Valence parton distribution function of pion from fine lattice

Taku Izbuchi1,2, Luchang Jin2,3, Christos Kallidonis4, Nikhil Karthik1, Swagato Mukherjee1, Peter Petreczky1, Charles Shugert1,4, Sergey Syritsyn1,4
1 Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA
2 RIKEN-BNL Research Center, Brookhaven National Lab, Upton, NY, 11973, USA
3 Physics Department, University of Connecticut, Storrs, Connecticut 06269-3046, USA
4 Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA
(Dated: May 17, 2019)

We present a lattice QCD study of 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 483 × 64 lattice at a fine lattice spacing a = 0.06 fm for this computation. We renormalize the quasi-PDF matrix element in the non-perturbative RI-MOM scheme. As a byproduct, we test the validity of 1-loop matching procedure by comparing the RI-MOM renormalized quasi-PDF 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 MS one-loop matching, implemented through a fit to phenomenologically motivated PDFs, we obtain the valence PDF 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

\[ f(x) = \frac{1}{2\pi} \int d\xi^- e^{i\xi P \cdot \xi^-} \langle H(P) | \bar{\psi}(\xi^-) \gamma^+ W(\xi^0, 0) \psi(0) | H(P) \rangle, \]

where \( W(\xi^0, 0) = \mathcal{P} \bar{\psi}(\xi) d^4 A^+ \) 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 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 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 large momentum \( P_z \):

\[ \tilde{q}(x, P_z) = \frac{1}{2\pi} \int d\xi^- e^{i\xi P \cdot \xi^-} \langle H(P_z) | \bar{\psi}(\xi^-) \Gamma \mathcal{W}(\xi, 0) \psi(0) | H(P_z) \rangle, \]

where \( \Gamma \) is either \( \gamma_5 \) or \( \gamma_\mu \) for the unpolarized parton distribution addressed in this paper. Here, \( \mathcal{W}(z, 0) \) is a straight spatial Wilson line joining the quark and anti-quark. For sufficiently boosted hadrons, one can use the Large Momentum Effective Theory (LaMET) [2] to relate the qPDF to PDF through a convolution with a matching kernel C as

\[ \tilde{q}(x, \mu_L, P_z) = \int_{-1}^{1} dy \left\{ C \left( \frac{x y}{y^2 \mu_L^2 - y^2 P_z^2} \right) f(y, \mu) \right\}. \]

Here \( \mu_L \) and \( \mu \) are the renormalization scales of the schemes in which the qPDF and PDF are defined. For the latter, MS scheme is used and \( \mu \) 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 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 Ref. [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 iso-vector quark distribution of 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 non-perturbative renormalization scheme because of the self-energy divergence of the Wilson line [19], and this is usually implemented using the 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 MS scheme through Eq. [3]. This is achieved through the convolution using the matching kernel between the RI-MOM and MS schemes that is perturbatively calculated in the continuum theory using dimensional regularization [8]. One could also define the qPDF operator in the 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 \( O(A_{SD}^2/(x^2 P_z^2)) \) 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 $O(1)$ fm. It is also important to ensure that $aP_z < 1$ to make sure we are not overcome with lattice artifacts\(^\text{[25]}\). Therefore, a closer look at this new methodology is warranted and is actively being studied.\(^\text{[7,29]}\). The aim of this paper is to explore these issues further by using finer lattices than what is being used in the qPDF literature, and use 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 pion since it can be accessed using the isosinglet $u - d$ PDF, and thereby, avoid mixing with the gluon sector. In our study, we will use the renormalization and matching strategy outlined in\(^\text{[7,29]}\). The pion valence PDF has been determined through leading order and next to leading order analysis of experimental data\(^\text{[27,35]}\), 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 section II, we discuss our lattices setup. In section 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 section 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 section V, we discuss the non-perturbative 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 section VI. Some technical aspects of the calculations are discussed in the Appendices. Preliminary results on this work have been reported in conference proceedings\(^\text{[36-38]}\).

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\(^\text{[39]}\) and Highly Improved Staggered Quark (HISQ) action\(^\text{[40]}\) 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\(^\text{[41]}\). In addition to this ensemble, we also used 64\(^3\) HISQ lattices\(^\text{[41]}\) with the lattice spacing $a = 0.04$ fm for the study of the non-perturbative 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.\(^\text{[7]}\)). 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, $m_{\pi}$, 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 separation $z$ than at larger ones; to be exact, we used 214, 162 and 100 gauge configurations for $|z|/a \in [0, 8], (8, 16]$ and $[16, 24]$ respectively. We further improved the statistics by using All-Mode Averaging (AMA)\(^\text{[42]}\) 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 iso-vector $u - d$ 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 Gaussian profile, typically implemented through a gauge covariant Wuppertal smearing\(^\text{[43]}\), are used for this purpose when the hadron is at rest. However, for 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\(^\text{[44]}\) 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 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 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.

III. TWO POINT FUNCTION OF THE BOOSTED PION

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

$$C_{2\text{pt}}(t, P_z) = \left\langle \left\{ \pi^+_s(t, \vec{P}) \right\} \left\{ \pi^+_s(0, \vec{P}) \right\} \right\rangle,$$

(4)

for a spatial pion momentum $\vec{P} = (0, 0, P_z)$ which is non-zero 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 $aP_z = \pm 2\pi /a_{x,y} n$ for $n$, 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 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 set-up 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 higher momenta, we followed Ref. [44] and used boosted sources in which the valence quarks are boosted to momentum \( k = \zeta P_z \), with \( \zeta \) being a tunable parameter. Naively, one might expect that the optimal choice to be \( \zeta = 0.5 \). However, we found that the optimal choice of \( \zeta \) for the pion in terms of the signal-to-noise ratio is between 0.6 and 0.75. For \( P_z = 0.86 \) GeV the signal-to-noise ratio is not very sensitive to the value of \( \zeta \). 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 \( \zeta \), 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 choices of 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 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 corresponding to the continuum dispersion relation

\[
E_{\text{eff}}(P_z) = \sqrt{P_z^2 + m_{\pi}^2},
\]

shown as the horizontal lines. The effective mass approach 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 multi-exponential fits on the SS and SP pion two point functions in the interval \( t \in [t_{\text{min}}, t_{\text{max}}] \) in order to extract the energy levels. For fixed \( t_{\text{min}} \) we varied \( t_{\text{max}} \) and checked the sensitivity of the result to \( t_{\text{max}} \). Then, we repeated the procedure for different values of \( t_{\text{min}} \). 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 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 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 \( E_\pi \) (blue circles) and \( E_1 \) (black circles) as a function of \( t_{\text{min}} \) used in the fits. For a given \( t_{\text{min}} \) the data points from two values of \( t_{\text{max}} \) have been clubbed together for \( P_z = 0 \) and 1.29 GeV respectively, and it demon-

FIG. 1. The effective masses \( E_{\text{eff}} \) from the pion two point functions with the boosted Coulomb gauge Gaussian source and sink for different momenta as a function of source-sink separation \( t \). The horizontal lines are the energy levels from the continuum dispersion relation with \( m_{\pi} = 300 \) MeV.
strates that there is no dependence on \( t_{\text{max}} \). 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 \( t_{\text{min}} \) 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 \( E_{Pz} = \sqrt{P_z^2 + m_{\pi}^2} \) with \( m_\pi = 300 \text{MeV} \), and imposing a prior on \( E_1 \) using its best estimate from the SP correlator – that is, we added the term \((E_1 - E_{1,\text{prior}})^2/\sigma^2_{\text{prior}}\) to the \( \chi^2 \) with \( E_{1,\text{prior}} \) and \( \sigma_{\text{prior}} \) being the mean and error of \( E_1 \), respectively, as determined from SP correlator. The \( t_{\text{min}} \) dependence of the resulting \( E_1 \) from 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 \( t_{\text{min}} \) whether or not priors are used, and there is a significant dependence on \( t_{\text{min}} \). One should use \( t_{\text{min}} \geq 6a \) to obtain reliable results for the first excited state from SS correlator. The three-state fits with priors on \( E_\pi \) and \( E_1 \) give energies of excited states that are the same within errors for the SS two point correlators and show almost no \( t_{\text{min}} \) dependence. In summary, we determined the lowest three energy levels using the SP correlator and then determined the corresponding amplitudes \( |\gamma_n\rangle = |\langle 0|\pi^+_3(0,\vec{P})|E_{\pi},P_z\rangle\rangle \) of these excited states in the SS correlator through a constrained fit analysis.

