

# Lattice QCD determination of the Collins-Soper kernel in the continuum and physical mass limits

Jin-Xin Tan,<sup>1,2</sup> Zhi-Chao Gong,<sup>1,2</sup> Jun Hua,<sup>3,4,\*</sup> Xiangdong Ji,<sup>5</sup> Xiangyu Jiang,<sup>6</sup> Hang Liu,<sup>7</sup> Andreas Schäfer,<sup>8,9</sup> Yushan Su,<sup>5</sup> Han-Zhang Wang,<sup>1,10</sup> Wei Wang,<sup>1,11,†</sup> Yi-Bo Yang,<sup>6,12,13,14</sup> Jun Zeng,<sup>15</sup> Jian-Hui Zhang,<sup>16</sup> Jia-Lu Zhang,<sup>2,1</sup> and Qi-An Zhang<sup>17,‡</sup>

<sup>1</sup>State Key Laboratory of Dark Matter Physics, Key Laboratory for Particle Astrophysics and Cosmology (MOE), Shanghai Key Laboratory for Particle Physics and Cosmology, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

<sup>2</sup>Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shanghai 201210, People's Republic of China

<sup>3</sup>State Key Laboratory of Nuclear Physics and Technology, Institute of Quantum Matter, South China Normal University, Guangzhou 510006, People's Republic of China

<sup>4</sup>Guangdong Basic Research Center of Excellence for Structure and Fundamental Interactions of Matter, Guangdong Provincial Key Laboratory of Nuclear Science, Guangzhou 510006, People's Republic of China

<sup>5</sup>Department of Physics, University of Maryland, College Park, Maryland 20742, USA

<sup>6</sup>CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, People's Republic of China

<sup>7</sup>Department of Physics, Shanghai Normal University, Shanghai 200234, People's Republic of China

<sup>8</sup>Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany

<sup>9</sup>Department of Physics, National Taiwan University, Taipei, Taiwan 106, People's Republic of China

<sup>10</sup>Zhiyuan College, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

<sup>11</sup>Southern Center for Nuclear-Science Theory (SCNT), Institute of Modern Physics, Chinese Academy of Sciences, Huizhou 516000, Guangdong Province, People's Republic of China

<sup>12</sup>School of Fundamental Physics and Mathematical Sciences, Hangzhou Institute for Advanced Study, UCAS, Hangzhou 310024, People's Republic of China

<sup>13</sup>International Centre for Theoretical Physics Asia-Pacific, Beijing/Hangzhou, People's Republic of China

<sup>14</sup>School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

<sup>15</sup>College of Physics and Electronic Engineering, Hainan Normal University, Haikou 571158, Hainan, People's Republic of China

<sup>16</sup>School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, People's Republic of China

<sup>17</sup>School of Physics, Beihang University, Beijing 102206, People's Republic of China



(Received 10 December 2025; accepted 6 February 2026; published 5 March 2026)

The Collins-Soper (CS) kernel governs the rapidity evolution of transverse-momentum-dependent (TMD) parton distributions, a cornerstone for QCD factorization and linking nucleon structure data across scales. Its nonperturbative behavior at large transverse separations ( $b_\perp$ ) remains weakly constrained due to phenomenological model dependencies. We present a first-principles determination of the CS kernel at the continuum limit and physical pion mass from lattice QCD in the large-momentum effective theory framework. Using  $(2 + 1)$ -flavor configurations (lattice spacings  $a \in [0.052, 0.105]$  fm, and pion mass  $m_\pi \approx (136, 230, 300, 320)$  MeV), we simulate the nonlocal equal-time correlation function and extract the quasi-TMD wave functions. Taking into account systematic improvements including hypercubic smearing, nonperturbative renormalization, and a  $b_\perp$ -unexpanded matching kernel, we obtain the CS kernel at the continuum, chiral, and infinite-momentum limits. Our results are determined up to  $b_\perp \sim 1$  fm, with

\*Contact author: junhua@scnu.edu.cn

†Contact author: wei.wang@sjtu.edu.cn

‡Contact author: zhangqa@buaa.edu.cn

controllable uncertainties, and agree with perturbative QCD at small  $b_\perp$  and global TMD phenomenological extractions. We conduct a global analysis integrated with phenomenological fits and demonstrate the impact of our results on such fits. This work yields the most precise nonperturbative constraint on the CS kernel's long-distance behavior from lattice QCD, which not only bridges lattice QCD, perturbation theory, and nucleon structure experiments for TMD studies, but also boosts the utility of our constraint for future global TMD analyses.

DOI: [10.1103/physrevd.113.054505](https://doi.org/10.1103/physrevd.113.054505)

## I. INTRODUCTION

Understanding the three-dimensional structure of the nucleon is a central objective of modern hadronic physics. While the collinear parton distribution functions (PDFs) provide crucial information on the longitudinal momentum fractions carried by quarks and gluons, they offer only a one-dimensional projection of the nucleon's internal structure. To achieve a more complete spatial and dynamical understanding, one must go beyond the collinear approximation and explore the transverse-momentum-dependent (TMD) regime. The TMD parton distribution functions (TMDPDFs) and TMD wave functions (TMDWFs) encode not only the longitudinal motion but also the intrinsic transverse dynamics of partons. In particular, TMDPDFs constitute essential theoretical ingredients for describing processes such as semi-inclusive deep-inelastic scattering (SIDIS) and Drell-Yan production, and TMDWFs play a key role in the study of exclusive production processes and hadron structures [1].

A central theoretical challenge in TMD physics lies in their rapidity evolution, which supplements the usual UV renormalization by connecting different rapidity scales. This evolution is governed by the Collins-Soper (CS) kernel [2–4], which encapsulates the interplay between soft and collinear dynamics. The CS kernel is perturbatively calculable at small transverse separations and has been determined at three loops and four loops in QCD [5–9], but becomes intrinsically nonperturbative at large  $b_\perp$ . Its precise determination is therefore crucial for achieving accurate TMD factorization and for linking data from different experiments.

Extensive experimental efforts over the past decades, including SIDIS at HERMES, COMPASS, and Jefferson Lab, as well as Drell-Yan and  $W/Z$  production at RHIC and the LHC, have provided rich information on TMD-sensitive observables [10–24]. Global phenomenological analyses [25–33] have extracted the CS kernel from these data, revealing consistent results in the perturbative region but leaving sizable uncertainties in the nonperturbative domain due to model dependencies and parametrization assumptions.

To overcome these limitations, a first-principles determination of the Collins-Soper kernel and the soft functions from lattice QCD has become an urgent task. However, direct lattice calculations of light-cone TMDs are prohibited by the presence of (two) lightlike Wilson lines, which

cannot be realized in Euclidean space. The large-momentum effective theory (LaMET) [34–37] provides a practical solution by establishing a systematic factorization between Euclidean and light-cone observables. In this framework, one considers Euclidean equal-time correlators of hadrons boosted to large longitudinal momentum, known as quasidistributions. These quantities can be directly computed on the lattice and are related to their light-cone counterparts through a perturbative matching at large  $P^z$ . Based on this framework, one can further utilize the dependence of quasidistributions on the momentum  $P^z$  to extract the CS kernel nonperturbatively [36,38,39]. Specifically, the variation of quasidistributions with respect to  $P^z$  encodes the same rapidity evolution that governs the TMDs in continuum QCD. By comparing results at different boosts, the Collins-Soper kernel can be determined from first principles. This approach thus provides a nonperturbative determination of the CS kernel and establishes a unified bridge connecting lattice observables with the phenomenological framework of TMD evolution.

Recent lattice studies [38,40–51] have convincingly demonstrated the feasibility of this strategy. By implementing improved operator constructions, refined renormalization schemes, and simulations at higher hadron momenta, these works have shown that the Collins-Soper kernel can be reliably extracted from lattice data with steadily increasing precision. The extracted CS kernels exhibit remarkable consistency with perturbative QCD predictions and phenomenological extractions in the overlapping kinematic regions. These developments highlight the growing predictive power of the lattice approach and its promising role in providing model-independent inputs for global analyses of TMD evolution.

However, existing calculations of the CS kernel were limited in precision, particularly at large transverse separations, where the signal rapidly deteriorates, making it difficult to access the long-range correlations relevant to nucleon-size scales. A recent study [50] performed the first continuum-extrapolated calculation at the physical pion mass using multiple lattice spacings in the range  $a \in [0.09, 0.15]$  fm. Their continuum extrapolation, however, relied on an explicit parametrization of the  $b_\perp$  dependence of the CS kernel, because discretization effects and power corrections could not be disentangled simultaneously at each  $b_\perp$  within the available data constraints.

In this work, we employ ensembles with multiple lattice spacings spanning  $a \in [0.052, 0.105]$  fm and multiple pion masses, including the physical one, together with substantially improved statistical precision. This enables controlled continuum and chiral extrapolations and allows us to determine the CS kernel from first principles up to transverse separations of  $b_\perp \simeq 1$  fm. Our results therefore provide, to date, the most precise determination of the long-distance CS kernel and the first continuum-extrapolated extraction achieved without introducing the explicit parametrization of its  $b_\perp$  dependence, thereby offering a robust nonperturbative input for future global TMD analyses.

The rest of this paper is organized as follows: in Sec. II, we present the theoretical formalism, including the definition of quasi-TMDWFs, their factorization properties, and the connection to the CS kernel. In Sec. III, we provide a detailed description of the lattice calculation of quasi-TMDWFs and the extraction of the CS kernel. We give our framework of the calculation from correlation functions to the bare quasi-TMDWFs and their subsequent renormalization. In addition, by comparing the renormalized quasi-TMDWFs, we demonstrate the significant improvement in the SNR achieved through hypercubic (HYP) smearing. Using results obtained from ensembles with different lattice parameters, we extract the CS kernel within the LaMET framework and perform extrapolations to the infinite momentum, continuum, and chiral limits to obtain the final physical results. Finally, in Sec. IV, we summarize our main results and discuss their implications for phenomenology, as well as prospects for future theoretical and experimental studies. Some details in the calculation and supplementary results are collected in the Appendixes.

## II. THEORETICAL FRAMEWORK

In this section, we present the theoretical framework underlying our calculation. We begin by defining the CS kernel and discussing its fundamental role in the rapidity evolution of transverse-momentum-dependent observables. We then introduce the quasi-TMDWFs within the LaMET framework and illustrate their factorization properties, highlighting how they enable the extraction of the CS kernel from lattice QCD.

### A. Collins-Soper kernel and rapidity evolution

The CS kernel is introduced to characterize the rapidity evolution of TMD distributions. Rapidity divergences arise from gluon radiation that is collinear to the lightlike gauge links, leading to an unbounded growth in the large-rapidity region and necessitating additional parameters and mechanisms to control this dependence. Within the TMD framework, the CS kernel  $K(b_\perp, \mu)$  functions as the rapidity anomalous dimension, governing the evolution of TMDs across different rapidity scales, and is defined as [2–4,52]

$$K(b_\perp, \mu) = 2\zeta \frac{d}{d\zeta} \ln f^{\text{TMD}}(x, b_\perp, \mu, \zeta), \quad (1)$$

where  $f^{\text{TMD}}(x, b_\perp, \mu, \zeta)$  denotes physical TMD distribution with longitudinal momentum fraction  $x$  and transverse displacement  $b_\perp$ , at rapidity scale  $\zeta$ , and renormalization scale  $\mu$ .

The CS kernel governs the universal rapidity evolution of transverse-momentum-dependent distributions, connecting the soft and collinear sectors in QCD. At small transverse separations  $b_\perp$ , it can be reliably computed perturbatively and is currently known at three and four loops [5–9]. In contrast, at large  $b_\perp$ , the CS kernel encodes essential nonperturbative physics that cannot be accessed perturbatively. A precise determination of  $K(b_\perp, \mu)$  is therefore crucial, both for relating TMD observables measured at different energy scales and for achieving high-precision resummation in QCD phenomenology.

The scale dependence of the CS kernel can be described by the renormalization group equation (RGE) [2,4,53–55]

$$\frac{d}{d \ln \mu} K(b_\perp, \mu) = -2\Gamma_{\text{cusp}}(\alpha_s), \quad (2)$$

where  $\Gamma_{\text{cusp}}$  denotes the cusp anomalous dimension. This equation reflects that the UV behavior of the CS kernel is governed by a universal quantity,  $\Gamma_{\text{cusp}}$ , which also appears in the evolution of Wilson lines and other QCD observables. Physically, the CS kernel characterizes how TMD distributions evolve with changes in the rapidity scale, providing a universal evolution mechanism that connects TMDs defined at different rapidity regimes and ensuring the consistency of QCD factorization across energy scales.

### B. Factorization in LaMET

Within the LaMET framework, the IR behavior of quasisdistributions can be systematically factorized and matched to their corresponding light-front quantities. In particular, the unsubtracted TMDs,  $f(x, b_\perp, \mu, \zeta)$ , together with the intrinsic soft function  $S_I(b_\perp, \mu)$ , can be factorized in terms of the quasi-TMDs computed at finite hadron momentum on the lattice [56–60], as follows:

$$\begin{aligned} \frac{f(x, b_\perp, \mu, \zeta)}{\sqrt{S_I(b_\perp, \mu)}} &= \exp \left[ -\frac{1}{2} K(b_\perp, \mu) \ln \left( \frac{-\zeta_z + i\epsilon}{\zeta} \right) \right] \\ &\times H(\zeta_z \bar{\zeta}_z, b_\perp, \mu) \tilde{f}(x, b_\perp, \mu, \zeta_z) \\ &+ \mathcal{O} \left( \frac{\Lambda_{\text{QCD}}^2}{\zeta_z}, \frac{1}{b_\perp^2 \zeta_z} \right), \end{aligned} \quad (3)$$

where  $\zeta = (2xP^+)^2$  and  $\zeta_z = (2xP^z)^2$  denote the rapidity scales of the TMDs and quasi-TMDs, respectively. Their evolution is connected through the exponential factor involving the CS kernel  $K(b_\perp, \mu)$ .  $H(\zeta_z, \bar{\zeta}_z, b_\perp, \mu)$  denotes

the perturbative matching kernel, while the terms  $\mathcal{O}(\Lambda_{\text{QCD}}^2/\zeta_z, 1/(b_\perp^2 \zeta_z))$  represent higher-twist corrections arising from finite-momentum and nonperturbative effects, which vanish in the large- $P^z$  limit.

Equation (3) thus provides the theoretical foundation for extracting the CS kernel from lattice QCD. By computing the quasi-TMDs at multiple hadron momenta and investigating their dependence on  $\zeta_z$ , one can eliminate the  $\zeta_z$ -independent factors,  $f(x, b_\perp, \mu, \zeta)$  and  $S_I(b_\perp, \mu)$ , from the ratio of quasi-TMDs at different  $\zeta_z$ . The CS kernel can then be extracted as

$$\begin{aligned} K(b_\perp, \mu, x, P_1^z, P_2^z) &= \frac{1}{\ln(P_1^z/P_2^z)} \left[ \ln \frac{H(\zeta_{z,2}, \bar{\zeta}_{z,2}, b_\perp, \mu) \tilde{f}(x, b_\perp, \mu, \zeta_{z,1})}{H(\zeta_{z,1}, \bar{\zeta}_{z,1}, b_\perp, \mu) \tilde{f}(x, b_\perp, \mu, \zeta_{z,2})} \right] \\ &+ \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}^2}{\zeta_z}, \frac{1}{b_\perp^2 \zeta_z}\right), \end{aligned} \quad (4)$$

with condition  $P_1^z \neq P_2^z \gg 1/b_\perp$ . At large  $P^z$ , power corrections are suppressed by powers of  $1/(P^z)^2$ , as will be discussed in Sec. III C.

It should be noted that the rapidity evolution of TMDs is universally valid for both TMDPDFs and TMDWFs. Consequently, Eq. (4) can be applied to extract the CS kernel either from the quasibeam function or from the unsubtracted quasi-TMDWF. Previous studies have shown that the quasi-TMDWF, which extracted from two-point correlation functions [41,44–46,48,50,51], exhibits a significantly better SNR than the quasibeam function obtained from three-point functions [40,42,43,47]. Therefore, in this work, we choose to determine the CS kernel using the pion quasi-TMDWF.

Furthermore, the factorization structure of the quasi-TMDWFs can be visualized through the leading-power reduced diagram shown in Fig. 1. In this picture, the quasi-TMDWF is decomposed into distinct collinear (C), soft (S), and hard (H) subdiagrams, each corresponding to a specific momentum region in QCD. In contrast to processes such as SIDIS or Drell-Yan, the hard-gluon exchange between the quark and antiquark inside a meson is power suppressed at

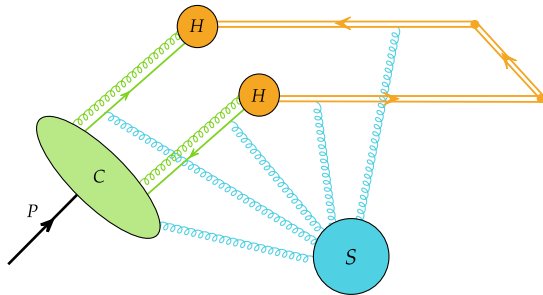


FIG. 1. Leading-power reduced graph for pseudoscalar meson quasi-TMDWFs.

large momentum. Consequently, the two hard subdiagrams become disconnected, and the quasi-TMDWF amplitude factorizes multiplicatively into the hard, collinear, and soft components [56,61]. This feature naturally leads to the multiplicative structure of the LaMET factorization formula given in Eq. (3).

### C. Definition of quasi-TMD wave functions

The quasi-TMDWF for a meson boosted to a large momentum  $P^z$  along the  $z$ -direction is defined as

$$\begin{aligned} \tilde{f}(x, b_\perp, \mu, \zeta_z) &= \lim_{L \rightarrow \infty} \int \frac{dz}{2\pi} e^{i(x-\frac{1}{2})zP^z} \\ &\times \frac{\tilde{\Phi}^0(z, b_\perp, P^z, L)}{Z_O(1/a, \mu) \sqrt{Z_E(2L+z, b_\perp)}}, \end{aligned} \quad (5)$$

where the bare matrix element  $\tilde{\Phi}^0(z, b_\perp, P^z, L)$  is constructed from a gauge-invariant nonlocal operator sandwiched between the vacuum and a highly boosted meson state [45,46,62]

$$\begin{aligned} \tilde{\Phi}^0(z, b_\perp, P^z, L) &= \langle 0 | \bar{q}(z\hat{n}_z/2 + b_\perp \hat{n}_\perp) \Gamma \\ &\times U_\square(L, z, b_\perp) q(-z\hat{n}_z/2) | P^z \rangle, \end{aligned} \quad (6)$$

where  $U_\square$  denotes the staple-shaped Wilson link connecting the quark and antiquark fields, consisting of both longitudinal ( $z$ ) and transverse ( $b_\perp$ ) spatial separations on the Euclidean lattice

$$\begin{aligned} U_\square(L, z, b_\perp) &\equiv U_z^\dagger \left[ \frac{z}{2} \hat{n}_z + b_\perp \hat{n}_\perp; \left( L + \frac{z}{2} \right) \hat{n}_z + b_\perp \hat{n}_\perp \right] \\ &\times U_\perp \left[ \left( L + \frac{z}{2} \right) \hat{n}_z + b_\perp \hat{n}_\perp; \left( L + \frac{z}{2} \right) \hat{n}_z \right] \\ &\times U_z \left[ \left( L + \frac{z}{2} \right) \hat{n}_z; -\frac{z}{2} \hat{n}_z \right], \end{aligned} \quad (7)$$

with the Euclidean gauge link along the  $\mu$ -direction is defined as

$$U_\mu(\vec{\xi} + \lambda \hat{n}_\mu; \vec{\xi}) \equiv \mathcal{P} \exp \left[ -ig \int_0^1 dt \lambda \hat{n} \cdot A(\vec{\xi} + t \lambda \hat{n}_\mu) \right]. \quad (8)$$

For a highly boosted pseudoscalar meson state, we employ an interpolating operator with the Dirac structure  $\Gamma = \gamma' \gamma_5$  to construct the quasi-TMDWFs. At sufficiently large hadron momenta, the leading-twist component of the boosted correlator is dominated by the  $\gamma^+ \gamma_5$  structure, which has a strong overlap with  $\gamma' \gamma_5$  in the large-momentum limit [63]. In our simulations, the pion momenta reach up to  $P^z \sim 10m_\pi$  on each ensemble, corresponding to a

sizable Lorentz boost, ensuring that the  $\gamma'\gamma_5$  projection accurately captures the leading-twist contribution.

The linear divergence in the bare matrix element  $\tilde{\Phi}^0$  originates from the self-energy correction of the Wilson link and can be removed by dividing by a Wilson link with identical geometry. To preserve gauge symmetry, we introduce a rectangular Wilson loop that is exactly formed by gluing together two identical staple-shaped Wilson links into a rectangle of longitudinal extent  $2L + z$  and transverse width  $b_\perp$ . We define

$$\begin{aligned} Z_E(2L + z, b_\perp) &= \frac{1}{N_c} \left\langle U_z \left[ \left( L + \frac{z}{2} \right) \hat{n}_z; \left( -L - \frac{z}{2} \right) \hat{n}_z \right] \right. \\ &\quad \times U_\perp^\dagger \left[ \left( -L - \frac{z}{2} \right) \hat{n}_z + b_\perp \hat{n}_\perp; \left( -L - \frac{z}{2} \right) \hat{n}_z \right] \\ &\quad \times U_z^\dagger \left[ \left( L + \frac{z}{2} \right) \hat{n}_z + b_\perp \hat{n}_\perp; \left( -L - \frac{z}{2} \right) \hat{n}_z + b_\perp \hat{n}_\perp \right] \\ &\quad \left. \times U_\perp \left[ \left( L + \frac{z}{2} \right) \hat{n}_z + b_\perp \hat{n}_\perp; \left( L + \frac{z}{2} \right) \hat{n}_z \right] \right\rangle. \end{aligned} \quad (9)$$

By construction, the square root  $\sqrt{Z_E}$  furnishes precisely the staple-shaped Wilson-line renormalization factor and thus cancels the linear divergence of  $\tilde{\Phi}^0$ .

Besides, the logarithmic divergence arises from vertex corrections involving the Wilson line and the light-quark field. We introduce the corresponding renormalization factor  $Z_O$ , which also serves to connect the lattice regularization used for quasi-TMDWFs to the perturbative  $\overline{\text{MS}}$  scheme. The parametrization of  $Z_O$  follows Refs. [62,64]:

$$\begin{aligned} \ln Z_O(1/a, \mu) &= \frac{\gamma_0}{\beta_0} \ln[\ln[1/(a\Lambda_{\text{QCD}})]] \\ &\quad + \frac{c_1}{\ln[1/(a\Lambda_{\text{QCD}})]} + d'(\mu), \end{aligned} \quad (10)$$

where  $a$  denotes the lattice spacing. The first two terms capture the characteristic logarithmic behavior, with  $\gamma_0$  the anomalous dimension of the operator and  $\beta_0$  the QCD  $\beta$ -function coefficient, and the constant  $c_1$  is determined from lattice data. All explicit  $\mu$ -dependence is absorbed into the finite term  $d'(\mu)$ . Although  $Z_O$  cancels in the extraction of

the CS kernel, we still determine it in order to obtain renormalized quasi-TMDWFs and to enable meaningful comparisons under different kinematic and lattice conditions.

### III. LATTICE SIMULATION AND RESULT

#### A. Lattice setup

The numerical simulations in this work are performed using the gauge configurations generated by the China Lattice QCD (CLQCD) Collaboration, employing  $N_f = 2 + 1$  flavor stout smeared clover fermions and the Symanzik gauge action [65]. One step of stout link smearing is applied to the gauge field used in the clover action to enhance the stability of the pion mass for a given bare quark mass. We utilize multiple ensembles with lattice spacings  $a = \{0.05187, 0.07746, 0.1053\}$  fm and pion masses  $m_\pi \simeq \{136, 230, 300, 320\}$  MeV. To explore the large-momentum limit and investigate the momentum dependence, we employ three different momenta,  $P^z \in \{4, 5, 6\} \times 2\pi/(n_s a)$  GeV for each ensemble. The detailed simulation parameters are listed in Table I.

To calculate the nonlocal matrix elements in Eq. (6), we generate the wall-source propagator

$$S_w(\vec{x}, t; t'; \vec{p}) = \sum_{\vec{y}} S(t, \vec{x}; t', \vec{y}) e^{i\vec{p} \cdot (\vec{y} - \vec{x})}, \quad (11)$$

on the Coulomb gauge-fixed configurations. This propagator starts from the source at  $t'$ , sums over all spatial positions  $\vec{y}$ , and propagates to the sink at  $(\vec{x}, t)$ , carrying momentum  $\vec{p}$ .

By extracting the ground-state energies  $E(P^z)$  at different momenta, we determine the pion dispersion relation to ensure that lattice discretization errors are under control. The dispersion relation is parametrized as

$$E(P^z) = \sqrt{m_\pi^2 + b_1(P^z)^2 + b_2(P^z)^4 a^2}, \quad (12)$$

where  $m_\pi$  denotes the pion mass,  $b_1$  accounts for deviations from continuum relativistic kinematics, and the term  $(P^z)^4 a^2$  parametrizes discretization effects. Figure 2 shows the pion dispersion relation across the five ensembles. The joint fit gives results of  $m_\pi = 136.0(7.2)$  MeV, consistent with the physical pion mass. The value  $b_1 = 1.0034(83)$  is in

TABLE I. Details of the parameters used for the calculations on each ensembles. The number of gauge configurations is denoted as  $N_{\text{cfg}}$ , and for each configuration,  $N_{\text{meas}}$  measurements are performed.

| Ensemble | $a$ (fm) | $n_s^3 \times n_t$ | $m_\pi$ (MeV) | $P^z$ (GeV)      | $N_{\text{cfg}} \times N_{\text{meas}}$ |
|----------|----------|--------------------|---------------|------------------|-----------------------------------------|
| C32P29   | 0.10530  | $32^3 \times 64$   | 292.4         | 1.47, 1.84, 2.21 | $984 \times 4$                          |
| C32P23   | 0.10530  | $32^3 \times 64$   | 228.0         | 1.47, 1.84, 2.21 | $448 \times 10$                         |
| C48P14   | 0.10530  | $48^3 \times 96$   | 135.5         | 0.98, 1.22, 1.47 | $204 \times 24$                         |
| F32P30   | 0.07746  | $32^3 \times 96$   | 303.2         | 2.00, 2.50, 3.00 | $1153 \times 6$                         |
| H48P32   | 0.05187  | $48^3 \times 144$  | 317.2         | 1.99, 2.49, 2.99 | $550 \times 8$                          |

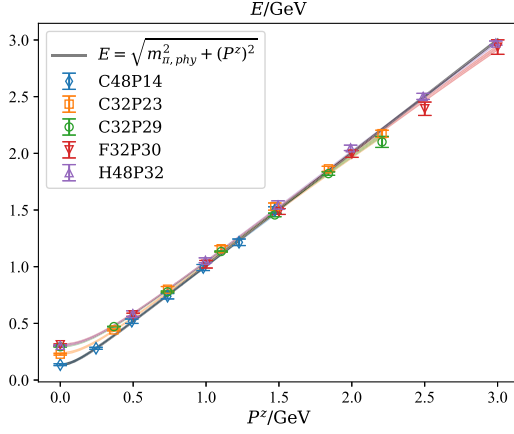


FIG. 2. Dispersion relation of the pion state. The data up to 3 GeV can be described with the formula in Eq. (12) with  $m_\pi = 136.0(7.2)$ ,  $b_1 = 1.0034(83)$ , and  $b_2 = -0.042(14)$ .

agreement with the square of the speed of light ( $c^2 = 1$  at the natural units) within the error range, and  $b_2 = -0.042(14)$  indicates small discretization effects. The fit quality is characterized by  $\chi^2 = 0.94$ .

