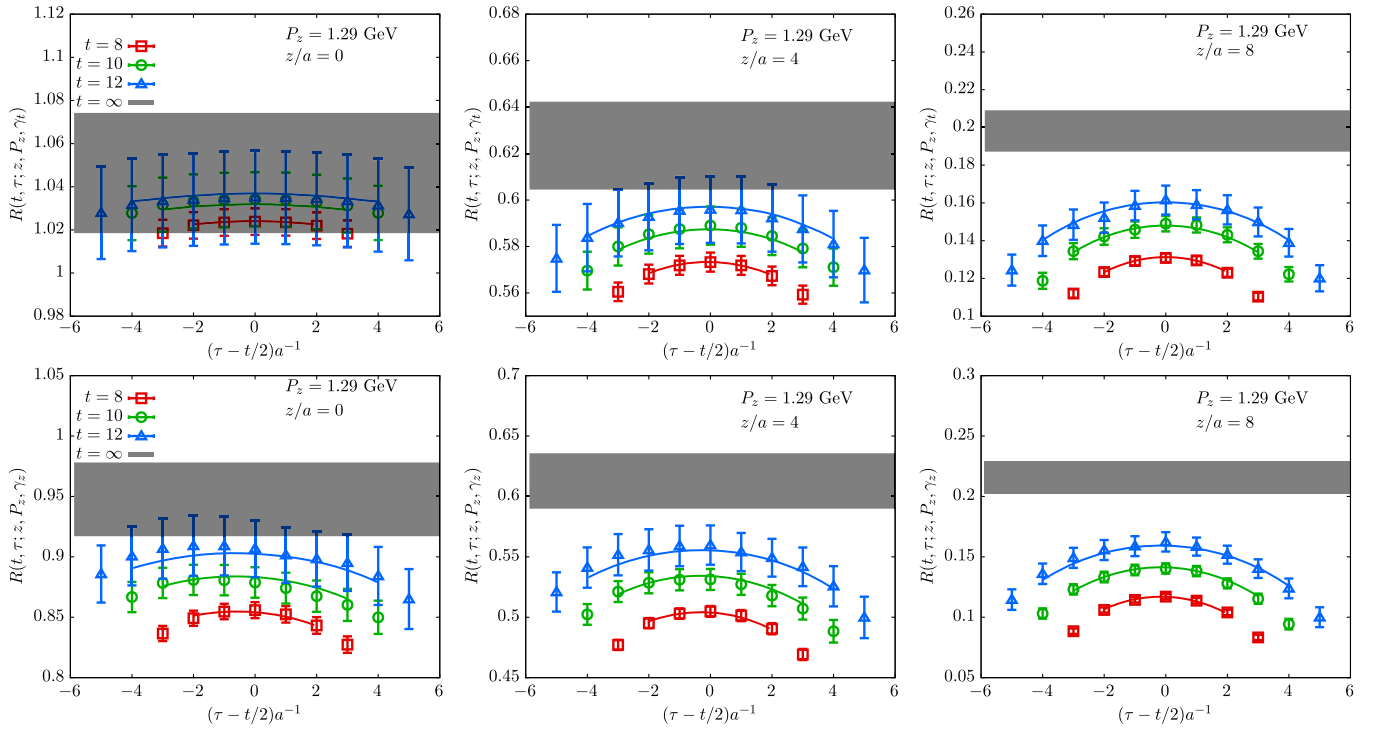

FIG. 4.   The ratio of the three-point function to the two-point function, $R(t, \tau; z, P_z, \Gamma)$ is shown as function of $\tau - t/2$ for γ_t (top row) and γ_z (bottom row) for $z/a = 0$, 4 and 8 (from left to right) and $P_z = 1.29$ GeV. The central values of the two-state fits of the lattice results for different source-sink separations are shown as the curves. The horizontal band corresponds to the extrapolated result for the infinite source-sink separation.
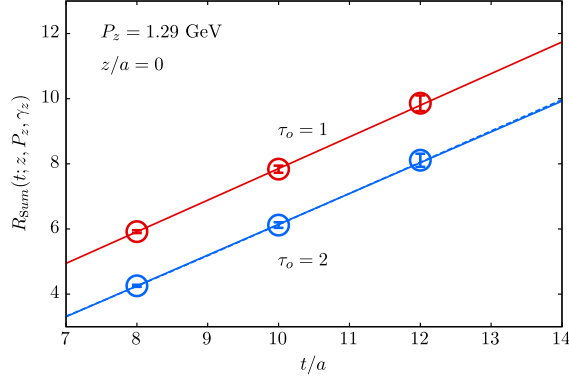
FIG. 5. The t dependence of $R_{sum}(t; z = 0, \gamma_z)$ for $P_z = 1.29$ GeV with $\tau_o = 1$ (red) and $\tau_o = 2$ (blue). The solid lines are the straight line fit $h^B_{\gamma_z}(z, P_z)t + \text{const}$ to the data. The dashed lines are the expected curve for $R_{sum}(t; z = 0, \gamma_z)$ using the `fit` (2,2) best fit parameters.

The advantage of this method is that the excited state contributions are suppressed as $\exp[-(E_n - E_\pi)t]$ instead of being suppressed as $\exp[-(E_n - E_\pi)t/2]$ in the fitting method. We show a sample result using `sum(1)` and `sum(2)` in Fig. 5 for $\Gamma = \gamma_z$ and $z = 0$. We see that $R_{sum}(t; z, \gamma_z)$ can be well fitted by a straight line in t, and the slope gives the value of the matrix element. As a cross check, we also show the expected curve for $R_{sum}(t; z, \Gamma)$ using our best fit from `Fit(2,2)` as the dashed curves. It can be seen that the difference between a simple straight line fit and the curve from `Fit(2,2)` is small. One can also note that `sum(1)` and `sum(2)` are almost parallel, meaning that the extracted matrix element is independent of τ_o confirming that the method works well.

To better understand the systematic effects due to excited state contaminations, one can look at the case $z = 0$ in detail, where the statistical errors are the smallest. The bare matrix element $h^B_{\gamma_z}(z = 0, P_z)$ after renormalization is expected to be proportional to the hadron velocity, $P_z/E_\pi(P_z)$. One can take the ratio of matrix elements $h^B_{\gamma_z}(z = 0, P_z)/h^B_{\gamma_t}(z = 0, P_z)$ to avoid the issues of renormalization. The results for this ratio of matrix elements is shown in Fig. 6 along with the curve for $P_z/\sqrt{P_z^2 + m_\pi^2}$. We see that our lattice results follow the expectations reasonably with small 3–4% deviations from the expected result at small P_z. A reason for this could be the large systematic uncertainty in γ_z matrix element due to the relatively larger excited state extrapolations required. We see that within errors, the two-state fit, three-state fit, and the summation methods are consistent.

In Fig. 7, we show the results for $h^B_\Gamma(z, P_z)$ as functions of z for the two highest momenta $P_z = 1.29$ and 1.72 GeV determined using the HYP smeared Wilson line. Since the real part is symmetric, about $z = 0$, we have only shown the data for $z \geq 0$. At each z/a, we have shown the resulting $t \to \infty$ extrapolated results using `Fit(3,1)`, `Fit(2,2)`,
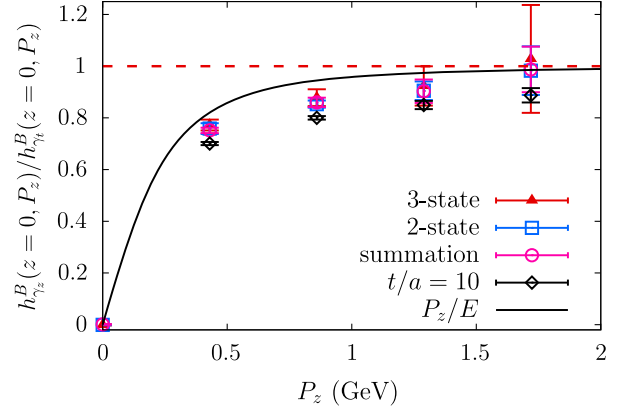

FIG. 6. The ratio of the matrix elements for γ_z to γ_t as a function of P_z. The curve shows the expected result, $P_z/E_\pi(P_z)$.

`sum(2)`, and `sum(2)` methods, and these points at a given z/a are slightly displaced for better visibility. We see that the results from all these methods agree with each other within the errors. For γ_z, some tension between the summation method, and the two- and three-state fits is observed at larger $|z|$. At larger values of z the matrix elements are suppressed partly because of the larger value of P_z and in part by the divergent self-energy contribution in the spatial Wilson line. The latter will be removed upon renormalization as we will see in the next section. Having demonstrated a robust determination of the matrix element using multiple $t \to \infty$ strategies, we will use the matrix elements obtained using `Fit(2,2)` in the rest of the paper.

So far we discussed results on the three-point function obtained using 1-HYP smearing for the spatial link. We also performed calculations using the unsmeared spatial Wilson line. In this case, the bare matrix element rapidly decreased with z due to the larger value of the Wilson line self-energy divergence. However, we found that the results, after nonperturbative renormalization (discussed next in Sec. V), were similar to those obtained with smeared Wilson lines within errors. The main difference between the renormalized three-point function obtained with the smeared and unsmeared Wilson line is that, for the latter, the statistical errors at large z are significantly larger.

V. RI-MOM NONPERTURBATIVE RENORMALIZATION AND ITS COMPARISON WITH 1-LOOP

In the last section, we discussed the extraction of the bare qPDF matrix element which has to be renormalized. The renormalizability of qPDF has been recently demonstrated to all orders of perturbation theory [50,51]. In addition to the quark wave function renormalization Z_q and the composite operator renormalization required for $z = 0$, the qPDF operator at nonzero z requires additional renormalization due to the UV divergence present in the Wilson line connecting the quark and antiquark [52]. When a
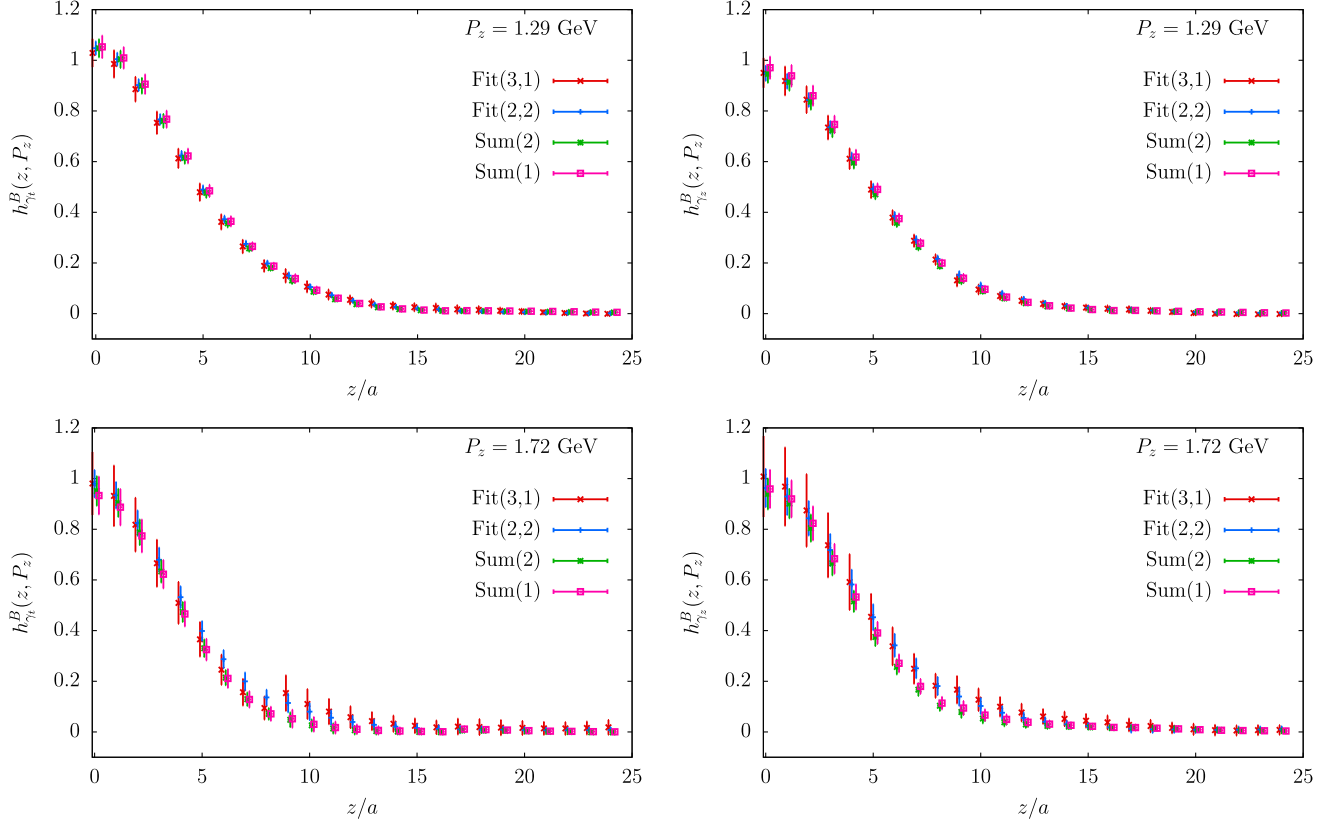
FIG. 7.   The bare matrix elements $h^B_{\gamma_t}(z, P_z)$ (left) and $h^B_{\gamma_z}(z, P_z)$ (right) as a function of quark-antiquark separation z. The panels in the top row show results for $P_z = 1.29$ GeV, while the panels in the bottom row show the results for $P_z = 1.72$ GeV. The different symbols are from various methods of $t \to \infty$ extrapolation.

lattice fermion that breaks chiral symmetry at finite lattice spacings is used, as is the case in this paper, it has been shown that only the renormalization of the γ_t qPDF operator is purely multiplicative, while the γ_z qPDF operator mixes with the scalar qPDF [6,20]. A renormalization scheme that is implementable on the Euclidean lattice is the RI-MOM scheme, and it is now standard in the lattice QCD literature. The corresponding RI-MOM counter-term for the qPDF operator in leading-order perturbation theory has been worked out using off-shell quark external states [8], and it is one of the ingredients used in the perturbative matching of the RI-MOM renormalized qPDF to the $\overline{\text{MS}}$ PDF. In this section, we discuss the renormalization procedure, and then compare the running of the renormalization constants as determined on the lattice with the corresponding perturbative expectations. This allows us to quantitate the validity of the leading-order perturbation theory and matching.