In Fig. 3 we show the three energy levels obtained from 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 PDG. 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 particle-like dispersion and this 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 \( E_\pi \) 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

\[
 h^R_t(z, P_z) = \langle E_{\pi}, P_z | O_t(z, \tau) | E_{\pi}, P_z \rangle, \tag{5}
\]

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

\[
 O_t(z; \tau) = \sum_{x_L} \left( \bar{u}_x W_{x,\tau+L} \Gamma_{\tau+L} \bar{d}_x W_{x+\tau+L} \Gamma_{\tau+L} \right), \tag{6}
\]

which is made gauge-invariant by the Wilson line \( W_{x,\tau+L} \). The Dirac \( \gamma \) matrices in the qPDF operator are in the Minkowskian convention. The state \( |E_{\pi}, P_z\rangle \) denotes the onshell ground state pion with momentum \( P_z \). In addition to the natural choices of \( \Gamma = \gamma_\tau \) and \( \gamma_z \) that approach \( \gamma^+ \) in the light-cone limit, we also considered \( \Gamma = 1 \). This choice of \( \Gamma \) is needed because under renormalization, \( O_{\gamma_\tau}(z) \) mixes with \( O_{\gamma_z}(z) \). We applied one-level of HYP smearing to the links entering \( W_{x,\tau+L} \) in order to reduce the noise. Since the qPDF calculation involves values of \( z \sim O(a) \), we checked that there is no significant difference between the renormalized matrix elements using smeared and unsmeared Wilson line. To obtain the bare matrix element \( h^R_t(P_z, z) \), we computed the three-point function at different source-sink separations \( \tau \) and operator insertion point \( \tau \)

\[
 C^{SS}_{3pt}(t, \tau; z, P_z) = \left( \langle \pi^+_3(t, \vec{P}) | O_t(z; \tau) | \pi^+_3(0, \vec{P}) \rangle \right), \tag{7}
\]

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

\[
 R(t, \tau; z, P_z, \Gamma) = \frac{C^{SS}_{3pt}(t, \tau; z, P_z)}{C^{SS}_{2pt}(t, P_z)}. \tag{8}
\]
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 $O_{\gamma}(z)$ in a pion external state is real at all $z$ for $\Gamma = \gamma_x, \gamma_t$ and purely imaginary for $\Gamma = 1$ (refer Appendix C). Inserting a complete set of states in the above equation,

$$R(t, \tau; z, P_z, \Gamma) = \sum_{n,n',} A_n A_{n'} \langle E_{n'}, P | O_{\gamma}(z) | E_{n}, P \rangle e^{-(E_{n'}-E_n)\tau-E_{n'}t},$$

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 $h^0_{\gamma}(z,P)$. The above equation holds for infinite time extent. For finite time extent, the effects of periodic boundary condition should be taken into account. This turns out to be important for $P_z = 0$, while for non-zero $P_z$ the effect is negligible as discussed in Appendix C. In practice, one truncates the sums in Eq. (9) at some value $n$, and then obtain $h^0_{\gamma}(z,P)$ by fitting the $t$ and the $\tau$ dependence of $R(t, \tau; z, P_z, \Gamma)$ using $\langle E_{n'}, P | O_{\gamma}(z) | E_{n'}, P \rangle$ 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 $\text{Fit}(n,\tau_o)$. In this method, the excited states are suppressed by $\exp(-E_{n'-E_n} t/2)$. For $z = 0$, it is easy to see that $R(t, \tau; z, P_z, \Gamma)$ is symmetric in $\tau$ around the mid-point $\tau-t/2$. For $z \neq 0$ and $P_z \neq 0$, we only have the following relation (see Appendix D),

$$\langle E_{n'}, P | O_{\gamma}(z) | E_{n'}, P \rangle^{*} = \phi_{\Gamma} \langle E_{n'}, P | O_{\gamma}(z) | E_{n}, P \rangle,$$

where $\phi_{\Gamma} = 1$ for $\Gamma = \gamma_x, 1$ and $\phi_{\Gamma} = -1$ for $\Gamma = \gamma_t$. Thus, generically the matrix elements $\langle E_{n'}, P | O_{\gamma}(z) | E_{n'}, P \rangle$ and $\langle E_{n'}, P | O_{\gamma}(z) | E_{n}, P \rangle$ are independent and the number of fit parameter is thereby increased. Based on the above relation, we constructed appropriate averages using both the positive and negative values of momenta to increase 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 ratio $R(t, \tau; z, P_z, \Gamma)$ is shown for $\Gamma = \gamma_x$ and $\gamma_t$ qPDF 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 C. In the figure, we show the data at $t/a = 8, 10$ and 12 along with the result of $\text{Fit}(t,2)$. Using the fit, the results for the $t \to \infty$ extrapolations are shown with the horizontal bands. We also performed three states fit of $R(t, \tau; z, P_z, \Gamma)$ and 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 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 $\Gamma = \gamma_x$ than for $\Gamma = \gamma_t$. Furthermore, the excited state contribution grows with increasing $z$. The non-symmetric 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 charge con-
FIG. 5. The \( t \) dependence of \( \mathcal{R}_{\text{sum}}(t; z = 0, \gamma_z) \) for \( P_z = 1.29 \text{ GeV} \) 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 \( \mathcal{R}_{\text{sum}}(t; z = 0, \gamma_z) \) using the fit(2,2) best fit parameters.

We will refer to this method as the sum(1) and sum(2) in Fig. 5 for \( \Gamma = \gamma_z \) and \( z = 0 \). We see that \( \mathcal{R}_{\text{sum}}(t; z, \gamma_z) \) 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 \( \mathcal{R}_{\text{sum}}(t; z, \Gamma) \) 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 \( \tau_o \) 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, \( P_z/E_z(P_z) \). One can take the ratio of matrix elements \( h^B_{\gamma_z}(z = 0, P_z)/h^B_{\gamma_z}(z = 0, P_z) \) to avoid issues of renormalization. The results for this ratio of matrix elements is shown in Fig. 6 along with the curve for \( P_z/\sqrt{P_z^2 + m^2_\pi} \). 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 \( \gamma_z \) matrix element due to 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 \( h^B_{\gamma_z}(z, P_z) \) as function 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), 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 \( \gamma_z \), some tension between the summation method and two and three state fits is observed at larger \( |z| \). At large value of \( z \) the matrix elements are suppressed partly because of the large 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 unsmear 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 non-perturbative renormalization (discussed next in Section V), were similar to those obtained with smeared Wilson lines within errors. The main difference between the renormalized three-point function obtained with smeared and unsmear Wilson line is that, for the latter the statistical errors at large \( z \) are significantly larger.
FIG. 7. The bare matrix elements $b_{\gamma}^0(z, P_z)$ (left) and $h_{\gamma}^0(z, P_z)$ (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.