To improve the statistical quality of the Wilson loop and nonlocal correlation functions, we apply the HYP smearing technique [66]. This method modifies the gauge links by replacing each original link with a new link constructed from its surrounding staple-shaped neighbors. Through iterative averaging of the gauge fields within local hypercubic neighborhoods, HYP smearing effectively suppresses short-distance UV fluctuations while preserving the long-distance physics and maintaining gauge invariance. As a result, the technique significantly enhances the SNR, and stabilizes operator renormalization and matching. It is worth noting that, as discussed later, the renormalized matrix elements are found to be insensitive to the number of HYP smearing steps applied.

## B. Quasi-TMD wave functions

### 1. Correlation functions

We construct the two-point correlation function corresponding to the bare matrix elements of the quasi-TMDWF as

$$C_2(z, b_\perp, P, p^z, L, t, t') = \frac{1}{n_s} \sum_{\vec{x}} e^{i\vec{P} \cdot \vec{x}} \text{Tr} \langle S_w(t'; \vec{x}_1, t; -p^z) \times U_\square(L, z, b_\perp) \Gamma_1 S_w(\vec{x}_2, t; t'; p^z) \Gamma_2 \rangle, \quad (13)$$

where  $\vec{x}_1 = \vec{x} + L\hat{n}_z + b\hat{n}_\perp$  and  $\vec{x}_2 = \vec{x} + (L+z)\hat{n}_z$ . The quark and antiquark are assigned momenta  $p^z$  and  $-p^z$  along the  $z$ -direction, respectively, and  $P$  denotes the hadron momentum, which satisfies  $P = (0, 0, 0, 2p^z)$ .

Since the Dirac structure of pseudoscalar meson interpolators at the source does not influence the matrix elements, we choose  $\Gamma_2 = (\gamma^t \gamma_5 + \gamma^z \gamma_5)/2$  [63]. At the sink, as mentioned earlier, we select  $\Gamma_1 = \gamma^t \gamma_5$ , which will projects onto the leading-twist light-cone operator  $\gamma^+ \gamma_5$  in the large-momentum limit. Employing the reduction formula, the parametrized form of the two-point correlation function can be expressed as

$$C_2(z, b_\perp, P^z, L, t) = \frac{A_w(P^z) A_p}{2E} \tilde{\Phi}^0(z, b_\perp, P^z, L) \times e^{-Et} [1 + c_0(z, b_\perp, P^z, L) e^{-\Delta Et}], \quad (14)$$

where  $A_w(P^z)$  and  $A_p$  denote the local matrix element of the meson interpolator from wall source and point sink, respectively, and these factors can be canceled by the local two-point correlation function from the ratio

$$R^\pm(z, b_\perp, P^z, t; L) \equiv \frac{C_2(z, b_\perp, P^z, L, t)}{C_2(0, 0, P^z, L, t)} = \tilde{\Phi}^{\pm 0}(z, b_\perp, P^z, L) \times [1 + c'(z, b_\perp, P^z, L) e^{-\Delta Et}], \quad (15)$$

in which the superscript  $\pm$  in  $R^\pm$  and  $\tilde{\Phi}^{\pm 0}$  denote the  $z$ -direction of the staple-shape Wilson link, and  $c'$  term denotes the contribution from excited states.

In this work, we calculate the nonlocal two-point correlation functions on all five configurations with the  $P^z$  settings listed in Table I. As shown in Eq. (15), we extract the ground-state matrix elements from the ratio of the nonlocal and local two-point correlation functions. In practice, we compare the results of the one-state fit and the two-state fit, as shown in Fig. 3. Typically, the one-state fit is applied when the ratio has reached a well-defined plateau region, while the two-state fit is used when residual excited-state contamination remains visible within the plateau. From the cases shown in Fig. 3, the results from the one-state and two-state fits are in good agreement. To conservatively estimate the statistical errors of the fit results, we adopt the two-state fit results for the subsequent analysis.

In the definition of physical TMDs in Minkowski space, the longitudinal direction of the staple-shaped Wilson link is nontrivial. The past-pointing and future-pointing link directions correspond to different physical processes, and they are not easily related through simple symmetries. However, in quasi-TMDs, the staple-shaped Wilson link is independent of time, and the  $O(3)$  spatial symmetry and isospin symmetry ensure the symmetry of the results when the staple-shaped Wilson link is oriented along different longitudinal directions ( $+z$  or  $-z$ ). This symmetry was observed in Refs. [45,46], where it was used to improve their analysis. In Appendix C, we provide a proof of this

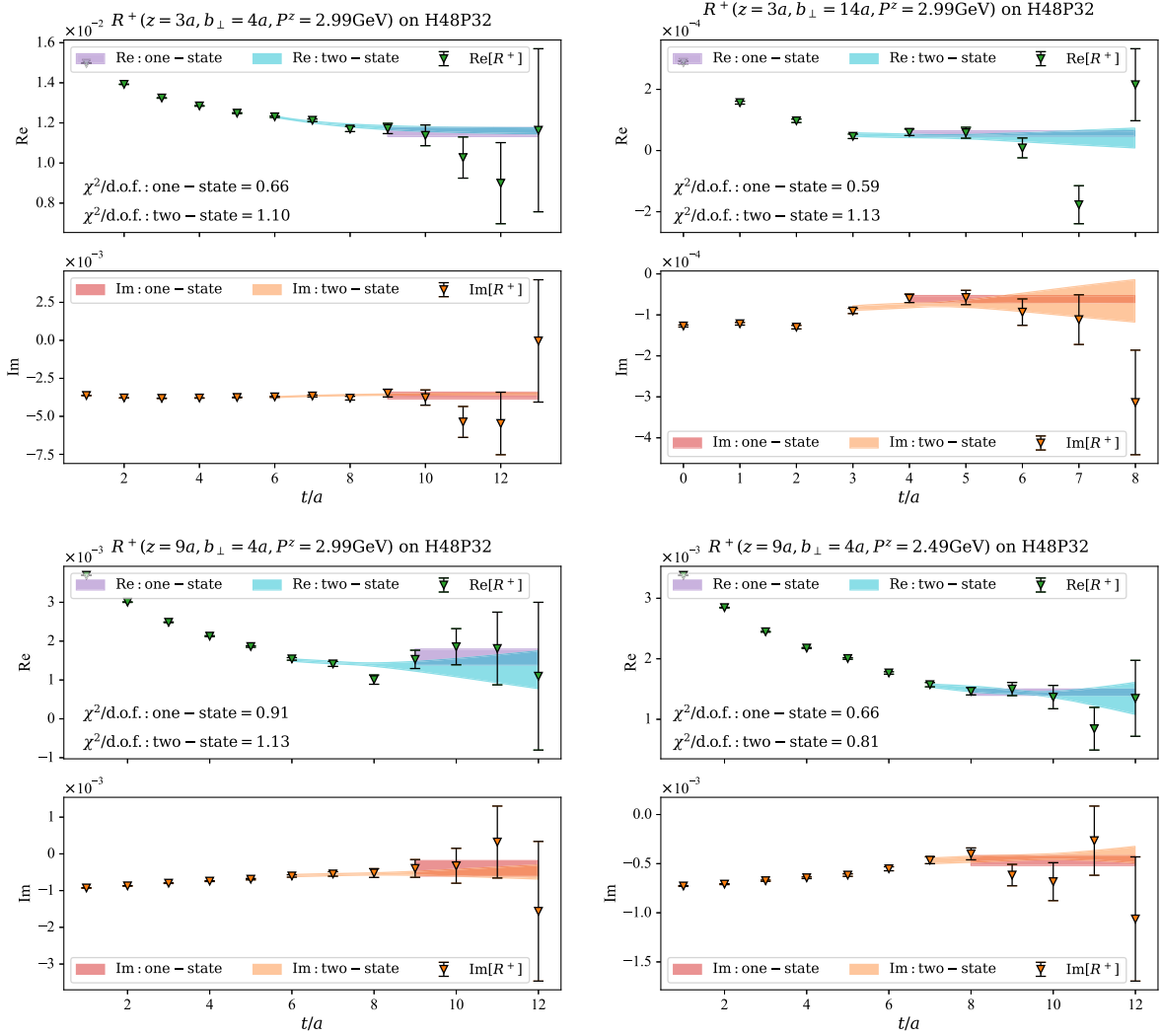


FIG. 3. Comparison of the lattice data  $R^+(z, b_\perp, P^z; t; L = 12a)$  along with the fit results of bare matrix element  $\tilde{\Phi}^0$  obtained using different fitting strategies. The figures present results from the H48P32 ensemble with 1-step HYP smearing, and compare the combinations for  $(z, b_\perp, P^z) = (3a, 4a, 2.99 \text{ GeV})$ ,  $(3a, 14a, 2.99 \text{ GeV})$ ,  $(9a, 4a, 2.99 \text{ GeV})$ , and  $(9a, 4a, 2.49 \text{ GeV})$  as examples to demonstrate the consistency of the results obtained from one-state and two-state fits.

symmetry: the coordinate-space matrix element defined in Eq. (6) is symmetric under  $\pm z$ , and correspondingly, the momentum-space distribution in Eq. (5) is symmetric about  $x = 1/2$ . Based on this symmetry, in the following discussion, we will no longer distinguish between quantities along the  $\pm z$  directions (e.g.,  $R^\pm$ ,  $\tilde{\Phi}^{\pm 0}$ ) or their averages ( $R$ ,  $\tilde{\Phi}^0$ ). In the final results, we use the averaged values to reduce statistical fluctuations.

In lattice calculations of quasi-TMDs, the most significant statistical error arises from the long Wilson link. In Eq. (13), the statistical fluctuations originating from the gauge field sampling accumulate will eventually dominate the statistical error of the nonlocal correlation functions as the length of the Wilson line increases. This is one of the main challenges in lattice calculations of TMDs. In this study, we employ the HYP smearing procedure to improve

the statistical accuracy of the gauge field. In practice, we apply different steps of HYP smearing to  $U_\square(L, z, b_\perp)$  (we test one step [HYP1] and two steps [HYP2], and compare these with the case without any HYP smearing [HYP0]) to examine the impact of HYP smearing on statistical fluctuations and the resulting measurements. Figure 4 shows a comparison of the bare matrix element results for different steps of HYP smearing. At the bare matrix element level, this comparison is not significant. Since HYP smearing modifies the short-distance UV fluctuations of the gauge field [66,67], one can observe noticeable differences in the bare matrix elements at different smearing steps, which arise from changes in the UV behavior. This difference will be corrected through nonperturbative renormalization, and the comparison of the renormalized results will be presented in the following section.

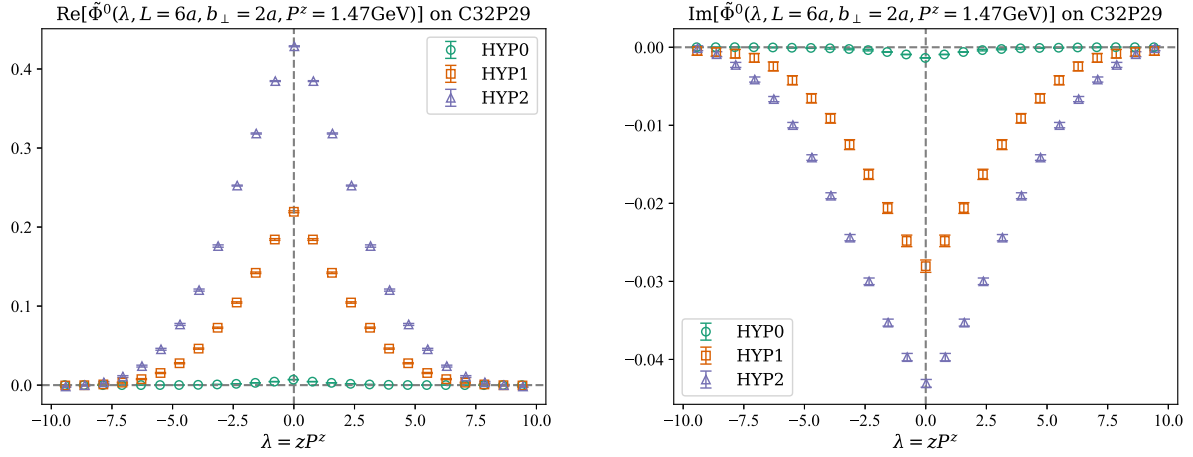


FIG. 4. Comparison of the real part (left panel) and imaginary part (right panel) of the bare matrix element from different HYP smearing steps. In this figure, we use the case with  $(L, b_{\perp}, P^z) = (6a, 1a, 1.47 \text{ GeV})$  from the C32P29 ensemble as an example, with similar results for other cases.

## 2. Nonperturbative renormalization and large- $L$ saturation

As mentioned earlier, to remove the linear divergence, the bare quasi-TMDWF matrix element must be divided by a Wilson line of the same shape and length. In this work, we employ the Wilson loop renormalization method [68–72], where the Wilson loop defined in Eq. (9) can be parametrized as proportional to the static potential  $V(n_{\perp})$ , as follows:

$$Z_E(n_{\perp}, n_z) \propto e^{-V(n_{\perp})n_z} (1 + \mathcal{O}(e^{-n_z \Delta E})). \quad (16)$$

This expression enables an extrapolation of the Wilson loop using a two-state fit, as described by the following equation:

$$Z_E(n_{\perp}, n_z) = c_0(n_{\perp}) e^{-E(n_{\perp})n_z} [1 + c_1(n_{\perp}) e^{-\Delta E(n_{\perp})n_z}]. \quad (17)$$

Figure 5 shows the lattice data and extrapolated results for the Wilson loop, both with and without HYP smearing, applied to the C32P29 gauge ensemble for  $n_{\perp} = \{1, 2, 3, 4, 5\}a$  as examples. The comparison of the results

from different steps of HYP smearing indicates that, for a fixed  $n_{\perp}$ , the Wilson loop without HYP smearing decreases more rapidly along  $n_z$ , accompanied by a more rapid deterioration of the SNR. This suggests that HYP smearing effectively improves the signal quality.

Based on the locality property of the gauge action, the influence of a gauge-field disturbance decays exponentially with distance, and large Wilson loops can be factorized into approximately independent subregions [73]. In quasi-TMDs, one also finds that for the nonlocal operator, the Wilson line segment that lies far away from the quark fields is essentially identical to the corresponding segment appearing in the Wilson loop used for renormalization. Therefore, once the bare quasi-TMDWF is divided by the Wilson loop of the same shape and length, the contributions from the distant Wilson-link segments (those far from the quark sources) saturate and no longer depend on the further extension of the Wilson link. In other words, the long-link contributions become independent of the Wilson-line length beyond some moderate distance, because the additional segments lie in a region whose gauge fluctuations no longer couple strongly to the quark-field insertion region.

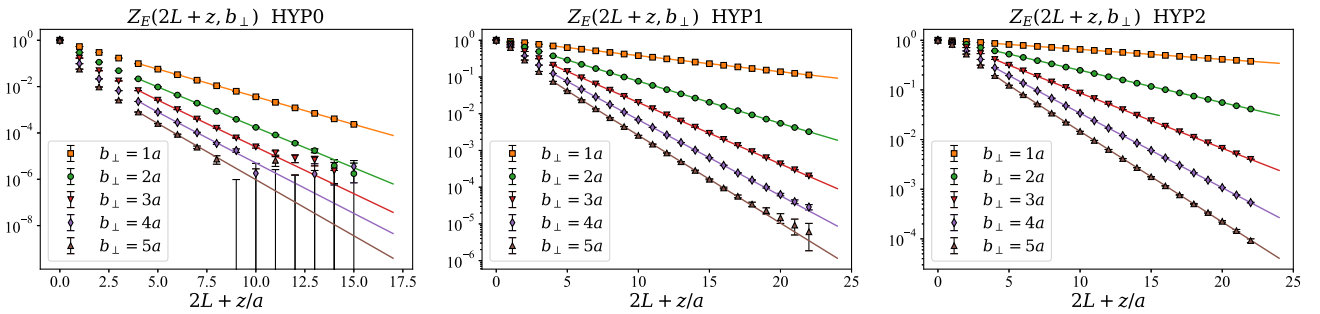


FIG. 5. Results of the Wilson loop for different shapes  $(2L+z) \times b_{\perp}$  on the C32P29 ensemble, with three panels showing a comparison of the results from different step of HYP smearing. The solid lines represent the extrapolated results based on Eq. (16). It can be seen that the extrapolated results agree well with the behavior of the data at large spatial separations.

TABLE II. Renormalization factor  $Z_O$  extracted from three ensembles.

| $a/\text{fm}$ | 0.1053     | 0.07746    | 0.05187    |
|---------------|------------|------------|------------|
| HYP0          | 1.072(22)  | 1.153(20)  | 1.251(19)  |
| HYP1          | 0.9221(62) | 0.9753(57) | 1.0397(53) |
| HYP2          | 0.8690(58) | 0.9230(54) | 0.9882(50) |

For the renormalization factor  $Z_O$ , as mentioned earlier, it is independent of the coordinates, so it can be factored out of the Fourier transform and then canceled in the extraction of the CS kernel. However, to ensure the completeness of the quasi-TMDWF results and to verify that the renormalized distributions are independent of HYP smearing, we still extract  $Z_O(1/a, \mu)$  using the self-renormalization scheme. Table II presents the results for  $Z_O(1/a, \mu)$  at  $\mu = 2 \text{ GeV}$  for different lattice spacings, with a detailed analysis provided in Appendix A.

Therefore, the renormalized matrix elements can be obtained by dividing the bare matrix elements by the Wilson loop  $Z_E$  as well as the renormalization factor  $Z_O$ , as follows:

$$\tilde{\Phi}(z, b_\perp, P^z, L) = \frac{\tilde{\Phi}^0(z, b_\perp, P^z, L)}{Z_O(1/a, \mu) \sqrt{Z_E(2L + z, b_\perp)}}, \quad (18)$$

ensuring proper accounting for the dependence on the UV divergences.

For a well-defined quasi-TMDWF, the length  $L$  of the Wilson link extending along the  $z$ -direction must be sufficiently large to ensure that the link reaches outside the hadron's spatial extent. Note that in this work  $L$  denotes the “additional” extension beyond the quark-antiquark separation along  $z$ , which convention differs from that adopted in several previous studies. Further details of this setup are provided in Appendix B.

The  $L$ -dependence of the bare matrix element is associated with the linear divergence of the Wilson link self-energy, which can be subtracted by the square root of the Wilson loop. Therefore, the  $\tilde{\Phi}$  defined in Eq. (18) should saturate to a plateau at  $L$  becomes large. To extract the asymptotic behavior as  $L \rightarrow \infty$ , we perform a constant fit of  $\tilde{\Phi}$  and define the deviation from the plateau as

$$\begin{aligned} \Delta\tilde{\Phi}(z, b_\perp, P^z, L) \\ = \tilde{\Phi}(z, b_\perp, P^z, L) - \tilde{\Phi}(z, b_\perp, P^z, L \rightarrow \infty). \end{aligned} \quad (19)$$

This allows us to investigate the value of  $L$  at which saturation occurs. As shown in Fig. 6, starting from  $L = 6a$  on the C32P29 ensemble,  $\Delta\tilde{\Phi}$  converges to zero, indicating that the staple-shaped Wilson link along the  $z$ -direction is sufficiently long. The corresponding length,  $L \simeq 0.6 \text{ fm}$ , is of the same order as the pion's strong-interaction (matter) radius, suggesting that once the link extends beyond the hadron's interaction region in a given direction, further

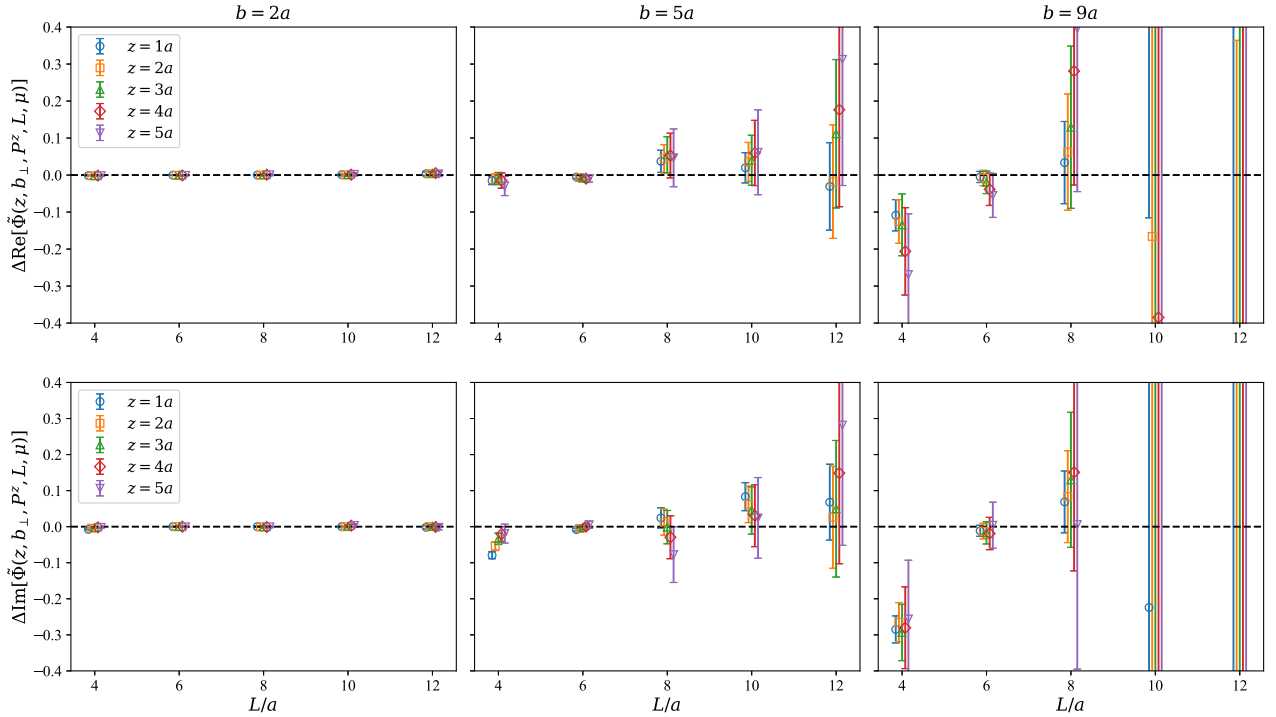


FIG. 6. Saturation of  $L$ -dependence for renormalized quasi-TMDWFs, illustrated using the results at  $z \in \{1, 2, 3, 4, 5\}a$ ,  $P^z = 1.47 \text{ GeV}$ , and  $b \in \{2, 5, 9\}a$  on the C32P29 ensemble. The upper panel shows the real part, and the lower panel shows the imaginary part.

increases in its length no longer affect the partonic distributions probed by the operator. Based on this, we choose  $L \gtrsim 0.6$  fm on the other ensembles: for C32P23 and C48P14, we take  $L = 6a$ ; for F32P30, we take  $L = 8a$ ; and for H48P32, we take  $L = 12a$ .

### 3. Renormalized quasi-TMDWFs

By subtracting the linear divergence in the bare matrix element from the Wilson loop  $Z_E$  and removing the remaining logarithmic divergence from  $Z_O$ , we obtain numerical results for renormalized matrix elements as function of  $\lambda = zP^z$ . Figure 7 compares the results of the renormalized matrix elements from different HYP smearing

steps. From the comparisons across the cases listed in the subplots, it is clear that the results are indeed consistent across different HYP smearing steps. Additionally, the noisy-to-signal (N/S) ratio shown below each subplot demonstrates that as the Wilson link length increases, using more HYP smearing steps significantly improves the SNR. For conservative estimates, we choose HYP1 as the smearing strategy in this work.