For nonperturbative renormalization, we compute the expectation value of qPDF operator between off-shell quark external states with momentum p. We refer to the momentum of quark in the direction of the Wilson line as p_z and the magnitude of the component perpendicular to the Wilson line as $p_\perp$. For these computations, we use Landau gauge fixing. Let $\Lambda_\Gamma(z, p)$ be the quark-line amputated bare qPDF,

$$\Lambda_\Gamma(z, p) = \langle Q(p) \rangle^{-1} \langle \bar{u}(p) \mathcal{O}_\Gamma(z, \tau) u(p) \rangle \langle Q(p) \rangle^{-1}, \quad (13)$$

where $Q(p)$ is the quark propagator $\langle u(p) \bar{u}(p) \rangle$ and $u(p) = \sum_x u_x e^{-ip.x}$. Let us define the bare qPDF after projection as

$$q_\Gamma(z, p) \equiv \text{Tr}[\mathcal{P} \Lambda_\Gamma(z, p)], \quad (14)$$

consistent with the definition used in perturbative calculations. Here, $\mathcal{P}$ is the operator used to project onto one of the γ-matrices $\Gamma = \gamma_\alpha$, and $\text{Tr}(...)$ is a trace over both color and Dirac indices. Based on previous works [8,20], we will use $\slashed{p}$-projection for which $\mathcal{P} = \slashed{p}/(12 p_\alpha)$. Alternatively, one can use $\mathcal{P} = \Gamma$ [21] or the minimal projection [8]. In the case of $\Gamma = \gamma_t$, since the renormalization is simply multiplicative, the renormalized quark qPDF is given by

$$q^R_{\gamma_t}(z, p, p^R) = Z_{\gamma_t \gamma_t}(z, p^R) q_{\gamma_t}(z, p), \quad (15)$$

where the z-dependent RI-MOM renormalization constant Z is determined using the renormalization condition set at momentum p^R as

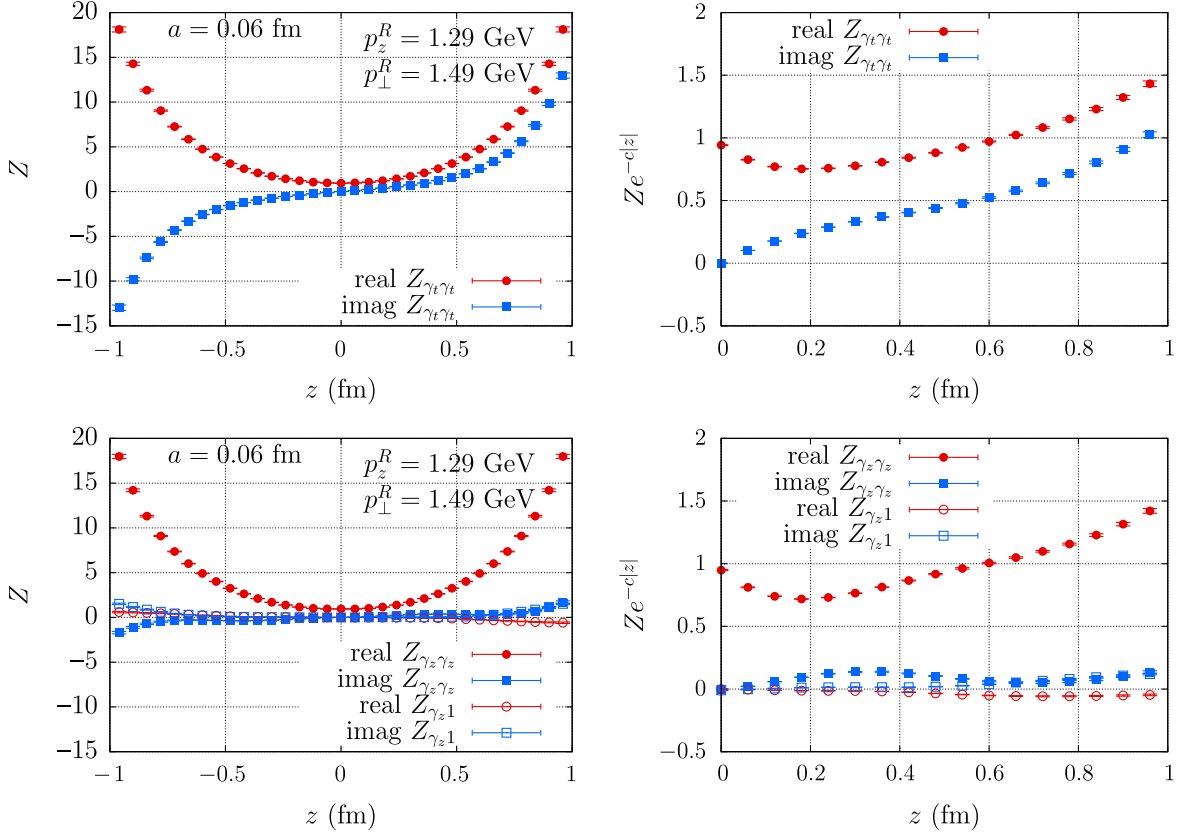$$q^R_{\gamma_t}(z, p, p^R)|_{p=p^R} \equiv e^{i p^R_z z}. \quad (16)$$

FIG. 8. The RI-MOM renormalization constants using $\not{p}$-projection at lattice spacing $a = 0.06$ fm at renormalization scale $p_z = 1.29$ GeV, $p_\perp = 1.49$ GeV are shown. In the top-left panel, the real and imaginary parts of the renormalization constant for γ_t qPDF operator are shown as a function of quark-antiquark separation z in physical units. On the top-right panel, the self-energy divergent part $e^{-c|z|}$ of the Wilson line is removed from renormalization constant for γ_t qPDF. Similarly, in the bottom-left panel, the diagonal part $Z_{\gamma_z\gamma_z}$ and the mixing term $Z_{\gamma_z 1}$ are shown, and the corresponding values after the removal of self energy divergence are shown in the bottom-right panel.

The right-hand side of the above equation is the tree-level value of q_{γ_t}. The renormalization constant so obtained is in general a complex number. For $\Gamma = \gamma_z$, we have to take care of mixing with the scalar $\Gamma = 1$. Hence, the renormalized qPDF is defined as

$$q_{\gamma_z}^R(z, p, p^R) = Z_{\gamma_z\gamma_z}(z, p^R)q_{\gamma_z}(z, p) + Z_{\gamma_z 1}(z, p^R)q_1(z, p).$$

$$(17)$$

The diagonal part $Z_{\gamma_z\gamma_z}$ and the mixing term $Z_{\gamma_z 1}$ are determined using the two RI-MOM conditions [20]:

$$q_{\gamma_z}^R(z, p, p^R)|_{p=p^R} \equiv e^{ip_z^R z},$$

$$Z_{\gamma_z\gamma_z}(z, p^R)\mathrm{Tr}\Lambda_{\gamma_z}(z, p) + Z_{\gamma_z 1}(z, p^R)\mathrm{Tr}\Lambda_1(z, p)|_{p=p^R} \equiv 0.$$

$$(18)$$

Using the renormalization constants Z determined above using quark external states, the renormalized pion qPDF can also be determined by

$$h_{\gamma_t}^R(z, P_z, p^R) = Z_q Z_{\gamma_t\gamma_t}(z, p^R)h_{\gamma_t}^B(z, P_z),$$

$$h_{\gamma_z}^R(z, P_z, p^R) = Z_q Z_{\gamma_z\gamma_z}(z, p^R)h_{\gamma_z}^B(z, P_z)$$
$$+ Z_q Z_{\gamma_z 1}h_1^B(z, P_z),$$

$$(19)$$

where Z_q is the quark renormalization, that can be determined using the condition [53]

$$Z_q(p^R)^{-1}\frac{1}{12}\mathrm{Tr}(\langle Q(p^R)\rangle^{-1}Q_{\mathrm{tree}}(p^R)) = 1,$$

$$(20)$$

where $Q(p)$ is the quark propagator determined using the Landau gauge, and Q_{tree} is the free quark propagator for which we use the free massless Wilson-Dirac propagator.

In Fig. 8, we show the renormalization factors using the above RI-MOM renormalization conditions on the 0.06 fm ensemble. On the top-left panel of Fig. 8, we show the real and imaginary parts of $Z_{\gamma_t\gamma_t}$ determined at $p_z^R = 1.29$ GeV and $p_\perp^R = 1.49$ GeV. The rapid, almost exponential, increase in Z with z is due to the self-energy divergence present in the bare Wilson line that connects the quark and antiquark in the qPDF operator. This divergent piece, $e^{c|z|}$,

cannot be captured perturbatively, and it needs to be determined nonperturbatively in a particular scheme. However, this might not be an issue for the one-loop matching if $e^{c|z|}$ cancels exactly between the renormalization factors and the bare qPDF operator. Therefore, we remove $e^{c|z|}$ from the renormalization constant that is shown in the top-left panel, and display the result in the top-right panel. The value of c for our $a = 0.04$ fm ensemble was determined in [54], and for 1-HYP Wilson line $ca = 0.1586$. This removal of Wilson line self-energy reduces the almost exponential dependence of $Z(z)$ to a weak dependence on z. In fact, we see that both the real and imaginary parts of $Z_{\gamma_t\gamma_t}$ remain $\mathcal{O}(1)$ even up to $z = 1$ fm, thereby providing a qualitative justification for the usage of leading-order perturbation theory to describe the lattice data at short z and at high quark momenta. We show similar data in the bottom-left and -right panels for the Z-factors for γ_z qPDF. In this case, we have the diagonal factor $Z_{\gamma_z\gamma_z}$ as well as the off-diagonal factor $Z_{\gamma_z 1}$ to take care of mixing with scalar on the lattice. We show $Z_{\gamma_z\gamma_z}$ and $Z_{\gamma_z 1}$ as the filled and unfilled symbols in the bottom panels, respectively. We observe that the imaginary part of $Z_{\gamma_z\gamma_z}$ is small compared to the real part. This is not the case for $Z_{\gamma_t\gamma_t}$, which in turn will affect the asymmetry of the $u - d$ qPDF $\tilde{q}_{u-d}(x)$ of pion, about $x = 0$. We also note that the mixing of γ_z with the scalar is a minor 5-10% effect, but we nevertheless take care of it in our calculation.

As we discussed in the last section, the matrix element at $z = 0$, h_{γ_t} is the local current operator which will be exactly conserved in the continuum limit. Hence, $Z_q Z_{\gamma_t\gamma_t}(z = 0)$ is the vector current renormalization factor Z_V and the dependence of Z_V on p will give us an idea of the leading $(pa)^2$ perturbative lattice artifacts for values of $p \gg \Lambda_{\text{QCD}}$ as well as the other higher order (or perhaps nonperturbative) contributions to this lattice correction to Z_V at smaller renormalization scales [53]. In Fig. 9, we show Z_q
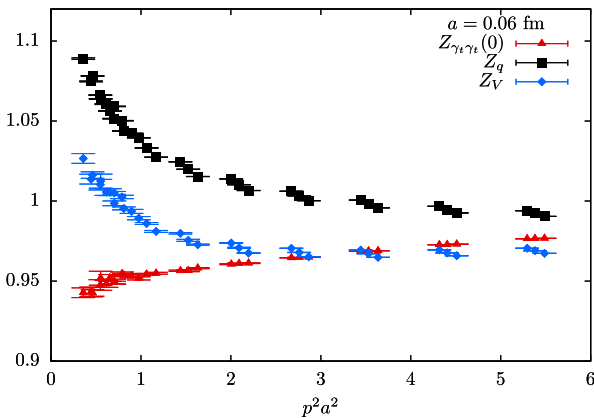

FIG. 9. The $(pa)^2$ dependence of the renormalization factors Z_q (black squares), $Z_{\gamma_t\gamma_t}(z = 0)$ (red triangles), and $Z_V = Z_q Z_{\gamma_t\gamma_t}(0)$ (blue diamonds) are shown for the $a = 0.06$ fm ensemble.

determined using Eq. (20), the value of $Z_{\gamma_t\gamma_t}$ at $z = 0$ as well as their product Z_V as a function of $(pa)^2$. One sees a reasonable plateau for $Z_V \approx 0.97$ only for $(pa)^2 > 2$. For comparison, the value of Z_V as obtained from the bare pion isospin charge $h^B_{\gamma_t}(z = 0, P_z = 0)$ is 0.961(3). The values of Z_V determined from $h^B_{\gamma_t}(z = 0, P_z)$ at the other nonzero P_z also give consistent values. With the uncertainties of choosing the scaling region in $(pa)^2$ to take the $(pa) \to 0$ limit of Z_V, we expect the Z_V to be in the range 0.97 to 0.99. For relatively smaller values of renormalization momenta $(pa) \approx 1$–1.5, chosen such that the renormalization scales lie in the vicinity of the pion momenta used in this paper, one sees noticeable, but small, 5% dependence on pa. We used the value of Z_q estimated at the same value of p as used in $Z_{\gamma_t\gamma_t}$ for renormalizing our pion qPDF.

A. Comparison with leading-order perturbation theory for $z < 0.3$ fm

We will now investigate in a quantitative way the agreement/disagreement of the lattice determination of the RI-MOM renormalized amputated quark qPDF at $z < 0.3$ fm which one can expect to be in the perturbative regime. For this, we construct a quantity $\zeta_\Gamma(z, p, p^R)$ in the following way:

$$\zeta_\Gamma(z, p, p^R) = \frac{q_\Gamma^R(z, p, p^R)}{q_\Gamma^R(z, p, p)} - 1, \qquad (21)$$

where $q_\Gamma^R(z, p, p) = e^{ip_z z}$ by the renormalization condition. In the case of $\Gamma = \gamma_t$, the above definition is simply

$$\zeta_{\gamma_t}(z, p, p^R) = \frac{Z_{\gamma_t\gamma_t}(z, p^R) - Z_{\gamma_t\gamma_t}(z, p)}{Z_{\gamma_t\gamma_t}(z, p)}, \qquad (22)$$

which is similar to a discrete scale-dependent anomalous dimension $\partial \log(Z_{\gamma_t\gamma_t}(z, p))/\partial p$. Through the dependence of ζ on p^R slightly away from p, we can understand how well the leading-order perturbation theory is able to describe the exact nonperturbative determination on the lattice. It is important to stress that apart from understanding nonperturbative renormalization of qPDF in this way, we are also essentially comparing one of the steps in the LaMET formalism that is calculable on the lattice. Hence, any agreement/disagreement we observe quantifies the limitations of the leading-order LaMET. In perturbation theory, ζ is the ratio of the one-loop perturbative correction to $q(z, p)$ to its tree-level value. This expression for ζ has been calculated, and it is given by[1]

---

[1]The formula differs from the one given in [8] due to the issue of the order of the $\epsilon = 0$ limit in dimensional regularization and the $z = 0$ limit. We thank Yong Zhao for communicating the corrected result to us.
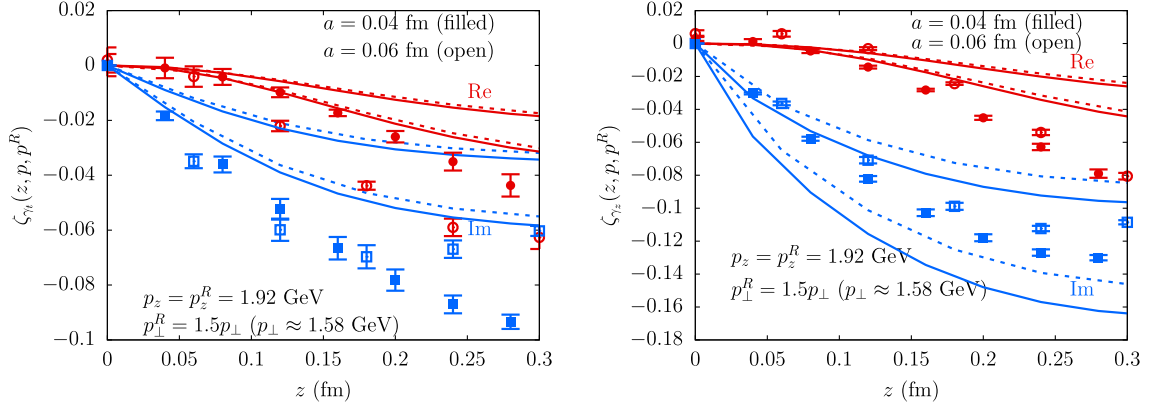
FIG. 10. The behavior of $\zeta(z)$ (symbols) with the quark-antiquark separation z, is compared with the expectation from 1-loop perturbation theory (bands) when $p_\perp^R$ is slightly away from $p_\perp$. The red symbol and bands are the real part of ζ while the blue ones correspond to the imaginary part. The bands enclosed by solid curves corresponds to the momenta $(p_z, p_\perp)$ for the 0.04 fm data, and similarly the band enclosed by the dashed curves corresponds to $(p_z, p_\perp)$ of the 0.06 fm data (see text). On the left and right panels, the comparisons are made for ζ_{γ_t} and ζ_{γ_z} respectively. In each of the panels, the data from two different lattice spacings are also shown ($a = 0.04$ fm as filled circles and $a = 0.06$ fm as open circles). For the data shown, $p_z = p_z^R = 1.92$ GeV. The transverse momentum of the quark $p_\perp = 1.58$ GeV, and the transverse renormalization momentum is chosen to be 1.5 times $p_\perp$.

$$\zeta_\Gamma(z, p, p^R) = \frac{\alpha_s C_F}{2\pi} \int_{-\infty}^{\infty} dx (e^{i(1-x)p_z} - 1) \left[H(x, p) \right.$$
$$\left. - \left| \frac{p_z}{p_R} \right| H \left(1 + \frac{p_z}{p_z^R}(x-1), p^R \right) \right], \quad (23)$$

where $H(x, p)$ is the 1-loop correction term to the bare qPDF, and the two terms in the right-hand side come from the bare and RI-MOM renormalization counter terms, respectively.[2] The functional forms of $H(x, p)$ for γ_z and γ_t isovector qPDFs are given in [7,8] and therefore we do not provide them here. The asymptotic $3/(2|x|)$ behavior of the bare and the RI-MOM counter term, that contributes to the UV divergence when integrated over x, gets exactly canceled and we obtain a UV finite and renormalized result for ζ. In the discussions below, we will consider the cases with $p_z = p_z^R$ and $p_z \neq p_z^R$ separately. In the above leading-order formula, the scale at which α_s has to be evaluated is arbitrary. Therefore, we vary α_s by changing the scale from $0.5 p_z^R$ to $2 p_z^R$ through the 1-loop running, and quote this variation as an uncertainty in the perturbative results below. On the lattice side, we determine $\zeta(z, p, p^R)$ using the nonperturbatively determined Z-factors. In order to estimate the lattice spacing effects, we determined ζ using two different lattice spacings; $a = 0.04$ fm is shown as filled symbols and $a = 0.06$ fm is shown as open symbols in the various plots that follow.