V. RI-MOM NON-PERTURBATIVE 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 [27, 48]. In addition to the quark wavefunction renormalization $Z_q$ and the composite operator renormalization required for $z = 0$, the qPDF operator at non-zero $z$ requires additional renormalization due to the UV divergence present in the Wilson line connecting the quark and anti-quark [49]. When a 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 $\gamma_z$ qPDF operator is purely multiplicative, while the $\gamma_z$ 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 counterterm for 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 RI-MOM renormalized qPDF to the $\overline{MS}$ 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 non-perturbative renormalization, we compute the expectation value of qPDF operator between offshell quark external states with momentum $p$. We refer to the momentum of quark in the direction of Wilson-line as $p$, and the magnitude of the component perpendicular to the Wilson-line as $p_\perp$. For these computations, we use Landau gauge fixing. Let $\Lambda_T(z, p)$ be the quark-line amputated bare qPDF,

$$\Lambda_T(z, p) = \langle Q(p) \rangle^{-1} \langle \bar{u}(p) O_T(z, \tau) u(p) \rangle \langle Q(p) \rangle^{-1},$$

(13)

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

$$q_T(z, p) \equiv \text{Tr} \left[ \mathcal{P} \Lambda_T(z, p) \right],$$

(14)

consistent with the definition used in perturbative calculations. Here, $\mathcal{P}$ is the operator used to project onto one of the $\gamma$-matrices $\Gamma = \gamma_\nu$, and $\text{Tr}(\ldots)$ is a trace over both color and Dirac indices. Based on previous works [8, 20], we will use $p$-projection for which $\mathcal{P} = \gamma_\nu / (2p_\nu)$. Alternatively, one can use $\mathcal{P} = \Gamma$ [21] or minimal projection [8]. In the case of $\Gamma = \gamma_t$, since the renormalization is simply multiplicative, the renormalized quark qPDF is given by

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

(15)
where the \( z \)-dependent RI-MOM renormalization constant \( Z \) is determined using the renormalization condition set at momentum \( p^R \) as

\[
q_{\gamma}^R(z, p, p^R) \bigg|_{p=p^R} = e^{\rho(z)}.
\]  

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

\[
q_{\gamma}^R(z, p, p^R) = Z_{\gamma \gamma}(z, p^R)q_{\gamma}(z, p) + Z_{\gamma 1}(z, p^R)q_1(z, p).
\]  

The diagonal part \( Z_{\gamma \gamma} \) and the mixing term \( Z_{\gamma 1} \) are determined using the two RI-MOM conditions

\[
q_{\gamma}^R(z, p, p^R) \bigg|_{p=p^R} \equiv e^{\rho_z(z)},
\]

\[
Z_{\gamma \gamma}(z, p^R)\text{Tr}A_\gamma(z, p) + Z_{\gamma 1}(z, p^R)\text{Tr}A_1(z, p) = 0.
\]  

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

\[
\begin{align*}
\hbar_\gamma^R(z, p, p^R) &= Z_{\gamma \gamma}(z, p^R)\hbar_\gamma(z, p) + Z_{\gamma 1}(z, p^R)\hbar_1(z, p), \\
\hbar_1^R(z, p, p^R) &= Z_{\gamma \gamma}(z, p^R)\hbar_\gamma(z, p) + Z_{\gamma 1}(z, p^R)\hbar_1(z, p),
\end{align*}
\]

where \( Z_{\gamma} \) is the quark renormalization, that can be determined using the condition

\[
Z_{\eta}(p^R)^{-1} \frac{1}{12} \text{Tr} \left( \langle Q(p^R) \rangle^{-1} Q_{\text{tree}}(p^R) \right) = 1,
\]

where \( Q(p) \) is the quark propagator determined using Landau gauge and \( Q_{\text{tree}} \) 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 \( Z_{\gamma \gamma} \) determined at \( p^R = 1.29 \text{ GeV} \) and \( p_1^R = 1.49 \text{ GeV} \). 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, \( e^{\rho(z)} \), 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 \( e^{\rho(z)} \) cancels exactly between the renormalization factors and the bare qPDF operator. Therefore, we remove \( e^{\rho(z)} \) 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 \( a = 0.04 \text{ fm} \) ensemble was determined in [51], 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 \( Z_{\gamma \gamma} \) remain \( O(1) \) 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 \( \gamma \) and \( \gamma \gamma \). In this case, we have the diagonal factor \( Z_{\gamma \gamma} \) as well as off-diagonal factor \( Z_{\gamma 1} \) to take care of mixing with scalar on the lattice. We show \( Z_{\gamma \gamma} \) and \( Z_{\gamma 1} \) as the filled and unfilled symbols in the bottom panels. We observe that the imaginary part of \( Z_{\gamma \gamma} \) is small compared to the real part. This is not the case for \( Z_{\gamma \gamma} \), which in turn will affect the asymmetry of the \( u - d \) qPDF \( q_{u-d}(x) \) of pion about \( x = 0 \). We also note that mixing of \( \gamma \) with 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 \( z = 0 \), \( h_\gamma \), is the local current operator which will be exactly conserved in the continuum limit. Hence, \( Z_q Z_{\gamma \gamma}(z = 0) \) 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 non-perturbative) contributions to this lattice correction to \( Z_V \) at smaller renormalization scales. In Fig. 9 we show \( Z_V \) determined using Eq. (20), the value of \( Z_{\gamma \gamma} \) 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 \( h_\gamma^R(z = 0, P_z = 0) = 0.961(3) \). The values of \( Z_V \) determined from \( h_\gamma^R(z = 0, P_z) \) at the other non-zero \( P_z \), 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 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 \( Z_{\gamma \gamma} \) for renormalizing our pion qPDF.

### A. Comparison with leading order perturbation theory for \( z < 0.3 \text{ fm} \)

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

\[
\xi(z, p, p^R) = \frac{q_{\gamma}^R(z, p, p^R)}{q_{\gamma}^R(z, p, p^R)} - 1,
\]

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

\[
\xi(z, p, p^R) = \frac{Z_{\gamma \gamma}(z, p^R) - Z_{\gamma \gamma}(z, p)}{Z_{\gamma \gamma}(z, p)}.
\]

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

$$
\zeta_t(z, p, p^R) = \frac{\alpha_s C_F}{2\pi} \int_{-\infty}^{\infty} dx (e^{(1-x)p_c} - 1) \left[ H(x, p) - \left| \frac{p_z}{p^R} \right| \left( 1 + \frac{p_z}{p^R} (x - 1) \right) \right],
$$

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.\[2\]

---

1 The formula differs from the one given in \[8\] due to the issue of order of $\epsilon = 0$ limit in dimensional regularization and the $\epsilon = 0$ limit. We thank Yong Zhao for communicating the corrected result to us.

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.
The functional form of $H(x, p)$ for $\gamma_t$ and $\gamma_l$ isovector qPDF is 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 leads to UV finite and renormalized result for $\zeta$. In the discussions below, we will consider the cases with $p_\perp = p_\perp^R$ and $p_\perp \neq p_\perp^R$ separately. In the above leading order formula, the scale at which $\alpha_s$ has to be evaluated is unclear. Therefore, we vary $\alpha_s$ by changing the scale from $0.5 p_\perp^R$ to $2 p_\perp^R$ through the 1-loop running, and quote this variation as an uncertainty in the perturbative results below. On the lattice side, we determine $\zeta(z, p, p_R^R)$ using non-perturbatively determined $Z$-factors. In order to estimate the lattice spacings, we determined $\zeta$ 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 various plots that follow.

In the left and right panels of Fig. 10 we show the typical dependence of $\zeta_{\|}(z)$ and $\zeta_{\perp}(z)$ respectively, as a function of $z$ when $p_\perp^L$ differs slightly from the transverse quark momentum $p_\perp$, while the longitudinal components $p_\| \neq p_\perp^R$ 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 $p_\|$ 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 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 $\zeta$ when $p^R_s$ is changed from $p_\perp$. Surprisingly, the 1-loop result seems to work better for $\zeta_R$ than for $\zeta_I$. In the case of $\gamma_R$, one can certainly see a large lattice spacing effect with the movement of data towards the 1-loop result as lattice spacing is reduced, while in the case of $\gamma_I$, 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 $\gamma_R$ qPDF fares worse compared to $\gamma_I$ qPDF simply due to the presence of mixing with scalar or whether $\gamma_I$ qPDF might eventually show better perturbative convergence and lesser lattice spacing dependence in spite of its other disadvantages.