According to our numerical simulations, the renormalized coordinate-space matrix elements  $\tilde{\Phi}$  as a function of  $\lambda = zP^z$  are shown in Fig. 8. Here, we take the results from H48P32 ensemble as examples, and additional cases are provided in the Appendix D. The figure shows the dependence of the

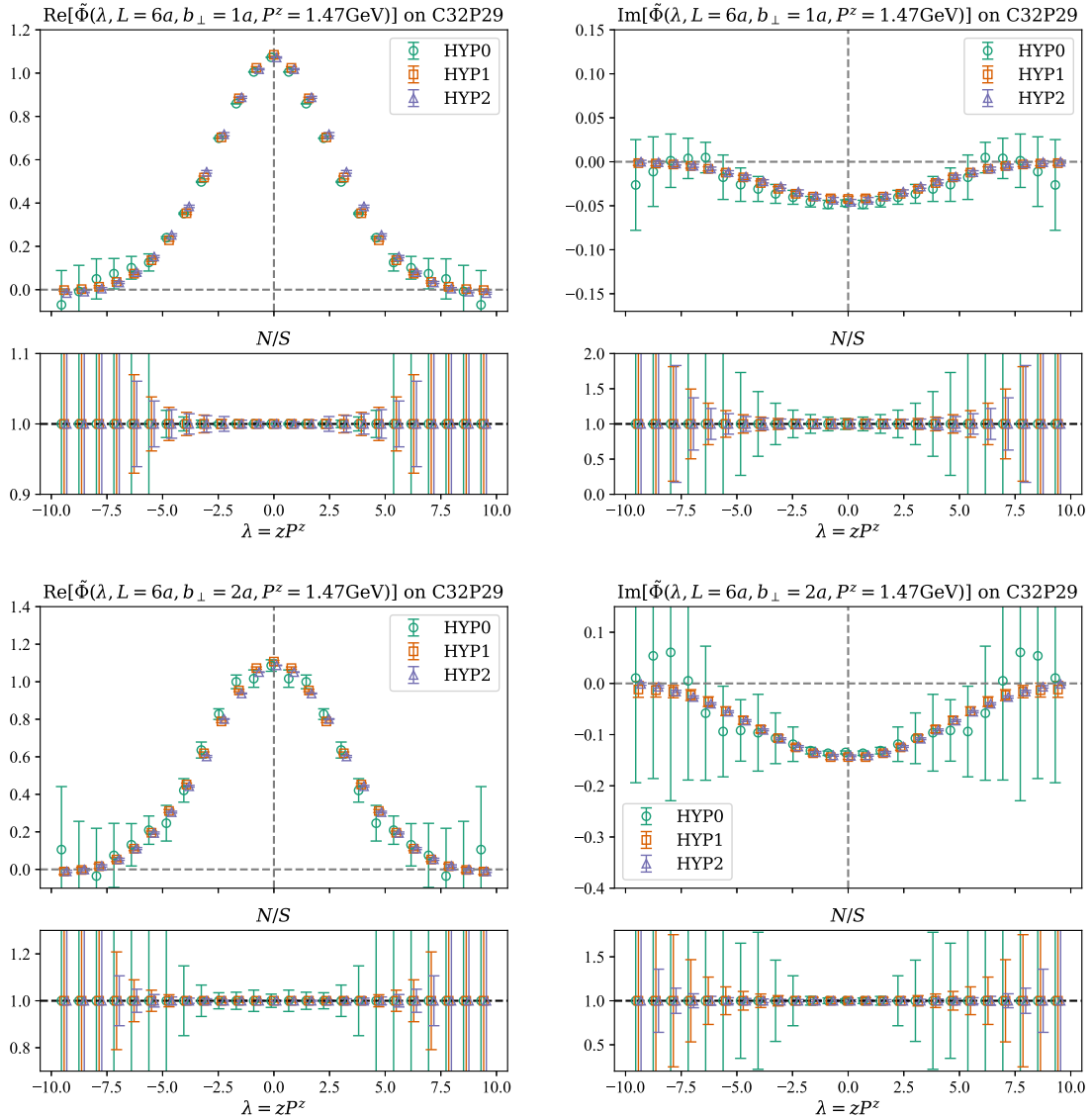


FIG. 7. Comparison of the real part (left panels) and imaginary part (right panels) of the renormalized matrix element from different HYP smearing steps. In this figure, we use the cases with  $(L, P^z) = (6a, 1.47 \text{ GeV})$  at  $b_\perp = 1a$  (upper panels) and  $2a$  (lower panels) from the C32P29 ensemble as examples. Similar results are observed for larger  $b_\perp$ , but the results for HYP0 will suffer large errors. Below each subplot, the N/S ratio for the different HYP smearing steps is displayed, showing that the SNR improves as the number of HYP smearing steps increases.

renormalized coordinate-space matrix elements on the boosted momentum  $P^z$  at fixed transverse separation  $b_\perp$ , as well as their  $b_\perp$  dependence at fixed  $P^z$ . Both the real and imaginary parts exhibit symmetry about  $z = 0$ , ensuring that the renormalized quasi-TMDWF after Fourier transformation is real. Exploiting this symmetry, we average the  $\pm z$  data to reduce statistical fluctuations. The observed  $P^z$  dependence at each  $b_\perp$  encodes the rapidity evolution, which in turn enables a numerical extraction of the CS kernel.

To determine the quasi-TMDWFs  $\tilde{f}(x, b_\perp, \mu, \zeta_z)$  in momentum space, we need to perform a Fourier transformation. Although Fig. 8 shows that  $\tilde{\Phi}$  converges to zero at large  $\lambda$ , however, a brute-force truncated Fourier transformation would introduce significant oscillations in the momentum-space distribution due to the errors [74–76]. Therefore, we extrapolate the renormalized matrix elements  $\tilde{\Phi}$  by using the following form [77]:

$$\tilde{\Phi}(\lambda, b_\perp) = \left[ \frac{c_1}{(-i\lambda)^{n_1}} + e^{i\lambda} \frac{c_2}{(i\lambda)^{n_2}} \right] e^{-\frac{\lambda}{\lambda_0}}, \quad (20)$$

in which all parameters  $c_{1,2}$ ,  $n_{1,2}$  and  $\lambda_0$  depend on the transverse separation  $b_\perp$ . The terms inside the square brackets account for the algebraic behavior and are motivated by the Regge behavior [78] of the light-cone distributions at endpoint regions, and the exponential term is motivated by the expectation that the correlation function has a finite correlation length at finite momentum, which denoted as  $\lambda_0$ .

To perform the extrapolation, a reasonable  $\lambda$ -range is required to constrain the parameters. In practice, we test the sensitivity of the extrapolation to the chosen fit window. The fit is carried out for  $\lambda \geq \lambda_L$ , where  $\lambda_L$  is a truncation parameter selected in the long-tail region of the coordinate-space distribution, corresponding to the endpoint region in momentum space. As shown in Fig. 9, the fitting bands obtained from different fit windows agree well with the original lattice data and are mutually consistent. This demonstrates the self-consistency of the extrapolation across fit intervals. To estimate the systematic uncertainty associated with the extrapolation form,

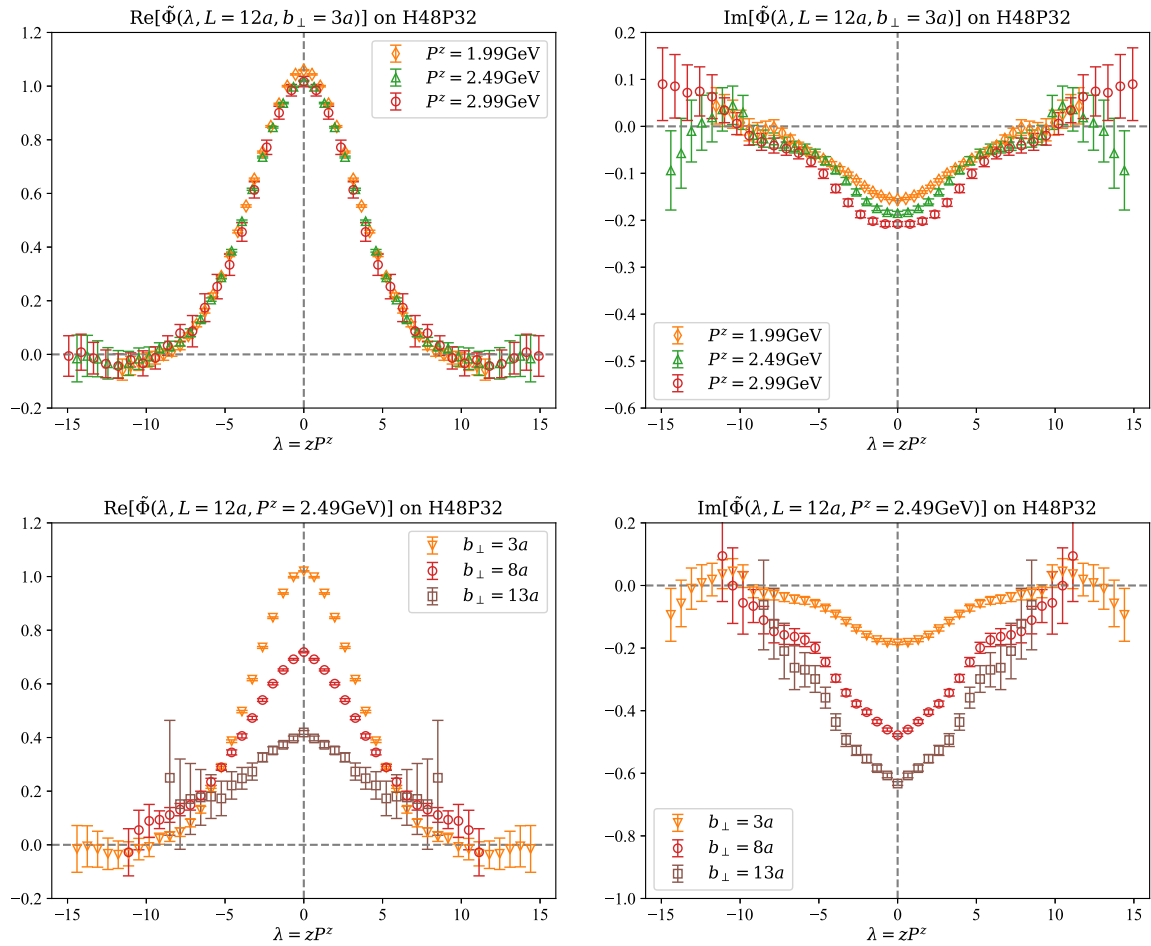


FIG. 8.  $\lambda$ -dependence of the renormalized matrix elements  $\tilde{\Phi}$ . Results from the H48P32 ensemble are shown as examples. The upper panel compares the results with  $P^z = \{1.99, 2.49, 2.99\}$  GeV at fixed  $b_\perp = 3a$ , and the lower panel compares compares the results with  $b_\perp = \{3, 8, 13\}a$  at fixed  $P^z = 2.49$  GeV. Additional cases are presented in the Appendix D.

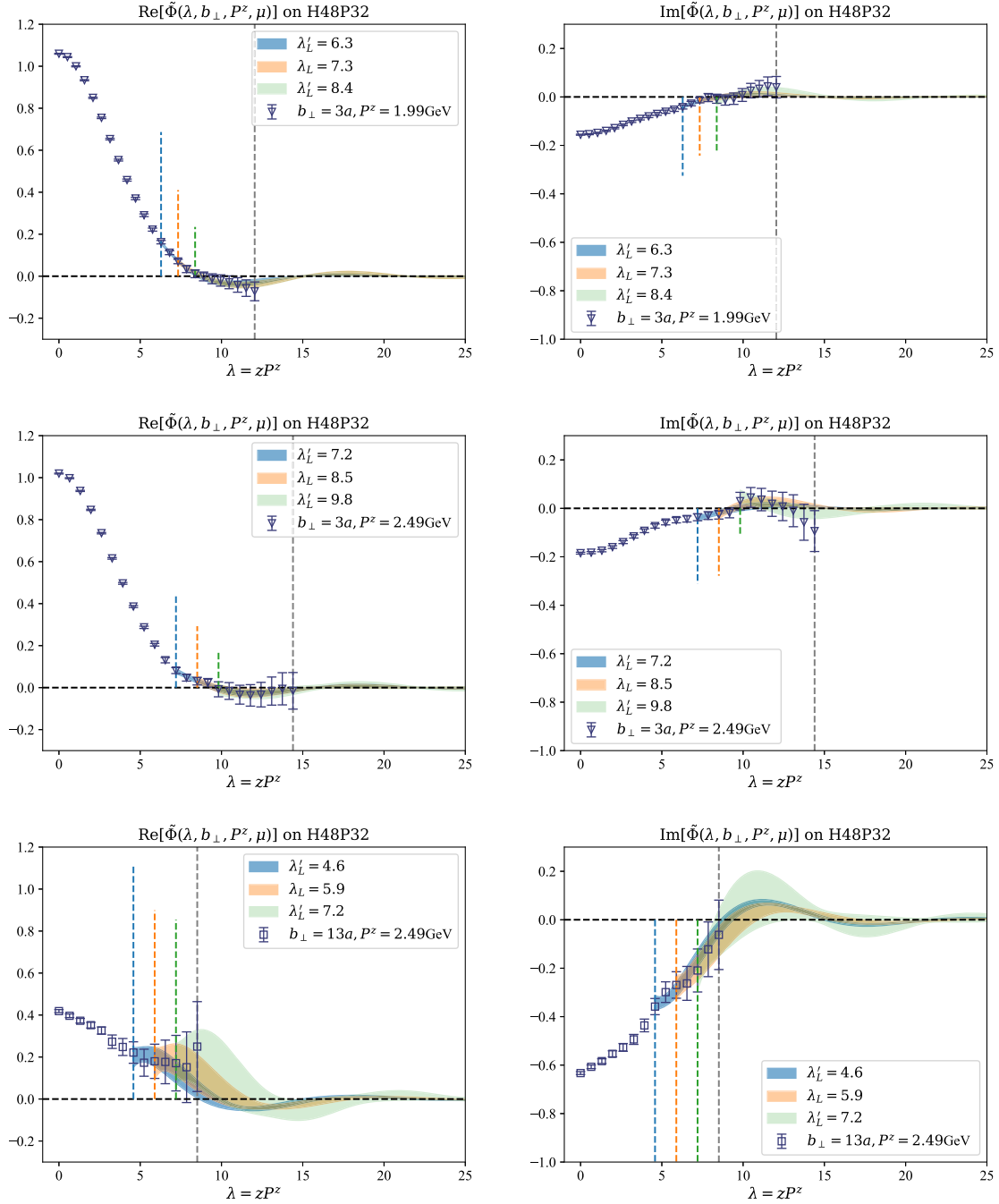


FIG. 9. Comparison of the original data (points) and the extrapolated results (colored bands) at  $(b_\perp, P^z) = (3a, 1.99 \text{ GeV})$ ,  $(3a, 2.49 \text{ GeV})$ , and  $(13a, 2.49 \text{ GeV})$  on the H48P32 ensemble. The renormalization scale is  $\mu = 2 \text{ GeV}$ . Extrapolations are performed over the fit window  $\lambda \geq \lambda_L$ , which gives the main results; systematic uncertainties are estimated by varying the window start to alternative values  $\lambda'_L$ .

we take the difference between the result obtained with the nominal  $\lambda_L$  window and those obtained by shifting  $\lambda_L$  forward and backward by two data points as the systematic error.

After Fourier transforming the coordinate-space matrix elements to momentum space, the quasi-TMDWF can be obtained as

$$\tilde{f}(x, b_\perp, P^z) = \int \frac{d\lambda}{2\pi} e^{i(x-\frac{1}{2})\lambda} \tilde{\Phi}(\lambda, b_\perp). \quad (21)$$

In Fig. 10, we present the results of the renormalized quasi-TMDWF in momentum space. Corresponding to Fig. 8, we compare the results for different HYP smearing steps, and the differences between the results obtained with different

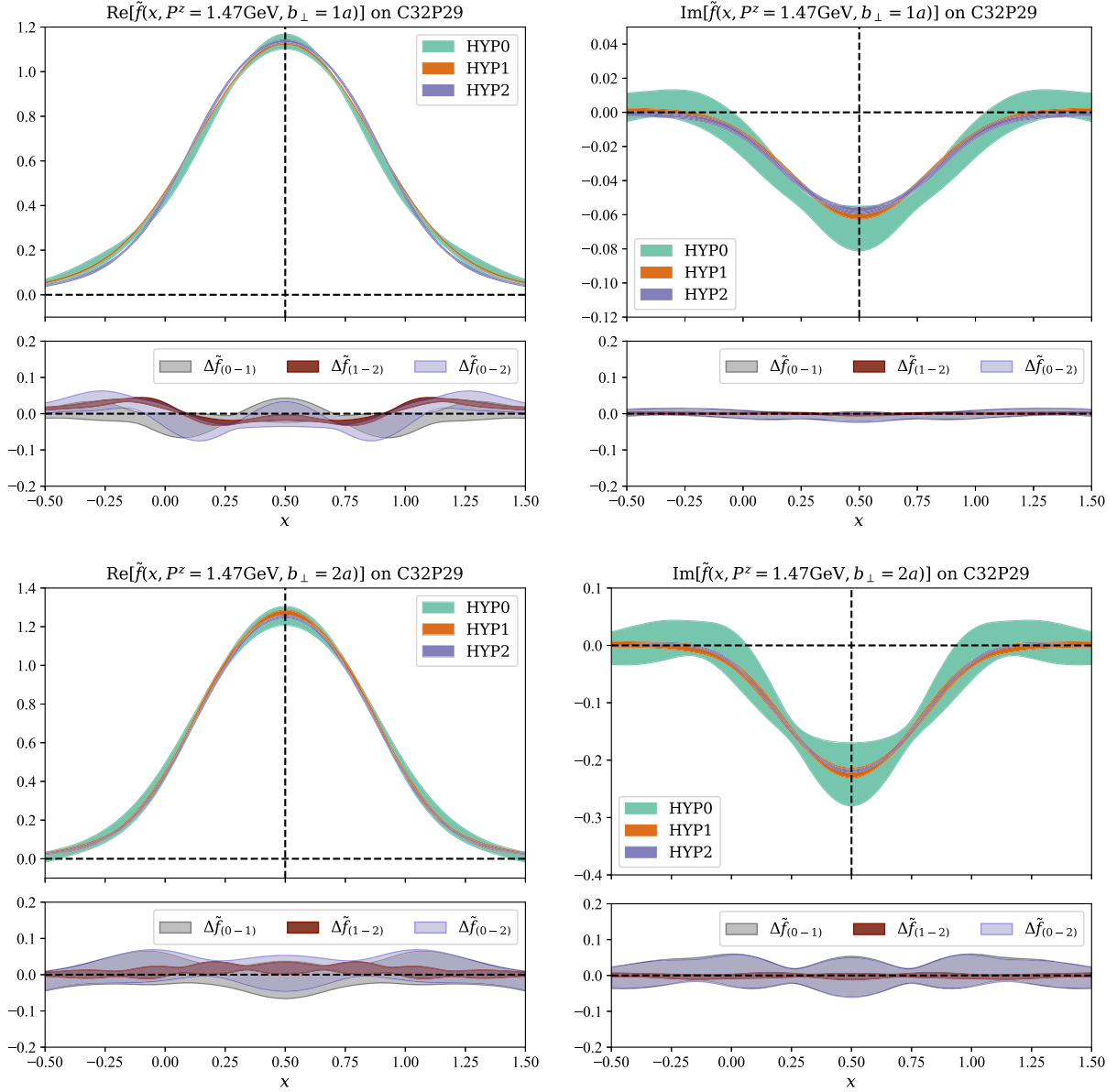


FIG. 10. Comparison of the real part (left panels) and imaginary part (right panels) of the renormalized quasi-TMDWF obtained with different HYP smearing steps. The settings are similar to Fig. 7. The panels beneath each subplot display the differences between smearing steps.

smearing steps are shown in the panels below each subplot. The difference,  $\Delta\tilde{f}(i-j) = \tilde{f}_{\text{HYP}i} - \tilde{f}_{\text{HYP}j}$ , remains consistent with zero across the region to extract CS kernel  $x \in [-0.5, 1.5]$ , indicating that while HYP smearing substantially improves statistical precision, its impact on the central values is statistically insignificant. Potential systematic effects induced by HYP smearing are well below the statistical uncertainties at our current precision and are therefore not considered in the subsequent analysis.

In Fig. 11, we present the renormalized quasi-TMDWF  $\tilde{f}(x, P^z, b_\perp, \mu)$  as a function of the momentum fraction  $x$ , together with its dependence on the longitudinal momentum  $P^z$  and the transverse separation  $b_\perp$ . The upper panels

show the  $P^z$  dependence at fixed  $b_\perp$ , which encodes the evolution in the rapidity scale  $\zeta_z = (xP^z)^2$ , and enables the extraction of the CS kernel. The lower panels display the  $b_\perp$  dependence at fixed  $P^z$ ; one can see that as  $b_\perp$  increases, the real part of  $\tilde{f}$  decreases while the imaginary part grows, indicating that the  $b_\perp$  dependence predominantly enters through a phase in  $\tilde{f}$ .

Furthermore, we compare ratios of renormalized quasi-TMDWFs at different  $P^z$ , which isolate the  $\zeta_z$  evolution. As indicated by Eq. (4), the connection between the rapidity evolution of  $\tilde{f}$  and that of the physical TMD  $f$ , that is the CS kernel, is perturbative. Notably, the ratios shown in Fig. 12 show a strong  $b_\perp$  dependence in the real part but

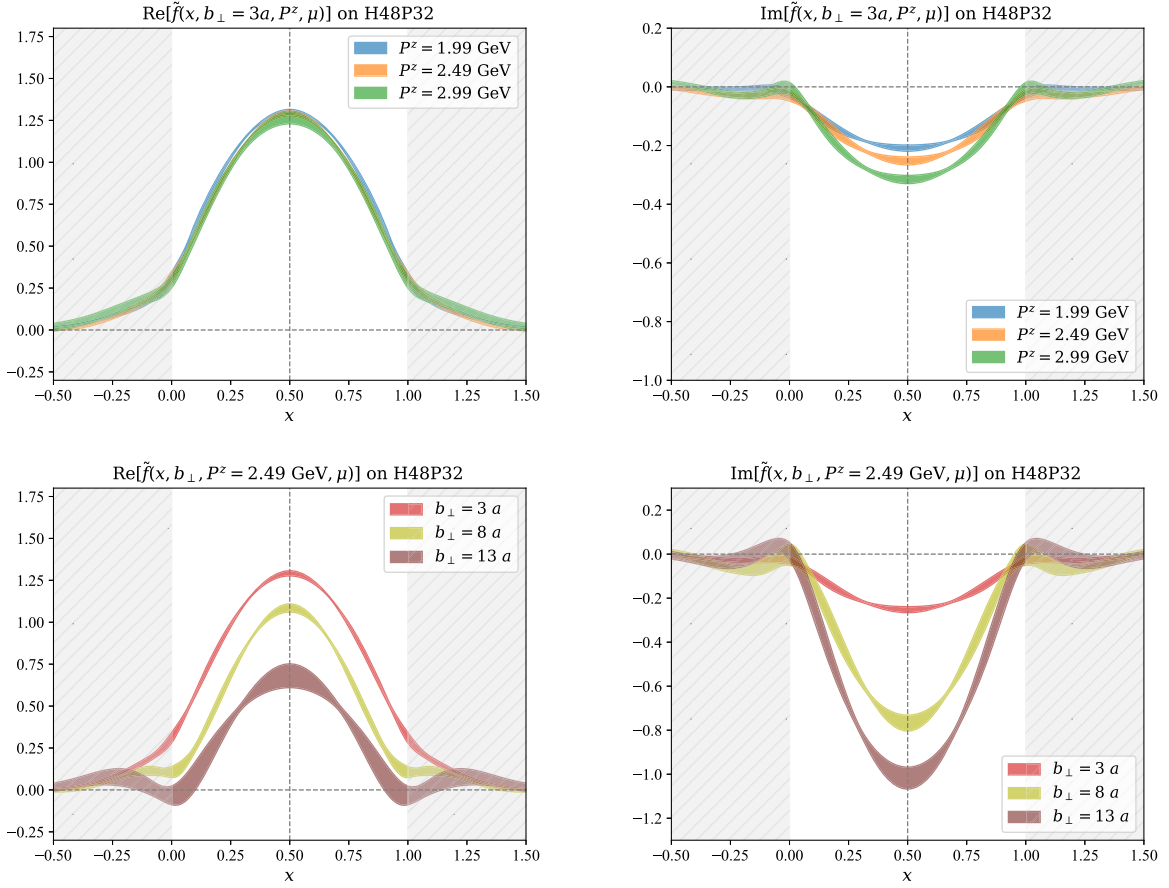


FIG. 11. Renormalized quasi-TMDWFs in momentum space. Upper panel: case of  $b_{\perp} = 3a$  and  $P^z = \{1.99, 2.49, 2.99\}$  GeV on H48P32 as examples. Lower panel: comparison of the quasi-TMDWF in momentum space with  $b_{\perp} \in \{3, 8, 13\}a$ .

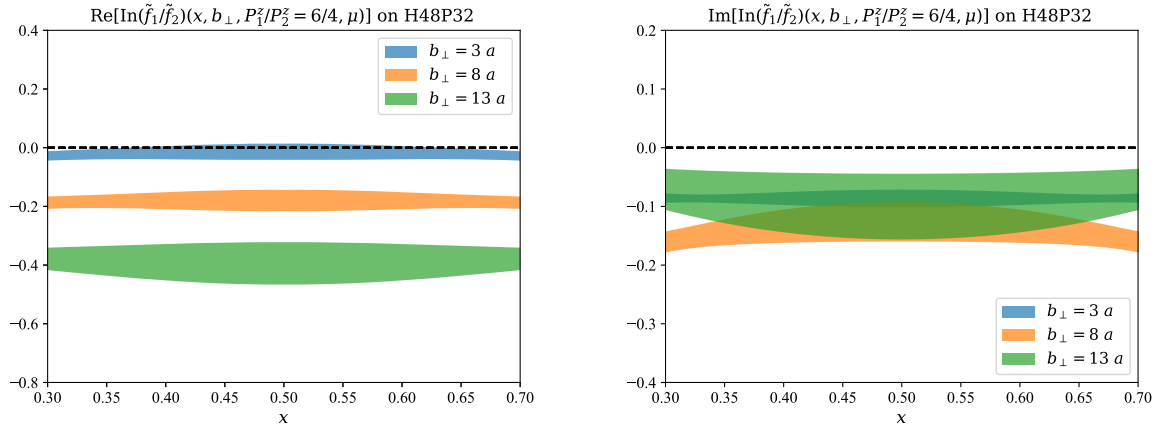


FIG. 12. Example of results for the real (left) and imaginary (right) part of the ratio for quasi-TMDWFs at different  $b_{\perp}$ ,  $\ln \frac{\tilde{f}(x, b_{\perp}, P_1^z, \mu)}{\tilde{f}(x, b_{\perp}, P_2^z, \mu)}$ , with  $P_1^z/P_2^z = 6/4$  as a function of momentum fraction  $x$ .

hardly any  $b_{\perp}$  dependence in the imaginary part, suggesting (within our current precision) that changing  $P^z$  mainly rescales the magnitude through  $\zeta_z$ , while the phase is nearly  $P^z$  independent and cancels in the ratio. Consequently, the real part is driven by the rapidity evolution kernel of  $\tilde{f}$ , producing the significant  $b_{\perp}$  dependence, whereas the

imaginary part reflects only a small, almost  $b_{\perp}$ - and  $x$  independent global phase offset. In the large- $P^z$  limit, this offset should vanish, and the residual imaginary signal is consistent with  $1/P^z$  power corrections. These features make momentum ratios a clean,  $x$ -insensitive handle for extracting the CS kernel  $K(b_{\perp}, \mu)$ .