In the left and right panels of Fig. 10, we show the typical dependence of $\zeta_{\gamma_t}(z)$ and $\zeta_{\gamma_z}(z)$, respectively, as a function of z when $p_\perp^R$ differs slightly from the transverse quark momentum $p_\perp$, while the longitudinal components p_z and p_z^R are the same. Using Eq. (23), we calculated the prediction from leading-order perturbation theory for $\zeta(z)$ at the same values of momenta. The uncertainty bands for the perturbative result are shown in Fig. 10 along with the actual lattice data at the two different lattice spacings that are shown using symbols. For the data shown in Fig. 10, the longitudinal components p_z for the two lattice spacings are exactly 1.92 GeV, but the transverse components $p_\perp$ are only approximately the same between the two due to the constraints of allowed momenta on the two different lattice volumes, i.e., $p_\perp = 1.49$ GeV for $a = 0.06$ fm and $p_\perp = 1.67$ GeV for $a = 0.04$ fm. To take care of this slight offset in $p_\perp$ between the two lattice spacings, we have distinguished the perturbative results corresponding to $a = 0.04$ fm as bands enclosed by solid lines, and similarly for $a = 0.06$ fm as bands enclosed by dashed lines. It can be seen that the two perturbative results are not very sensitive to this difference in $p_\perp$ assuring us that whatever change we observe between the data at two different a is mainly due to the change in a. We observe from the plots that the leading-order perturbation theory captures the qualitative z-dependence of both the real and imaginary parts of ζ when $p_\perp^R$ is changed from $p_\perp$. Surprisingly, the 1-loop result seems to work better for ζ_{γ_z} than for ζ_{γ_t}. In the case of γ_t, one can certainly see a large lattice spacing effect with the movement of data towards the 1-loop result as the lattice spacing is reduced, while in the case of γ_z, one can already see a consistency with the one loop result at the lattice spacings that we use. Thus, it opens up a question on whether the γ_z qPDF fares worse compared to the γ_t qPDF simply due to the presence of mixing with the scalar or whether γ_z qPDF might eventually show better perturbative convergence and

---

[2] The function $H(x, p)$ here is referred to as $h(x, p)$ in Ref. [8]. We reserve h to refer to qPDF matrix element as is the convention.
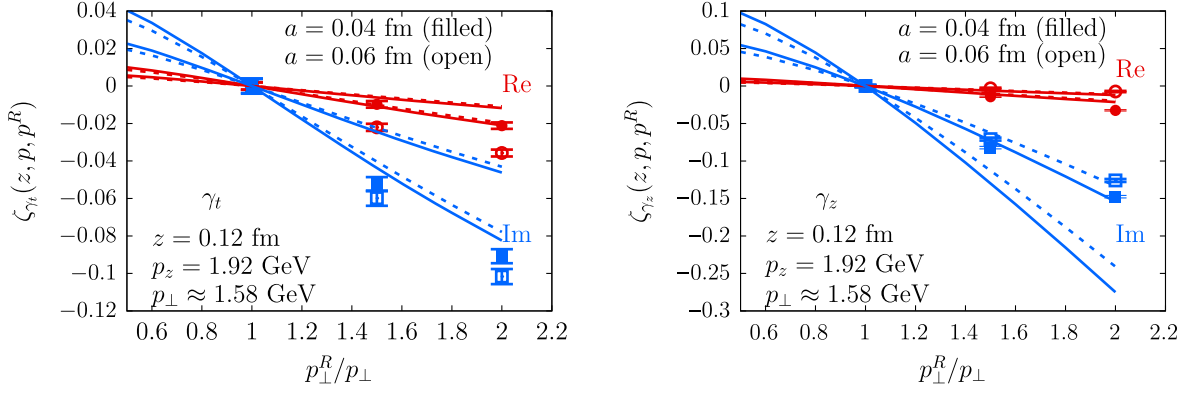
FIG. 11.   The renormalization flow of ζ_{γ_t} (left panel) and ζ_{γ_z} (right panel) with the transverse renormalization scale, $p_\perp^R$ are shown at fixed $z = 0.12$ fm. The data from two different lattice spacings, $a = 0.04$ fm (filled circles) and $a = 0.06$ fm (open circles), shown at fixed $z = 0.12$ fm and $p_z = p_z^R = 1.92$ GeV. The transverse momenta of the quark at $a = 0.04$ and 0.06 fm are $p_\perp = 1.67$ and 1.48 GeV, respectively, and they are chosen to be roughly equal for this comparison. The real part of ζ is shown in red while the imaginary part is shown in blue. The band enclosed by the solid red (blue) curves corresponds to the 1-loop result for real (imaginary) parts of ζ at $(p_z, p_\perp)$ for the 0.04 fm data, and similarly the band enclosed by the dashed curves corresponds to $(p_z, p_\perp)$ of the 0.06 fm data.

lesser lattice spacing dependence in spite of its other disadvantages.

In Fig. 11, we concentrate on the renormalization flow of ζ_Γ at the fixed small value of $z = 0.12$ fm. The two panels show the dependence of ζ_{γ_t} and ζ_{γ_z} on $p_\perp^R$ which is changed around $p_\perp$. As before, we keep $p_z = p_z^R = 1.92$ GeV. The 1-loop result is able to capture the qualitative trend of the flow in both γ_z and γ_t. For both the cases, we can see that the reduction of lattice spacing leads to a better agreement with the 1-loop result. Having discussed the cases where $p_z = p_z^R$, we now study the dependence of ζ on $p_z^R \neq p_z$, while keeping $p_\perp = p_\perp^R$. We show the z-dependence of ζ_{γ_z} when $p_z^R = 1.5 p_z$ in Fig. 12. We find the perturbative result
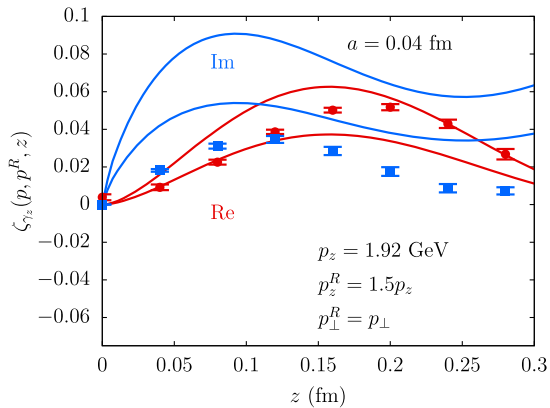

FIG. 12.   The dependence of ζ on the longitudinal momentum p_z^R. ζ_{γ_z} is shown as a function of z for a specific choice of p_z and $p_\perp = p_\perp^R$. The uncertainty bands for the real and imaginary parts for the leading-order expectation are shown using bands enclosed by solid lines. The symbols are the lattice data determined at lattice spacing $a = 0.04$ fm.

to have the same qualitative behavior as the lattice data. Putting together the various observations in this section, we found only an overall qualitative agreement between the lattice results on ζ and the one-loop perturbative results. When the lattice spacing is reduced, we found the agreement to get better. It remains to be seen what the effect is of including higher order corrections in the perturbative result for ζ.

B. A way to classify quark-antiquark separations as perturbative or nonperturbative

Physically, one would expect that for a well-separated quark-antiquark with $z > 1$ fm, one would start seeing traces of nonperturbative physics in the qPDF. Quantifying the advent of nonperturbative physics for large enough z at finite quark/hadron momentum is important with regard to the extraction of PDF since the real-space qPDF at all z enter the computation of its Fourier transform. A simple first approximation to study this effect is the following. In free theory, the qPDF with external quark states is a pure wave $e^{ip_z z}$. We expect, to a first approximation, that the effect of nonperturbative physics is to damp this pure wave via an inverse screening length $m_{\rm scr} \sim \mathcal{O}(\Lambda_{\rm QCD})$. Thus, we model the bare quark qPDF as

$$q(z, p) = A e^{i\omega z} e^{-m_{\rm scr}|z|} e^{-c|z|}, \qquad (24)$$

where we have removed the UV divergent piece $e^{-c|z|}$ from the qPDF and defined the left-over exponent $m_{\rm scr}$ as a physical scale. We have also accounted for $\omega \neq p_z$ in the interacting theory since the quark can lose momentum by emitting gluons. There could be remnant nontrivial dependence of the amplitude A on z, which we assume to be
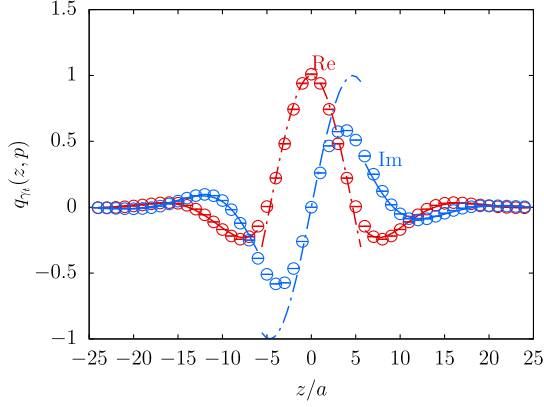
FIG. 13. The real and imaginary parts of the lattice data for $q_{\gamma_t}(z, p)$ are compared with the purely oscillatory model (dashed curves) at short-distances, and the damped oscillatory ansatz (solid curves) at larger $|z|$. The data correspond to quark momentum $(p_z, p_\perp) = (1.29, 1.49)$ GeV.

subleading compared to the leading damped oscillatory behavior and ignore it in the discussion here. There is an ambiguity in m_{scr} depending on the scheme used to determine the divergent piece c. Since the values of c determined from the static quark potential method ensure that the renormalization factors after the removal of $e^{-c|z|}$ are $\mathcal{O}(1)$ at smaller z in Fig. 8, the choice of separation of the exponential suppression factor into divergent and physical scales as defined in Eq. (24) is well motivated in this scheme. In Fig. 13, we show the bare quark qPDF $q_{\gamma_t}(z, p)$ for $p_\perp = 1.49$ GeV and $p_z = 1.29$ GeV determined on the $a = 0.06$ fm ensemble. The short distance can simply be described by a pure oscillatory $e^{i\omega z}$ behavior which is shown using the dashed curves (with $\omega = 0.85 p_z$ for the case shown). The solid curves in the figure correspond to the ansatz in Eq. (24) which describes the data at larger $|z|$ well. Without dwelling further on finding the best parametrization of the lattice data that asymptotically behaves like Eq. (24), we simply define an effective z-dependent ω and m_{scr} through

$$m_{\mathrm{scr}}(z) - i\omega(z) \equiv -\frac{1}{a}\log\left(\frac{q(z + a, p)}{q(z, p)}\right) - c. \quad (25)$$

In Fig. 14, we show the behavior of m_{scr} and ω as a function of z as extracted from $q_{\gamma_t}(z, p)$. We have chosen a different set of p_z and $p_\perp$ to show the dependence on p_z at fixed $p_\perp$ and vice versa. From the top panel, we see that ω/p_z is below 1 for $z < 0.4$ fm and seems to approach a plateau closer to 1 for $z > 0.4$ fm. While the values of ω at short distances depend on p_z and $p_\perp$, the approach to $\omega \approx p_z$ is universal. We observed this behavior when we used $q_{\gamma_z}(z, p)$ as well. A physical reasoning for this observation could be that at shorter z, the quark has the ability to radiate
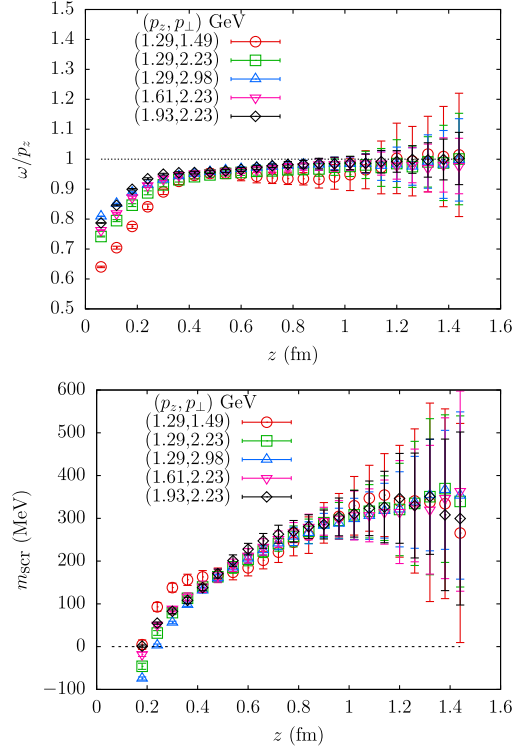

FIG. 14. The effective frequency of oscillations ω (top) and effective screening mass m_{scr} (bottom) as extracted from $q_{\gamma_t}(z, p)$ are shown. The various values of $p = (p_z, p_\perp)$ for the data are tabulated in the plots. The values of ω are normalized with respect to p_z. A 1-HYP smeared Wilson line was used in all the cases shown in the plots, and the corresponding self-energy $ca = 0.1586$ was subtracted to obtain m_{scr}.

a gluon, and at distances $z > 0.4$ fm there is effectively a dressed quark carrying all the momentum. In the bottom panel of Fig. 14, we have shown the effective screening mass m_{scr}. In the plots, we have only shown the data where 1-HYP smeared Wilson line was used. For this case, we subtracted $ca = 0.1586$ in Eq. (25) to get m_{scr}. One can clearly see the emergence of nonzero $m_{\mathrm{scr}} \approx 300$ MeV for $|z| > 0.5$ fm which is in the typical Λ_{QCD} scale. When we repeated this using quark qPDF with an unsmeared Wilson line, we found the results to be consistent with the data shown in Fig. 14 after we subtracted out $ca = 0.3687$ corresponding to the unsmeared Wilson line. This assures us that the observed $m_{\mathrm{scr}} \approx 300$ MeV is a real physical scale independent of the self-energy divergence of the Wilson line. This signals the significant presence of a confinement scale beyond $z \approx 0.5$ fm. Also, the near plateauing of both ω and m_{scr} for these larger z indicates that a simple physically motivated ansatz in Eq. (25) offers a surprisingly good description of the actual nonperturbative data. One could have expected this simply from observing the large $|z|$ part of Fig. 13. It remains to be seen if this observation can be used advantageously in improving the LaMET matching at finite moderately large p_z.

VI. FROM RENORMALIZED QUASI-PDF TO PDF

A. On obtaining the valence PDF using isovector $u - d$ qPDF of pion

Having determined the renormalized qPDF we can now discuss the matching between qPDF and PDF as well as the determination of pion PDF from our lattice results. We computed the $u - d$ qPDF matrix element of a pion which in practice we obtained from the real part of the connected piece of the u quark qPDF matrix element. Now, we discuss how the $u - d$ qPDF and PDF are related to the valence PDF of pion.