In Fig. 11, we concentrate on the renormalization flow of $\zeta_T$ at fixed small value of $z = 0.12$ fm. The two panels show the dependence of $\zeta_R$ and $\zeta_I$ on $p_R$ which is changed around $p_\perp$. As before, we keep $p_z = p^R_z = 1.92$ GeV. The 1-loop result is able to capture the qualitative trend of the flow in both $\gamma_R$ and $\gamma_I$. For both the cases, we can see that the reduction of lattice spacing leads to better agreement with 1-loop result. Having discussed the cases where $p_z = p^R_z$, we now study the dependence of $\zeta$ on $p^R_z \neq p_z$, while keeping $p_\perp = p^R_\perp$. We show the $z$-dependence of $\zeta_R$, when $p^R_z = 1.5 p_z$, in Fig. 12. We find the perturbative result 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 $\zeta$ 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 of including higher order corrections in the perturbative result for $\zeta$.

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

Physically, one would expect that for well-separated quark-antiquark with $z > 1$ fm, one would start seeing traces of nonperturbative physics in the qPDF. Quantifying the advent of non-perturbative 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 $e^{ipz}$. We expect, to a first approximation, that the effect of nonperturbative physics is to damp this pure wave via an inverse screening length $m_{scr} \sim O(\Lambda_{QCD})$. Thus, we model the bare quark qPDF as

$$q(z, p) = A e^{ipz} e^{-m_{scr} |z|} e^{-\epsilon |z|},$$

where we have removed the UV divergent piece $e^{-\epsilon |z|}$ from the qPDF and defined the left-over exponent $m_{scr}$ 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 non-trivial dependence of the amplitude $A$ on $z$, which we assume to be sub-leading compared to the leading damped oscillatory behavior and ignore them in the discussion here. There is an ambiguity in $m_{scr}$ depending on the scheme used to determine the divergent piece $\epsilon$. Since the values of $\epsilon$ determined from static quark potential method ensures that the renormalization factors after the removal of $e^{-\epsilon |z|}$ are $O(1)$ at smaller $z$ in Fig. 8, the choice of

![Figure 12](image12.png)

FIG. 12. The dependence of $\zeta$ on the longitudinal momentum $p_R$. $\zeta_R$ is shown as a function of $z$ for a specific choice of $p_z$ and $p_\perp = p^R_\perp$. 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.

![Figure 13](image13.png)

FIG. 13. The real and imaginary parts of the lattice data for $q_R(z, p)$ is compared with purely oscillatory model (dashed curves) at short-distances, and damped oscillatory ansatz (solid curves) at larger $|z|$. The data corresponds to quark momentum $(p_z, p_\perp) = (1.29, 1.49)$ GeV.
separation of the exponential suppression factor into a divergent and physical scales as defined in Eq. (24) is well motivated in this scheme. In Fig. [13] we show the bare quark qPDF $q_{\pi}(z, p)$ for $p_z = 1.49$ GeV and $p_t = 1.29$ 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 $\omega = 0.85p_t$ 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 $\omega$ and $m_{\text{scr}}$ through

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

In Fig. [14] we show the behavior of $m_{\text{scr}}$ and $\omega$ as a function of $z$ as extracted from $q_{\pi}(z, p)$. We have chosen different set of $p_z$ and $p_t$ to show the dependence on $p_t$ at fixed $p_z$ and vice versa. From the top panel, we see that $\omega/p_t$ 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 $\omega$ at short distances is dependent on $p_z$ and $p_t$, the approach to $\omega \approx p_t$ is universal. We observed this behavior when we used $q_{\pi}(z, p)$ as well. A physical reasoning for this observation could be that

at shorter $z$, the quark has the ability to radiate 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 $m_{\text{scr}}$. 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 $m_{\text{scr}}$. One can clearly see the emergence of non-zero $m_{\text{scr}} \approx 300$ MeV for $|z| > 0.5$ fm which is in the typical $\Lambda_{\text{QCD}}$ scale. When we repeated this using quark qPDF with 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 unsmeared Wilson line. This assures us that the observed $m_{\text{scr}} \approx 300$ 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 $\omega$ and $m_{\text{scr}}$ for these larger $z$ indicates that a simple physically motivated ansatz in Eq. (25) surprisingly offers a very good description of the actual non-perturbative 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 $p_t$.

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 connection with the conventional, separately defined quark distributions $Q_{u,d}(x)$ and the anti-quark distributions $Q_{d,u}(x)$ that are non-zero only between $x = 0$ and 1, through the relation

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

(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, $Q_{d}(x) = Q_{u}(x)$ and $Q_{u}(x) = Q_{d}(x)$. Therefore, for the positively charged pion $Q_{d}(x) - Q_{u}(x) = Q_{u}(x) - Q_{d}(x) = f^{u,u}_{\pi}(x)$ is the valence u-quark distribution. Again, due to isospin symmetry the $u$ and $d$ valence distributions are the same as $f^{u,d}_{\pi}(x) = f^{u,d}_{\pi}(x) = f^{d,u}_{\pi}(x)$. However, unlike the valence quark distribution, the isoritpplet $u - d$ PDF $f_{u-d}(x)$ satisfies $f_{u-d}(x) = f_{u-d}(x)$
and it has support from $-1$ to $1$. That is,

$$
\begin{align*}
    f_{u-d}(x) &= \begin{cases} 
        Q_u(x) - Q_d(x), x > 0 \\
        Q_u(-x) - Q_d(-x), x < 0 
    \end{cases}, \\
    f_{V}^{c}(x) &= \begin{cases} 
        Q_u(x) - Q_d(x), x > 0 \\
        0, x < 0.
    \end{cases}
\end{align*}
$$

(27)

Therefore, one can obtain the $u - d$ quark distribution, and from it, one can obtain $f_{V}^{c}(x)$ 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 $y P_2 / P_2^R$, $(P^R / P_2^R)\gamma^2$ and factorization scale $\mu$ implicit, the one-loop contribution to matching kernel $C(x/y)$ from RI-MOM to $\overline{\text{MS}}$ consists of two terms: $\mathcal{F}_1(x/y)$ and $\mathcal{F}_2(1 + \eta'(x - y))$ with $\eta' = P_2 / P_2^R$. The expressions $\mathcal{F}_1, \mathcal{F}_2$ depend on the choice of $\Gamma (\gamma, \gamma_\text{Reg})$. Furthermore, $\mathcal{F}_2$ depends on the projection method of RI-MOM scheme. Using the matching formula \cite{7, 8} on $f_u(x)$ and $f_d(x)$ to obtain the qPDFs $\tilde{q}_u(x)$ and $\tilde{q}_d(x)$,

$$
\begin{align*}
    \tilde{q}_{u-d}(x, P_2, P_2^R) &= f_{u-d}(x, \mu) + \frac{\alpha_s C_F}{2\pi} \int_{-1}^{1} dy \mathcal{F}_1 \left( \frac{x}{y} \right) f_{u-d}(y) \\
    &\quad - \frac{\alpha_s C_F}{2\pi} \int_{-1}^{1} dy \eta'[\mathcal{F}_2 (1 + \eta'(x - y))] f_{u-d}(y) + \cdots. \quad (28)
\end{align*}
$$

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

$$
\begin{align*}
    \mathcal{F}_{1,2}(\xi, \epsilon) = \mathcal{F}_{1,2}(\xi) - \delta(1 - \xi) \int_{\text{Reg}} d\xi' \mathcal{F}_{1,2}(\xi'). \quad (29)
\end{align*}
$$