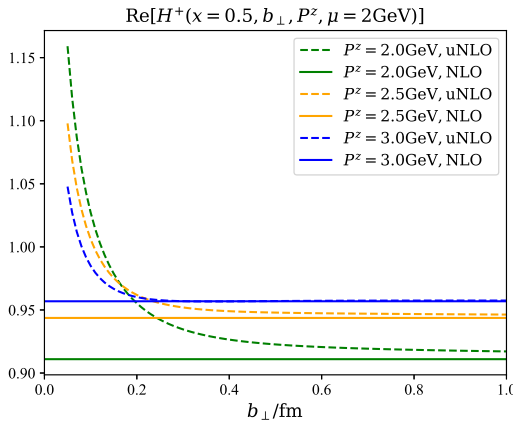
### C. CS kernel from quasi-TMDWFs

As indicated in Eq. (3), the factorization of the renormalized quasi-TMDWF requires sufficiently large hadron momenta, such that either  $xP^z b_\perp \gg 1$  or  $(1-x)P^z b_\perp \gg 1$ . However, in practice, these conditions are not always satisfied, especially in the small- $b_\perp$  region. Therefore, following Ref. [48], we adopt the  $b_\perp$ -unexpanded matching kernel, in which the full  $b_\perp$  dependence of the perturbative kernel is retained rather than expanded in powers of  $b_\perp$ . This treatment avoids relying on the small- $b_\perp$  expansion and thus provides a more reliable description over a broader range of transverse separations.

Figure 13 compares the results of  $b_\perp$ -unexpanded next-to-leading order (uNLO) matching kernel with the conventional next-to-leading order (NLO) results [61]. A significant discrepancy is observed at small  $b_\perp$ , which diminishes as  $b_\perp$  increases. The uNLO treatment improves the systematics in the small- $b_\perp$  region and helps suppress unphysical imaginary parts in the extracted CS kernel. This improvement stems from avoiding the fixed-order expansion in  $b_\perp$ , which implicitly assumes the hierarchy  $1/b_\perp \gg P^z$  and is not strictly valid in typical lattice kinematics. Retaining the full  $b_\perp$  dependence preserves a more faithful analytic continuation between coordinate and momentum space, leading to a more stable CS kernel extraction and better perturbative behavior in the nonperturbative regime.

Based on Eq. (4), we extract the CS kernel from the renormalized quasi-TMDWFs using the  $b_\perp$ -unexpanded matching kernel at different hadron momenta. As illustrated by the data points with uncertainties in Fig. 14, we obtain  $K(b_\perp, \mu; x, P_1^z, P_2^z)$  from ratios formed at various large-momentum pairs  $(P_1^z, P_2^z)$ , and then perform an infinite-momentum extrapolation using

$$K(b_\perp, \mu; x, P_1^z, P_2^z, a, m_\pi) = K(b_\perp, \mu; a, m_\pi) + A(b_\perp, \mu; x, a, m_\pi) \left[ \frac{1}{(P_1^z)^2} - \frac{1}{(P_2^z)^2} \right], \quad (22)$$



which is motivated by factorization analysis of the TMDWF in Eq. (3). The coefficient  $A(b_\perp, \mu; x, a, m_\pi)$  parametrizes next-to-leading power effects, which can be numerically enhanced at the endpoint regions in  $x$ .

The extractions of  $K(b_\perp, \mu; a, m_\pi)$  in the infinite-momentum limit are shown in Fig. 14. We find that the results obtained from different momentum pairs are mutually consistent, indicating that the chosen hadron momenta are sufficiently large and that residual higher-power contributions are well suppressed. Based on Eq. (22), we therefore perform the infinite-momentum extrapolation in the intermediate- $x$  window,  $x \in [0.3, 0.7]$ , where power corrections are expected to be minimized. The resulting leading power CS kernel  $K(b_\perp, \mu; a, m_\pi)$  are presented as the gray band in Fig. 14.

Following the procedure outlined above, we extract the CS kernel on the ensembles with different lattice spacings and pion masses, as shown in Fig. 15. The real part of the CS kernels, illustrated in the left panel, exhibit remarkable consistency across the five ensembles, indicating that the discretization effects and mass corrections have been well controlled.

Theoretically, the CS kernel is purely real, and this expectation is borne out by the right panel of Fig. 15. The imaginary parts from all ensembles are basically consistent with zero, especially for large  $b_\perp$  cases. The residual unphysical imaginary contributions primarily arise from the perturbative matching kernel, which will mainly contribute to the small- $b_\perp$  region [45]. To estimate the impact of this effect on the final results, we quantify the systematic uncertainty associated with the residual imaginary part as

$$\delta_{\text{sys}}^{\text{Im}} = \sqrt{K(b_\perp, \mu)^2 + \text{Im}[K(b_\perp, \mu)]^2} - |K(b_\perp, \mu)|, \quad (23)$$

which measures the deviation between the complex magnitude and the real part of  $K(b_\perp, \mu)$ , and thus isolates the contribution from the imaginary part under the expectation that the physical CS kernel is real. A detailed discussion of this systematic appears in Sec. III D.

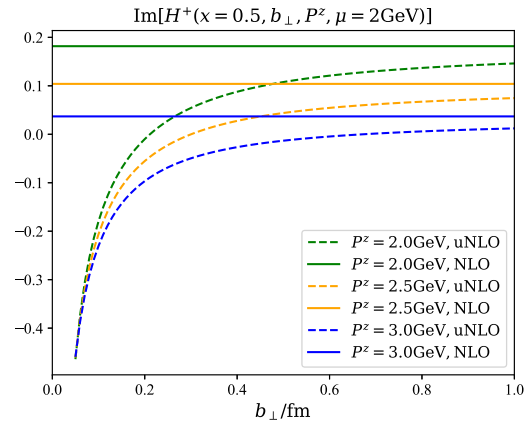


FIG. 13. Comparison of matching kernels at NLO (solid) and  $b_\perp$ -unexpanded NLO (uNLO, dotted) for fixed  $x = 0.5$  and  $\mu = 2$  GeV.

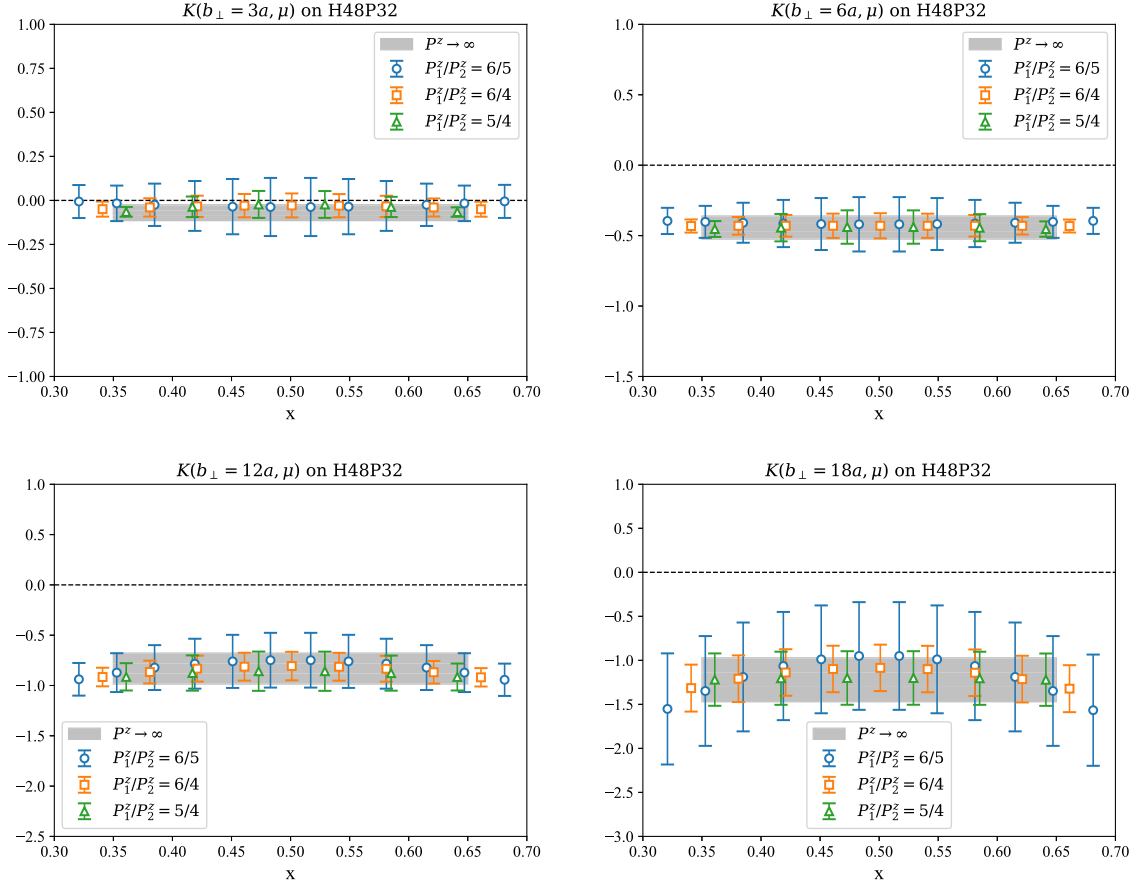


FIG. 14. CS kernel extracted from quasi-TMDWFs and its extrapolation to the large-momentum limit. Results are shown for  $b_\perp = 3, 6, 12, 18a$  on the H48P32 ensemble, using three distinct momentum pairs  $P_1^z/P_2^z$ . Data points with errors denote  $K(b_\perp, \mu; x, P_1^z, P_2^z)$  from individual momentum combinations, and the band indicates the large-momentum extrapolated result of  $K(b_\perp, \mu)$ .

To perform the continuum extrapolation, results obtained at different lattice spacings are first interpolated to a common spacing  $a' = 0.05187$  fm. We then carry out, for each fixed  $b_\perp$ , a combined continuum and chiral extrapolation using

$$K(b_\perp, \mu; a, m_\pi) = K(b_\perp, \mu) + a^2 B(b_\perp, \mu) + (m_\pi^2 - m_{\pi, \text{phy}}^2) C(b_\perp, \mu), \quad (24)$$

where the term  $B$  encodes discretization effects and  $C$  parametrizes deviations from unphysical pion masses. The fits are illustrated in Fig. 16, which compares data from all ensembles with the extrapolated curves at several  $b_\perp$ . Across all shown  $b_\perp$  values, the dependence on the pion mass is negligible, consistent with the expected universality of the CS kernel at sufficiently large boost factors  $P^z/m_\pi$ . The lattice spacing dependence is likewise modest, indicating that discretization effects are under good control. An exception occurs at very small transverse separations,  $b_\perp \lesssim 0.1$  fm, where  $b_\perp$  approaches the lattice cutoff ( $b_\perp \sim a$ ) and discretization artifacts become significant. Since this region is beyond the reliable range of LaMET, we omit these points from the final predictions.

Beyond the leading  $1/(P^z)^2$  corrections included in Eq. (22), there also may exist potential lattice-spacing-dependent discretization corrections proportional to  $aP^z$  [50]. Such contributions may originate from discretization artifacts at large hadron momentum and could, in principle, affect the infinite-momentum extrapolation.

Beyond the leading  $1/(P^z)^2$  corrections included in Eq. (22), additional lattice-spacing-dependent power corrections proportional to  $aP^z$  may also arise, as suggested in Ref. [50]. Such terms can originate from discretization artifacts at large hadron momenta and could, in principle, affect the infinite-momentum extrapolation in Eq. (22). In the present analysis, we observe that these corrections are small for two reasons. First, as shown in Fig. 2, the boosted hadron energies on all ensembles are consistent with the continuum dispersion relation over the range of momenta used, indicating that discretization effects that grow with  $P^z$  are well under control. Secondly, as illustrated in the left panel of Fig. 15, the CS kernels extracted at the same momentum combinations but at different lattice spacings  $a$  agree within uncertainties, providing direct evidence that the residual  $aP^z$ -dependent effects are swamped by the statistical errors.

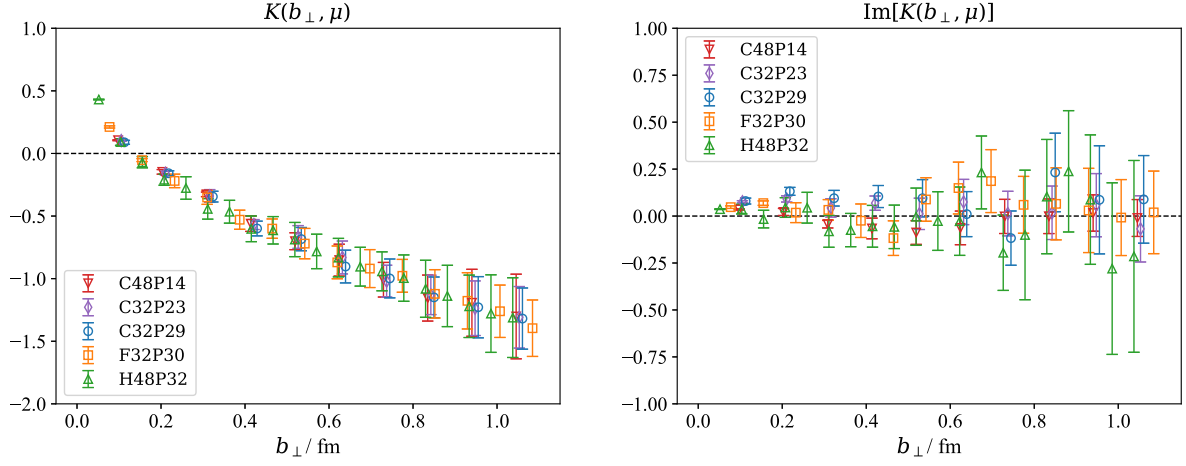


FIG. 15. Real (left) and imaginary (right) parts of the CS kernel extracted from five ensembles. Error bars denote statistical uncertainties.

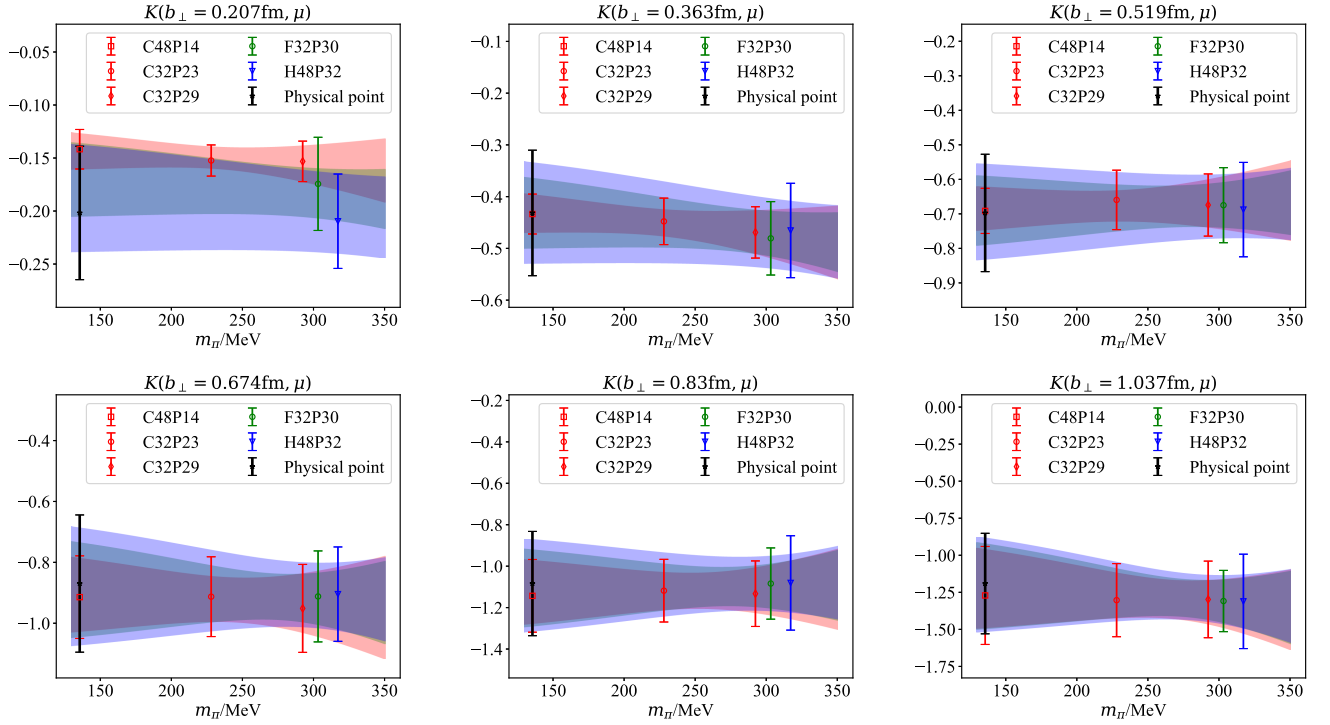


FIG. 16. Continuum and chiral extrapolations of the CS kernel. Colored points denote results from each ensembles, and black points indicate the extrapolated values. Error bars reflect statistical uncertainties only.

To further validate this assumption, we combine the extrapolation *Ansätze* in Eqs. (22) and (24), and explicitly include a mixed term proportional to  $aP^z$  to parametrize lattice-spacing-dependent power corrections suggested in Ref. [50], as follows:

$$K(b_\perp, \mu; x, P_1^z, P_2^z, a, m_\pi) = K(b_\perp, \mu) + A(x, b_\perp, \mu) \left[ \frac{1}{(P_1^z)^2} - \frac{1}{(P_2^z)^2} \right] + a^2 B(b_\perp, \mu) + (m_\pi^2 - m_{\pi, \text{phy}}^2) C(b_\perp, \mu) + D(b_\perp, \mu) a(P_1^z - P_2^z), \quad (25)$$

where the term proportional to  $D(b_\perp, \mu)$  is introduced to capture potential discretization artifacts that grow linearly with the hadron momentum.

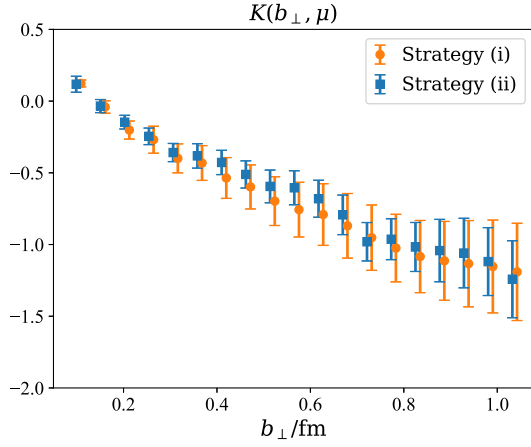


FIG. 17. Comparison of the extrapolated CS kernel  $K(b_\perp, \mu)$  obtained with two fitting strategies: (i) a chained fit based on Eqs. (22) and (24) and (ii) a joint fit using Eq. (25). Only statistical uncertainties are shown.

The comparison of the extrapolated result of  $K(b_\perp, \mu)$  from the fitting strategies (i) a chained fit based on Eqs. (22) and (24) and (ii) a joint fit from Eq. (25) are displayed in Fig. 17. One can see that within the current statistical precision, the results for two strategies are consistent within uncertainties. In particular, no statistically significant  $aP^z$  term can be resolved, indicating that its contribution cannot be reliably disentangled from the leading power corrections with the present data. For the remainder of this work, we adopt the results obtained from the chained fitting strategy for our final analysis.

#### D. Estimation of systematic uncertainties and final results

For a reliable prediction, we evaluate the main sources of systematic uncertainties in lattice QCD calculation and analysis as follows:

- (1)  $\delta_{\text{sys}}^{\text{Im}}$ : from the imaginary part.

As mentioned above, the physical CS kernel is purely real, and the nonzero imaginary component primarily originates from truncation effects in the perturbative matching kernel (particularly in the small- $b_\perp$  region). Although adopting the  $b_\perp$ -unexpanded matching kernel substantially reduces this artifact (compared with our earlier analysis [45,46]), the residual imaginary part, which now consistent with zero within uncertainties as illustrated in Fig. 15, can still inflate statistical fluctuations and thus affect the final precision. To quantify the associated systematic, we evaluate  $\delta_{\text{sys}}^{\text{Im}}$  by taking  $\text{Im}[K]$  as the ensemble-averaged imaginary part of the CS kernel over the five ensembles and into Eq. (23) to quantify the systematic uncertainty from the unphysical imaginary component.

- (2)  $\delta_{\text{sys}}^\lambda$ : from  $\lambda$  extrapolation.

On the lattice, matrix elements can only be computed up to a finite spatial separation, whereas

the Fourier transform to momentum space requires integration over  $|\lambda| = |zP^z| \rightarrow \infty$ . To solve this, we employ the extrapolation form in Eq. (20) for the dimensionless variable  $\lambda$ . Physically, a reliable extrapolation should be performed in the asymptotic region where the renormalized coordinate-space matrix elements have entered the exponential-decay region at large  $\lambda$ . The residual ambiguity in choosing the fit window induces a systematic uncertainty. In practice, we assess this by varying the truncation parameter  $\lambda_L$  by shifting  $\lambda_L$  forward and backward by  $\sim 0.1$  fm in separation (corresponding to 1–2 lattice spacings), repeating the fits, and taking the discrepancy of the resulting momentum-space observables as  $\delta_{\text{sys}}^\lambda$ .

- (3)  $\delta_{\text{sys}}^{a,m_\pi}$ : from continuum and chiral extrapolations.

Lattice QCD is formulated at finite lattice spacings, and the results must be extrapolated to the continuum. Similarly, ensembles with unphysical quark masses require a chiral extrapolation to the physical point. In this work, we employ five ensembles spanning three lattice spacings  $a \simeq (0.105, 0.077, 0.052)$  fm and a range of pion masses including the physical one. For each fixed  $b_\perp$ , we fit the data using Eq. (22) and obtain the physical result  $K(b_\perp, \mu)$ . To estimate the associated systematic uncertainty, we take the single data point (from any ensemble) whose central value lies closest to the extrapolated result and use the absolute difference between this point and the extrapolated value as  $\delta_{\text{sys}}^{a,m_\pi}$ . This choice provides a conservative estimate of the residual model dependence of the extrapolation.

- (4)  $\delta_{\text{sys}}^{P^z}$ : from infinite-momentum extrapolation.

LaMET requires sufficiently large hadron boosts such that power corrections of order  $1/(P^z)^2$  are under control. In practice, only finite momenta are accessible, and we therefore extrapolate to  $P^z \rightarrow \infty$  using Eq. (22) to remove the leading  $1/(P^z)^2$  effects. We estimate the associated systematic uncertainty by comparing the extrapolated value  $K(b_\perp, \mu)$  with the result obtained at the largest available lattice momentum, and take the absolute difference as  $\delta_{\text{sys}}^{P^z}$ .

- (5)  $\delta_{\text{sys}}^{aP^z}$ : from lattice-spacing-dependent power corrections.

As discussed in Sec. III C, results from the joint-fit strategy based on Eq. (25) are consistent with those from the chained-fit strategy. For a conservative error budget, we adopt the latter one in the final analysis, which allows us to estimate the systematic uncertainties from the continuum extrapolation and the infinite-momentum extrapolation separately. Accordingly, we do not introduce an independent uncertainty  $\delta_{\text{sys}}^{aP^z}$ ; instead, any residual  $aP^z$ -dependent effects are treated as subleading and are effectively covered by the systematic variations already included in  $\delta_{\text{sys}}^{P^z}$  and  $\delta_{\text{sys}}^a$ .

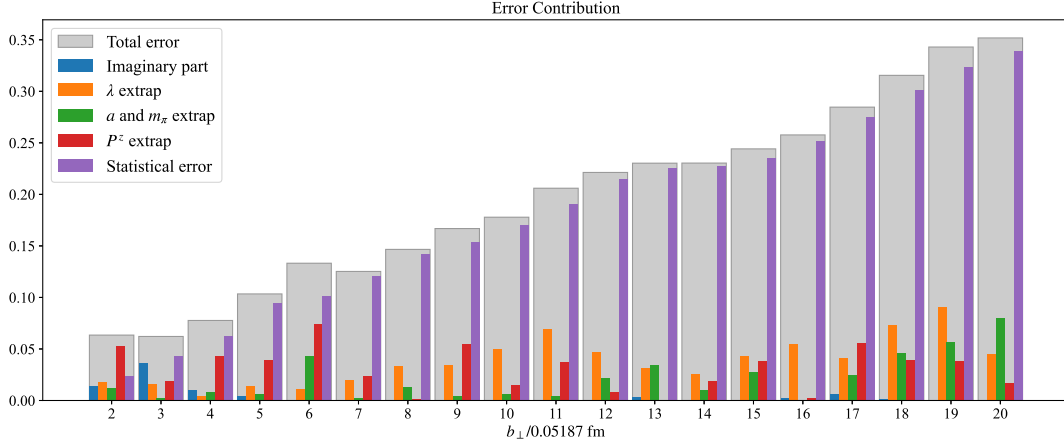


FIG. 18. Error budget for the CS kernel, showing statistical uncertainties and systematic ones from different sources. Total uncertainties are obtained by adding these contributions in quadrature.