The u and d quark distributions, $f_u(x)$ and $f_d(x)$, as determined using Eq. (1) has support from $x = -1$ to 1. One can make a connection with the conventional, separately defined quark distributions $Q_{u,d}(x)$ and the antiquark distributions $Q_{\bar{u},\bar{d}}(x)$ that are nonzero only between $x = 0$ and 1, through the relation

$$f_{u,d}(x) = Q_{u,d}(x)\theta(x) - Q_{\bar{u},\bar{d}}(-x)\theta(-x). \quad (26)$$

Therefore, $f_{u,d}(x)$ contains information on both the quark as well as the antiquark distributions in the positive and negative regions of x respectively. Let us first focus on $x \geq 0$. In the isospin symmetric case we are considering, $Q_u(x) = Q_{\bar{d}}(x)$ and $Q_{\bar{u}}(x) = Q_d(x)$. Therefore, for the positively charged pion $Q_u(x) - Q_d(x) = Q_u(x) - Q_{\bar{u}}(x) = f_v^{\pi,u}(x)$ is the valence u-quark distribution. Again, due to the isospin symmetry, the u and $\bar{d}$ valence distributions are the same as $f_v^{\pi,u}(x) = f_v^{\pi,\bar{d}}(x) = f_v^\pi(x)$. However, unlike the valence quark distribution, the isotriplet $u - d$ PDF $f_{u-d} = f_u(x) - f_d(x)$ satisfies $f_{u-d}(|x|) = f_{u-d}(-|x|)$ and it has support from -1 to 1. That is,

$$f_{u-d}(x) = \begin{cases} Q_u(x) - Q_d(x), & x > 0 \\ Q_u(-x) - Q_d(-x), & x < 0; \end{cases}$$

$$f_v^\pi(x) = \begin{cases} Q_u(x) - Q_d(x), & x > 0 \\ 0, & x < 0. \end{cases} \quad (27)$$

Therefore, one can obtain the $u - d$ quark distribution, and, from it, one can obtain $f_v^\pi(x)$ from $x \in [0, 1]$, or, equivalently, from $[-1, 0]$.

By applying the matching formula on $f_u(x)$ and $f_d(x)$ separately and taking the difference to obtain the $u - d$ RI-MOM qPDF, we now try to learn what is expected for this qPDF. Writing down only the x/y dependence for the sake of brevity and keeping the dependence on yP_z/P_z^R, $(P^R/P_z^R)^2$ and factorization scale μ implicit, the one-loop contribution to the matching kernel $C(x/y)$ from the RI-MOM to $\overline{\text{MS}}$ scheme consists of two terms: $\mathcal{F}_1(x/y)$ and $\mathcal{F}_2(1 + \eta'(x - y))$ with $\eta' = P_z/P_z^R$. The expressions[3]

---

[3] In [8], the terms $\mathcal{F}_1$ and $\mathcal{F}_2$ are referred to as f_1 and f_2, respectively.

for $\mathcal{F}_{1,2}$ depend on the choice of Γ (γ_z or γ_t) [7,8]. Furthermore, $\mathcal{F}_2$ depends on the projection method of the RI-MOM scheme. Using the matching formula [7,8] on $f_u(x)$ and $f_d(x)$ to obtain the qPDFs $\tilde{q}_u(x)$ and $\tilde{q}_d(x)$,

$$\tilde{q}_{u,d}(x, P_z, p^R)$$
$$= f_{u,d}(x, \mu) + \frac{\alpha_s C_F}{2\pi} \int_{-1}^1 \frac{dy}{|y|} \mathcal{F}_1\left(\frac{x}{y}\right)_+ f_{u,d}(y)$$
$$- \frac{\alpha_s C_F}{2\pi} \int_{-1}^1 dy|\eta'|\mathcal{F}_2(1 + \eta'(x - y))_+ f_{u,d}(y) + \cdots. \quad (28)$$

The above equation includes both the sea and valence quarks, and there will be mixing with the gluon PDF which is included in the "$\cdots$" part. In the above convolution, the vector current conservation is ensured by the plus function defined as

$$\mathcal{F}_{1,2}(\xi)_+ = \mathcal{F}_{1,2}(\xi) - \delta(1 - \xi) \int_{\text{reg}} d\xi \mathcal{F}_{1,2}(\xi), \quad (29)$$

such that any extra variable that $\mathcal{F}_{1,2}$ will depend on are held fixed in the above integral. Since the matching between qPDF and PDF is linear, the $\tilde{q}_{u-d} = \tilde{q}_u(x) - \tilde{q}_d(x)$ is simply obtained as

$$\tilde{q}_{u-d}(x, P_z, P^R)$$
$$= f_{u-d}(x, \mu) + \frac{\alpha_s C_F}{2\pi} \int_{-1}^1 \frac{dy}{|y|} \mathcal{F}_1\left(\frac{x}{y}\right)_+ f_{u-d}(y)$$
$$- \frac{\alpha_s C_F}{2\pi} \int_{-1}^1 dy|\eta'|\mathcal{F}_2(1 + \eta'(x - y))_+ f_{u-d}(y), \quad (30)$$

with the terms in "$\cdots$" in Eq. (28) exactly canceled between the u and d terms. This is the matching relation we use to obtain the $u - d$ PDF from $u - d$ qPDF. Using the $u - d$ PDF, we obtained the valence PDF as discussed above. While $f_{u-d}(|x|) = f_{u-d}(-|x|)$, it is also true that $\tilde{q}_{u-d}(|x|) \neq \tilde{q}_{u-d}(-|x|)$ in the RI-MOM scheme. One way to understand this is from the fact that the bare qPDF matrix element is purely real while the RI-MOM renormalization factor is in general complex, thereby making the renormalized qPDF matrix element complex. One can see this by starting from the matching convolution above, and find that

$$\tilde{q}(|x|, P_z, P^R) - \tilde{q}(-|x|, P_z, P^R) = \frac{\alpha_s C_F |\eta'|}{2\pi}$$
$$\times \int_{-1}^1 dy[-\mathcal{F}_2(1 + \eta'(x - y)) + \mathcal{F}_2(1 - \eta'(x - y))] f(y, \mu), \quad (31)$$

is nonzero due to an RI-MOM specific term $\mathcal{F}_2$, while the terms containing $\mathcal{F}_1$ cancel due to their dependence only on $|P_z|$. In other schemes such as the $\overline{\text{MS}}$, this symmetry

about $x = 0$ would be preserved by matching because the corresponding factorization formulas depend on renormalization/regularization scales through combinations such as $\mu^2 z^2$ [55]. In the RI-MOM scheme there are two renormalization scales, P^R and P_z^R, and since the z-direction is special the above statement does not hold. Thus, it is important to capture this asymmetry in the qPDF, or equivalently to describe both the real and imaginary parts of the RI-MOM renormalized pion qPDF from matching. We use the matching kernel corresponding to the $\not{p}$-projection in the results to be discussed next.

B. Numerical results on pion valence PDF from matching

The one-loop perturbative matching relates the Fourier transform $\tilde{q}(x, P_z, P^R)$ of the renormalized RI-MOM real-space qPDF matrix element $h^R(z, P_z, P^R)$, and the $\overline{\text{MS}}$ PDF $f(x, \mu)$ at factorization scale μ. The relation is through the convolution in Eq. (3). There are two approaches to consider here:

(1) One can parametrize the real space data $h^R(z, P_z, P^R)$ over the range z where one has the lattice data and then model the dependence of $h^R(z, P_z, P^R)$ over z extending to infinity where data does not exist (c.f., [56,57]). Using such a parametrization, one can obtain its Fourier transform $\tilde{q}(x, P_z, P^R)$. Since the matching is only up to $\mathcal{O}(\alpha_s)$, one can invert the relation Eq. (3) by replacing $f \leftrightarrow \tilde{q}$ and $\alpha_s \rightarrow -\alpha_s$. Thereby, one can obtain $f(x, \mu)$. In this method, one does not control what values of z enter the Fourier transform and one could question the validity of perturbation theory for $z > 1$ fm.

(2) One can start from a phenomenologically motivated n-parameter family of PDFs $f(x, \mu; a_1, \ldots a_n)$. Through Eq. (3), one can obtain qPDF $\tilde{q}(x, \mu; a_1, \ldots a_n)$, and, thereby, obtain a family of real space qPDF matrix elements $h^R(z, P_z, P^R; a_1 \ldots a_n)$. Using this, one can fit the parameters $(a_1, \ldots, a_n)$ so as to best describe the real space lattice data over a range z. This method was used in the case of the lattice cross-section approach in [13]. Since the model PDFs are not predictions from QCD, the model dependence enters the analysis and one has to rely on the prior that experimentally determined PDFs are indeed very well described by such a family of PDFs. However, the advantage of this method is that one can precisely control the range of z that enters the analysis, and one also does not have to invert the matching convolution.

From our observation on how the 1-loop perturbation theory fails to capture the quark qPDF quantitatively, even at short distances, and from the observation of significant nonperturbative screening effects beyond $z = 1$ fm, we think it is important to be in control of what values of z enter the convolution and, hence, in this paper we take the

second approach. Also, due to the loss of signal to noise ratio for $z > 1$ fm, we found Fourier transforming the noisy data to be challenging without introducing unwanted wiggles in $\tilde{q}(x)$ at larger x.

To be on par with the experimental extraction of PDFs, one should use sophisticated methods such as the usage of neural networks to choose the set of model PDFs to start with (c.f., [58]). We defer such an analysis to a future work and, instead, we use a simple two-parameter phenomenologically motivated functional form for the valence PDF:

$$f_v^\pi(x; a, b) = A x^a (1 - x)^b, \tag{32}$$

for $x \in [0, 1]$ and zero elsewhere. As we will see below, such a form is enough to describe our lattice data. One can fix the coefficient A through a stringent condition $\int_0^1 f_v^\pi(x) dx = 1$. Instead, we use a more conservative constraint on A using $\int_0^1 f_v^\pi(x) dx = h^R(z = 0, P_z, P^R)$ to allow for sample by sample fluctuations in $h^R(z = 0, P_z, P^R)$ close to 1 and fold this into the error estimate. It should be noted that the valence PDF of the pion determined from the experimental data by the JAM collaboration [59] can be well described by such a two parameter ansatz, for example with $a = -0.407$ and $b = 1.12$ at $\mu = 3.2$ GeV.

Using the above valence PDF, we construct the $u - d$ PDF as

$$f_{u-d}(x; a, b) = \frac{1}{2} (f_v^\pi(x; a, b) + f_v^\pi(-x; a, b)), \tag{33}$$

with $x \in [-1, 1]$ and zero elsewhere. Through the convolution of $f_{u-d}(x)$ with the matching kernel, we obtain $\tilde{q}_{u-d}(x; a, b)$, which in turn we use to construct the real space qPDFs $h(z; a, b) = \int_{-\infty}^\infty \tilde{q}(x; a, b) e^{ixP_z z} dx$. We will refer to these functions $h(z; a, b)$ as the two-parameter family of phenomenologically motivated qPDF matrix elements. With the set of $h(z; a, b)$ from a range of a and b, we can fit the parameters a and b to the data by minimizing either χ_r^2 or χ_{ri}^2 below:

$$\chi_r^2 = \sum_{z=-z_{\max}}^{z_{\max}} \frac{(\text{Re}(h^R(z)) - \text{Re}(h(z; a, b)))^2}{\sigma_r(z)^2 + \sigma_r^{\text{pert}}(z)^2},$$

$$\chi_i^2 = \sum_{z=-z_{\max}}^{z_{\max}} \frac{(\text{Im}(h^R(z)) - \text{Im}(h(z; a, b)))^2}{\sigma_i(z)^2 + \sigma_i^{\text{pert}}(z)^2},$$

$$\chi_{ri}^2 = \chi_r^2 + \chi_i^2. \tag{34}$$

In the above equations, $[-z_{\max}, z_{\max}]$ specifies the fit range. The statistical errors on the real and imaginary parts of the lattice data $h^R(z)$ is $\sigma_r(z)$ and $\sigma_i(z)$, respectively. To account for any systematic errors coming from the higher order corrections in α_s in the matching kernel, we determine $h(z; a, b)$ from $f_v^\pi(x; a, b)$ by varying the value of α_s

in the matching kernel from $\alpha_s(\mu/2)$ to $\alpha_s(2\mu)$ though the 1-loop running. The corresponding changes in the real and imaginary parts of $h(z; a, b)$ are denoted as $\sigma_r^{\text{pert}}(z)$ and $\sigma_i^{\text{pert}}(z)$, respectively, and we include these uncertainties in the matched result in the χ^2. If the matching was exact, then by fitting only the real part by minimizing χ_r^2 would automatically guarantee that the imaginary part also agrees with the data. Therefore at any finite order matching, the fits obtained by minimizing χ_r^2 and χ_{ri}^2 will in general be different. For the results shown below, we used χ_{ri}^2 in order to obtain the PDF that best describes both the real and imaginary parts of the real space qPDF, but we also used χ_r^2 and found it to lead to consistent results but with larger uncertainties. We did not include the correlations between the data at different z for the primary reason that it is difficult to keep these correlations intact in the process of excited state extrapolations. It also helps us to easily incorporate the effect of σ^{pert} from nonstatistical origin in the analysis, and in treating the real and imaginary parts of the renormalized matrix elements as two distinct pieces of data as is the case in the context of matching. We determined the errors on the fit parameters through the bootstrap analysis.

In Fig. 15, we show the fitting procedure for γ_t qPDF. In the top panels, we show the $P_z = 0.86$, 1.29, and 1.72 GeV real-space RI-MOM pion qPDF matrix elements from left to right. The symbols are the actual lattice data. The solid and patterned red (blue) bands are $1 - \sigma$ error-bands of the real (imaginary) parts of the fitted real space qPDF matrix element that best fits the data over the range $[-z_{\max}, z_{\max}]$

for $z_{\max} = 1.44$ fm and 0.72 fm, respectively. The agreement with both the real and imaginary parts of the lattice data is noteworthy. In fact, we find the qPDF matrix element as inferred from the JAM PDF [59] is able to explain the lattice data well for the entire range of z at the two largest momenta. In the bottom panels, we show the process leading from model PDF to the real space qPDF matrix elements shown in the top panels. In order to avoid cluttering the figure, we have shown only the mean value of $f_{u-d}(x)$ (shown as dashed lines) while we have shown the error bands for the qPDF $\tilde{q}_{u-d}(x)$ as obtained through the 1-loop matching. The colors red and blue in the bottom panels correspond to the fits with $z_{\max} = 0.72$ and 1.44 in the top panels, respectively. As one can see, we started from a symmetric $u - d$ PDF by construction and matching introduces an $x \to -x$ asymmetry. After Fourier transformation, this asymmetry leads to the imaginary part in the real space data in the top panels which captures the lattice data to a good accuracy. For both the real-space as well as in x space, we find no significant difference between using $z_{\max} = 1.44$ fm and 0.72 fm in the fits. We could infer that within the precision of our numerical results, the nonperturbative effects at $z \approx 1$ fm that we found using quark qPDFs is not important. Therefore, we show results for an intermediate $z_{\max} = 0.98$ fm in the results below. When we repeated this analysis by minimizing χ_r^2, we found the estimates to be consistent with the above, but with larger uncertainties.