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

$$
\begin{align*}
    \tilde{q}_{u-d}(x, P_2, P_2^R) &= f_{u-d}(x, \mu) + \frac{\alpha_s C_F}{2\pi} \int_{-1}^{1} dy \mathcal{F}_1 \left( \frac{x}{y} \right) f_{u-d}(y) \\
    &\quad - \frac{\alpha_s C_F}{2\pi} \int_{-1}^{1} dy \eta'[\mathcal{F}_2 (1 + \eta'(x - y))] f_{u-d}(y), \quad (30)
\end{align*}
$$

with the terms in “…” 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 making the renormalized qPDF matrix element complex. One can see this by starting from the matching convolution above and find that,

$$
\begin{align*}
    \tilde{q}(x, P_2, P_2^R) - \tilde{q}(-x, P_2, P_2^R) &= \frac{\alpha_s C_F |\eta'|}{2\pi} \int_{-1}^{1} dy \left[ -\mathcal{F}_2 (1 + \eta'(x - y)) + \mathcal{F}_2 (1 - \eta'(x - y)) \right] f(y, \mu),
\end{align*}
$$

(31)

is non-zero due to an RI-MOM specific term $\mathcal{F}_2$, while the terms containing $\mathcal{F}_1$ cancel due to its dependence only on $|P_2|$. 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 $\mu^2 z^2$. \cite{52}. In RI-MOM scheme there are two renormalization scales, $P_2^R$ and $P_2$, 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 $p$-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 $\tilde{q}(x, P_2, P_2^R)$ of the renormalized RI-MOM real-space qPDF matrix element $h^R(z, P_2, P_2^R)$, and the $\overline{\text{MS}}$ PDF $f(x, \mu)$ at factorization scale $\mu$. The relation is through the convolution in Eq. (3). There are two approaches to consider here:

1. One can parametrize the real space data $h^R(z, P_2, P_2^R)$ over the range $z$ where one has the lattice data and then model the dependence of $h^R(z, P_2, P_2^R)$ over $z$ extending to infinity where data does not exist (c.f., \cite{53, 54}). Using such a parametrization, one can obtain its Fourier transform $\tilde{q}(x, P_2, P_2^R)$. Since, the matching is only up to $O(\alpha_s)$, 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 space qPDF matrix elements $h^R(z, P_2, P_2^R; a_1, \ldots, a_n)$. 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 lattice cross-section approach in \cite{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 family of

---

3 In \cite{8}, the terms $F_1$ and $F_2$ are referred to as $f_1$ and $f_2$, respectively.
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., [53]). 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:

$$f^v_{0}(x; a, b) = Ax^a(1-x)^b,$$

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 $\int_0^1 f^v_{0}(x)dx = 1$. Instead, we use a more conservative constraint on $A$ using $\int_0^1 f^v_{0}(x)dx = h^v_{0}(z = 0, P_z, P^b) \approx 1$ and fold this into the error estimate. It should be noted that the valence PDF of pion determined from the experimental data by the JAM collaboration [56] 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

$$f_{u-d}(x; a, b) = \frac{1}{2} \left( f^v_{0}(x; a, b) + f^v_{0}(-x; a, b) \right),$$

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 $h_{\alpha}(z; a, b) = \int_{-\infty}^{\infty} \tilde{q}(x; a, b) e^{iP_z x} dx$. We will refer to these functions $h_{\alpha}(z; a, b)$ as the two-parameter family of phenomenologically motivated qPDF matrix elements. With the set of $h_{\alpha}(z; a, b)$ from a range of $a$ and $b$, we can fit the parameters $a$ and $b$ to the data by minimizing either $\chi^2_{r}$ or $\chi^2_{i}$ below:

$$\chi^2_{r} = \sum_{z=x_{\text{min}}}^{x_{\text{max}}} \frac{\left( \text{Re}(h^R_{\alpha}(z)) - \text{Re}(h_{\alpha}(z; a, b)) \right)^2}{\sigma_i(z)^2 + \sigma_{\text{pert}}^2(z)^2},$$

$$\chi^2_{i} = \sum_{z=x_{\text{min}}}^{x_{\text{max}}} \frac{\left( \text{Im}(h^R_{\alpha}(z)) - \text{Im}(h_{\alpha}(z; a, b)) \right)^2}{\sigma_i(z)^2 + \sigma_{\text{pert}}^2(z)^2},$$

$$\chi^2_{ri} = \chi^2_{r} + \chi^2_{i}.$$  

In the above equations, $[-x_{\text{max}}, x_{\text{max}}]$ specifies the fit range. The statistical errors on the real and imaginary parts of the lattice data $h^R_{\alpha}(z)$ is $\sigma_i(z)$ and $\sigma_i(z)$ respectively. To account for systematic errors coming from higher order corrections in $\alpha_s$ in the matching kernel, we determine $h_{\alpha}(z; a, b)$ from $f^v_{0}(x; a, b)$ by varying the value of $\alpha_s$ in matching kernel from $\alpha_s(\mu/2)$ to $\alpha_s(2\mu)$ though the 1-loop running. The corresponding change in real and imaginary parts of $h_{\alpha}(z; a, b)$ is denoted as $\sigma_{\text{pert}}^i(z)$ and $\sigma_{\text{pert}}^r(z)$ respectively, and we include this uncertainty in the match the result in the $\chi^2$. If the matching was exact, then by fitting only the real part by minimizing $\chi^2_{r}$ would automatically guarantee that the imaginary part also agree with the
data. Therefore at any finite order matching, the fits obtained by minimizing $\chi^2_z$ and $\chi^2_{\pi}$ will in general be different. For the results shown below, we used $\chi^2_z$ in order to obtain the PDF that best describes both the real and imaginary parts of the real space qPDF, but we also used $\chi^2_{\pi}$ 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 $\sigma_{pert}$ from non-statistical 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 bootstrap analysis.

In Fig. 15, we show the fitting procedure for $\gamma_z$ 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 $z_{max} = 1.44$ and 0.72 fm respectively. The agreement with both real and imaginary parts of the lattice data is noteworthy. In fact, we find the qPDF matrix element as inferred from the JAM PDF\cite{55} is able to explain the lattice data well for the entire range of $z$ at the two largest momentum. 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 $f_{\pi}(x)$ (shown as dashed lines) while we have shown the error bands for the qPDF $\tilde{q}_{\pi}(x)$ as obtained through 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 it can be seen, we started from a symmetric $u-d$ PDF by construction and matching introduces an $x \rightarrow -x$ 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 non-perturbative 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 $\chi^2_{\pi}$, we found estimates to be consistent with the above, but with larger uncertainties.

In Fig. 16, we show our results for $f_{\pi}(x,\mu)$ and $x f_{\pi}(x,\mu)$ at the factorization scale $\mu = 3.2$ GeV using the procedure described above at our two largest pion momenta $P_z = 1.29$ GeV and 1.72 GeV starting from $\gamma_z$ qPDF. For each case, we overlay results from two different RI-MOM scales $P^R$ in or-
FIG. 17. The 1-σ 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 are 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 \( \rho_R \). The black point is the JAM value \([56]\) for valence pion PDF. The black straight line is the line of constant first moment of valence PDF, \( \langle x \rangle = 0.215 \).