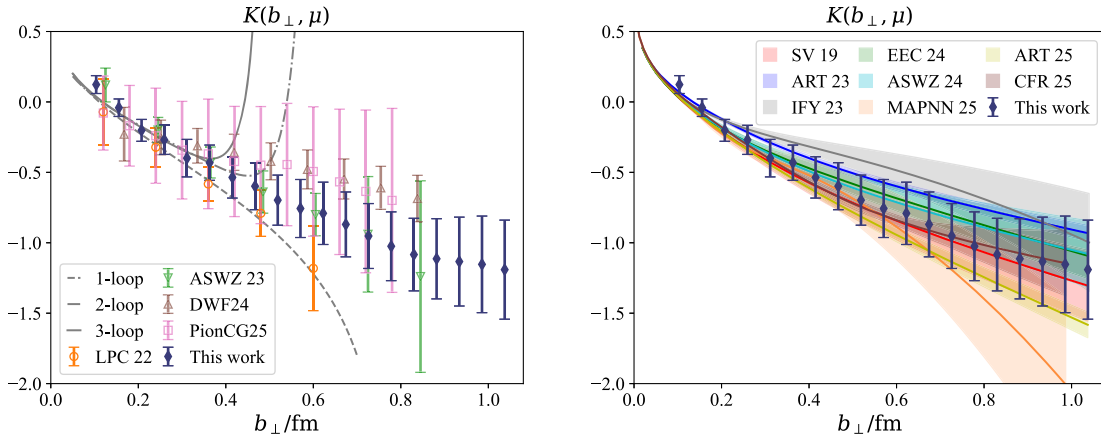


FIG. 19. Left panel: comparison of our result  $K(b_\perp, \mu)$  the lattice calculations by LPC22 [45], ASWZ23 [48], DWF24 [79], PionCG25 [80], as well as the perturbative calculations up to three loops [5,6]. Right panel: comparison of our result with phenomenological extractions from SV19 [25], ART23 [27], IFY23 [22], EEC24 [29], ASWZ24 [50], MAPNN25 [32], ART25 [31], and CFR25 [33].

In Fig. 18, we compare the statistical uncertainty of the final CS kernel with the systematic ones discussed above. As noted above, our predictions for the CS kernel are quoted for  $b_\perp \gtrsim 0.1$  fm. The comparison shows that at small  $b_\perp$  various systematics remain non-negligible, whereas at larger  $b_\perp$  the statistical error becomes dominant. We combine uncertainties in quadrature,

$$\begin{aligned} \delta_{\text{tot}} &= \sqrt{\delta_{\text{stat}}^2 + \delta_{\text{sys}}^2} \\ &= \sqrt{\delta_{\text{stat}}^2 + (\delta_{\text{sys}}^{\text{Im}})^2 + (\delta_{\text{sys}}^\lambda)^2 + (\delta_{\text{sys}}^{P^z})^2 + (\delta_{\text{sys}}^{a, m_\pi})^2}, \quad (26) \end{aligned}$$

where  $\delta_{\text{stat}}$  denotes the statistical uncertainty. The statistical uncertainty dominates at large  $b_\perp$  because the signal of quasi-TMDWF decays exponentially with increasing of spatial separation, while the noise, set by correlators of local operators, remains approximately constant, leading to a rapid deterioration of the SNR. Overall, the results shown

indicate that the uncertainties remain under control up to  $b_\perp \sim 1$  fm, this allows us to make reliable predictions for the CS kernel in the large- $b_\perp$  region.

Combining statistical and systematic uncertainties, Fig. 19 presents our results (black data points) for the CS kernel at  $\mu = 2$  GeV. As shown in the left panel, in the perturbative region where  $1/b_\perp \gtrsim 1$  GeV, our results agree well with perturbation theory up to three-loop accuracy [5,6]. In the nonperturbative region, we compare with previous first-principles lattice determinations, including LPC22 [45], ASWZ23 [48], ASWZ24 [50], DWF24 [79], and PionCG25 [80]. Notably, our results provide improved precision; even for  $b_\perp \gtrsim 1$  fm, the uncertainties remain under control.

In addition, we compare our results with those from global analyses, as shown in the right panel of Fig. 19. Specifically, we include SV19 [25], ART23 [27], IFY23 [22], EEC24 [29], MAPNN25 [32], ART25 [31], and CFR25 [33]. Usually, global analyses typically achieve

smaller uncertainties than direct lattice determinations by imposing explicit model parametrizations, in particular for the  $b_\perp$  dependence (see, e.g., Ref. [50]). Remarkably, our determination reaches a comparable precision, especially at large  $b_\perp$ , without introducing a dedicated model parametrization for the  $b_\perp$  dependence in the continuum-extrapolated CS kernel. In our analysis, the independence of a  $b_\perp$  Ansatz is enabled by the infinite-momentum extrapolation in Eq. (22), where we find no statistically significant sensitivity to the mixed discretization term proportional to  $aP^z$  and therefore neglect it at the current level of precision. The resulting agreement between our CS kernel and the global-fit determinations demonstrates full compatibility with phenomenological extractions within uncertainties.

### E. Some attempted global fits using phenomenology and lattice inputs

To quantitatively assess the potential influence of lattice data on phenomenological determinations of the CS kernel, we perform a global one-parameter fit combining phenomenological and lattice results. The analysis follows the standard perturbative-nonperturbative decomposition of the CS kernel [25,27,29,31],

$$K(b_\perp, \mu) = -2 \int_{\mu_*}^{\mu} \frac{d\mu'}{\mu'} \Gamma_{\text{cusp}}[\alpha_s(\mu')] - 2D_{\text{pert}}(b_*, \mu_*) - 2D_{\text{NP}}(b_\perp), \quad (27)$$

where  $\Gamma_{\text{cusp}}$  is the cusp anomalous dimension, and  $D_{\text{pert}}$  and  $D_{\text{NP}}$  denote the perturbative and nonperturbative contributions, respectively. The variable  $b_*(b_\perp) = b_\perp / \sqrt{1 + b_\perp^2 / b_{\text{max}}^2}$  and  $\mu_*(b_\perp) = 2e^{-\gamma_E} / b_*(b_\perp)$  are introduced to regulate the Landau pole [54,81], and we choose  $b_{\text{max}} = 2e^{-\gamma_E} \text{ GeV}^{-1}$ . In our analysis,  $D_{\text{pert}}(b_*, \mu_*)$  is evaluated at two-loop level [82,83], and the three-loop cusp anomalous dimension [84–87] is used in the evolution kernel to achieve a consistent perturbative accuracy.

For the nonperturbative part, we adopt the following models [25,25,88]:

$$\text{Model 1 : } D_{\text{NP}}(b_\perp) = g_2 b_\perp b_*, \quad (28)$$

$$\text{Model 2 : } D_{\text{NP}}(b_\perp) = g_2 b_\perp^2, \quad (29)$$

$$\text{Model 3 : } D_{\text{NP}}(b_\perp) = g_2 \ln(b_\perp / b_*), \quad (30)$$

which introduce a single free parameter  $g_2$  that governs the long-distance behavior of the CS kernel. This parametrization effectively captures the mild  $b_\perp$ -dependence of nonperturbative effects at large  $b_\perp$ , and has been widely used in recent global TMD analyses.

Using this model, we perform three fitting scenarios to assess how lattice inputs constrain the CS kernel:

TABLE III. Summary of the global one-parameter fits to the Collins-Soper kernel. The fitted parameter  $g_2$  and the corresponding  $\chi^2/\text{d.o.f.}$  are also listed.

|         | Fit scenario                      | $g_2$       | $\chi^2/\text{d.o.f.}$ |
|---------|-----------------------------------|-------------|------------------------|
| Model 1 | Only phenomenological extractions | 0.0702(28)  | 2.12                   |
|         | Lattice QCD                       | 0.0711(36)  | 0.80                   |
|         | Joint analysis                    | 0.0705(22)  | 1.46                   |
| Model 2 | Only phenomenological extractions | 0.1570(33)  | 3.34                   |
|         | Lattice QCD                       | 0.1435(46)  | 1.02                   |
|         | Joint analysis                    | 0.1521(27)  | 2.26                   |
| Model 3 | Only phenomenological extractions | 0.1972 (81) | 2.73                   |
|         | Lattice QCD                       | 0.210 (10)  | 0.97                   |
|         | Joint analysis                    | 0.1996 (64) | 1.86                   |

- (1) Exp only: fit using phenomenological extractions from global TMD analyses (SV19 [25], IFY23 [22], EEC24 [29], MAPNN25 [32]);
- (2) Lattice only: fit using recent lattice calculations (ASWZ23 [48], DWF24 [79], PionCG25 [80], LPC25); and
- (3) Combined: simultaneous fit including both phenomenological and lattice data.

To ensure a uniform comparison, all available determinations of the CS kernel are first interpolated with a fixed spacing of 0.1 fm over the range [0.2, 1.0] fm, and the global fits are then performed on the interpolated dataset. Each fitting scenario contains only one free parameter,  $g_2$ , and the quality of fit is characterized by the resulting  $\chi^2/\text{d.o.f.}$

As summarized in Table III, the inclusion of lattice results in the combined analysis produces a visible reduction in the statistical uncertainty of  $g_2$ , while the central value shifts moderately compared with the phenomenology-only fit. This demonstrates that lattice determinations provide complementary constraints on the long-distance behavior of the CS kernel, particularly in the large- $b_\perp$  region where phenomenological extractions alone tend to be less precise.

It should be pointed out that this study is only intended to illustrate the potential of fits that combine lattice data and phenomenological input. We fully acknowledge that doing so correctly is a highly subtle endeavor that will require a comprehensive joint analysis, far beyond the simple combination of individual results we have presented here. At present, we lack the capability to conduct such a rigorous analysis, which would also need to account for correlations among experimental datasets (such as in Ref. [88]). Our primary aim here is merely to highlight that incorporating lattice QCD input into global TMD analyses could represent a promising direction for future research.

#### IV. SUMMARY AND OUTLOOK

In order to achieve a first-principles determination of the Collins-Soper kernel on the lattice, we have carried out a comprehensive calculation within the framework of LaMET using five  $(2 + 1)$ -flavor QCD ensembles from CLQCD, with three different lattice spacings and four valence pion masses, and boosted the hadron momentum up to 3 GeV. The quasi-TMDWF matrix elements were extracted from carefully constructed nonlocal two-point correlation functions. The Wilson-loop renormalization procedure was employed to remove the linear divergence of the matrix elements. We have implemented a renormalization approach inspired by the self-renormalization, which allowed us to investigate in detail the logarithmic divergences of the quasi-TMDWF matrix elements and successfully match the lattice data to perturbative results in the  $\overline{\text{MS}}$  scheme.

To optimize the statistical precision of our simulations, we adopted HYP smearing for the gauge fields, and explicitly demonstrated that the renormalized matrix elements remain unaffected by the smearing procedure. By performing large- $\lambda$  extrapolations and Fourier transformations, we obtained the quasi-TMDWFs in momentum space. To further suppress systematic uncertainties from the imaginary part of the CS kernel, we adopted the  $b_\perp$ -unexpanded matching kernel at NLO. In addition, the large-momentum extrapolation of the CS kernel under different momentum combinations was used to systematically eliminate power corrections.

Finally, through a combined extrapolation across all four ensembles, we achieved, for the first time, a determination of the CS kernel up to 1 fm in the continuum limit and at the physical pion mass without introducing any explicit parametrization of its  $b_\perp$  dependence. All major sources of systematic uncertainty have been carefully analyzed and quantified, with their contributions explicitly presented. Our results are in good agreement with both previous lattice QCD calculations and phenomenological extractions from experimental data, demonstrating the reliability of this first-principles approach.

Looking ahead, the methods established in this work provide a solid foundation for further lattice QCD studies of TMD observables. With increasing computational resources and future extensions to higher statistics, finer lattice spacings, and larger hadron momenta, it will become possible to achieve even more precise determinations of the CS kernel and related quantities. Such efforts will play an essential role in connecting lattice QCD, perturbative QCD, and experimental measurements, thereby deepening our understanding of the three-dimensional structure of hadrons.

The computations in this paper were run on the Siyuan-1 cluster supported by the Center for High Performance Computing at Shanghai Jiao Tong University, and Advanced Computing East China Sub-center. The LQCD simulations were performed using the PyQUDA software suite [89] and QUDA [90–92] through HIP programming model [93].

#### ACKNOWLEDGMENTS

We thank Huey-Wen Lin, Tian-Bo Liu, and Yong Zhao for useful discussions. We thank the CLQCD Collaboration for providing us the gauge configurations with dynamical fermions [65], which are generated on the HPC Cluster of ITP-CAS, the Southern Nuclear Science Computing Center (SNSC), and the Siyuan-1 cluster supported by the Center for High Performance Computing at Shanghai Jiao Tong University and the Dongjiang Yuan Intelligent Computing Center. This work is supported in part by Natural Science Foundation of China under Grants No. 12375069, No. 12125503, No. 12305103, No. 12565014, No. 12405101, No. 12293060, No. 12205106, No. 12575084, No. 12293062, No. 12525504, No. 12435002, and No. 12447101 by the Fundamental Research Funds for the Central Universities. Q. A. Z. is also supported by the Fundamental Research Funds for the Central Universities. J. H. is also supported by Guang-dong Major Project of Basic and Applied Basic Research No. 2025A1515012199. H. L. is also supported by the GHfund B (No. 202407028788). Y. Y. is also supported by supported in part by National Key R&D Program of China No. 2024YFE0109800. J. L. Z. is also supported by the Tsung-Dao Lee scholarship. J. Z. is also supported by the Talent Research Startup Foundation of Hainan Normal University Grant No. HSZK-KYQD-202523, and Opening Foundation of Shanghai Key Laboratory of Particle Physics and Cosmology under Grant No. 22DZ2229013-5.

#### DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

#### APPENDIX A: SELF-RENORMALIZATION

The RGEs allow us to establish the factorization between the UV divergence scale (represented by  $\mu$  or  $1/a$ ) and the relevant physical scales. This leads to the definition of a multiplicative renormalization factor  $Z_O$ , which serves to convert the lattice results into the  $\overline{\text{MS}}$  scheme [94]. Building upon these first-principle insights, phenomenologically motivated renormalization strategies are often employed in practical lattice computations to effectively incorporate such conversions.

We need to eliminate the  $a$ -dependence of the subtracted quasi-TMDWF matrix elements as defined in Eq. (18), and match the results to those in the  $\overline{\text{MS}}$  scheme. Motivated by the idea of self-renormalization, we propose the following new scheme to extract the renormalization factor  $Z_O$  [62].

When the external hadron momentum is zero, the renormalization condition for the subtracted quasi-TMDWFs can be expressed as [61,95]

$$\begin{aligned} \tilde{\Phi}(z, P^z = 0, b_\perp, 1/a, m) \\ = Z_O(1/a, \mu) \times \tilde{\Phi}^{\overline{\text{MS}}}(\mu, z, P^z = 0, b_\perp, m), \end{aligned} \quad (\text{A1})$$

where the finite quark mass  $m$  is introduced to regulate the IR divergence. The perturbative result up to one loop for zero momentum matrix element can be found in Ref. [95], as follows:

$$\begin{aligned} \tilde{\Phi}^{\overline{\text{MS}}}(\mu, z, P^z = 0, b_\perp, m) \\ \equiv \tilde{\Phi}_0^{\overline{\text{MS}}}(\mu, z, b_\perp) \\ = 1 + \frac{\alpha_{\overline{\text{MS}}}(\mu) C_F}{2\pi} \left[ \frac{3}{2} \ln \frac{\mu^2(z^2 + b_\perp^2)}{4e^{-2\gamma_E}} + \frac{1}{2} - \frac{2z}{b_\perp} \arctan \frac{z}{b_\perp} \right]. \end{aligned} \quad (\text{A2})$$

The parametrization of the matrix element  $\tilde{\Phi}$  is written as [62]

$$\begin{aligned} \ln \tilde{\Phi}(z, P^z = 0, b_\perp, 1/a) \\ = \frac{\gamma_0}{\beta_0} \ln[\ln[1/(a\Lambda_{\text{QCD}})]] + \frac{c_1}{\ln[1/(a\Lambda_{\text{QCD}})]} + g(z, b_\perp), \end{aligned} \quad (\text{A3})$$

where the first two terms correspond to logarithmic divergences, while  $g(z, b_\perp)$  denotes the intrinsic nonperturbative physics.

It should be noted that, for clarity of presentation, the lattice data and fit results shown below are based on data with HYP smearing once (HYP1). Results for HYP0, HYP2, and others for HYP1 are presented in Appendix D 2. After

interpolating the lattice data at three different lattice spacings to a common spacing of  $a' = 0.05$  fm, a joint fit allows us to constrain the form of the logarithmic divergence. In the fit, we fix  $\mu = 2$  GeV and  $\Lambda_{\text{QCD}} = 0.2$  GeV, and we introduce the associated systematic uncertainty

$$\delta_{\text{all}}^{\ln \tilde{\Phi}} = \sqrt{(\delta_{\text{stat}}^{\ln \tilde{\Phi}})^2 + (\delta_{\text{sys}} a \mu)^2}, \quad (\text{A4})$$

where we set  $\delta_{\text{sys}} = 0.007$ , and from the joint fit for all lattice matrix elements ( $b_\perp \geq 0.1$  fm), we obtain  $c_1 = 0.020(15)$  with  $\chi^2/\text{d.o.f.} = 1.0$ .

Besides taking into account that discretization effects may play a role in the region of small  $b_\perp$  and  $z$ , we fit the lattice data in the perturbative range using the following formula ( $\sqrt{z^2 + b_\perp^2} \leq 0.2$  fm), as follows:

$$\begin{aligned} \ln \tilde{\Phi}(z, P^z = 0, b_\perp, 1/a) \\ = \frac{\gamma_0}{\beta_0} \ln[\ln[1/(a\Lambda_{\text{QCD}})]] + \frac{c_1}{\ln[1/(a\Lambda_{\text{QCD}})]} \\ + g'(z, b_\perp) + f(z, b_\perp) a^2, \end{aligned} \quad (\text{A5})$$

where the  $g'(z, b_\perp)$  term will correspond to the result of the scheme  $\overline{\text{MS}}$ , and the  $f(z, b_\perp)$  term denotes the contribution of the discretization error. In addition, considering the renormalization factor  $Z_O(1/a, \mu)$  can be parametrized as

$$\begin{aligned} \ln Z_O(1/a, \mu) = \frac{\gamma_0}{\beta_0} \ln[\ln[1/(a\Lambda_{\text{QCD}})]] + \frac{c_1}{\ln[1/(a\Lambda_{\text{QCD}})]} \\ + d'(\mu), \end{aligned} \quad (\text{A6})$$

where the  $d'(\mu)$  term absorbs all the scale dependence in  $Z_O$ . In order to match  $g'$  to the zero-momentum matrix element calculated in the  $\overline{\text{MS}}$  scheme, we include the parameter  $d'$  in the fit within the perturbative regime and tune its value such that

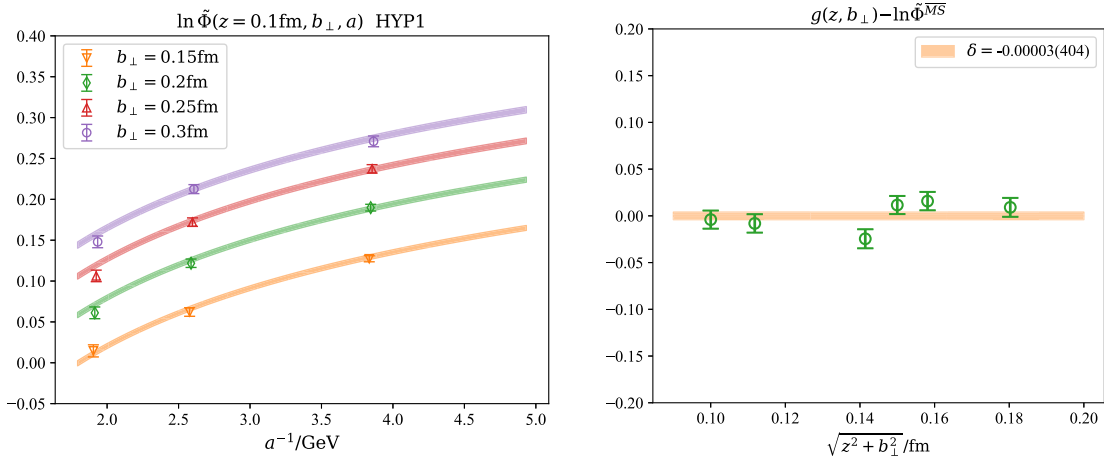


FIG. 20. Left panel: result of joint fit for the self-renormalization, taking  $b_\perp = \{0.15, 0.2, 0.25, 0.3\}$  fm and  $z = 0.1$  fm as an example. Right panel: matching result of the  $g(z, b_\perp)$  to the  $\overline{\text{MS}}$  scheme with HYP1.

$$\exp[g'(z, b_\perp) - d'(\mu)] = \tilde{\Phi}^{\overline{\text{MS}}}(\mu, z, P^z = 0, b_\perp, m). \quad (\text{A7})$$

As shown in the right panel of Fig. 20, when  $d'(\mu) = -0.448$ , the average deviation  $\delta$  between the  $g'(z, b_\perp)$  term and the  $\overline{\text{MS}}$  result is consistent with zero within uncertainties. This allows us to determine the form of the logarithmic divergence, eliminate discretization effects, and ensure consistency between the lattice results and the  $\overline{\text{MS}}$  results in the perturbative region.

Finally, after determining the form of the logarithmic divergence and the renormalization scale dependence, the results for different lattice spacings and HYP smearing steps are obtained, as shown in Table II.

### APPENDIX B: CONVENTION ON THE LENGTH OF THE WILSON LINE IN LATTICE SIMULATION

In lattice simulation, we adopt the convention on the length of the Wilson line shown in Fig. 21, where the direction of the relative displacement between the quark and antiquark is taken as the  $z$ -direction, while maintaining the requirement of a large Wilson link length  $L$ . The left panel illustrates the convention when  $z > 0$ , which follows directly from Eq. (7). The right panel shows the convention for  $z < 0$ . To ensure that the Wilson link remains sufficiently long in this case, one can redefine the length as  $L' = L + |z| \geq L$ , which is also consistent with Eq. (7). With this convention, the Wilson loop has a total longitudinal extent of  $(2L + |z|)$  for both positive and negative  $z$ , and a transverse width of  $b_\perp$ . This convention guarantees

that the Wilson line is sufficiently long to ensure the saturation of the final results for quasi-TMDWFs.

### APPENDIX C: SYMMETRY IN QUASI-TMDWFs

In this Appendix, we will show that the quasi-TMDWF of the pion is symmetric around  $x = 1/2$  under the assumption of isospin symmetry. To make it transparent, we label the color and spinor indices explicitly, and the staple-shaped Wilson link is denoted as  $U_{ab}(z\hat{n}_z/2 + b_\perp\hat{n}_\perp, -z\hat{n}_z/2)$ . So the bare matrix element is now expressed as

$$\tilde{\Phi}^0(z, b_\perp, P^z, L) = \langle 0 | \bar{q}_{i,b}(z\hat{n}_z/2 + b_\perp\hat{n}_\perp) \Gamma_{ij} U_{ab}(z\hat{n}_z/2 + b_\perp\hat{n}_\perp, -z\hat{n}_z/2) \psi_{j,b}(-z\hat{n}_z/2) | P^z \rangle, \quad (\text{C1})$$

where  $a, b$  are color indices in the fundamental representation and  $i, j$  are spinor indices. To proceed, we need the charge conjugation of the Wilson line,

$$\begin{aligned} CU_{ab}(x, 0)C^{-1} &= (e^{igP \int_0^x n \cdot CA(nt) C^{-1} dt})_{ab} \\ &= (e^{-igP \int_0^x n \cdot A^T(nt) dt})_{ab} \\ &= (e^{ig\bar{P} \int_0^x (-n) \cdot A^T(x-nt) dt})_{ab} \\ &= (e^{igP \int_0^x (-n) \cdot A(x-nt) dt})_{ba} \\ &= U_{ba}(0, x), \end{aligned} \quad (\text{C2})$$

where  $C$  is the charge conjugation operator. The effect of charge conjugation is basically to reverse the direction of the Wilson line.

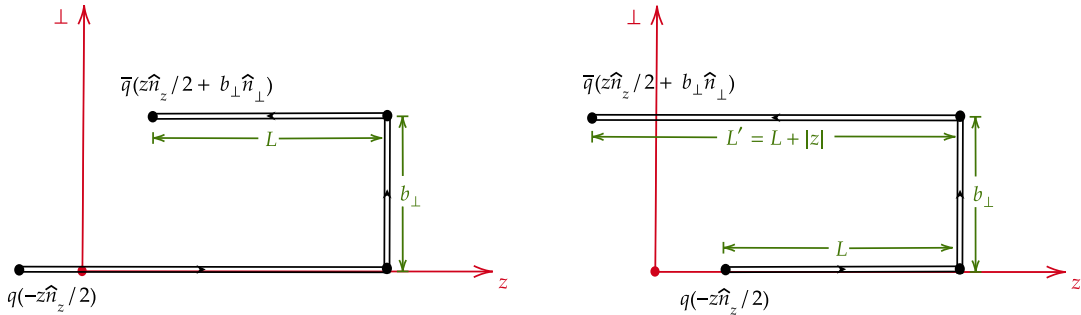


FIG. 21. Construction of the Wilson link in the lattice simulation for  $z > 0$  (left panel) and  $z < 0$  (right panel).

Then, we insert  $I = C^{-1}C$  into Eq. (C1) and obtain

$$\begin{aligned}
 \tilde{\Phi}^0(z, b_\perp, P^z, L) &= \langle 0 | C^{-1} C \bar{q}_{i,b}(z\hat{n}_z/2 + b_\perp \hat{n}_\perp) C^{-1} \Gamma_{ij} C U_{ab}(z\hat{n}_z/2 + b_\perp \hat{n}_\perp, -z\hat{n}_z/2) C^{-1} C q_{j,b}(-z\hat{n}_z/2) C^{-1} C | P^z \rangle, \\
 &= -\langle 0 | q_{i,a}(z\hat{n}_z/2 + b_\perp \hat{n}_\perp) C_{ij} \Gamma_{jk} U_{ba}(-z\hat{n}_z/2, z\hat{n}_z/2 + b_\perp \hat{n}_\perp) C_{kl}^{-1} \bar{q}_{l,b}(-z\hat{n}_z/2) | P^z \rangle, \\
 &= \langle 0 | \bar{q}_{l,b}(-z\hat{n}_z/2) C_{ij} \Gamma_{jk} C_{kl}^{-1} U_{ba}(-z\hat{n}_z/2, z\hat{n}_z/2 + b_\perp \hat{n}_\perp) q_{i,a}(z\hat{n}_z/2 + b_\perp \hat{n}_\perp) | P^z \rangle, \\
 &= \langle 0 | \bar{q}_{l,b}(-z\hat{n}_z/2) \Gamma'_{li} U_{ba}(-z\hat{n}_z/2, z\hat{n}_z/2 + b_\perp \hat{n}_\perp) q_{i,a}(z\hat{n}_z/2 + b_\perp \hat{n}_\perp) | P^z \rangle,
 \end{aligned} \tag{C3}$$

where

$$\Gamma' = (C \Gamma C^{-1})^T = (C \gamma' C^{-1} C \gamma_5 C^{-1})^T = -(\gamma'^T \gamma_5^T)^T = \gamma' \gamma_5 = \Gamma. \tag{C4}$$

In the third line of Eq. (C6), we use the fact that

$$C|\pi^0\rangle = |\pi^0\rangle, \quad C|\pi^-\rangle = |\pi^+\rangle, \quad C|\pi^+\rangle = |\pi^-\rangle. \tag{C5}$$

For  $\pi^+$  and  $\pi^-$ , further using the isospin symmetry brings us back to the original meson.