In Fig. 16, we show our results for $f_v^\pi(x, \mu)$ and $x f_v^\pi(x, \mu)$ at the factorization scale $\mu = 3.2$ GeV using the procedure described above at our two largest pion momenta
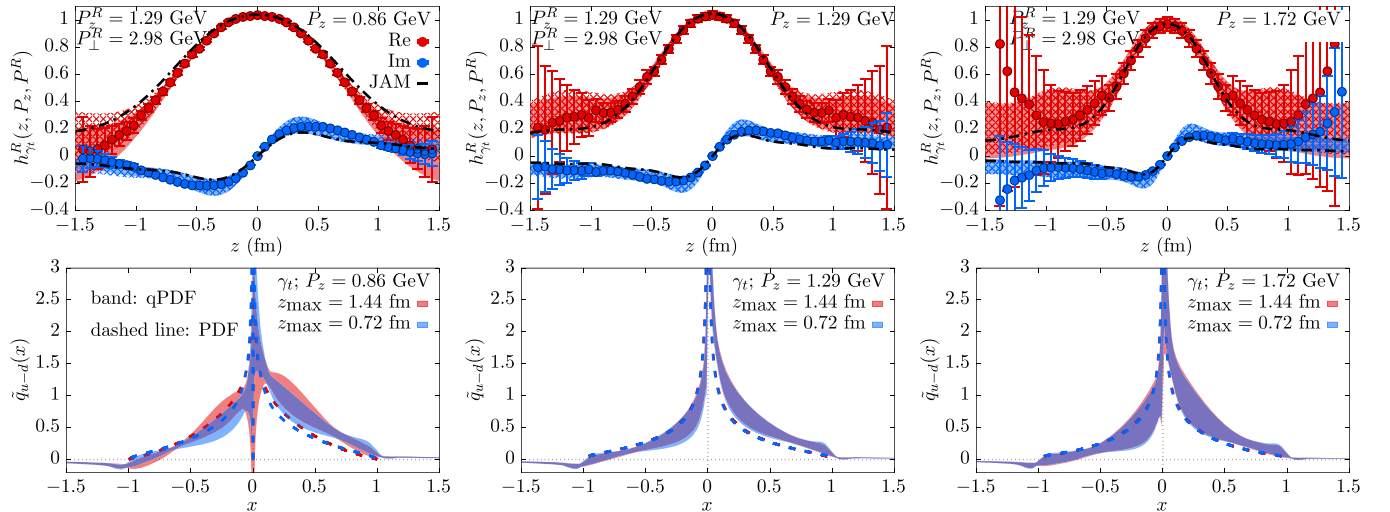

FIG. 15.   Top panels: The RI-MOM renormalized qPDF matrix element in real space $h_{\gamma_t}^R(z, P_z, P^R)$ at pion momenta $P_z = 0.82$, 1.29, and 1.72 GeV are shown at fixed RI-MOM renormalization scale $(P_z^R, P_\perp^R) = (1.29, 2.98)$ GeV. The red and blue points are the real and imaginary parts of the actual data, respectively. The bands were obtained by fitting the two-parameter phenomenologically motivated real space qPDF matrix element to the data over a range $z \in [-z_{\max}, z_{\max}]$—the solid band is for $z_{\max} = 1.44$ fm and patterned one for $z_{\max} = 0.72$ fm. Bottom panels: The two parameter $u - d$ PDF $f_{u-d}(x)$ (dashed lines) at $\mu = 3.2$ GeV, and the matched qPDFs $\tilde{q}_{u-d}(x)$ ($1 - \sigma$ error bands) that describe the real space qPDF on the top panels are shown. To avoid clutter, only the central values of $f_{u-d}(x)$ are shown as dashed lines. The results from different $z_{\max}$ are shown in red and blue.
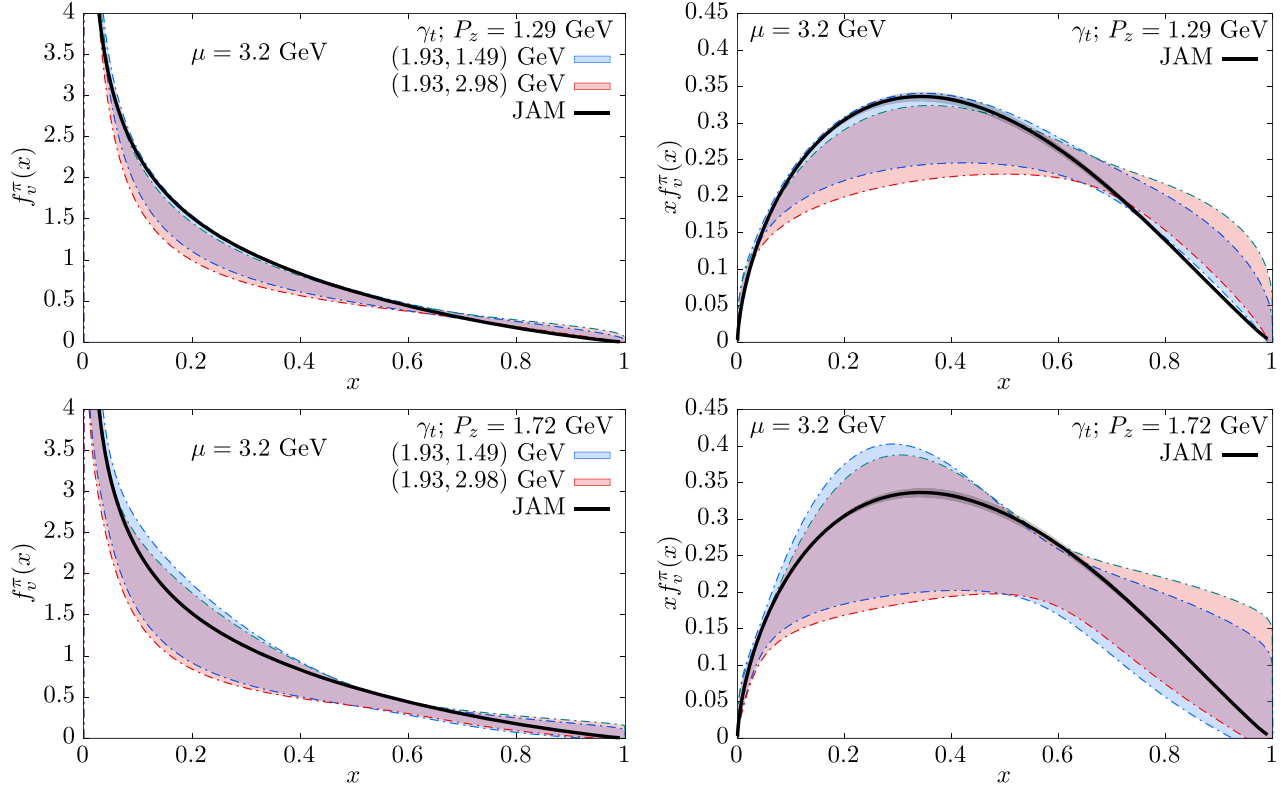
FIG. 16. The top and bottom panels show our estimated pion valence PDF at $\mu = 3.2$ GeV using γ_t qPDF at $P_z = 1.29$ and 1.72 GeV, respectively. The results using multiple RI-MOM scales $(P_z^R, P_\perp^R)$ are shown using different colored error bands. On the left panels, the results for $f_v^\pi(x)$ are shown, while on the right panels the results for $xf_v^\pi(x)$ are shown. For all the cases shown, the fit range was held fixed at $z_{\max} = 0.98$ fm. The solid line (with a small error band around it) is the JAM result [59] for pion valence PDF at the same μ.

$P_z = 1.29$ GeV and 1.72 GeV starting from γ_t qPDF. For each case, we overlay the results from two different RI-MOM scales P^R in order to show the scatter as a systematic error in our estimates. We find the P^R dependence to be minor compared to the error bands (we repeated the analysis with multiple other values of P^R that are not shown and only minor scatter with respect to P^R was seen). We also show the result from the JAM collaboration [59] for the pion valence PDF at the same factorization scale as the black solid line, which lies within the statistical and systematic uncertainties of our estimates. In the left panels showing $f_v^\pi(x)$, this overall agreement can be seen even up to smaller x, but one has to be cautious of this agreement for $x \lesssim \Lambda_{\text{QCD}}/P_z \approx 0.2$ for the two highest pion momenta we use. By construction, in our fitting procedure $f_v^\pi(x)$ has support only from 0 to 1 without any necessity to recover this condition in the infinite P_z limit. However, the values of exponent b closer to zero are also allowed thereby leading to a wider error band closer to $x = 1$. This seems to be consistent with the observation in Ref. [18] that the PDF obtained from qPDF through the inverse one-loop matching (approach-1) vanishes at about $x \approx 1.2$. We see our $P_z = 1.29$ GeV and 1.72 GeV estimates to be consistent, albeit with a significant increase in error at the largest momentum.

In Fig. 17, we summarize the information in Fig. 16 by showing the $1 - \sigma$ ellipses (whose x and y projections give the marginal 68% confidence intervals of the exponents a and b, respectively). In this figure, the dashed and continuous ellipses are for $P_z = 1.29$ and 1.72 GeV, respectively. The ellipses for different P^R are distinguished by the colors, with the color code being the same as in Fig. 16. The $P_z = 1.29$ GeV data offer a stronger constraint on the allowed region of (a, b) than the noisier $P_z = 1.72$ GeV. In this plot, the JAM estimate is the black point. The JAM data are well within the $P_z = 1.72$ GeV ellipses while the $P_z = 1.29$ GeV data seem to favor the slightly smaller exponent b. However, these differences are well within 2σ. Even though our lattice data have large errors on the exponents a and b individually, the data offers a tight constraint on the combined allowed region. In particular, the principal component of this correlation between a and b points directly at the JAM data implying that if one fixes the exponent a to be from the experiment, then the best value of b would also be closer to that from the experiment. To understand this better, we have also shown the line of constant value of first moment of the valence PDF, $\langle x \rangle = \int_0^1 xq_v^\pi(x)dx$, set to 0.215 as inferred from the JAM data. It is clear that the 1σ ellipses are oriented along this line, which means that qPDF determines $\langle x \rangle$ robustly
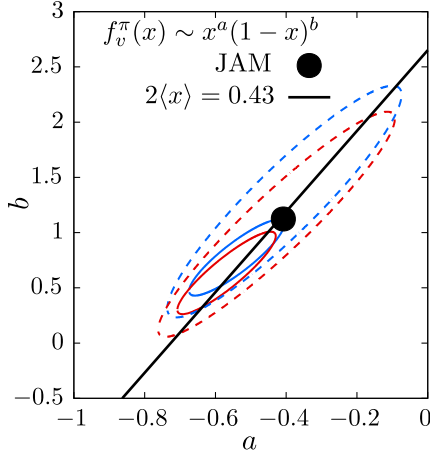
FIG. 17. The $1 - \sigma$ confidence region ellipse of the exponents a and b in the model PDF at $\mu = 3.2$ GeV that best describes the real space RI-MOM qPDF is shown. The solid lines and dashed lines correspond to $P_z = 1.29$ GeV and $P_z = 1.72$ GeV. For each of these pion momenta, the different colored lines correspond to different RI-MOM scale P^R. The black point is the JAM value [59] for valence pion PDF. The black straight line is the line of constant first moment of valence PDF, $\langle x \rangle = 0.215$.

and this in turn provides a strong constraint in the allowed PDFs. Not surprisingly, we do find consistent values of $\langle x \rangle = 0.21(2)$ and $0.22(3)$ from the $P_z = 1.29$ and 1.72 GeV estimates. It should be noted that the moments of pion PDF have also been directly determined without the usage of LaMET formalism [60–64] and similar values for the first moment for the pion were obtained, but at slightly different values of μ^2 than used here.

The exponents a and b were also recently obtained using the lattice cross-section approach [13] which used current-current correlators, with the matching implemented at tree-level. Here, the exponents were estimated as $a = -0.34(31)$ and $b = 1.93(68)$ which are consistent with the region allowed at the largest momentum in Fig. 17. It is worth noting that there are indications from next-to-leading-logarithmic soft gluon resummation calculation [35], the Dyson-Schwinger equation [65–67] and light-front holographic QCD [68] that the value of exponent b could be approximately 2 as expected from perturbative counting rule (c.f., [69]), whereas a chiral quark model analysis [70,71] suggests a value of b closer to 1. It will be interesting to see if a similar implementation of an improved matching kernel could lead to a softer large x behavior for the pion than what is observed using the 1-loop qPDF matching here and perhaps in [18]. In fact, a general consideration of power corrections to qPDF [72] revealed the presence of the form $\Lambda^2_{QCD}/((1 - x)x^2 P_z^2)$ implying higher values of P_z might be required in order to correctly describe physics close to $x = 1$, and this might be the effect which we are finding. Similar conclusions have also been obtained in 2d QCD [73].
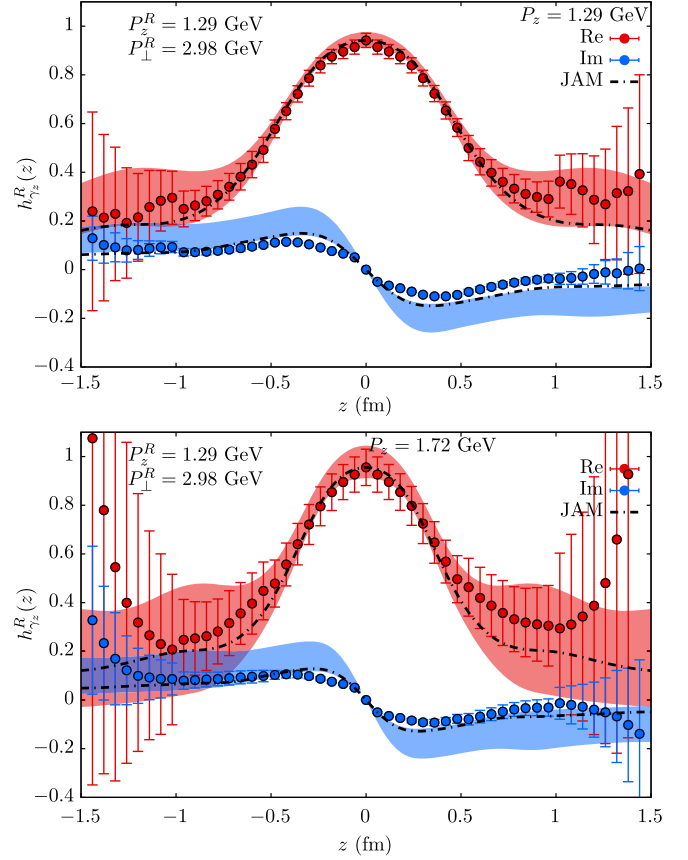

FIG. 18. In the top and bottom panels, the real (red) and imaginary (blue) parts of the renormalized real space γ_z qPDF matrix element are shown for pion momenta $P_z = 1.29$ GeV and 1.72 GeV, respectively. The data points are the actual lattice data. The bands are the expected matched γ_z qPDF matrix element starting from our best estimate for valence pion PDF obtained using γ_t qPDF analysis.

Due to the larger errors in the γ_z qPDF attributed mostly to the steep excited state extrapolations, we use the γ_z qPDF to provide a consistency check of our calculations instead. For this, we use our above best estimates of the PDF obtained using the γ_t qPDF to get the corresponding prediction for the real space γ_z qPDF matrix element through a convolution with the appropriate matching kernel. In the top and bottom panels of Fig. 18, we show such a comparison between the actual real space data of γ_z qPDF (data points) along with the prediction from our estimated PDF (bands) for pion momenta $P_z = 1.29$ and 1.72 GeV. We find good descriptions of the real part of the RI-MOM γ_z qPDF at both the pion momenta with a slight tension between the imaginary parts. From our discussion on the excited state contamination, it is important to first gain better control of the larger excited state contamination in the γ_z qPDF before one can investigate the effect of one-loop matching on this rather small discrepancy.

VII. CONCLUSIONS

We studied pion PDF in the framework of LaMET, which relates the qPDF to PDF through the matching convolution in Eq. (3). For this, we used a small lattice spacing $a = 0.06$ fm. We carefully examined the effects of excited states using the two- and three-state exponential fits of the relevant 2-point and 3-point functions as well as the summation method. For our final analysis we used two momenta $P_z = 1.29$ GeV and $P_z = 1.72$ GeV. We found the qPDF defined using $\Gamma = \gamma_t$ was better determined compared to the γ_z qPDF in the lattice calculation for the following reasons: smaller statistical error, relatively smaller excited state extrapolation leading to a more robust result for the matrix element as well as due to the absence of mixing. Therefore, we focused on the analysis of the γ_t matrix element.