Consider to show the scatter as a systematic error in our estimates. We find the \( P^b \) dependence to be minor compared to the error bands (we repeated the analysis with multiple other values of \( P^b \) that are not shown and only minor scatter with respect to \( P^b \) was seen). We also show the result from the JAM collaboration \([56]\) 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 \( f_\pi^b(x) \), this overall agreement can be seen even up to smaller \( x \), but one has to be cautious of this agreement for \( x \lesssim \Lambda_{QCD}/P_z \approx 0.2 \) for the two highest pion momenta we use. By construction, in our fitting procedure \( f_\pi^b(x) \) 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 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-σ 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 \( P^b \) are distinguished by the colors, with the color code being the same as in Fig. 16. The \( P_z = 1.29 \) GeV data offers a stronger constraint on 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 is well within the \( P_z = 1.72 \) GeV ellipses while the \( P_z = 1.29 \) GeV data seems to favor slightly smaller exponent \( b \). However, these differences are well within 2σ. Even though our lattice data has 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, \( \langle x \rangle = 0.215 \) 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 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 \([57, 61]\) and similar values for the first moment of the pion PDF were obtained, but at slightly different values of \( \mu^2 \) 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 matching implemented at tree-level. Here, the exponents were estimated as \( a = -0.34(31) \) and \( b = 1.93(68) \) which are consistent with 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 \([35]\) and from the Dyson-Schwinger equation \([62–64]\) that the value of exponent \( b \) could be ap-
proximately 2 as expected from perturbative counting rule (c.f., [65]). It will be interesting to see if a similar implementation of an improved matching kernel could lead to a softer behavior for the pion than what is observed using 1-loop qPDF matching here and perhaps in [13]. In fact, a general consideration of power corrections to qPDF [66] revealed the presence of the form $\Lambda_{QCD}^2 / \left((1-x)^2 P_z^2\right)$ 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 were obtained in 2d QCD also [67].

Due to the larger errors in the $\gamma_z$ qPDF attributed mostly to the steep excited state extrapolations, we use the $\gamma_z$ qPDF to provide a consistency check of our calculations instead. For this, we use our above best estimates of the PDF obtained using the $\gamma_z$ qPDF to get the corresponding prediction for the real space $\gamma_z$ 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 $\gamma_z$ 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 $\gamma_z$ qPDF at both the pion momenta with a slight tension between the imaginary parts. From our discussion on excited state contamination, it is important to first gain better control of the larger excited state contamination in $\gamma_z$ 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 small lattice spacing $a = 0.06$ fm. We carefully examined the effects of excited states using 2 and 3 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 $P_0 = 1.29$ GeV and $P_0 = 1.72$ GeV. We found the qPDF defined using $\Gamma = \gamma_z$ was better determined compared to the $\gamma_z$ qPDF in the lattice calculation for the following reasons: smaller statistical error, relatively smaller excited state extrapolation leading to more robust result for the matrix element as well as due to the absence of mixing. Therefore, we focused on the analysis of the $\gamma_z$ matrix element.

The pion qPDF was non-perturbatively renormalized using RI-MOM scheme by calculating the matrix elements of qPDF operator with off-shell quark states in Landau gauge for different separations $z$. For these calculations we also used finer lattices with lattice spacing $a = 0.04$ fm. We performed the comparison of this matrix element in Landau gauge with 1-loop perturbative calculations in RI-MOM scheme and found 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 non-perturbative 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 asymmetry in the iso-vector pion qPDF $\bar{q}(x, P_0, p^2, \mu_R)$ 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 1-loop perturbative matching of the $\gamma_z$ qPDF, which we implemented through a fit to 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 [55] 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 $\gamma_z$ 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 $x$.

ACKNOWLEDGMENTS

We thank P.C. Barry for sharing the JAM valence PDF data, and J. H. Weber for providing us the Wilson line data. We thank R. S. Sufian and Y. Zhao for fruitful discussions. This material is based upon work supported by: (i) The U.S. Department of Energy, Office of Science, Office of Nuclear Physics and High Energy Physics through the Contract No. de-sc0012704, (ii) The U.S. Department of Energy, Office of Science, Office of Nuclear Physics and Office of Advanced Scientific Computing Research within the framework of Scientific Discovery through Advance Computing (SciDAC) award Computing the Properties of Matter with Leadership Computing Resources; (iii) The Brookhaven National Laboratory’s Laboratory Directed Research and Development (LDRD) project No. 16-37. S.S. is supported by the National Science Foundation under CAREER Award PHY-1847893. This research used awards of computer time provided by: (i) The INCITE program at Oak Ridge Leadership Computing Facility, a DOE Office of Science User Facility operated under Contract No. DE-AC05-00OR22725; (ii) The USQCD consortium at its Brookhaven National Laboratory and Jefferson Laboratory computing facilities. The gauge configurations used in this study were generated using awards of computing time provided by: (i) The INCITE program at Argonne Leadership Computing Facility, a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-06CH11357.
Appendix A: Coulomb-gauge Gaussian and momentum (boosted) smearing

In order to create hadron interpolating operators that have 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 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,

\[ S_{CG, \mathbf{x}, \mathbf{y}} = \mathcal{O}_C \mathbf{C}_x \mathbf{C}_y \mathbf{e}^{-\frac{(\mathbf{x} - \mathbf{y})^2}{2w_{CG}^2}}. \] (A1)

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

\[ \psi_x^C = \Omega_C \psi_x, \quad U_{x,\mu}^C = U_{x,\mu}^\Omega \Omega_C U_{x,\mu} \Omega^\dagger_C, \] (A2)

where \(\Omega_C\) is the gauge transformation to the Coulomb gauge, which minimizes the functional

\[ \Omega_C = \Omega : \quad \min_{\Omega_C} F_{C}[U^\Omega] = \min_{\Omega_C} \sum_x \sum_{\mu, \nu} \left[ -\text{Re} \text{Tr} U_{x,\mu}^\Omega U_{x,\nu}^C \Omega^\dagger_C \right], \] (A3)

(for the Coulomb gauge, \(\mu \neq \nu\) 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

\[ [S_{CG} \psi]_x = \mathcal{O}_C \sum_{\beta} \mathcal{e}^{i \mathbf{p} \cdot \mathbf{b}} \mathcal{e}^{-\frac{1}{2} w_{CG}^{\mathbf{b}}^2} \left[ \frac{1}{V} \sum_{y} \mathcal{e}^{-i \mathbf{p} \cdot \mathbf{y}} \Omega_{\mathbf{y}}^{\mathbf{C}} \psi_{\mathbf{y}} \right], \] (A4)

which is accelerated with offloading matrix-matrix products to GPU.

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

\[ (S^{(\vec{\xi})} \psi)_{x,\mathbf{y}} = e^{i \vec{\xi} \cdot \mathbf{y}} [S^{(\vec{0})}]_{x,\mathbf{y}} e^{-i \vec{\xi} \cdot \mathbf{x}} \psi_{\mathbf{x}} = e^{i \vec{\xi} \cdot \mathbf{x}} [S^{(\vec{0})}]_{x,\mathbf{y}} \psi_{\mathbf{y}}. \] (A5)