Furthermore, we exploit the translational and rotational invariance in the  $\hat{n}_\perp$  plane, as follows:

$$\begin{aligned}
 \tilde{\Phi}^0(z, b_\perp, P^z, L) &= \langle 0 | \bar{q}_{l,b}(-z\hat{n}_z/2 - b_\perp \hat{n}_\perp) \Gamma'_{li} U_{ba}(-z\hat{n}_z/2 - b_\perp \hat{n}_\perp, z\hat{n}_z/2) q_{i,a}(z\hat{n}_z/2) | P^z \rangle, \\
 &= \langle 0 | \bar{q}_{l,b}(-z\hat{n}_z/2 + b_\perp \hat{n}_\perp) \Gamma'_{li} U_{ba}(-z\hat{n}_z/2 + b_\perp \hat{n}_\perp, z\hat{n}_z/2) q_{i,a}(z\hat{n}_z/2) | P^z \rangle, \\
 &= \tilde{\Phi}^0(-z, b_\perp, P^z, L).
 \end{aligned} \tag{C6}$$

The renormalized matrix elements should have the same symmetry property as the bare matrix element, since the renormalization factor  $Z_O(1/a, \mu)$  and the Wilson loop  $Z_E(2L + z, b_\perp)$  are the same under  $z \leftrightarrow -z$ . Therefore, we arrive at

$$\tilde{\Phi}(z, b_\perp, P^z, L) = \tilde{\Phi}(-z, b_\perp, P^z, L). \tag{C7}$$

Transforming to momentum space, we have

$$\begin{aligned}
 \tilde{f}(x, b_\perp, \mu, \zeta_z) &= \lim_{L \rightarrow \infty} \int \frac{dz}{2\pi} e^{i(x-1/2)zP^z} \tilde{\Phi}(z, b_\perp, P^z, L), \\
 &= \lim_{L \rightarrow \infty} \int \frac{dz}{2\pi} e^{i(x-1/2)zP^z} (\tilde{\Phi}(z, b_\perp, P^z, L) + \tilde{\Phi}(-z, b_\perp, P^z, L))/2, \\
 &= \lim_{L \rightarrow \infty} \int \frac{dz}{2\pi} e^{i(1/2-x)zP^z} (\tilde{\Phi}(-z, b_\perp, P^z, L) + \tilde{\Phi}(z, b_\perp, P^z, L))/2, \\
 &= \lim_{L \rightarrow \infty} \int \frac{dz}{2\pi} e^{i(1-x-1/2)zP^z} \tilde{\Phi}(z, b_\perp, P^z, L), \\
 &= \tilde{f}(1-x, b_\perp, \mu, \zeta_z),
 \end{aligned} \tag{C8}$$

which implies that the quasi-TMDWF is symmetric around  $x = 1/2$  under isospin symmetry.

## APPENDIX D: MORE RESULTS FOR SEC. III

## 1. Results for Wilson loop

More results for the Wilson loop on different lattice ensembles are given in Fig. 22.

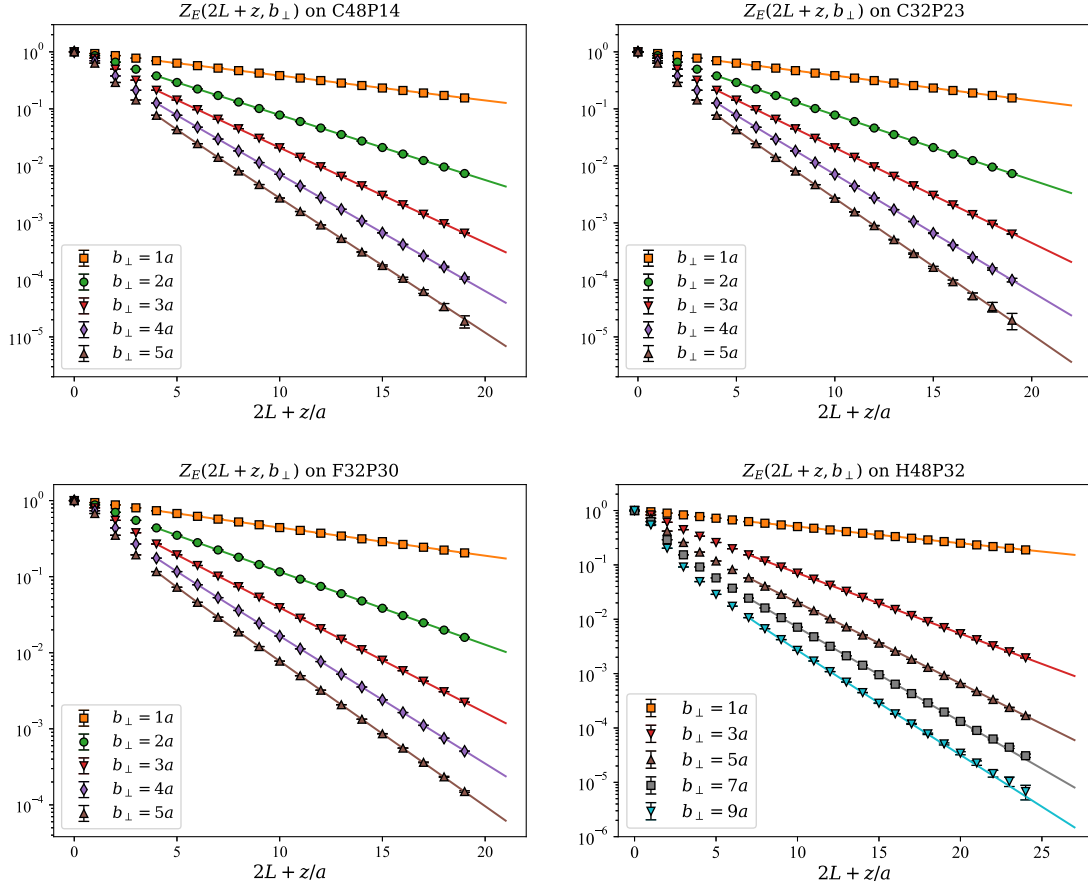


FIG. 22. Results of the Wilson loop for different shapes  $(2L+z) \times b_\perp$  on the different ensembles, with three panels showing a comparison of the results from HYP1. The solid lines represent the extrapolated results based on Eq. (16).

## 2. Results for self-renormalization

More results for the self-renormalization on different HYP smearing steps are given in Figs. 23–26.

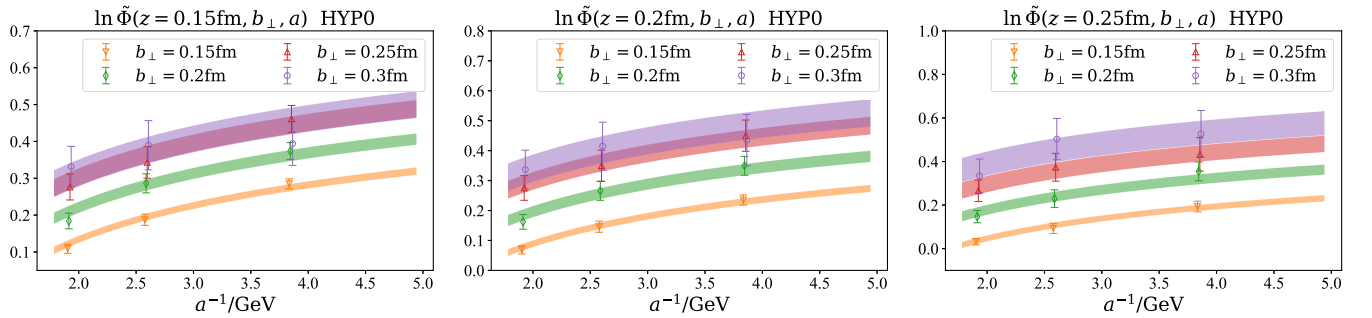
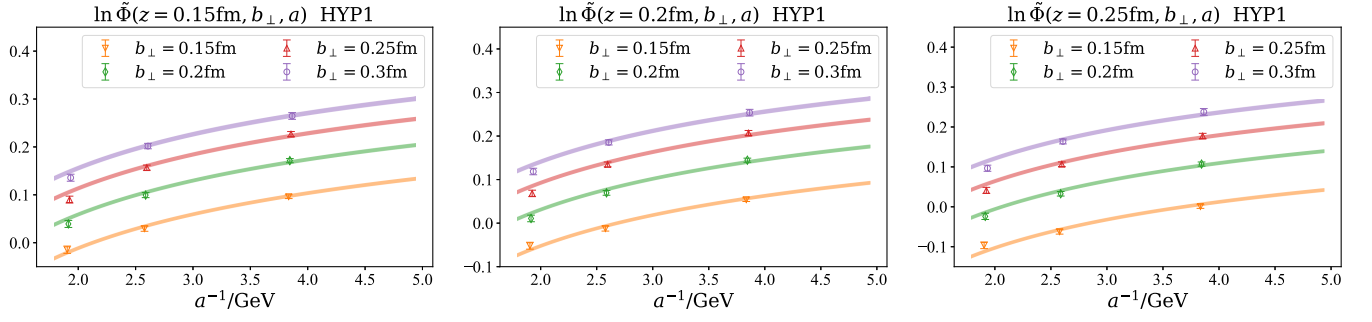
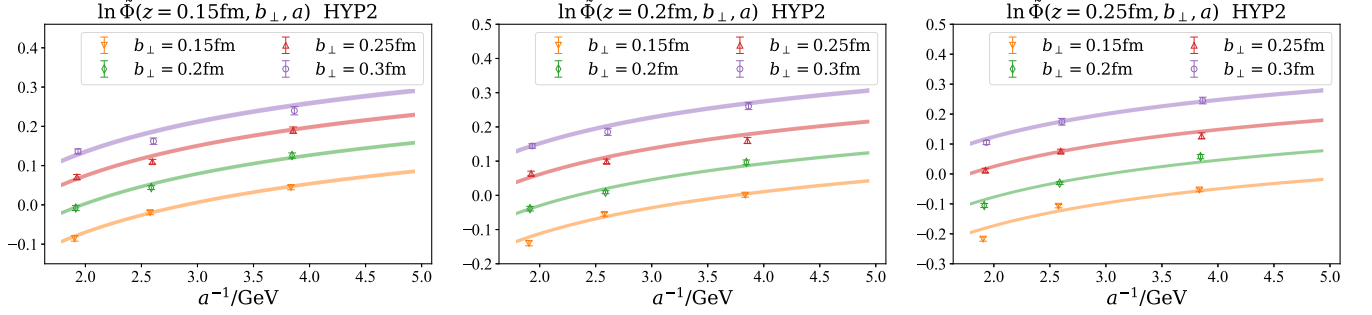
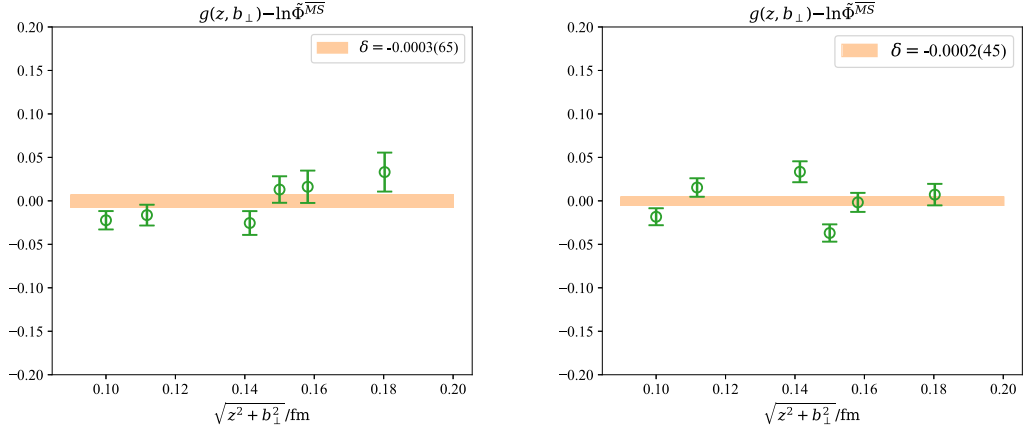


FIG. 23. Result of joint fit for the self-renormalization with HYP0,  $b_\perp = \{0.15, 0.2, 0.25, 0.3\}$  fm, and  $z = \{0.15, 0.2, 0.25\}$  fm.

FIG. 24. Result of joint fit for the self-renormalization with HYP1,  $b_{\perp} = \{0.15, 0.2, 0.25, 0.3\} \text{ fm}$ , and  $z = \{0.15, 0.2, 0.25\} \text{ fm}$ .FIG. 25. Result of joint fit for the self-renormalization with HYP2,  $b_{\perp} = \{0.15, 0.2, 0.25, 0.3\} \text{ fm}$ , and  $z = \{0.15, 0.2, 0.25\} \text{ fm}$ .FIG. 26. Result of matching to the  $\overline{\text{MS}}$  scheme for the self-renormalization with HYP0 (left panel) and HYP 2 (right panel).

### 3. Results for fits to the bare quasi-TMDWF matrix elements

More results for the fits to the bare quasi-TMDWF matrix elements on different ensembles are given in Figs. 27–30.

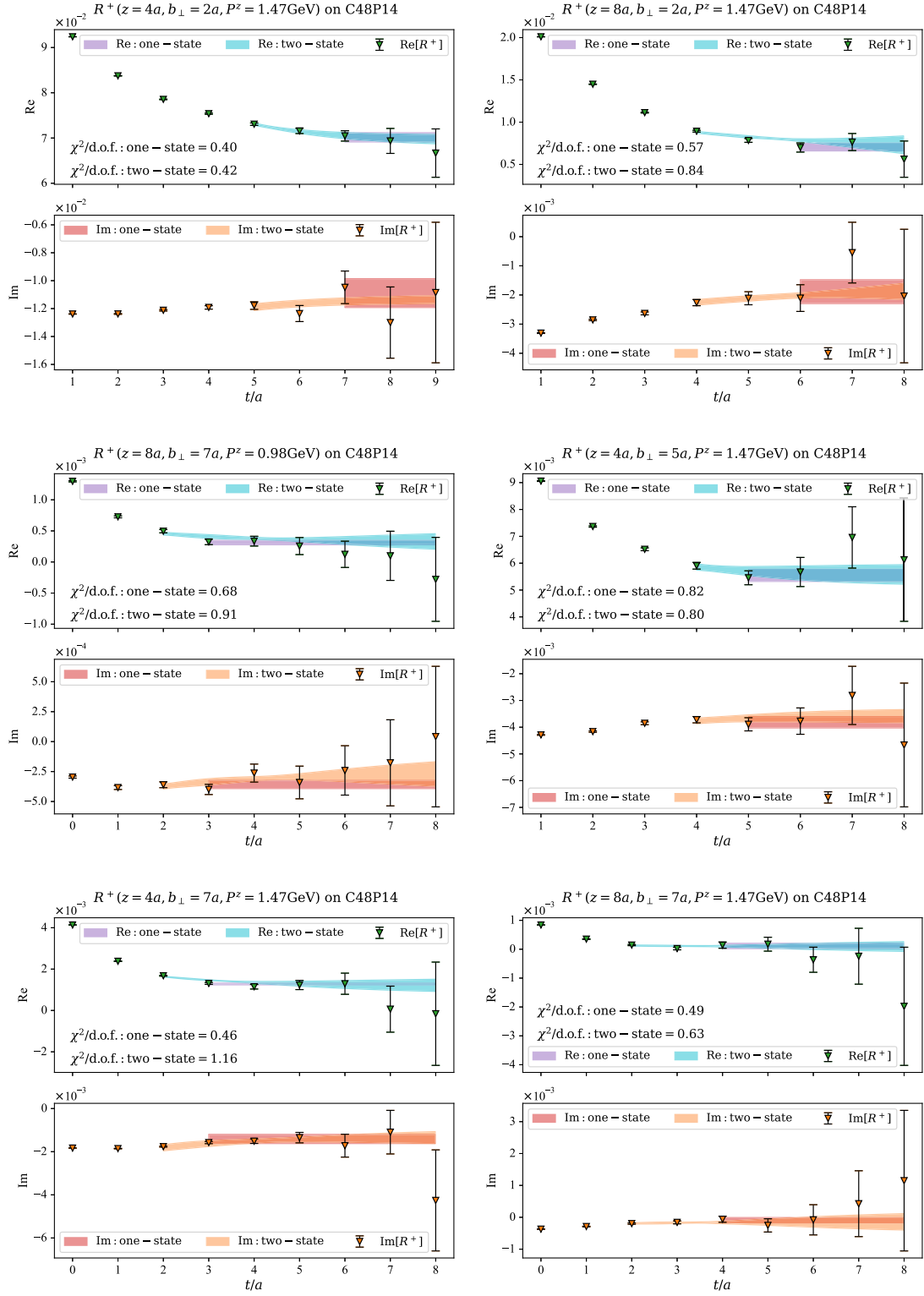


FIG. 27. Results from the C48P14 ensemble with one-step HYP smearing, and comparison of the combinations for  $\{z, b_\perp, P^z\} = \{4a, 2a, 1.47\text{ GeV}\}$ ,  $\{8a, 2a, 1.47\text{ GeV}\}$ ,  $\{8a, 7a, 0.98\text{ GeV}\}$ ,  $\{4a, 5a, 1.47\text{ GeV}\}$ ,  $\{4a, 7a, 1.47\text{ GeV}\}$ , and  $\{8a, 7a, 1.47\text{ GeV}\}$ .

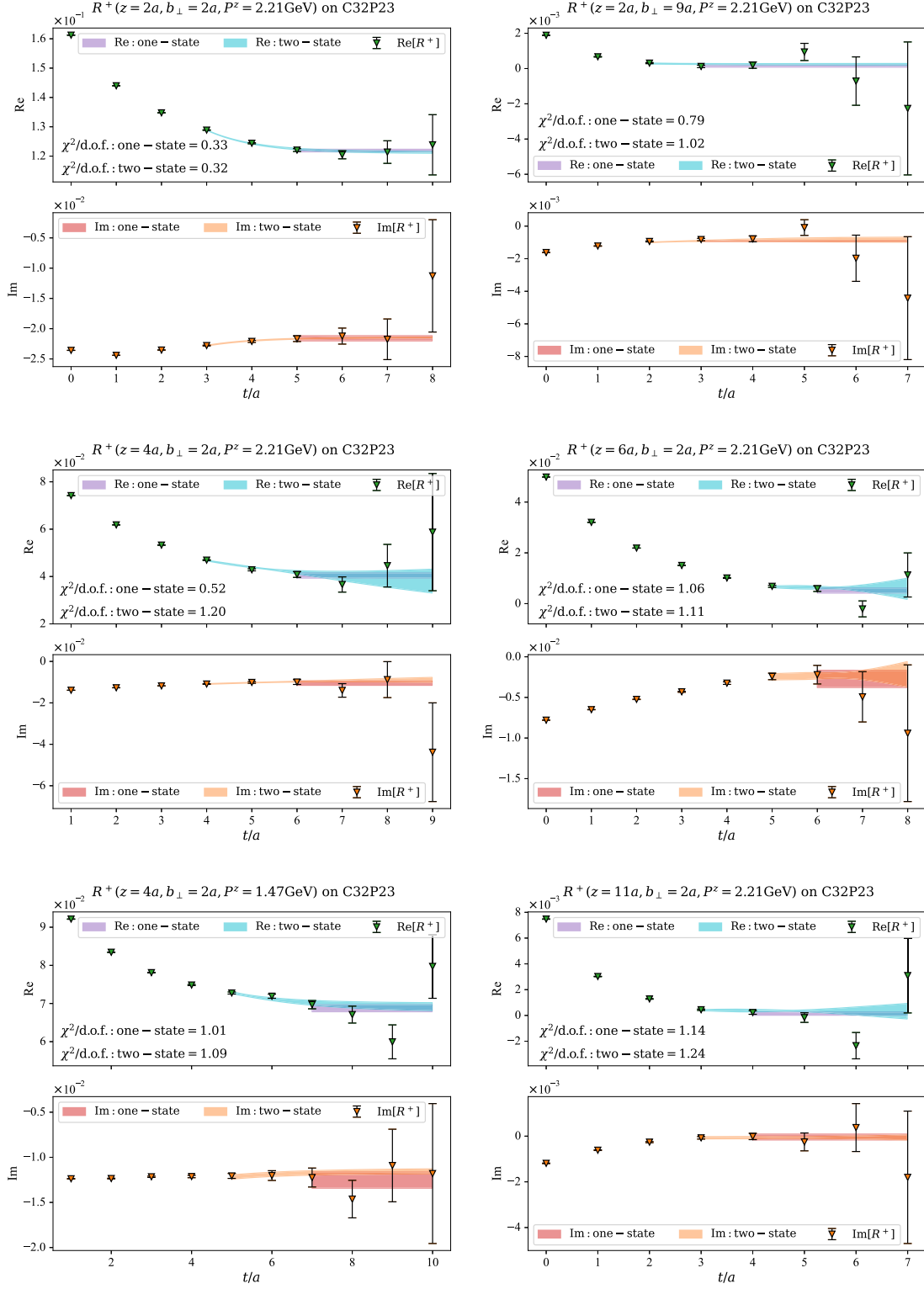


FIG. 28. Results from the C32P23 ensemble with one-step HYP smearing, and comparison of the combinations for  $\{z, b_\perp, P^z\} = \{2a, 2a, 2.21 \text{ GeV}\}, \{2a, 9a, 2.21 \text{ GeV}\}, \{4a, 2a, 2.21 \text{ GeV}\}, \{6a, 2a, 2.21 \text{ GeV}\}, \{4a, 2a, 1.47 \text{ GeV}\},$  and  $\{11a, 2a, 2.21 \text{ GeV}\}.$

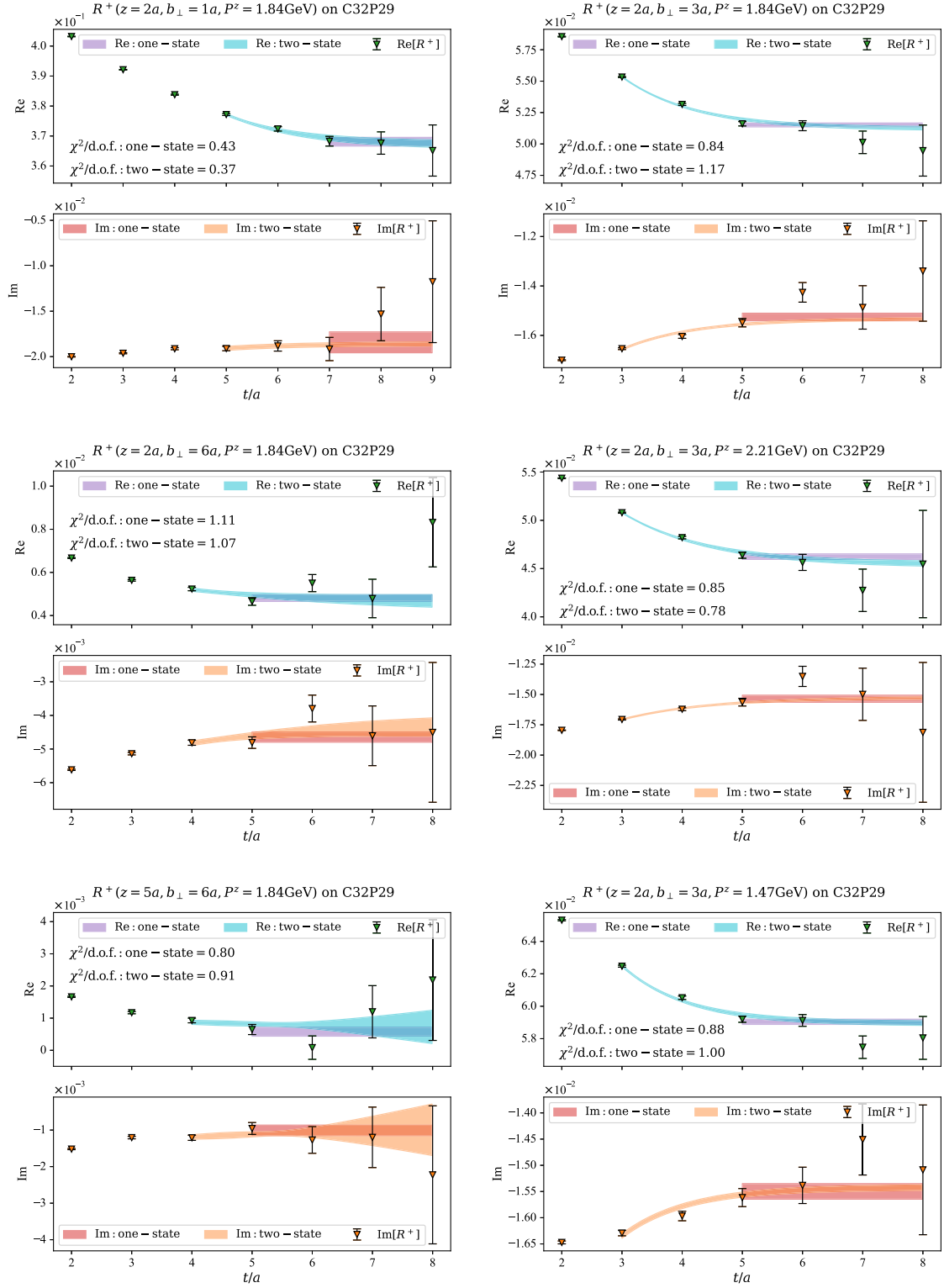


FIG. 29. Results from the C32P29 ensemble with one-step HYP smearing, and comparison of the combinations for  $\{z, b_{\perp}, P^z\} = \{2a, a, 1.84 \text{ GeV}\}, \{2a, 3a, 1.84 \text{ GeV}\}, \{2a, 6a, 1.84 \text{ GeV}\}, \{2a, 3a, 2.21 \text{ GeV}\}, \{5a, 6a, 1.84 \text{ GeV}\},$  and  $\{2a, 3a, 1.47 \text{ GeV}\}$ .