The pion qPDF was nonperturbatively renormalized using the RI-MOM scheme by calculating the matrix elements of qPDF operator with off-shell quark states in the Landau gauge for different separations z. For these calculations we also used finer lattices with the lattice spacing $a = 0.04$ fm. We performed the comparison of this matrix element in the Landau gauge with 1-loop perturbative calculations in the RI-MOM scheme and found a qualitative agreement for $z < 0.3$ fm. For the smaller lattice spacings, $a = 0.04$ fm we even found quantitative agreement with the 1-loop result for sufficiently small z. We also explored the role of nonperturbative effects in the calculation of the off-shell matrix element. The real part of the RI-MOM renormalization coefficient is close to one, while the imaginary part is close to zero once the divergent self energy part of the Wilson line is removed. We pointed out that the RI-MOM renormalization procedure leads to an asymmetry in the isovector pion qPDF $\tilde{q}(x, P_z, p_z^R, \mu_R)$ around $x = 0$, while other renormalization procedures lead to qPDF that is symmetric around $x = 0$.

From the renormalized qPDF, we determined the valence quark pion PDF using the 1-loop perturbative matching of the γ_t qPDF, which we implemented through a fit to the phenomenologically motivated $x^a(1 - x)^b$ functional form for the valence PDF. We found our results for the pion valence PDF using the two largest pion momenta were consistent with each other, though the statistical errors are rather large. An overall agreement with the results obtained recently by the JAM collaboration [59] was seen. We found our result for the PDF to capture the first moment $\langle x \rangle$ more robustly than the small-x and large-x exponents, a and b themselves. We used the γ_z qPDF matrix elements to provide an internal consistency check by comparing to the expectation from our estimates of the PDF and a satisfactory agreement was seen. From our analysis it is clear that the dominant source of errors in the PDF determination is the statistical error of the lattice calculations. It will be necessary to significantly increase the statistics in the future lattice calculations. Future high statistics lattice calculations will be important for an accurate determination of the pion PDF as well as testing of the LaMET approach around small x.

APPENDIX A: COULOMB-GAUGE GAUSSIAN AND MOMENTUM (BOOSTED) SMEARING

In order to create hadron interpolating operators that have a good overlap with the corresponding ground states, quark field smearing is typically required. The amount of applied smearing is tuned to produce spatial quark distributions of roughly the same spatial size as the hadron. Gauge-covariant Wuppertal (Gaussian) smearing [74] is commonly used for this purpose. However, calculation of quasi- and pseudo-PDFs and high-momentum hadron structure and spectrum in general requires lattices with small lattice spacing. Keeping the physical size of smeared quark distributions the same becomes a numerical challenge on finer lattices because it requires larger numbers of smearing iterations. For this reason, we use Gaussian shape smearing in a fixed (Coulomb) gauge that can be performed efficiently through a convolution with a Gaussian profile kernel,

$$\mathcal{S}_{x,y}^{\mathrm{CG}} \propto e^{\frac{(\vec{x}-\vec{y})^2}{2w_{\mathrm{CG}}^2}} \propto \sum_{\vec{p}} e^{i\vec{p}(\vec{x}-\vec{y})} e^{-\frac{1}{2}w_{\mathrm{CG}}^2 \vec{p}^2}. \qquad \text{(A1)}$$

In the free-field case, this kernel corresponds to the Wuppertal smearing operator $(1 + \frac{w^2}{4N}\Delta_{\rm sp})^N$, where $\Delta_{\rm sp}$ is the spatial Laplacian and $w^2 = 2w_{\rm CG}^2$. The value for the width $w_{\rm CG}$ is chosen to match the mean-squared radius $\langle r^2 \rangle = 3w_{\rm CG}^2$ to that of the optimal Wuppertal-smeared quark sources. First, we fix the Coulomb gauge

$$\psi_x^{\rm C} = \Omega_x^{\rm C}\psi_x, \qquad U_{x,\mu}^{\rm C} = U_{x,\mu}^{\Omega^{\rm C}} = \Omega_x^{\rm C}U_{x,\mu}\Omega_{x+\hat{\mu}}^{\rm C\dagger}, \quad \text{(A2)}$$

where $\Omega_x^{\rm C}$ is the gauge transformation to the Coulomb gauge, which minimizes the functional

$$\Omega^{\rm C} = \Omega\colon \min_{\Omega_x}F^{\rm C}[U^{\Omega}] = \min_{\Omega_x}\sum_x\sum_{\mu\neq t}[-\Re{\rm Tr}U_{x,\mu}^{\Omega}] \quad \text{(A3)}$$

(for the Coulomb gauge, $\mu \neq t$ and the functional is minimized independently on each time slice). The numerical implementation is identical to the algorithm used for fixing the Landau gauge in NPR calculations. Application of the smearing kernel requires two 3D Fourier transformations:

$$[\mathcal{S}^{\rm CG}\psi]_x = \Omega_x^{\rm C\dagger}\sum_{\vec{p}}e^{i\vec{p}\cdot\vec{x}}e^{-\frac{1}{2}w_{\rm CG}^2\vec{p}^2}\left[\frac{1}{V}\sum_{\vec{y}}e^{-i\vec{p}\cdot\vec{y}}\Omega_y^{\rm C}\psi_y\right], \quad \text{(A4)}$$

which is accelerated with off-loading matrix-matrix products to GPU.

Incorporating momentum (boosted) into Coulomb-gauge Gaussian smearing amounts to translation of the kernel in the momentum space,

$$(\mathcal{S}^{(\vec{k})}\psi)_x = e^{i\vec{k}\cdot\vec{x}}[\mathcal{S}^{(\vec{0})}]_{x,y}e^{-i\vec{k}\cdot\vec{y}}\psi_y = [e^{i\vec{k}\cdot\vec{x}}\mathcal{S}e^{-i\vec{k}\cdot\vec{y}}]_{x,y}\psi_y. \quad \text{(A5)}$$

In a periodic finite volume, care must be taken to avoid spatial discontinuities in the boosted smearing kernel [Eq. (A5)]. Such discontinuities may arise because the optimal boosted smearing momentum $\vec{k}$ typically does not conform to finite-volume momentum quantization $\vec{k} = 2\pi\frac{\vec{n}}{L}$ and the phase factors $e^{i\vec{k}\cdot\vec{x}}$, $e^{i\vec{k}\cdot\vec{y}}$ do not satisfy periodic boundary conditions. The solution is to define the smearing kernel in the momentum space as $\mathcal{S}_{\vec{x},\vec{y}}^{(\vec{k})} = \sum_{\vec{p}}e^{i\vec{p}\cdot(\vec{x}-\vec{y})}e^{-\frac{1}{2}w_{\rm CG}^2(\vec{p}-\vec{k})^2}$, where the momentum difference $(\vec{p}-\vec{k})$ is understood as the shortest distance between $\vec{p}$ and $\vec{k}$ in the Brillouin zone. Such choice leads to a smooth distribution in the momentum space and, respectively, smooth and continuous smearing kernel in the coordinate space.

Finally, it is important to note that the smearing kernel in Eq. (A5) is Hermitian (as an operator acting in the [coordinate $\otimes$ color] space),

$$\mathcal{S}_{x,y}^{(\vec{k})\dagger} = [\mathcal{S}_{y,x}^{(\vec{k})}]^{\dagger} = \mathcal{S}_{x,y}^{(\vec{k})}, \quad \text{(A6)}$$

which is similar to the (boosted) Wuppertal smearing operator and important for computing symmetric hadron correlation functions.

APPENDIX B: MESON CORRELATION FUNCTIONS WITH BOOSTING

We use the interpolating operator for the $\pi^+ = \bar{d}u$ meson

$$[\pi^{+,(2\vec{k})}]_x = \bar{\tilde{d}}_x\Gamma_M\tilde{u}_x = \bar{d}_{x''}\mathcal{S}_{x'',x}^{(-\vec{k})}\Gamma_M\mathcal{S}_{x,x'}^{(\vec{k})}u_{x'}, \quad \text{(B1)}$$

which is constructed from smeared quark fields

$$\bar{\tilde{d}}_x = \bar{d}_{x''}\mathcal{S}_{x'',x}^{(-\vec{k})}, \qquad \tilde{u}_x = \mathcal{S}_{x,x'}^{(\vec{k})}u_{x'}, \quad \text{(B2)}$$

where the spinor matrix $\Gamma_M = \gamma_5$. The Hermitian-conjugated (creation) meson operator is

$$[\pi^{+,(2\vec{k})}]_x^{\dagger} = \tilde{u}_x^{\dagger}\Gamma_M^{\dagger}\gamma_4\tilde{d}_x = \bar{u}_{x'}\mathcal{S}_{x',x}^{(\vec{k})}\bar{\Gamma}_M\mathcal{S}_{x,x''}^{(-\vec{k})}d_{x''}, \quad \text{(B3)}$$

where $\bar{\Gamma}_M = \gamma_4\Gamma_M\gamma_4 = (-\gamma_5)$. The meson two-point correlation function with boost-smeared source and sink and momentum projection at the sink is[4]

$$\begin{aligned}
&C_{\rm 2pt}(y_4, \vec{p}; x) \\
&= \sum_{\vec{y}}e^{-i\vec{p}(\vec{y}-\vec{x})}\langle[\pi^{+,(2\vec{k})}]_y[\pi^{+,(2\vec{k})}]_x^{\dagger}\rangle \\
&= \sum_{\vec{y}}e^{-i\vec{p}(\vec{y}-\vec{x})}(-{\rm Tr}[\mathcal{S}_{x,x''}^{(-\vec{k})}Q_{x'',y''}^d\mathcal{S}_{y'',y}^{(-\vec{k})}\Gamma_M\mathcal{S}_{y,y'}^{(\vec{k})}Q_{y',x'}^u\mathcal{S}_{x',x}^{(\vec{k})}\bar{\Gamma}_M]) \\
&= \sum_{\vec{y}}e^{-i\vec{p}(\vec{y}-\vec{x})}{\rm Tr}[\tilde{Q}_{x,y}^{d,(-\vec{k})}\Gamma_M\tilde{Q}_{y,x}^{u,(\vec{k})}(-\bar{\Gamma}_M)], \quad \text{(B4)}
\end{aligned}$$

where $Q_{x,y}^q = \langle q_x\bar{q}_y\rangle$ and $\tilde{Q}^{q,(\pm\vec{k})} = \mathcal{S}^{(\pm\vec{k})}Q^q\mathcal{S}^{(\pm\vec{k})}$ are unsmeared and smeared quark propagators, respectively. Note that the meson two-point function is constructed from the u-quark propagator $y \leftarrow x$ and the d-quark propagator $x \leftarrow y$ smeared with momenta $(\vec{k})$ and $(-\vec{k})$, respectively. Therefore, separate propagators for u and d quarks are required to construct meson correlation functions "boosted" with the total momentum $(2\vec{k})$,

$$\begin{aligned}
\tilde{Q}_{y,x}^{u,(\vec{k})} &= \mathcal{S}_{y,y'}^{(\vec{k})}Q_{y',x'}^u\mathcal{S}_{x',x}^{(\vec{k})} \propto e^{i\vec{k}(\vec{y}-\vec{x})}, \\
\tilde{Q}_{x,y}^{d,(-\vec{k})} &= \mathcal{S}_{x,x'}^{(-\vec{k})}Q_{x',y'}^d\mathcal{S}_{y',y}^{(-\vec{k})} \propto e^{i\vec{k}(\vec{y}-\vec{x})}, \quad \text{(B5)}
\end{aligned}$$

where "$\propto$" sign stands for additional coordinate dependence due to the boosting. The d-quark $x \leftarrow y$ propagator, as usual, is computed using γ_5-Hermiticity of the Dirac operator,

---

[4]Unless explicit summation is performed, implicit summation over repeated coordinate indices x', x'', y', y'' is assumed, as well as all over omitted spin and color indices.

$$\tilde{Q}_{x,y}^{d,(-\vec{k})} = \gamma_5 [\tilde{Q}_{y,x}^{d,(-\vec{k})}]^\dagger \gamma_5 = \gamma_5 [\mathcal{S}_{y,y'}^{(-\vec{k})} Q_{y',x'}^d \mathcal{S}_{x',x}^{(-\vec{k})}]^\dagger \gamma_5, \quad \text{(B6)}$$

where the sign of the boosting momentum is preserved due to the Hermiticity of the (boosted) smearing operator $\mathcal{S}$ in Eq. (A6).

Repeating similar steps for the meson three-point function with the insertion of the operator $[\bar{u}W\Gamma u]_z$ with arbitrary Γ-matrix and Wilson line $W_{z,z+\hat{\mathcal{L}}} = (\prod_{\mathcal{L}} U)_{z,z+\hat{\mathcal{L}}}$ along path $\mathcal{L}$, we get

$$C_{3\mathrm{pt}}^{W\Gamma}(y_4, \vec{p}'; z_4, \vec{q}; x)$$

$$= \sum_{\vec{y},\vec{z}} e^{-i\vec{p}'(\vec{y}-\vec{x})+i\vec{q}\vec{z}} \langle [\pi^{+,(2\vec{k})}]_y [\bar{u}_z W_{z,z+\hat{\mathcal{L}}} \Gamma u_{z+\hat{\mathcal{L}}}] [\pi^{+,(2\vec{k})}]_x^\dagger \rangle$$

$$= \sum_{\vec{z}} e^{i\vec{q}\vec{z}} \mathrm{Tr}[B_{x,z}^{\bar{d}\Gamma_M u(y_4,\vec{p}')} W_{z,z+\hat{\mathcal{L}}} \Gamma F_{z+\hat{\mathcal{L}},x}^u], \quad \text{(B7)}$$

where the forward propagator $F^u = Q^u S^{(\vec{k})}$ and the meson sink-sequential (backward) propagator $B^{\bar{d}\Gamma_M u(y_4,\vec{p}')}$ is defined as

$$B_{x,z}^{\bar{d}\Gamma_M u(y_4,\vec{p}')} = \sum_{\vec{y}} e^{-i\vec{p}'(\vec{y}-\vec{x})}(-\bar{\Gamma}_M)\tilde{Q}_{x,y}^{d,(-\vec{k})}\Gamma_M \mathcal{S}_{y,y'}^{(-\vec{k})} Q_{y',z}^d,$$

$$\text{(B8)}$$

which is also computed using the γ_5-conjugation.

APPENDIX C: EXPLICIT CALCULATION TO SHOW THAT BARE PION $u-d$ THREE-POINT FUNCTION IS PURELY REAL OR IMAGINARY

In the previous appendix, we constructed the connected piece of the three-point function of $\bar{u}W\Gamma u$ operator in π^+. If one repeats the computation using the $\bar{d}W\Gamma d$ operator, one will find the disconnected piece to be the same as the one in the full $\bar{u}W\Gamma u$ three-point function and hence such quark line disconnected terms will cancel in the $\bar{u}W\Gamma u - \bar{d}W\Gamma d$ isospin nonsinglet operator that we are interested in. Below, we further explain why only the real part of the connected $\bar{u}W\Gamma u$ three-point function for $\Gamma = \gamma_t, \gamma_z$ and the imaginary part for $\Gamma = 1$ contributes to the total isospin nonsinglet three-point function. For the sake of simplicity let us take the case of point-source and point-sink, and take $\Gamma = \gamma_t$. The full expression for the $u-d$ qPDF three-point function is

$$C_{3\mathrm{pt}}^{u-d}(t, \tau, \mathcal{L})$$

$$= \sum_{\vec{y},\vec{z}} e^{-i\vec{p}\cdot\vec{y}} \mathrm{Tr}[(-\gamma_5)Q_{x,y}\gamma_5 Q_{y,z}W_{z,z+\mathcal{L}}\gamma_t Q_{z+\mathcal{L},x}]$$

$$+ \sum_{\vec{y},\vec{z}} e^{-i\vec{p}\cdot\vec{y}} \mathrm{Tr}[(-\gamma_5)Q_{x,y}\gamma_5 Q_{y,z+\mathcal{L}}W_{z,z+\mathcal{L}}^\dagger \gamma_t Q_{z,x}]^*,$$

$$\equiv \sum_{\vec{y},\vec{z}} (e^{-i\vec{p}\cdot\vec{y}} T_1 + e^{-i\vec{p}\cdot\vec{y}} T_2^*), \quad \text{(C1)}$$

where we do not make distinctions between the u and d quark propagators due to isospin symmetry. Let us call the trace in first term on the right-hand side as T_1 and the second trace before being conjugated as T_2. One can go from T_2 to T_1 by parity transformation $x = (\vec{x}, x_4) \to x_p = (-\vec{x}, x_4)$, followed by a spatial translation $x \to x + \mathcal{L}$ by making use of the transformation of the Dirac propagator to be $Q_{x,y} \to \gamma_t Q_{x_p,y_p}\gamma_t$ and $W_{x,x+\mathcal{L}} \to W_{x_p-\mathcal{L},x_p}^\dagger$ under parity. In this case, the γ_t from parity transformation for Q commutes with $\Gamma = \gamma_t$. In other cases, one should take care of the $\pm$ factor. Thus, $C_{3\mathrm{pt}}$ becomes

$$C_{3\mathrm{pt}}^{u-d}(t, \tau, \mathcal{L}) = \sum_{\vec{y},\vec{z}}(T_1 e^{-i\vec{p}\cdot\vec{y}} + T_1^* e^{i\vec{p}\cdot\vec{y}}), \quad \text{(C2)}$$

and therefore proportional to the connected piece of $\bar{u}\Gamma W u$, which is the first term in the above equation. We normalize the three-point function such that the $u - d$ isospin charge of the pion is 1. By going through the similar calculation, one can show that the three-point function is real also for $\Gamma = \gamma_z$ while it is purely imaginary for $\Gamma = 1$ $u - d$ pion qPDF.