In a periodic finite volume, care must be taken to avoid spatial discontinuities in the boosted smearing kernel \((\vec{\xi} \cdot \mathbf{y} = \mathbf{x})\) such discontinuities may arise because the optimal boosted smearing momentum \(\vec{k}\) typically does not conform to finite-volume momentum quantization \(\vec{k} = \frac{2\pi}{L} \). The phase factors \(e^{i \vec{\xi} \cdot \mathbf{x}}\) and \(e^{i \vec{\xi} \cdot \mathbf{y}}\) do not satisfy periodic boundary conditions. The solution is to define the smearing kernel in the momentum space as

\[ S^{(\vec{\xi})}_{x,\mathbf{y}} = \sum_{\beta} e^{i \mathbf{p} \cdot \mathbf{b}} e^{-\frac{1}{2} w^2_{CG} \mathbf{p} \cdot \mathbf{n} \mathbf{p} \cdot \mathbf{n}} \mathcal{e}^{i \mathbf{p} \cdot \mathbf{y}} \mathcal{e}^{-i \mathbf{p} \cdot \mathbf{x}}, \] (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 \(\pi^+\) meson

\[ [\pi^+]_{x,\mathbf{y}} = \bar{u}_x \Gamma_{\pi^+} \psi_{\mathbf{y}}, \] (B1)

which is constructed from smeared quark fields

\[ \bar{d}_x \Gamma_{\pi^+} \bar{u}_x = \bar{u}_x \psi_{\mathbf{x}} \psi_{\mathbf{y}} \Gamma_{\pi^+} \bar{d}_y, \] (B2)

where the spinor matrix \(\Gamma_{\pi^+} = \gamma_5\). The Hermitian-conjugated (creation) meson operator is

\[ [\pi^+]_{x,\mathbf{y}}^\dagger = \bar{u}_x^\dagger \Gamma_{\pi^+} \psi_{\mathbf{x}} \psi_{\mathbf{y}} \Gamma_{\pi^+} \bar{d}_y, \] (B3)

where \(\Gamma_{\pi^+} = \gamma_4 \gamma_5\). The meson two-point correlation function with boost-smeared source and sink and momentum projection at the sink is

\[ C_{2\pi}(\gamma_4, \mathbf{p}; \mathbf{x}) = \sum_y e^{-i \mathbf{p} \cdot (\mathbf{y} - \mathbf{x})} \langle \pi^+]_{x,\mathbf{y}}^\dagger \pi^+_{\mathbf{y},\mathbf{x}} \rangle = \sum_y e^{-i \mathbf{p} \cdot (\mathbf{y} - \mathbf{x})} \text{Tr} [S^{(\vec{\xi})}_{x,\mathbf{y}} \Omega_{\mathbf{x},\mathbf{y}}^\dagger \Gamma_{\pi^+} \psi_{\mathbf{y}} \psi_{\mathbf{x}}] \] (B4)

where \(\Omega_{\mathbf{x},\mathbf{y}}^\dagger = \langle \mathbf{q}_x, \tilde{q}_y \rangle\) and \(\Omega^{(\pm \vec{\xi})}_{\mathbf{x},\mathbf{y}} = S^{(\vec{\xi})}_{\mathbf{x}} \Omega^{(\pm \vec{\xi})}_{\mathbf{y}} S^{(\vec{\xi})}_{\mathbf{y}}\) 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

---

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.
propagators for $u$ and $d$ quarks are required to construct meson correlation functions “boosted” with the total momentum $(2\vec{k})$,

$$Q^{d(-\vec{k})}_{x,y} = S_{x,y}^{d(-\vec{k})} O^{d(-\vec{k})}_{x,y} S_{y,x}^{d(-\vec{k})} \propto e^{i\vec{k} \cdot \vec{r}} \,,$$

$$\bar{Q}^{d(-\vec{k})}_{x,y} = S_{x,y}^{d(-\vec{k})} \bar{O}^{d(-\vec{k})}_{x,y} S_{y,x}^{d(-\vec{k})} \propto e^{i\vec{k} \cdot \vec{r}} \,,$$  

(B5)

where “$\propto$” sign stands for additional coordinate dependence due to the boosting. The $d$-quark $x \rightarrow y$ propagator, as usual, is computed using $\gamma_5$-Hermiteicity of the Dirac operator,

$$\bar{Q}^{d(-\vec{k})}_{x,y} = \gamma_5 \left[ Q^{d(-\vec{k})}_{x,y} \right]^\dagger \gamma_5 = \gamma_5 \left[ S_{x,y}^{d(-\vec{k})} O^{d(-\vec{k})}_{x,y} S_{y,x}^{d(-\vec{k})} \right]^\dagger \gamma_5$$  

(B6)

where the sign of the boosting momentum is preserved due to the Hermiteicity of the (boosted) smearing operator $S_{x,y}$.

Repeating similar steps to for the meson three-point function with the insertion of the operator $[\bar{u} W \Gamma u]$, with arbitrary $\Gamma$-matrix and Wilson line $W_{z+,z+L}$ along path $L$, we get

$$C^{\Pi U}_{3pt}(y_4, \mathbf{p}; z_4, \mathbf{q}; x) = \sum_{\vec{z}} e^{-i\mathbf{p} \cdot \vec{z}} e^{i\mathbf{q} \cdot \vec{z}} \langle \prod_{\mathbf{r}+L} \left( \Gamma U \right)_{z+,z+L} \rangle_{\mathbf{r}} = \sum_{\vec{z}} e^{i\mathbf{q} \cdot \vec{z}} \mathrm{Tr} \left[ e^{i\mathbf{p} \cdot \vec{z}} \bar{Q}^\dagger_{z+,z+L} \Gamma F^\dagger_{z+,z+L} \right],$$  

(B7)

where the forward propagator $F^u = Q^u S^d(\vec{k})$ and the meson sink-sequential (backward) propagator $B^\Pi U(y_i,\mathbf{p})$ is defined as

$$B^\Pi U(y_i,\mathbf{p}) = \sum_{\vec{y}} e^{-i\mathbf{p} \cdot \vec{y}} \left( -\Gamma \right)_{\vec{y},\vec{z}} R^\Pi U(y_i,\mathbf{p}) R U(y_i,\mathbf{p})^\dagger,$$

which is also computed using the $\gamma_5$-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 $\bar{u} WT u$ operator in $\pi^+$. If one repeats the computation using $\bar{u} WT d$ operator, one will find the disconnected piece to be the same as the one in the full $\bar{u} WT u$ three point function and hence such quark line disconnected terms will cancel in the $\bar{u} WT u - \bar{u} WT d$ isospin nonsinglet operator that we are interested in. Below, we further explain as to why only the real part of the connected $\bar{u} WT u$ three point function for $\Gamma = \gamma_i, \gamma_5$ and 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 $\Gamma = \gamma_5$. The full expression for the $u - d$ qPDF three point function is

$$C^{\Pi U}_{3pt}(t, \mathbf{r}, L) = \sum_{\vec{y},\vec{z}} e^{-i\mathbf{p} \cdot \vec{z}} \mathrm{Tr} \left[ \left( -\gamma_5 \right)_{\vec{y},\vec{z}} \gamma_5 Q_{\vec{y},\vec{z}} W_{z+,z+L} \gamma_5 Q_{\vec{z},\vec{x}} \right] + \sum_{\vec{y},\vec{z}} e^{-i\mathbf{p} \cdot \vec{z}} \mathrm{Tr} \left[ \left( -\gamma_5 \right)_{\vec{y},\vec{z}} \gamma_5 Q_{\vec{y},\vec{z}} W_{z+,z+L} \gamma_5 Q_{\vec{z},\vec{x}} \right]^\dagger, \quad (C1)$$

where we do not make distinctions between $u$ 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 $x = (\vec{x}, x_4) \rightarrow x_p = (\vec{x}, x_4)$, followed by a spatial translation $x \rightarrow x + \mathbf{L}$ by making use of the transformation of the Dirac propagator to be $Q_{\vec{x},y} \rightarrow \gamma_5 Q_{\vec{x},y} y_i$ and $W_{z+,z+L} \rightarrow W_{z+,z+L}^y$ under parity. In this case, the $y_i$ from parity transformation for $Q$ commutes with $\Gamma = \gamma_5$. In other cases, one should take care of the $\pm$ factor. Thus $C^{\Pi U}_{3pt}$ becomes

$$C^{\Pi U}_{3pt}(t, \mathbf{r}, L) = \sum_{\vec{y},\vec{z}} \left( e^{-i\gamma_5 T_1} + e^{-i\gamma_5 T_2} \right),$$  

(C2)

and therefore proportional to the connected piece of $\bar{u} WT u$, 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 $\Gamma = \gamma_5$, while it is purely imaginary for $\Gamma = 1$.