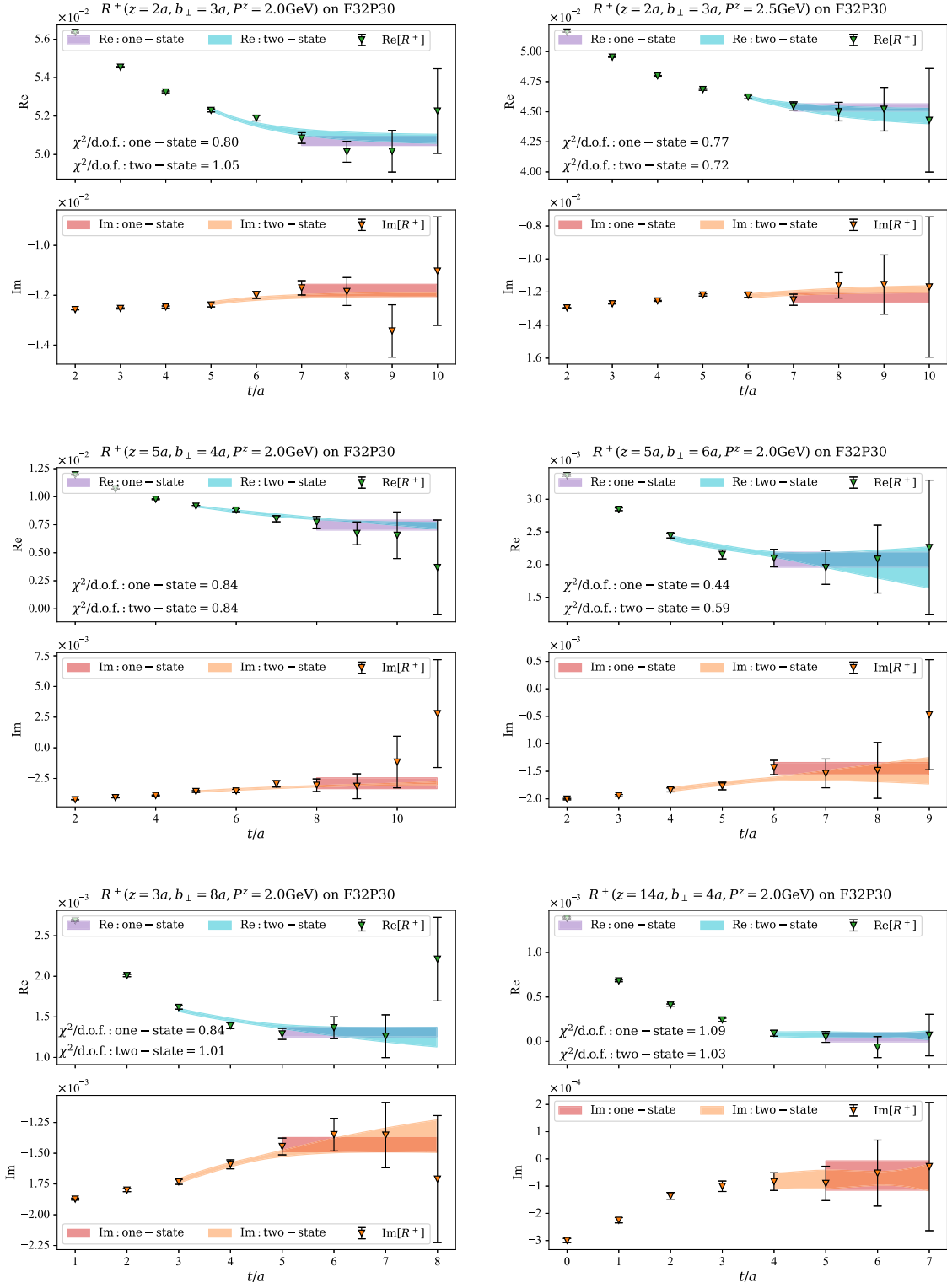


FIG. 30. Results from the F32P30 ensemble with one-step HYP smearing, and comparison of the combinations for  $\{z, b_{\perp}, P^z\} = \{2a, 3a, 2.0 \text{ GeV}\}$ ,  $\{2a, 3a, 2.5 \text{ GeV}\}$ ,  $\{5a, 4a, 2.0 \text{ GeV}\}$ ,  $\{5a, 6a, 2.0 \text{ GeV}\}$ ,  $\{3a, 8a, 2.0 \text{ GeV}\}$ , and  $\{14a, 4a, 2.0 \text{ GeV}\}$ .

#### 4. Results for renormalized quasi-TMDWF matrix elements

More results for the renormalized quasi-TMDWF matrix elements on different ensembles are given in Figs. 31–34.

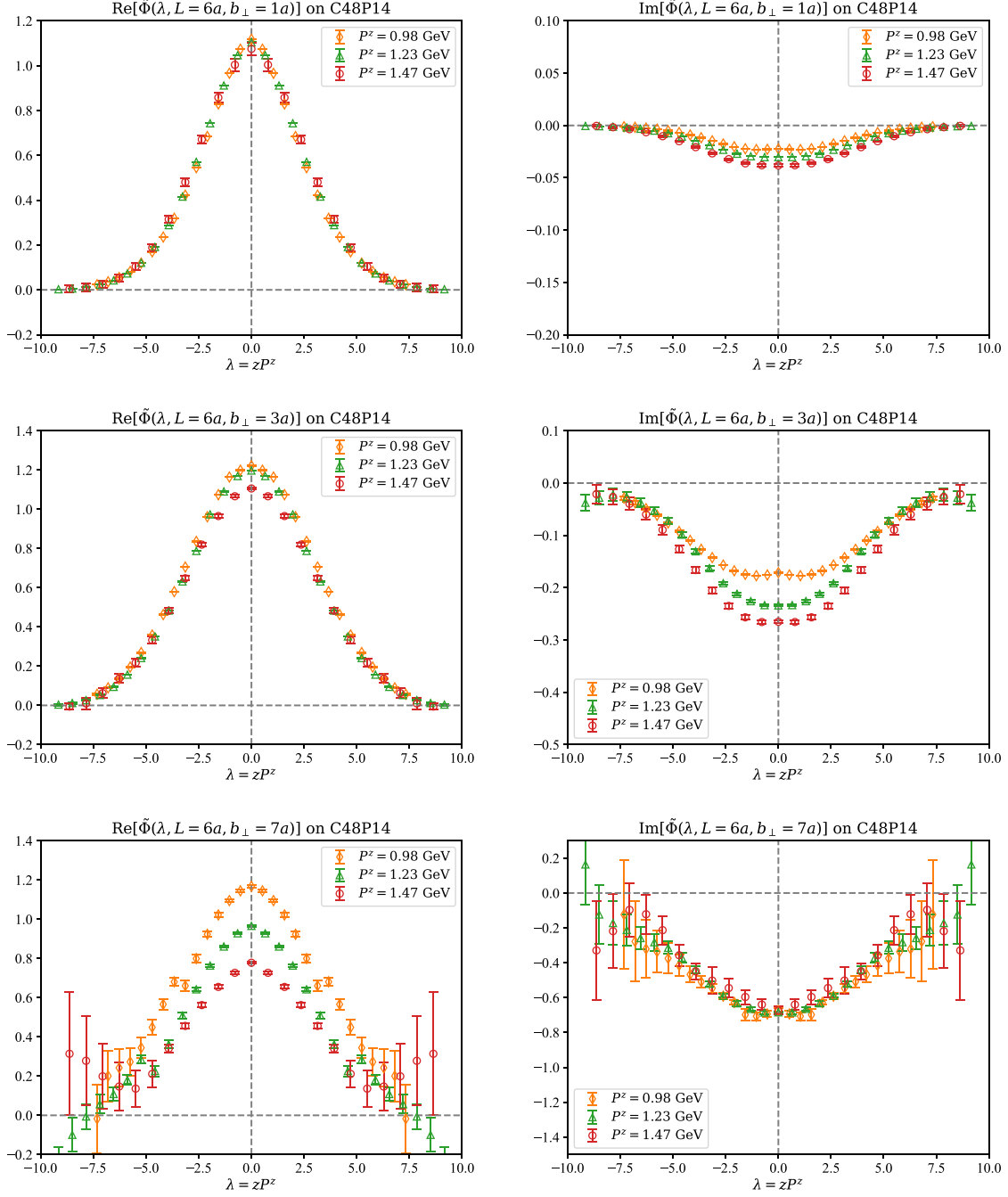


FIG. 31. Results from the C48P14 ensemble. A comparison of the results for  $P^z = \{0.98, 1.23, 1.47\}$  GeV at different  $b_\perp$  values is presented in the three panels.

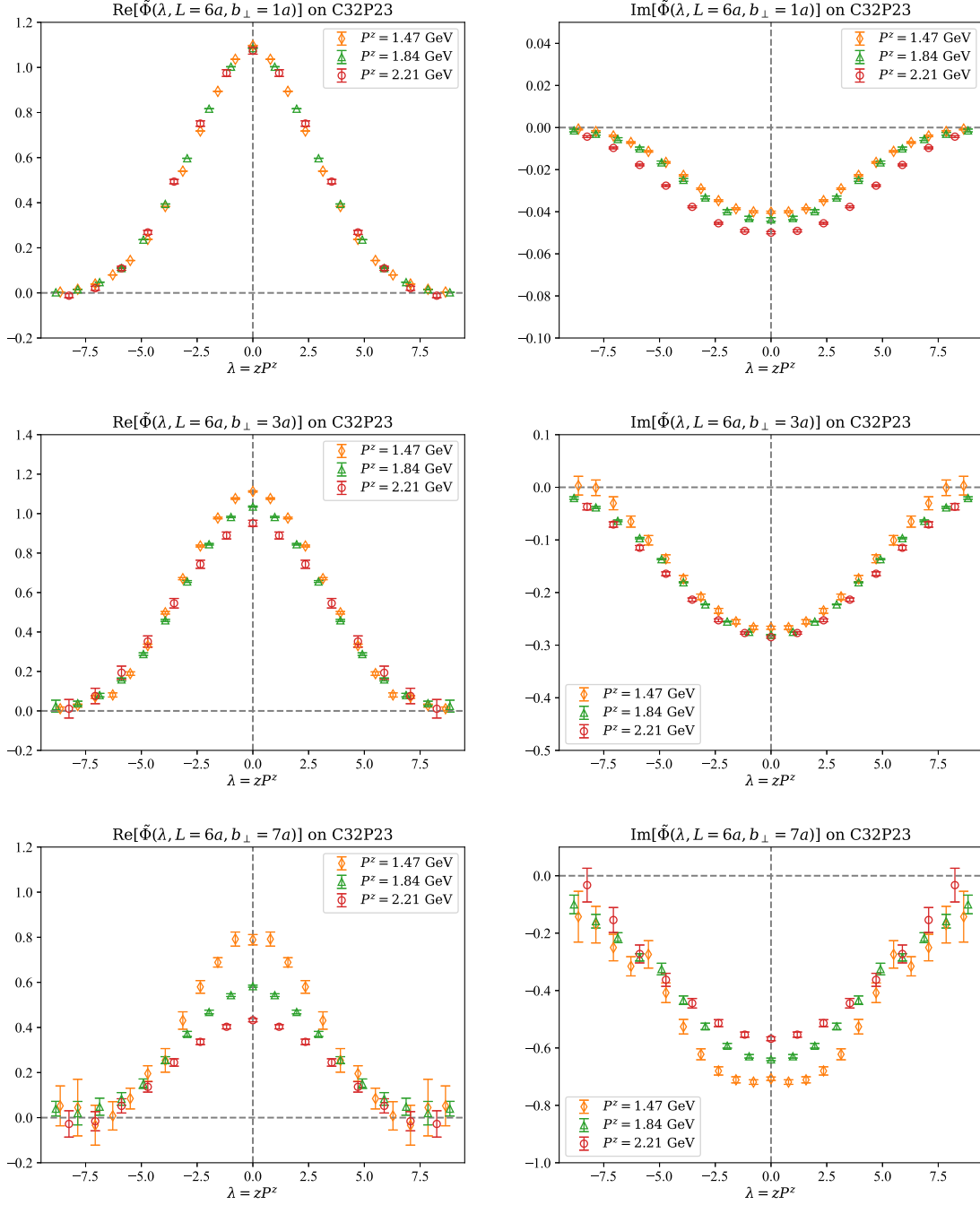


FIG. 32. Results from the C32P23 ensemble. A comparison of the results for  $P^z = \{1.47, 1.84, 2.21\}$  GeV at different  $b_{\perp}$  values is presented in the three panels.

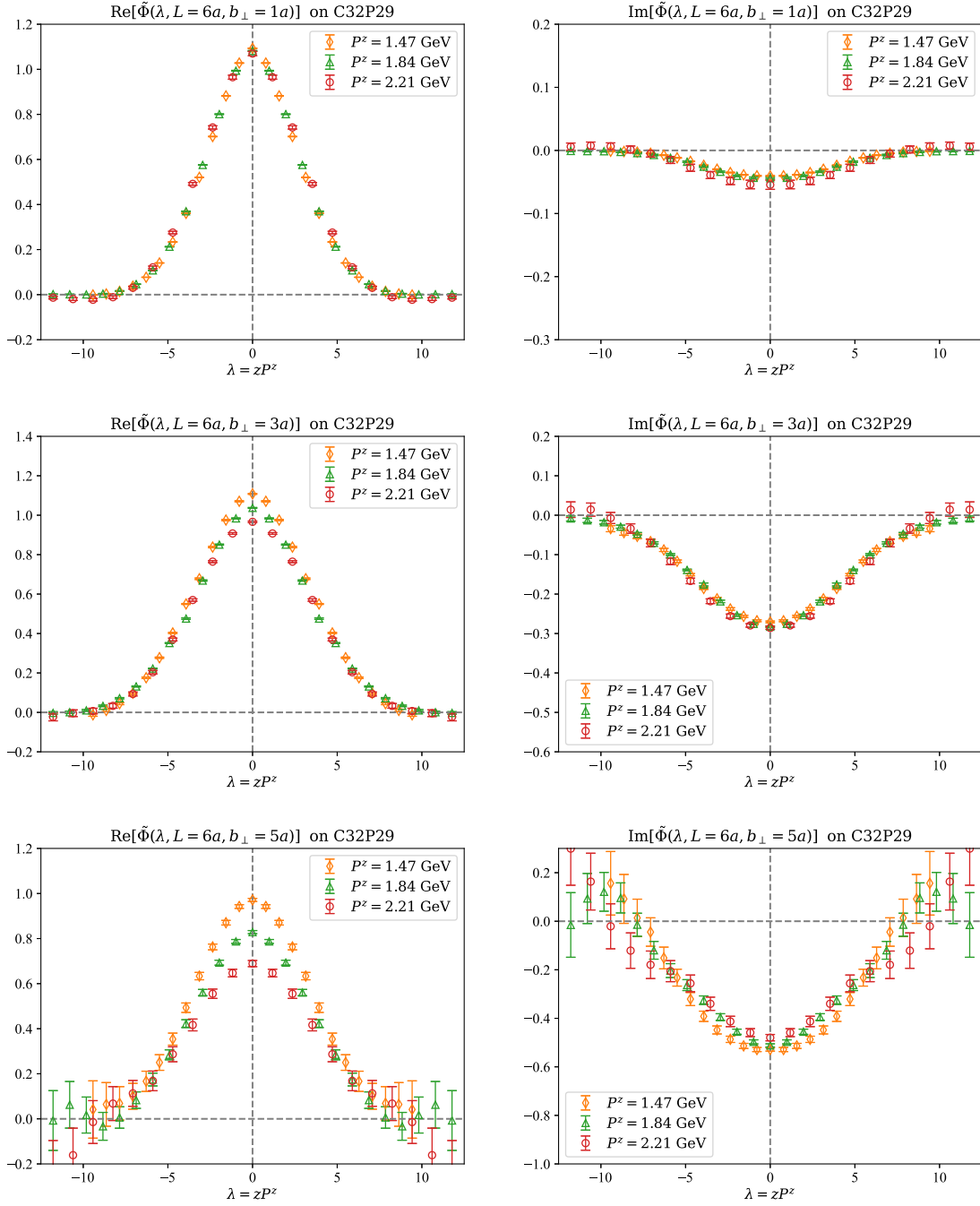


FIG. 33. Results from the C32P29 ensemble. A comparison of the results for  $P^z = \{1.47, 1.84, 2.21\}$  GeV at different  $b_\perp$  values is presented in the three panels.

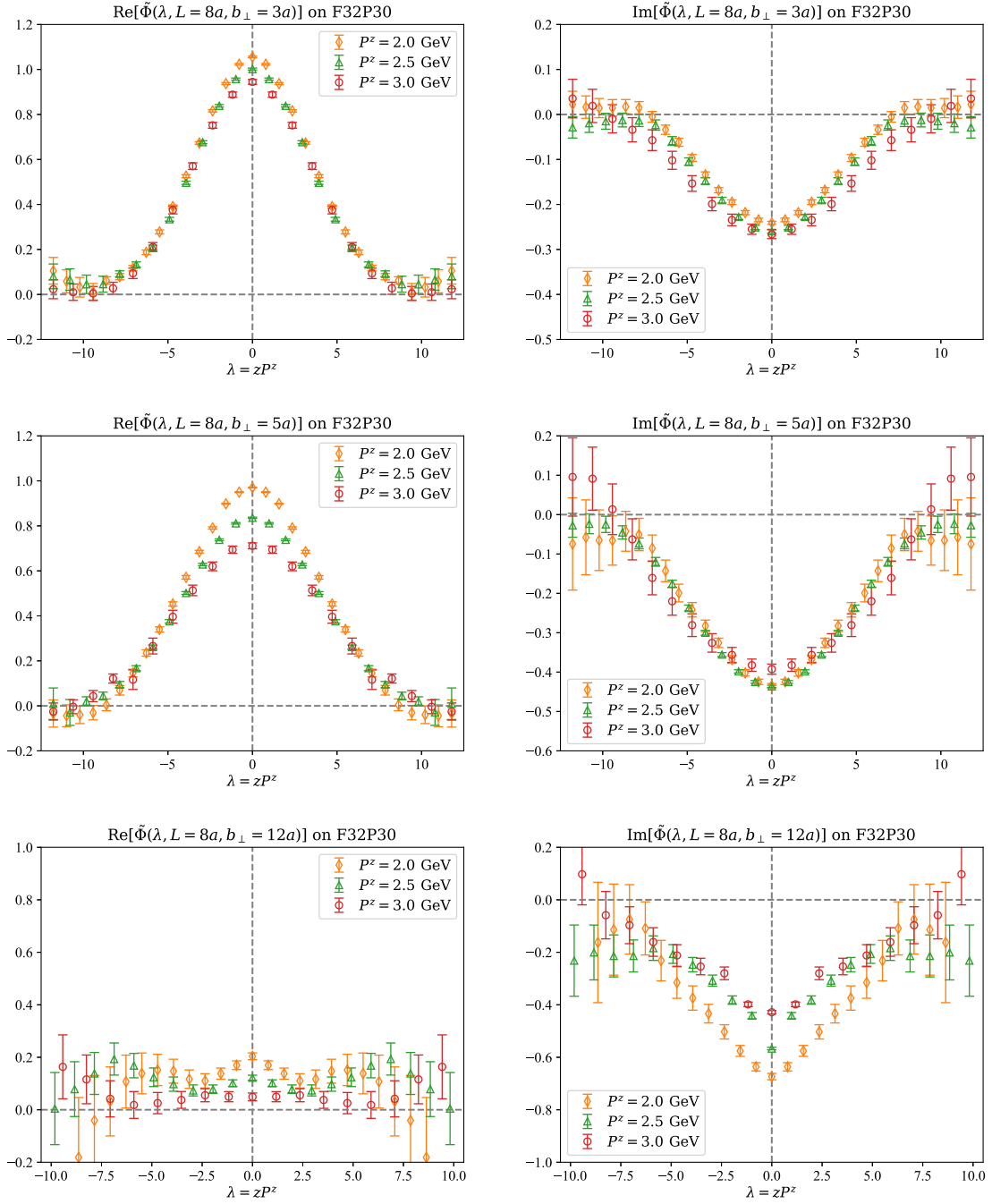


FIG. 34. Results from the F32P30 ensemble. A comparison of the results for  $P^z = \{2.0, 2.5, 3.0\}$  GeV at different  $b_\perp$  values is presented in the three panels.

### 5. Results for quasi-TMDWFs

More results for the quasi-TMDWFs on different ensembles are given in Figs. 35–38.

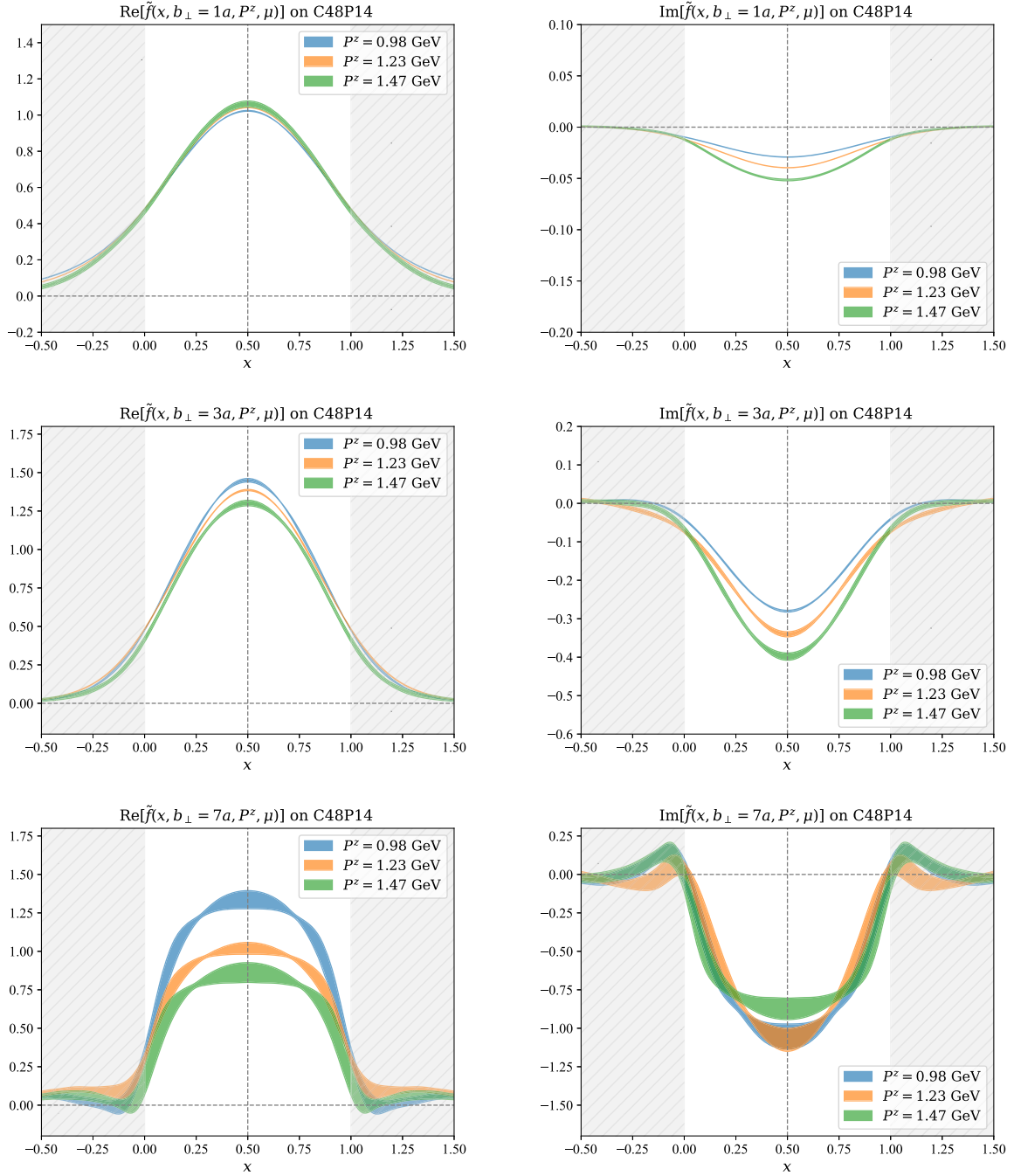


FIG. 35. Results from the C48P14 ensemble. A comparison of the results for  $P^z = \{0.98, 1.23, 1.47\}$  GeV at different  $b_\perp$  values is presented in the three panels.