APPENDIX D: RELATION BETWEEN P_z AND-P_z MATRIX ELEMENTS

In this appendix, we derive the relation between conjugates of the matrix elements $\langle E_{n'}, P_z | \mathcal{O}_\Gamma(z;\tau) | E_n, P_z \rangle$ that enter the excited state contributions to the qPDF three-point function. For this, let us consider the conjugate of the simplest component of the qPDF matrix element:

$$\sum_z \langle E_{n'}, P_z | \bar{u}_z \Gamma W_{z,z+\mathcal{L}} u_{z+\mathcal{L}} | E_n, P_z \rangle^*$$

$$= \sum_z \langle E_n, P_z | \bar{u}_{z+\mathcal{L}} \Gamma W_{z,z+\mathcal{L}}^\dagger u_z | E_{n'}, P_z \rangle, \quad \text{(D1)}$$

for $\Gamma = \gamma_t, \gamma_z, 1$. Using the parity operator Π, the right-hand side of the above equation becomes

$$\sum_z \langle E_n, P_z | \bar{u}_{z+\mathcal{L}} \Gamma W_{z,z+\mathcal{L}}^\dagger u_z | E_{n'}, P_z \rangle$$

$$= \sum_z \langle E_n, -P_z | (\Pi \bar{u}_{z+\mathcal{L}} \Pi)(\Pi W_{z,z+\mathcal{L}}^\dagger \Pi)(\Pi u_z \Pi) | E_{n'}, -P_z \rangle$$

$$= \sum_z \langle E_n, -P_z | \bar{u}_{z-\mathcal{L}} \gamma_t \Gamma W_{z-\mathcal{L},z} \gamma_t u_z | E_{n'}, -P_z \rangle$$

$$= \sum_z \langle E_n, -P_z | \bar{u}_z \gamma_t \Gamma \gamma_t W_{z,z+\mathcal{L}} u_{z+\mathcal{L}} | E_{n'}, -P_z \rangle. \quad \text{(D2)}$$

Defining, $\gamma_t \Gamma \gamma_t = \Phi_\Gamma \Gamma$ with $\Phi_\Gamma = \pm 1$, we have the relation

$$\sum_z \langle E_{n'}, P_z | \bar{u}_z \Gamma W_{z,z+\mathcal{L}} u_{z+\mathcal{L}} | E_n, P_z \rangle^*$$

$$= \Phi_\Gamma \sum_z \langle E_n, -P_z | \bar{u}_z \Gamma W_{z,z+\mathcal{L}} u_{z+\mathcal{L}} | E_{n'}, -P_z \rangle, \quad \text{(D3)}$$

with $\Phi_\Gamma = 1$ for $\Gamma = \gamma_t$, 1 and $\Phi_\Gamma = -1$ for $\Gamma = \gamma_z$ with all Γ being in the Minkowskian convention. Thus, we can average over P_z and $-P_z$ data after taking care of appropriate factor of Φ_Γ.

APPENDIX E: PION TWO-POINT FUNCTIONS AND ENERGY LEVELS

In this appendix we discuss some details of the calculations of the pion two-point function. We tested several different sources for the pion. In these tests we used 50 gauge configurations. We used Gaussian sources with several steps of Wuppertal smearings as well as in the Coulomb gauge (see main text). In Fig. 19 we show the effective mass for 40 and 90 steps of Wuppertal smearings as well as the Coulomb gauge Gaussian sources of size 0.3 fm. We see that 90 steps of Wuppertal smearings and Coulomb gauge Gaussian sources give similar effective masses, while the excited state contamination is larger for 40 steps of Wuppertal smearings. We also studied the
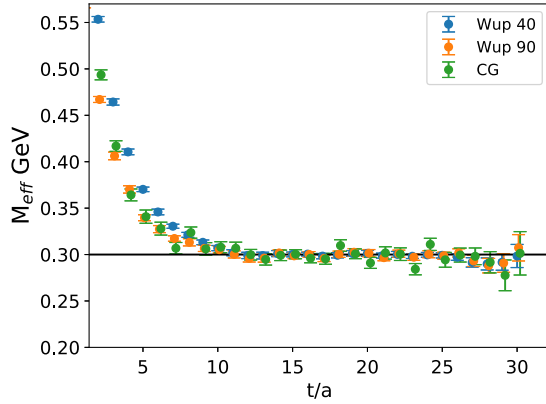

FIG. 19. Effective masses for $P_z = 0$ using Gaussian sources with 40 steps of Wuppertal smearings, 90 steps of Wuppertal smearings, and Coulomb gauge.
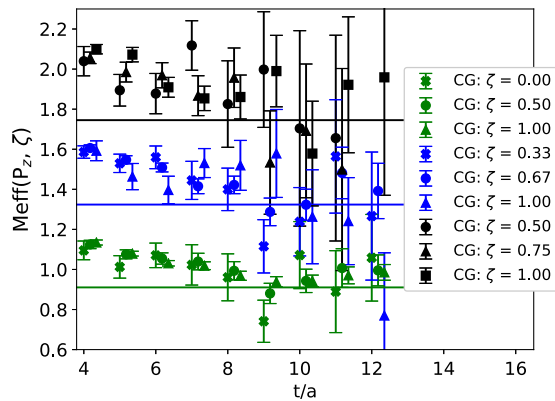

FIG. 20. Effective masses for different values of ζ with 50 configurations; green, blue, and black points correspond to momentum 0.86, 1.29, and 1.72 GeV, respectively.

two-point functions for different boosted Gaussian sources, with momentum boost k_z. The corresponding effective masses are shown in Fig. 20 for $P_z = 0.86$, 1.29, and 1.72 GeV for different values of $\zeta = k_z/P_z$. We clearly see that the nonzero value of ζ improves the signal for all P_z. We also see that $\zeta = 0.5$ is too small, while $\zeta = 1.0$ is too large for $P_z = 1.29$, but works well for $P_z = 0.86$ GeV.

APPENDIX F: IMPLEMENTATION OF MATCHING CONVOLUTION

Here, we describe the implementation of the plus function in the matching formula such as to ensure current conservation. The matching kernel is of the form

$$C\left(\frac{x}{y}, yP_z\right) = \delta\left(\frac{x}{y} - 1\right) + \frac{\alpha_s C_F}{2\pi} C_+^{(1)}\left(\frac{x}{y}, yP_z\right), \quad (F1)$$

where the dependence on P^R and μ are implicit. The first perturbative correction is a plus function that ensures the vector current conservation. The property we know of the plus function is that $\int_{-\infty}^{\infty} dx C_+^{(1)}(\frac{x}{y}, yP_z) = 0$, since the second dependence of the function is independent of x. In order to implement the plus function correctly, we can use the following procedure:

$$C_+^{(1)}(\xi, yP_z) = C^{(1)}(\xi, yP_z) - N(yP_z)\delta(\xi - 1). \quad (F2)$$

The x-independent but momentum dependent coefficient $N(yP_z)$ is

$$N(yP_z) \equiv \int_{\text{reg}} C^{(1)}(\xi, yP_z)d\xi, \quad (F3)$$

where $\int_{\text{reg}} d\xi$ involves an integration over the intervals $[-\Lambda, -\epsilon] \cup [\epsilon, 1 - \epsilon] \cup [1 + \epsilon, \Lambda]$ for some upper cutoff Λ and a small exclusion parameter ϵ. The above definition gives the usual result that

$$\int_{-\infty}^{\infty} d\xi C_+^{(1)}(\xi, yP_z)f(\xi) = \int_{-\infty}^{\infty} C^{(1)}(\xi, yP_z)(f(\xi) - f(1)), \quad (F4)$$

with y held fixed as $\Lambda \to \infty$ and $\epsilon \to 0$. The following is then true for any function f:

$$\int \int dxdy C_+^{(1)}(x/y, yP_z)f(y)$$
$$= \int dy \left(\int dx C_+^{(1)}(x/y, yP_z)\right)f(y)$$
$$= 0, \quad (F5)$$

leading to the vector current conservation or equivalently to the total area preservation between the qPDF and PDF. With this prescription, the matching formula becomes

$$\int_{-\infty}^{\infty} \frac{dy}{|y|} C_{+}^{(1)}(x/y, yP_z) q(y)$$
$$= \int_{\text{reg}} \frac{dy}{|y|} C^{(1)}(x/y, yP_z) q(y) - N(xP_z) q(x). \quad \text{(F6)}$$

It is convenient to write the above formula in an explicitly vector current conservation preserving form as

$$\int_{-\infty}^{\infty} \frac{dy}{|y|} C_{+}^{(1)}(x/y, yP_z) q(y)$$
$$= \int_{\text{reg}} \frac{dy}{|y|} C^{(1)}\left(\frac{x}{y}, yP_z\right) q(y) - \int_{\text{reg}} \frac{dy'}{|x|} C^{(1)}\left(\frac{y'}{x}, xP_z\right) q(x).$$
$$\text{(F7)}$$

However, care has to be taken in the numerical regularization of the above equation to be consistent with the one in Eq. (F3). That is, in the above equation, $\int_{\text{reg}} dy$ in the first integral in the right-hand side involves the range $y \in [-x/\epsilon, -x/\Lambda] \cup [x/\Lambda, x/(1+\epsilon)] \cup [x/(1-\epsilon), x/\epsilon]$ when $x > 0$, and the range reversed when $x < 0$. A consistent prescription for $\int_{\text{reg}} dy'$ in the second integral in the right-hand side involves $y' \in [-\Lambda x, -\epsilon x] \cup [\epsilon x, x(1-\epsilon) \cup [x(1+\epsilon), x\Lambda]$.

APPENDIX G: RESULTS ON TWO-STATE EXTRAPOLATIONS TO OBTAIN THE MATRIX ELEMENTS AT ALL P_z

In Fig. 4 in the main text, we showed some sample results for the $t - \tau/2$ behavior and the $t \to \infty$
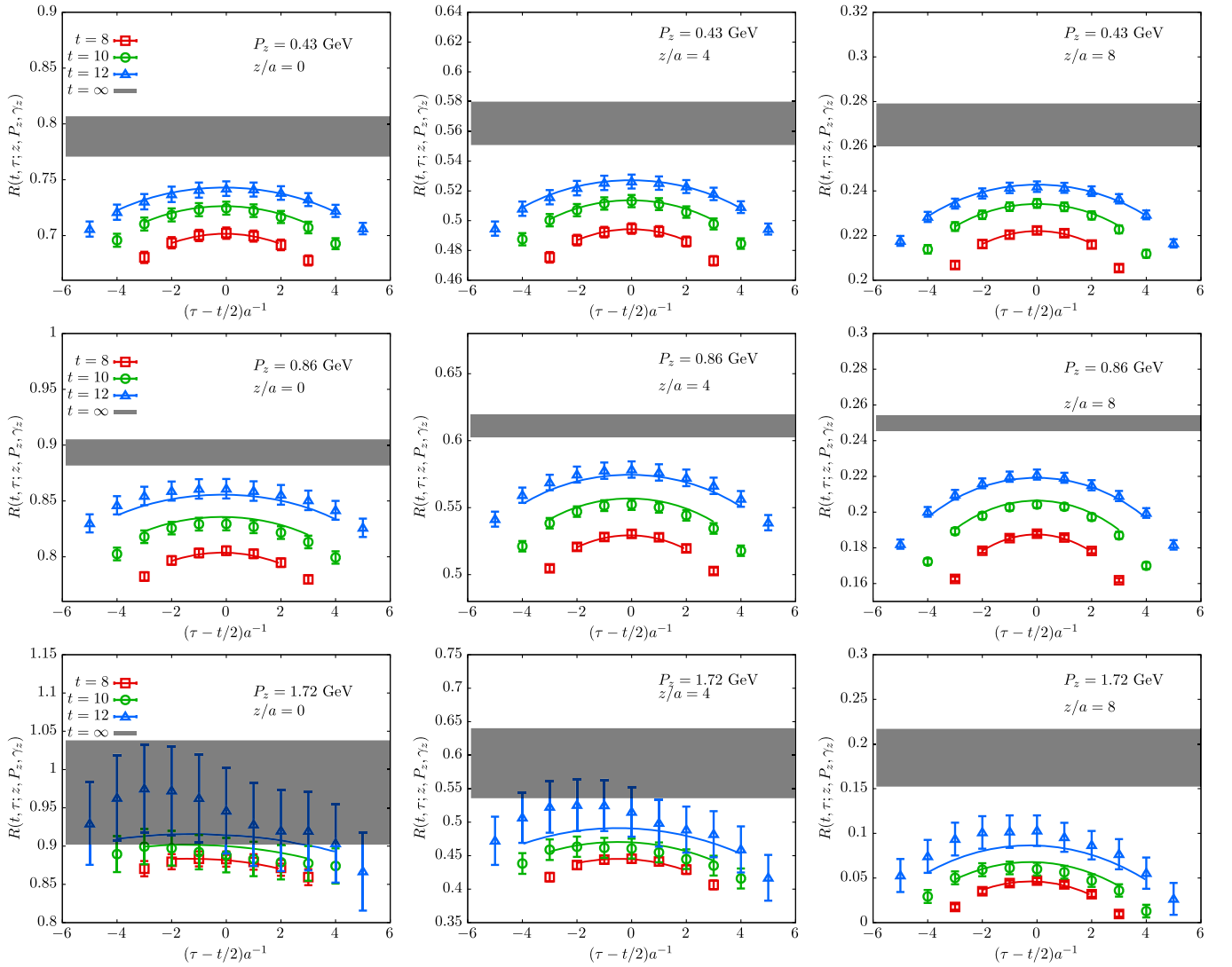