**Appendix D: Relation between $P_c$ and $-P_c$ matrix elements**

In this appendix, we derive the relation between conjugates of the matrix elements $\langle E_{n'}, P_c | \mathcal{O}(t, \mathbf{r}) | E_{n}, P_c \rangle$ 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:

$$\sum_{\vec{z}} \langle E_{n'}, P_c | \bar{u} \gamma_5 \Gamma W_{z+,z+L} | E_{n}, P_c \rangle^\dagger = \sum_{\vec{z}} \langle E_{n}, P_c | \bar{u} \gamma_5 \Gamma W_{z+,z+L}^\dagger | E_{n'}, P_c \rangle,$$

(D1)

for $\Gamma = \gamma_i, \gamma_5, 1$. Using the parity operator $\Pi$, the right-hand-side of the above equation becomes

$$\sum_{\vec{z}} \langle E_{n}, P_c | \bar{u} \gamma_5 \Gamma W_{z+,z+L} \Pi | E_{n'}, P_c \rangle = \sum_{\vec{z}} \langle E_{n}, -P_c | \Pi \bar{u} \gamma_5 \Gamma W_{z+,z+L} | E_{n'}, P_c \rangle = \sum_{\vec{z}} \langle E_{n}, -P_c | \bar{u} \gamma_5 \Gamma W_{z+,z+L} \gamma_5 | E_{n'}, -P_c \rangle = \sum_{\vec{z}} \langle E_{n}, -P_c | \bar{u} \gamma_5 \gamma_5 \Gamma W_{z+,z+L} \gamma_5 | E_{n'}, -P_c \rangle.$$  

(D2)

Defining, $\gamma_i \Gamma \gamma_i = \Phi_T \Gamma$ with $\Phi_T = \pm 1$, we have the relation

$$\sum_{\vec{z}} \langle E_{n'}, P_c | \bar{u} \gamma_5 \Gamma W_{z+,z+L} | E_{n}, P_c \rangle^\dagger = \Phi_T \sum_{\vec{z}} \langle E_{n}, -P_c | \bar{u} \gamma_5 \Gamma W_{z+,z+L} | E_{n'}, -P_c \rangle,$$

(D3)

with $\Phi_T = 1$ for $\Gamma = \gamma_i, 1$ and $\Phi_T = -1$ for $\Gamma = \gamma_5$ with all $\Gamma$ being in the Minkowskian convention. Thus, we can average over $P_c$ and $-P_c$ data after taking care of appropriate factor of $\Phi_T$. 

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

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

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

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

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

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

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

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

$$\int_{-\infty}^{\infty} \frac{dy}{\left| y \right|} C_{\mu}^{(1)}(x/y, yP_z)q(y) =$$

$$\int_{\text{reg}} \frac{dy}{\left| y \right|} C_{\mu}^{(1)}(x/y, yP_z)q(y) - N(xP_z)q(x).$$  (F6)

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

$$\int_{-\infty}^{\infty} \frac{dy}{\left| y \right|} C_{\mu}^{(1)}(x/y, yP_z)q(y) =$$

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

However, care has to be taken in the numerical regularization of the above equation to be consistent with the one
in Eq. (33). That is, in the above equation, \( \int_{E_{reg}} dy \) 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 \( \int_{E_{reg}} dy' \) in the second integral in the right-hand side involves \( y' \in [-\Lambda x, -x] \cup [x, x(1+\epsilon)] \cup [x(1+\epsilon), x\Lambda] \).

**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 \( t \to \infty \) extrapolations of the three-point function to two-point function ratio \( R(t, \tau; z, P_z, \Gamma) \) for \( \Gamma = \gamma_1 \) and \( \gamma_2 \) at a specific intermediate value of \( P_z = 1.29 \) GeV. In Fig. 21 and 22 of this appendix, we show similar results at all \( P_z \) for \( \Gamma = \gamma_1 \) and \( \gamma_2 \) respectively using \( \Gamma(t,2,2) \).

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 \( e^{-E_{reg}(t-\tau)} \) in the denominator of Eq. (9) is justified when \( e^{-E_{reg}} \ll e^{-E_{reg}(L_t-\tau)} \). But, one should include the effect of boundary condition if the terms become comparable. For the largest source-sink separation \( t/a = 12 \) we use, the contribution from the wrapping-around term, \( e^{-E_{reg}(L_t-\tau)} \) relative to \( e^{-E_{reg}} \) for \( P_z = 0 \) is 2.7\%, whereas for higher \( P_z \), it is negligible e.g., for the smallest non-zero momentum \( P_z = 0.43 \) GeV, this effect is 0.2\%. Hence, we included the term \( e^{-E_{reg}(L_t-\tau)} \) 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 non-zero \( P_z \) 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 \( R(t, \tau; z, P_z, \gamma_1) \) 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 \( e^{-E_{reg}(L_t-\tau)} \) 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, \( C_{2pt}(t; P_z) - A_0e^{-E_{reg}(L_t-\tau)} \), 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 \( E_{reg} \) were obtained by the two-state fit as described in the main text. Now, the trend with increasing \( t \) is clearer.

![Figure 21](image-url)

**FIG. 21.** The ratio of the three point function to the two-point function, \( R(t, \tau; z, P_z, \gamma) \) for \( \Gamma = \gamma_1 \), is shown as function of \( \tau - t/2 \) for \( z/a = 0, 4 \) and 8 (from left to right) and \( P_z = 0, 0.43, 0.86, \) 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 $P_z = 0, 0.438, 0.86,$ 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 [6].

[1] J. C. Collins, D. E. Soper, and G. F. Sterman, Adv. Ser. Direct. High Energy Phys. 5, 1 (1989) arXiv:hep-ph/0409313 [hep-ph]
[2] D. E. Soper, Phys. Rev. D15, 1141 (1977)
[3] X. Ji, Phys. Rev. Lett. 110, 262002 (2013) arXiv:1305.1539 [hep-ph]
[4] X. Ji, Sci. China Phys. Mech. Astron. 57, 1407 (2014) arXiv:1404.6680 [hep-ph]
[5] X. Xiong, X. Ji, J.-H. Zhang, and Y. Zhao, Phys. Rev. D90, 044051 (2014) arXiv:1310.7471 [hep-ph]
[6] M. Constantinou and H. Panagopoulos, Phys. Rev. D96, 054506 (2017) arXiv:1705.11193 [hep-lat]
[7] Y.-S. Liu, J.-W. Chen, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, (2018), arXiv:1807.06566 [hep-lat]
G. Koutsou, B. Kostrzewa, F. Steffens, C. Urbach, and S. Zafeiropoulos, Phys. Rev. D99, 014508 (2019), arXiv:1810.09743 [hep-lat].

[62] T. Nguyen, A. Bashir, C. D. Roberts, and P. C. Tandy, Phys. Rev. C83, 062201 (2011), arXiv:1102.2448 [nucl-th].

[63] C. Chen, L. Chang, C. D. Roberts, S. Wan, and H.-S. Zong, Phys. Rev. D93, 074021 (2016), arXiv:1602.01502 [nucl-th].

[64] K. D. Bednar, I. C. Clot, and P. C. Tandy, (2018), arXiv:1811.12310 [nucl-th].

[65] S. J. Brodsky, Proceedings, 13th International Workshop on Deep Inelastic Scattering (DIS 2005): Madison, Wisconsin, 27 April-1 May 2005, AIP Conf. Proc. 792, 977 (2005).

[66] V. M. Braun, A. Vladimirov, and J.-H. Zhang, Phys. Rev. D99, 014013 (2019), arXiv:1810.00048 [hep-ph].

[67] X. Ji, Y. Liu, and I. Zahed, Phys. Rev. D99, 054008 (2019), arXiv:1807.07528 [hep-ph].

[68] S. Gusken, Nucl. Phys. Proc. Suppl. 17, 361 (1990).