FIG. 21. The ratio of the three-point function to the two-point function, $R(t, \tau; z, P_z, \gamma_z)$ for $\Gamma = \gamma_z$ is shown as function of $\tau - t/2$ for $z/a = 0$, 4, and 8 (from left to right), and $P_z = 0$, 0.483, 0.86, and 1.72 GeV (top to bottom). The corresponding plots for $P_z = 1.29$ GeV are shown in Fig. 4 in the main text. The central values of the two-state fits to the lattice results for different source-sink separations are shown as the curves. The horizontal band corresponds to the extrapolated result for infinite source-sink separation.
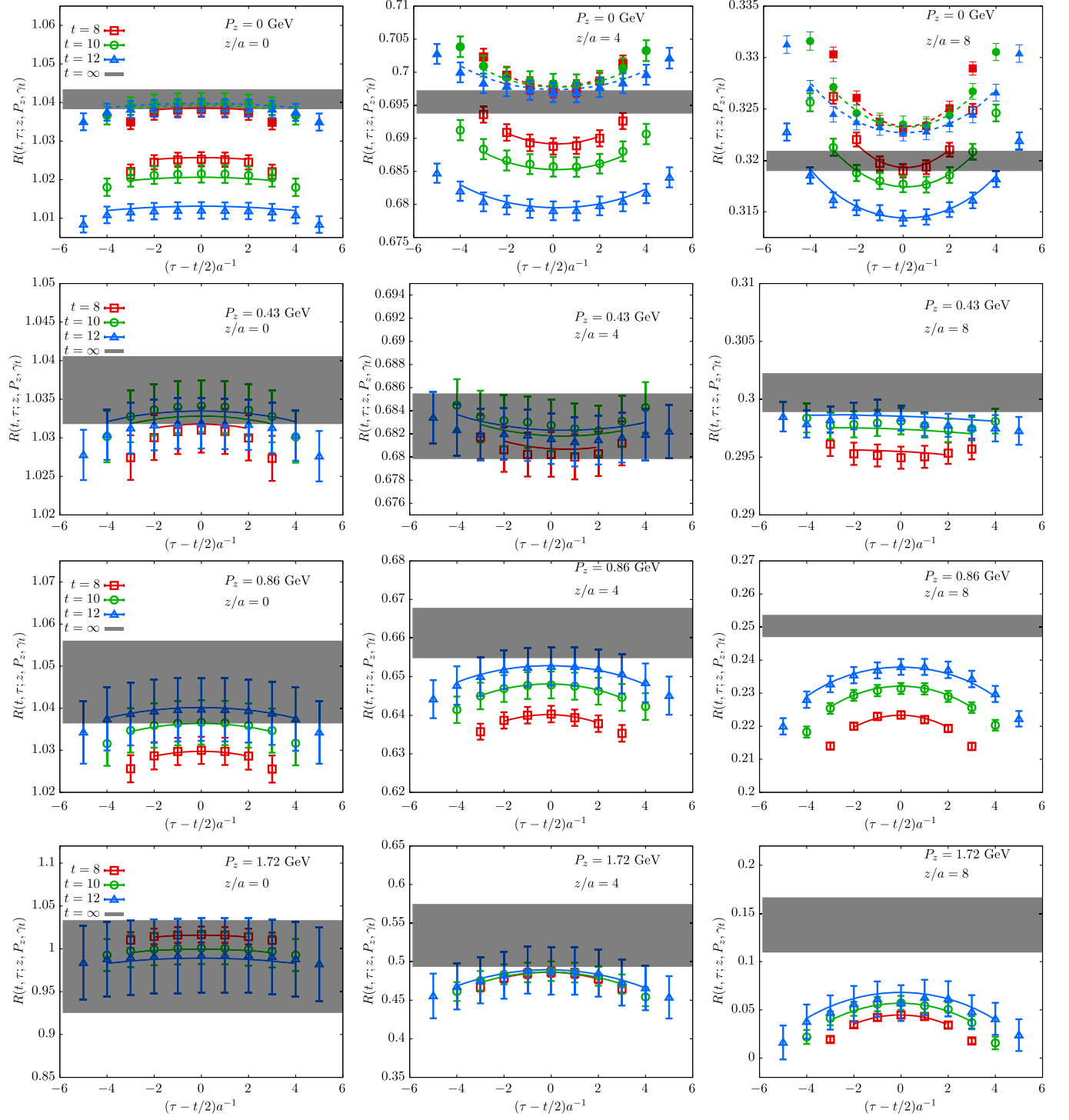
FIG. 22.   The ratio of the three-point function to the two-point function, $R(t, \tau; z, P_z, \gamma_t)$ for $\Gamma = \gamma_t$ is shown as function of $\tau - t/2$ for $z/a = 0$, 4, and 8 (from left to right), and $P_z = 0$, 0.483, 0.86, and 1.72 GeV (top to bottom). The corresponding plots for $P_z = 1.29$ GeV are shown in Fig. 4 in the main text. The central values of the two-state fits to the lattice results for different source-sink separations are shown as the curves. The horizontal band corresponds to the extrapolated result for infinite source-sink separation. The case of $P_z = 0$, in the top-most panels, is special due to the presence of the effect of lattice periodicity, and, hence, the various symbols and curves for the top-most panels are explained in detail in the text of Appendix G.

extrapolations of the three-point function to two-point function ratio $R(t, \tau; z, P_z, \Gamma)$ for $\Gamma = \gamma_t$ and γ_z at a specific intermediate value of $P_z = 1.29$ GeV. In Fig. 21 and Fig. 22 of this appendix, we show similar results at all P_z for $\Gamma = \gamma_z$ and γ_t, respectively, using Fit(2,2).

For the case of $P_z = 0$, special care needs to be taken. For a finite temporal extent L_t of the lattice, ignoring the effect of periodicity due to the presence of the terms $e^{-E_\pi(L_t - t)}$ in the denominator of Eq. (9) is justified when $e^{-E_\pi t} \ll e^{-E_\pi(L_t - t)}$. But one should include the effect of boundary condition if the two terms become comparable. For the largest source-sink separation $t/a = 12$ we use, the contribution from the wrapping-around term, $e^{-E_\pi(L_t - t)}$, relative to $e^{-E_\pi t}$ for $P_z = 0$ is 2.7%, whereas for higher P_z it is negligible; e.g., for the smallest nonzero momentum $P_z = 0.43$ GeV, this effect is 0.2%. Hence, we included the term $e^{-E_\pi(L_t - t)}$ in the denominator of Eq. (9) for the extrapolation of $R(t, \tau, z; P_z)$ for $P_z = 0$, and we also

checked that the effect of the periodicity of lattice was indeed negligible for any of the nonzero P_z we used.

For the case of $P_z = 0$ displayed in the top-most panels of Fig. 22, we have shown the data in two ways to make the fits and the extrapolated value easier to understand. The unfilled symbols are the data for $R(t, \tau; z, P_z, \gamma_t)$ defined as the ratio of $C_{3pt}(t, \tau; z, P_z)$ to $C_{2pt}(t; P_z)$, and the solid curves are the fits including the $e^{-E_\pi(L_t - t)}$ term in the denominator of Eq. (9). While the fits describe the data well, the trend in the data with increasing t can be seen to be away from the extrapolated value. To make the reason clearer, we have shown the modified ratio of $C_{3pt}(t, \tau; z, P_z)$ to the two-point function without the wrap-around term, $C_{2pt}(t; P_z) - A_0 e^{-E_\pi(L_t - t)}$, as the filled symbols. The dashed curves are now the fits using just Eq. (9). The values of the amplitude A_0 and the energy E_π were obtained by the two-state fit as described in the main text. Now, the trend with increasing t is clearer.

[1] J. C. Collins, D. E. Soper, and G. F. Sterman, Adv. Ser. Dir. High Energy Phys. **5,** 1 (1989).
[2] D. E. Soper, Phys. Rev. D **15,** 1141 (1977).
[3] X. Ji, Phys. Rev. Lett. **110,** 262002 (2013).
[4] X. Ji, Sci. China Phys. Mech. Astron. **57,** 1407 (2014).
[5] X. Xiong, X. Ji, J.-H. Zhang, and Y. Zhao, Phys. Rev. D **90,** 014051 (2014).
[6] M. Constantinou and H. Panagopoulos, Phys. Rev. D **96,** 054506 (2017).
[7] Y.-S. Liu, J.-W. Chen, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, arXiv:1807.06566.
[8] I. W. Stewart and Y. Zhao, Phys. Rev. D **97,** 054512 (2018).
[9] A. V. Radyushkin, Phys. Rev. D **96,** 034025 (2017).
[10] K. Orginos, A. Radyushkin, J. Karpie, and S. Zafeiropoulos, Phys. Rev. D **96,** 094503 (2017).
[11] Y.-Q. Ma and J.-W. Qiu, Phys. Rev. D **98,** 074021 (2018).
[12] Y.-Q. Ma and J.-W. Qiu, Phys. Rev. Lett. **120,** 022003 (2018).
[13] R. S. Sufian, J. Karpie, C. Egerer, K. Orginos, J.-W. Qiu, and D. G. Richards, Phys. Rev. D **99,** 074507 (2019).
[14] H.-W. Lin, J.-W. Chen, L. Jin, Y.-S. Liu, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Phys. Rev. Lett. **121,** 242003 (2018).
[15] J.-W. Chen, L. Jin, H.-W. Lin, Y.-S. Liu, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, arXiv:1803.04393.
[16] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato, and F. Steffens, Phys. Rev. D **98,** 091503 (2018).
[17] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato, and F. Steffens, Phys. Rev. Lett. **121,** 112001 (2018).
[18] J.-W. Chen, L. Jin, H.-W. Lin, Y.-S. Liu, A. Schfer, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, arXiv:1804.01483.
[19] T. Ishikawa, Y.-Q. Ma, J.-W. Qiu, and S. Yoshida, arXiv:1609.02018.

[20] J.-W. Chen, T. Ishikawa, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Phys. Rev. D **97,** 014505 (2018).
[21] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, H. Panagopoulos, and F. Steffens, Nucl. Phys. **B923,** 394 (2017).
[22] Y. Zhao, Int. J. Mod. Phys. A **33,** 1830033 (2018).
[23] K. Cichy and M. Constantinou, Adv. High Energy Phys. **2019,** 3036904 (2019).
[24] C. Monahan, Proc. Sci., LATTICE2018 (2018) 018.
[25] X. Xiong, T. Luu, and U.-G. Meiner, arXiv:1705.00246.
[26] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, A. Scapellato, and F. Steffens, Phys. Rev. D **99,** 114504 (2019).
[27] J. Badier et al. (NA3 Collaboration), Z. Phys. C **18,** 281 (1983).
[28] B. Betev et al. (NA10 Collaboration), Z. Phys. C **28,** 9 (1985).
[29] J. S. Conway et al., Phys. Rev. D **39,** 92 (1989).
[30] J. F. Owens, Phys. Rev. D **30,** 943 (1984).
[31] P. J. Sutton, A. D. Martin, R. G. Roberts, and W. J. Stirling, Phys. Rev. D **45,** 2349 (1992).
[32] M. Gluck, E. Reya, and A. Vogt, Z. Phys. C **53,** 651 (1992).
[33] M. Gluck, E. Reya, and I. Schienbein, Eur. Phys. J. C **10,** 313 (1999).
[34] K. Wijesooriya, P. E. Reimer, and R. J. Holt, Phys. Rev. C **72,** 065203 (2005).
[35] M. Aicher, A. Schafer, and W. Vogelsang, Phys. Rev. Lett. **105,** 252003 (2010).
[36] P. Petreczky, T. Izubuchi, L. Jin, C. Kallidonis, N. Karthik, S. Mukherjee, C. Shugert, and S. Syritsyn, arXiv:1812.04334.

[37] N. Karthik, T. Izubichi, L. Jin, C. Kallidonis, S. Mukherjee, P. Petreczky, C. Shugert, and S. Syritsyn, Proc. Sci., LATTICE2018 (**2018**) 109.

[38] C. Shugert, T. Izubichi, L. Jin, C. Kallidonis, N. Karthik, S. Mukherjee, P. Petreczky, and S. Syritsyn, Proc. Sci., LATTICE2018 (**2018**) 110.

[39] A. Hasenfratz and F. Knechtli, Phys. Rev. D **64**, 034504 (2001).

[40] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD and UKQCD Collaborations), Phys. Rev. D **75**, 054502 (2007).

[41] A. Bazavov *et al.* (HotQCD Collaboration), Phys. Rev. D **90**, 094503 (2014).

[42] E. Shintani, R. Arthur, T. Blum, T. Izubichi, C. Jung, and C. Lehner, Phys. Rev. D **91**, 114511 (2015).

[43] S. Gusken, U. Low, K. H. Mutter, R. Sommer, A. Patel, and K. Schilling, Phys. Lett. B **227**, 266 (1989).

[44] G. S. Bali, B. Lang, B. U. Musch, and A. Schfer, Phys. Rev. D **93**, 094515 (2016).

[45] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, Comput. Phys. Commun. **181**, 1517 (2010).

[46] R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower, and S. Gottlieb, arXiv:1109.2935.

[47] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. Brower, arXiv:1612.07873.

[48] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. D **98**, 030001 (2018).

[49] L. Maiani, G. Martinelli, M. L. Paciello, and B. Taglienti, Nucl. Phys. **B293**, 420 (1987).

[50] T. Ishikawa, Y.-Q. Ma, J.-W. Qiu, and S. Yoshida, Phys. Rev. D **96**, 094019 (2017).

[51] X. Ji and J.-H. Zhang, Phys. Rev. D **92**, 034006 (2015).

[52] A. M. Polyakov, Nucl. Phys. **B164**, 171 (1980).

[53] C. Alexandrou, M. Constantinou, T. Korzec, H. Panagopoulos, and F. Stylianou, Phys. Rev. D **83**, 014503 (2011).

[54] A. Bazavov, N. Brambilla, P. Petreczky, A. Vairo, and J. H. Weber (TUMQCD Collaboration), Phys. Rev. D **98**, 054511 (2018).

[55] T. Izubichi, X. Ji, L. Jin, I. W. Stewart, and Y. Zhao, Phys. Rev. D **98**, 056004 (2018).

[56] J.-W. Chen, T. Ishikawa, L. Jin, H.-W. Lin, A. Schfer, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, Sci. China Phys. Mech. Astron. **62**, 991021 (2019).

[57] H.-W. Lin, J.-W. Chen, T. Ishikawa, and J.-H. Zhang (LP3 Collaboration), Phys. Rev. D **98**, 054504 (2018).

[58] J. Karpie, K. Orginos, A. Rothkopf, and S. Zafeiropoulos, J. High Energy Phys. 04 (2019) 057.

[59] P. C. Barry, N. Sato, W. Melnitchouk, and C.-R. Ji, Phys. Rev. Lett. **121**, 152001 (2018).

[60] C. Best, M. Gockeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. E. L. Rakow, A. Schafer, G. Schierholz, A. Schiller, and S. Schramm, Phys. Rev. D **56**, 2743 (1997).

[61] M. Guagnelli, K. Jansen, F. Palombi, R. Petronzio, A. Shindler, and I. Wetzorke (Zeuthen-Rome Collaboration), Eur. Phys. J. C **40**, 69 (2005).

[62] S. Capitani, K. Jansen, M. Papinutto, A. Shindler, C. Urbach, and I. Wetzorke, Phys. Lett. B **639**, 520 (2006).

[63] A. Abdel-Rehim *et al.*, Phys. Rev. D **92**, 114513 (2015); **93**, 039904(E) (2016).

[64] M. Oehm, C. Alexandrou, M. Constantinou, K. Jansen, G. Koutsou, B. Kostrzewa, F. Steffens, C. Urbach, and S. Zafeiropoulos, Phys. Rev. D **99**, 014508 (2019).

[65] T. Nguyen, A. Bashir, C. D. Roberts, and P. C. Tandy, Phys. Rev. C **83**, 062201 (2011).

[66] C. Chen, L. Chang, C. D. Roberts, S. Wan, and H.-S. Zong, Phys. Rev. D **93**, 074021 (2016).

[67] K. D. Bednar, I. C. Clot, and P. C. Tandy, arXiv:1811.12310.

[68] G. F. de Teramond, T. Liu, R. S. Sufian, H. G. Dosch, S. J. Brodsky, and A. Deur (HLFHS Collaboration), Phys. Rev. Lett. **120**, 182001 (2018).

[69] S. J. Brodsky, AIP Conf. Proc. **792**, 977 (2005).

[70] E. Ruiz Arriola, Acta Phys. Pol. B **33**, 4443 (2002).

[71] W. Broniowski and E. Ruiz Arriola, Phys. Lett. B **773**, 385 (2017).

[72] V. M. Braun, A. Vladimirov, and J.-H. Zhang, Phys. Rev. D **99**, 014013 (2019).

[73] X. Ji, Y. Liu, and I. Zahed, Phys. Rev. D **99**, 054008 (2019).

[74] S. Gusken, Nucl. Phys. B, Proc. Suppl. **17**, 361 (1990).