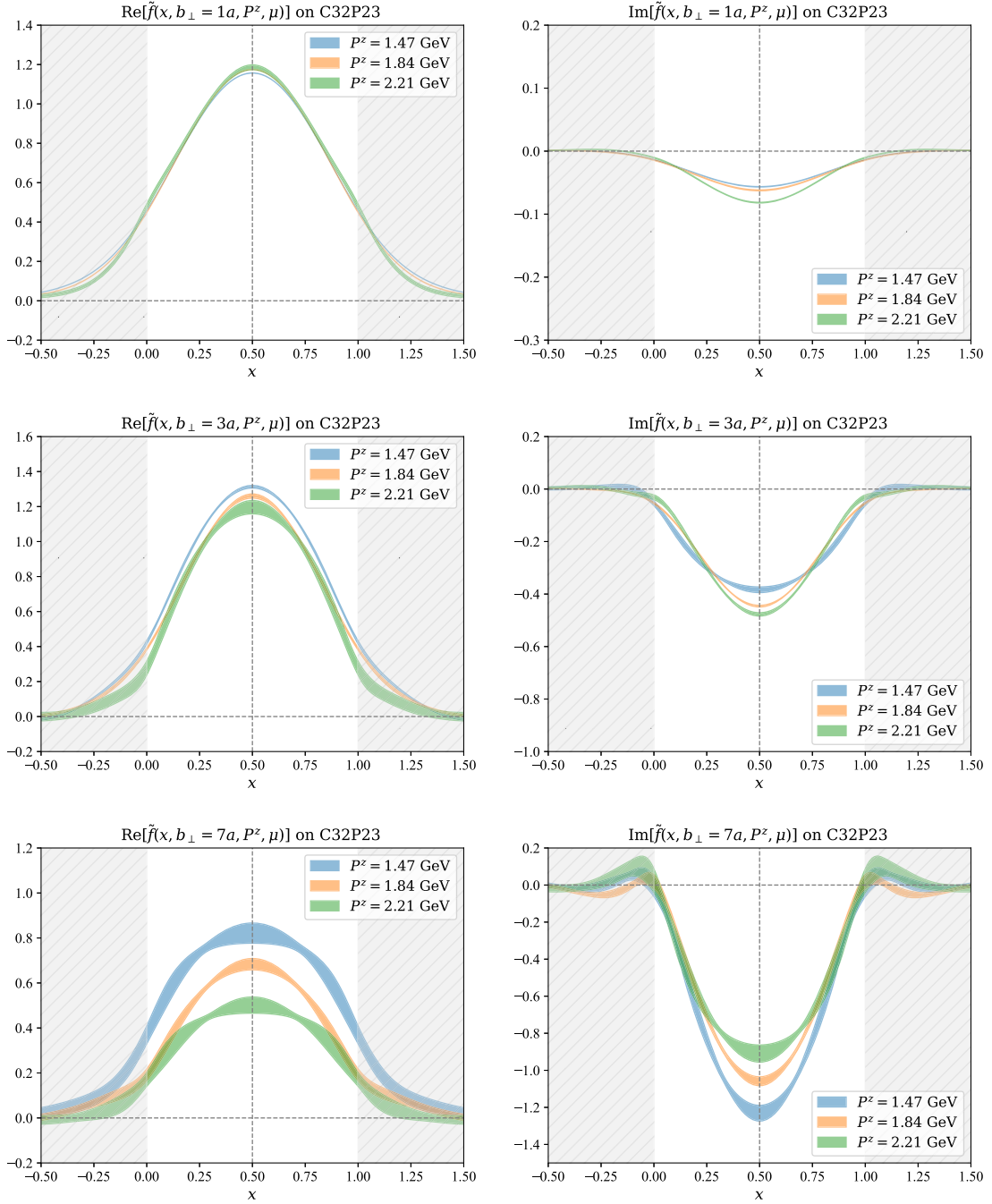


FIG. 36. Results from the C32P23 ensemble. A comparison of the results for  $P^z = \{1.47, 1.84, 2.21\}$  GeV at different  $b_\perp$  values is presented in the three panels.

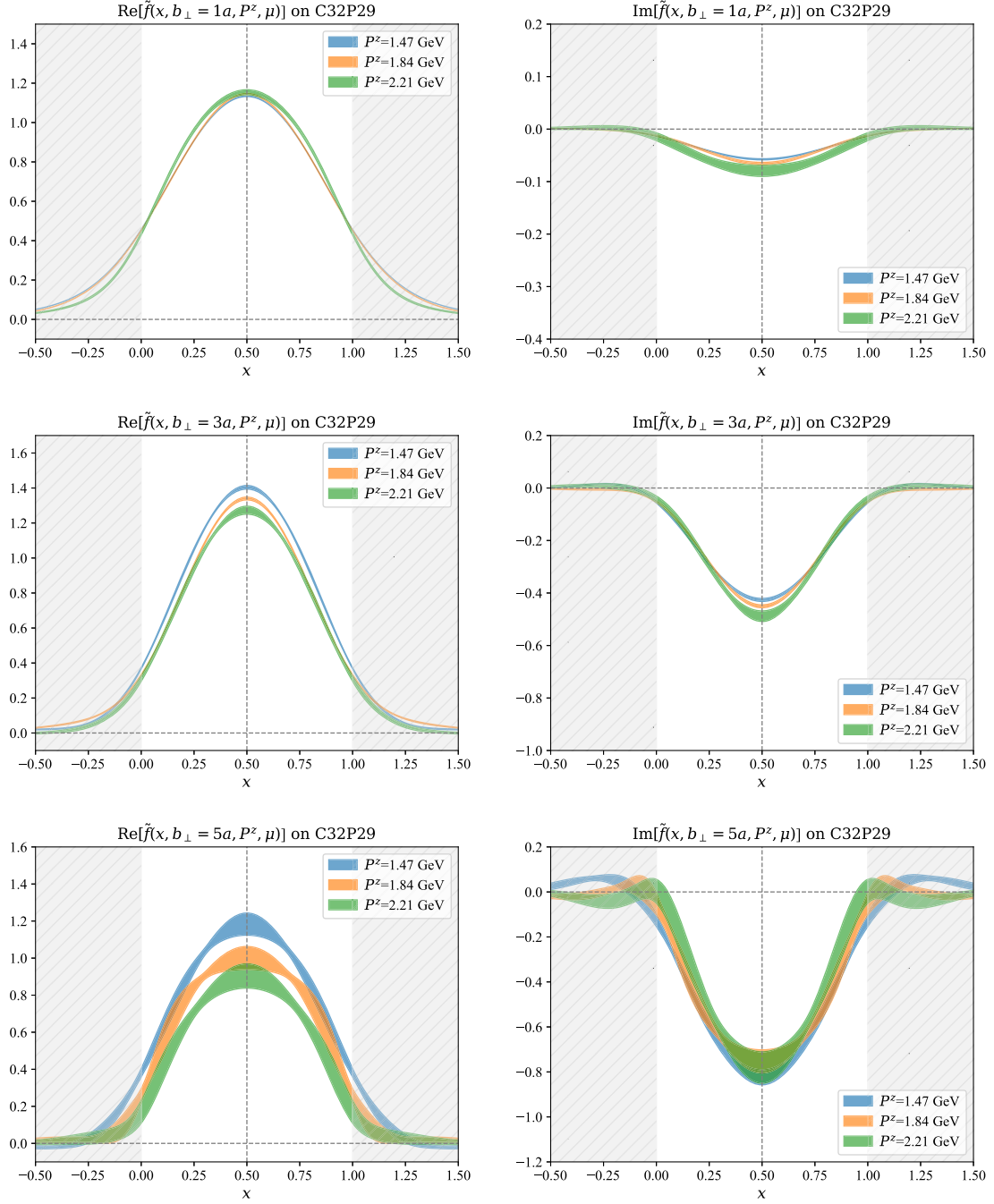


FIG. 37. Results from the C32P29 ensemble. A comparison of the results for  $P^z = \{1.47, 1.84, 2.21\}$  GeV at different  $b_\perp$  values is presented in the three panels.

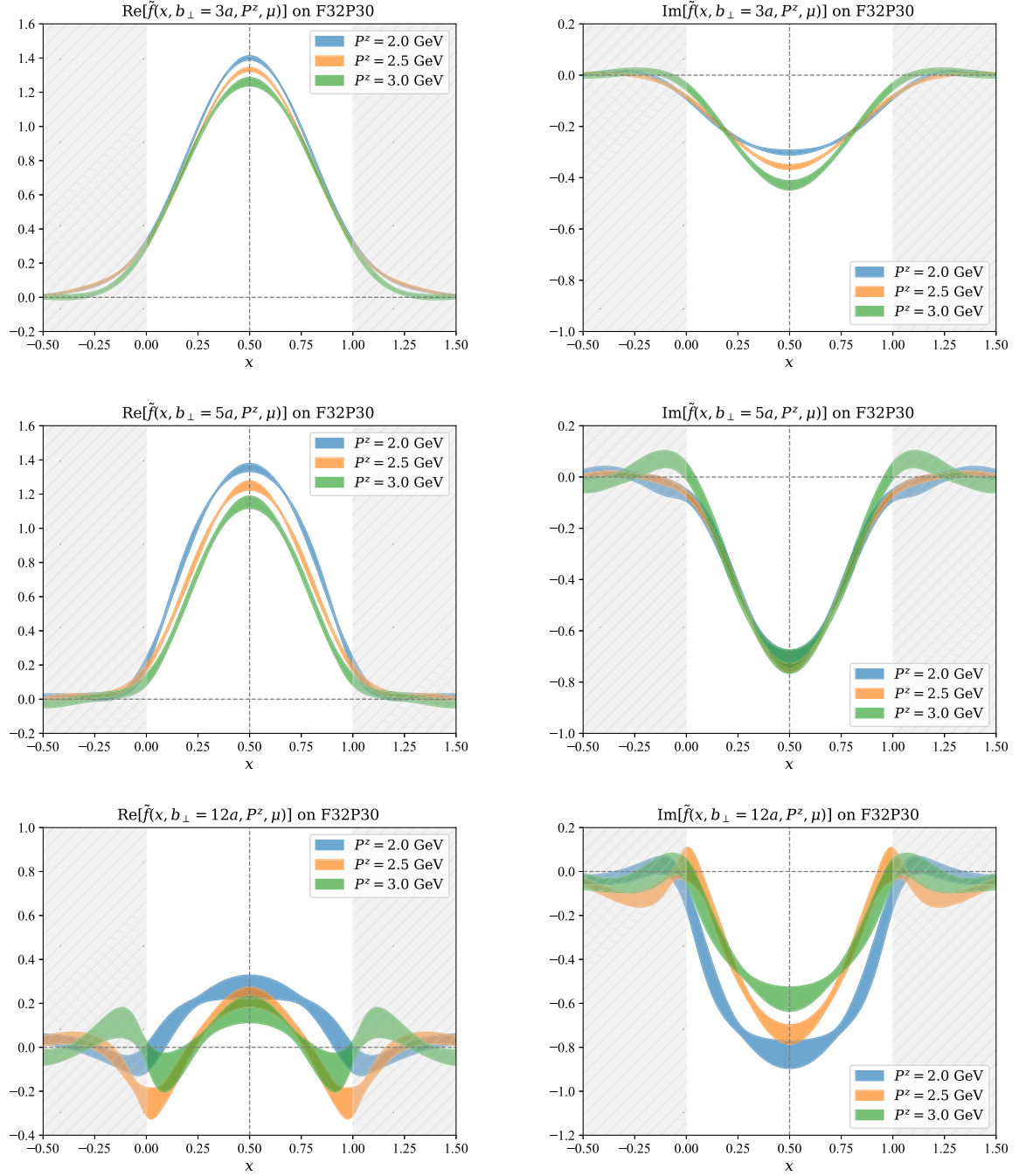


FIG. 38. Results from the F32P30 ensemble. A comparison of the results for  $P^z = \{2.0, 2.5, 3.0\}$  GeV at different  $b_\perp$  values is presented in the three panels.

## 6. Results for CS kernel

More results for the CS kernel on different ensembles are given in Figs. 39–42.

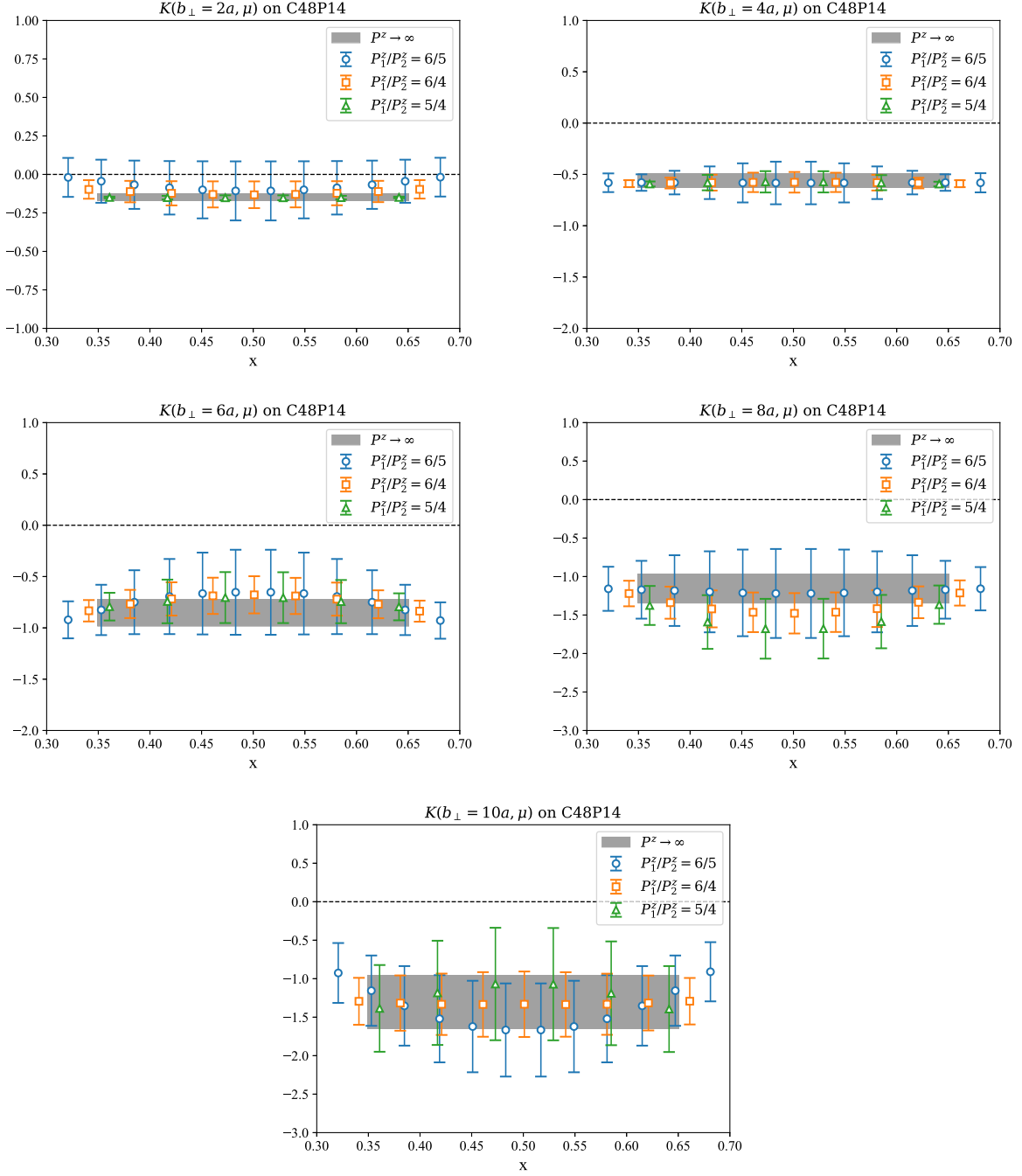
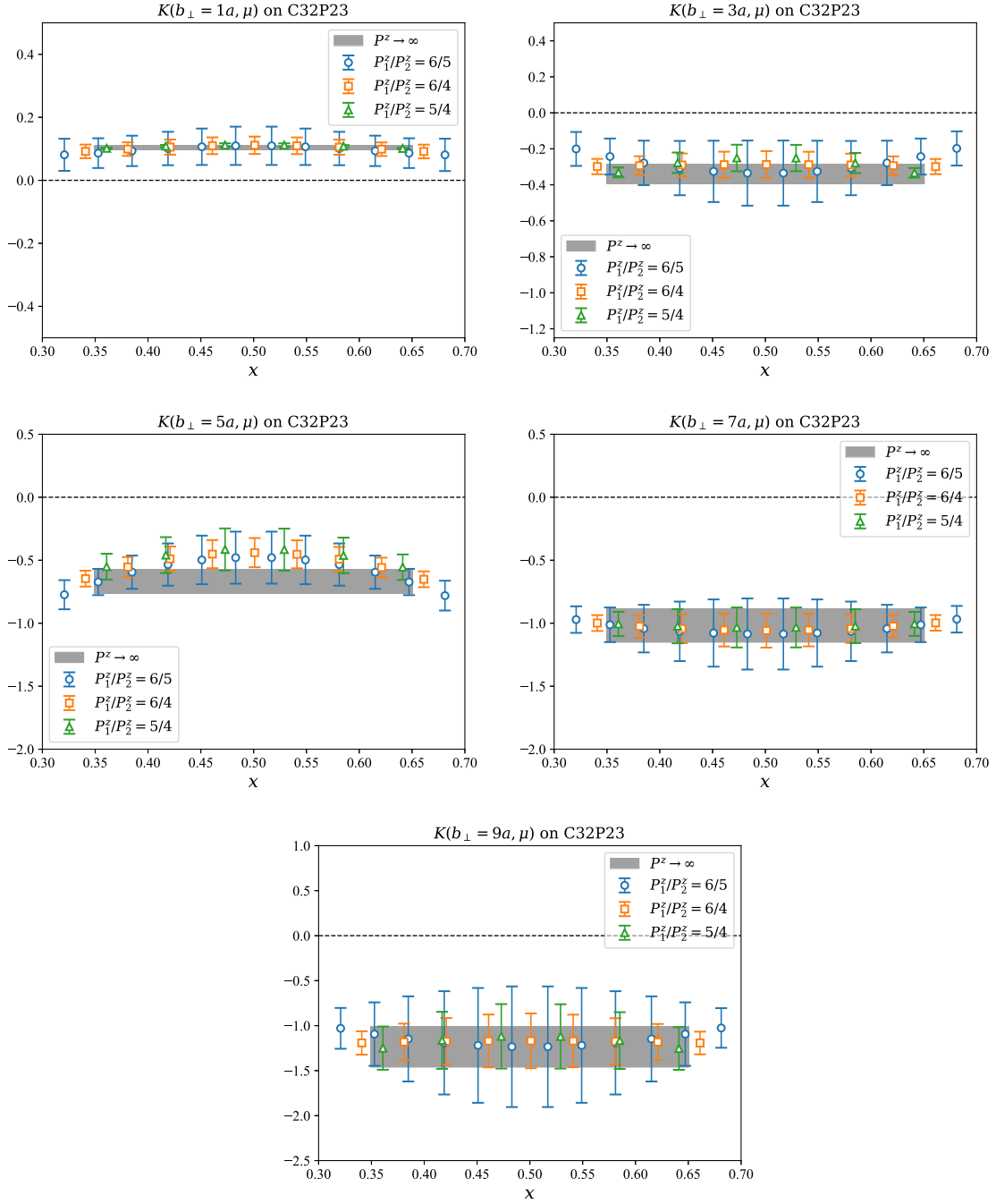
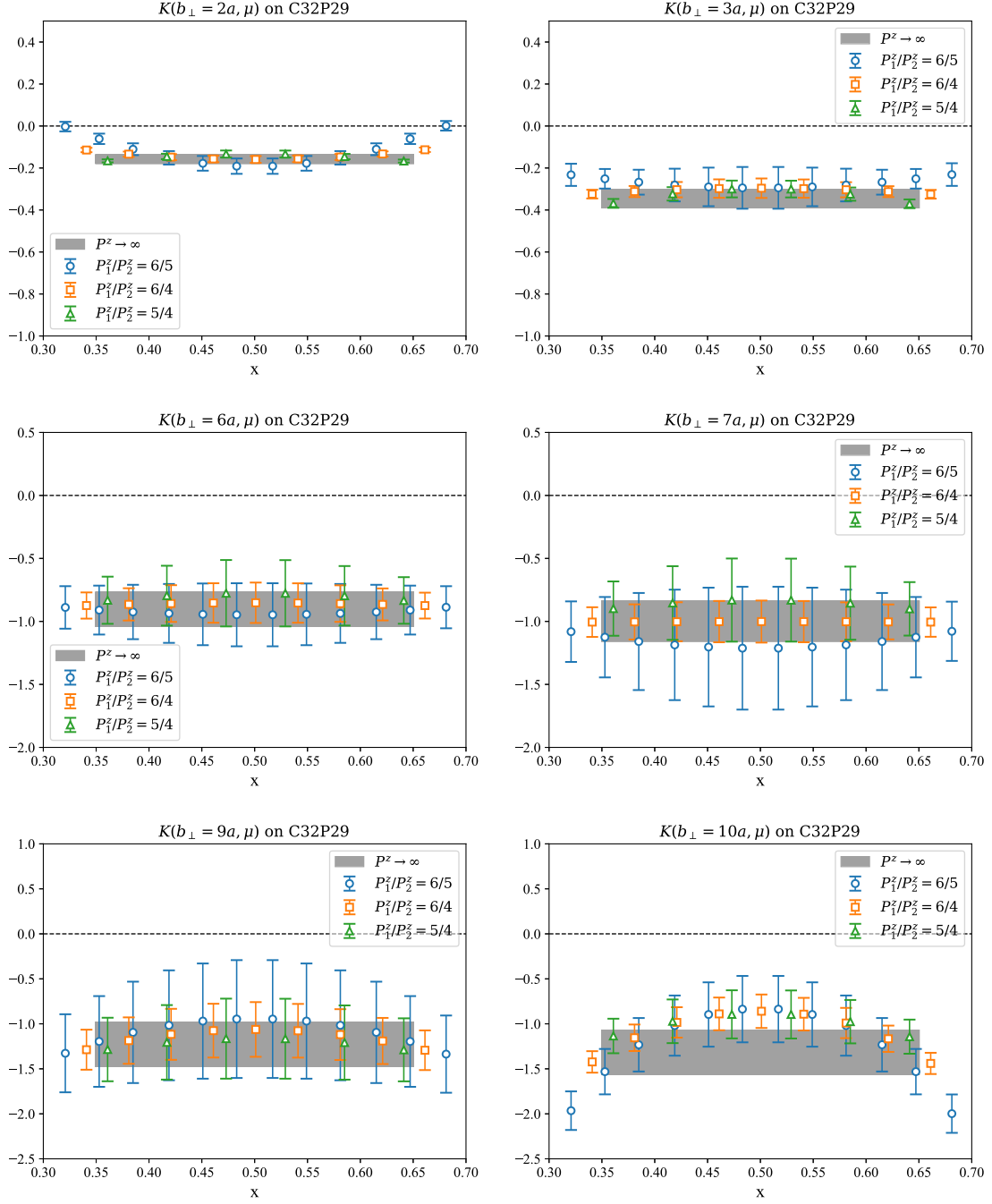


FIG. 39. Results of  $b_\perp = \{2, 4, 6, 8, 10\}a$  on the C48P14 ensemble for three different momentum combinations.

FIG. 40. Results of  $b_{\perp} = \{1, 3, 5, 7, 9\}a$  on the C23P23 ensemble for three different momentum combinations.

FIG. 41. Results of  $b_\perp = \{2, 3, 6, 7, 9, 10\}a$  on the C32P29 ensemble for three different momentum combinations.

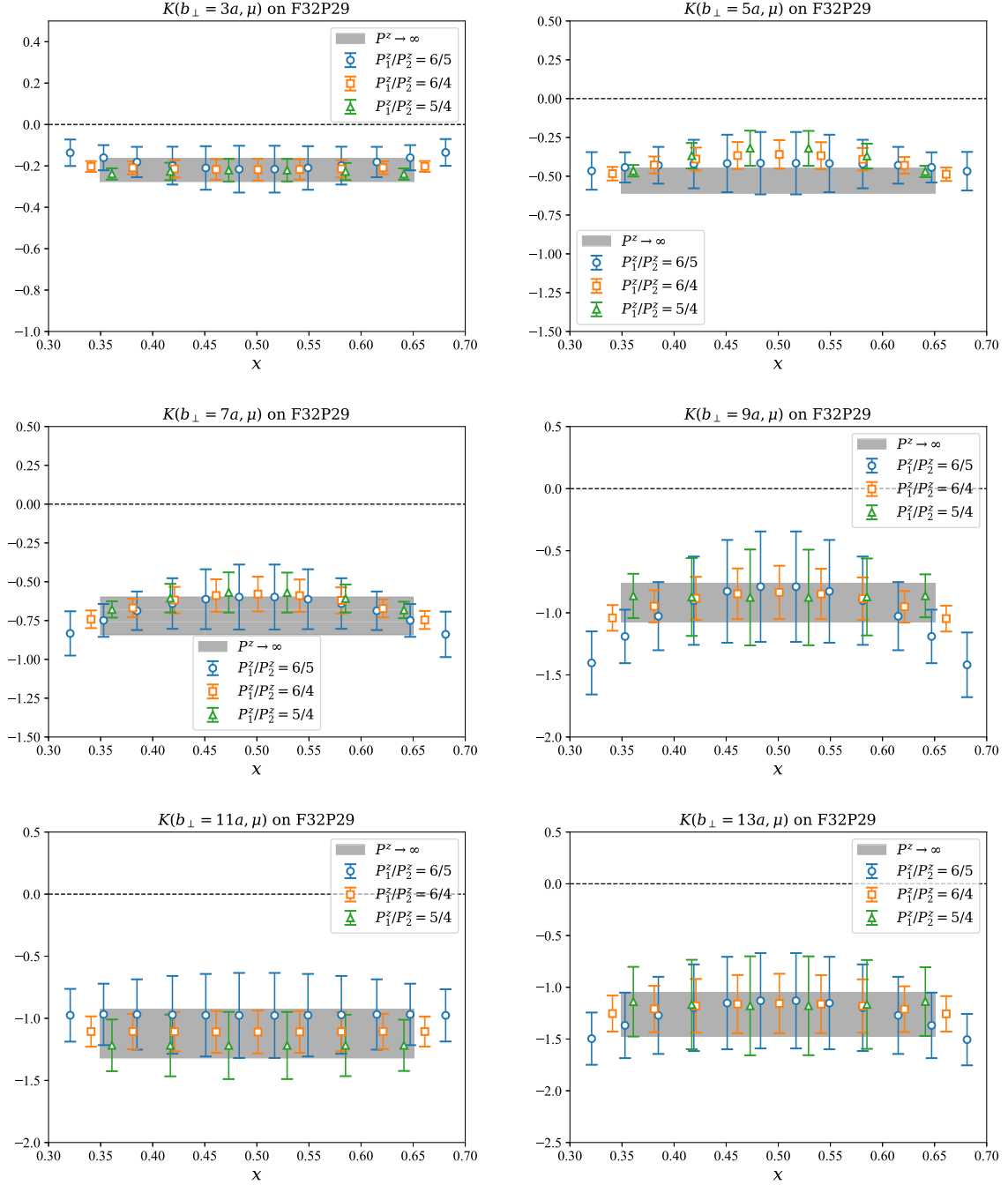


FIG. 42. Results of  $b_{\perp} = \{3, 5, 7, 9, 11, 13\}a$  on the F32P30 ensemble for three different momentum combinations.

[1] R. Boussarie *et al.*, arXiv:2304.03302.  
[2] J. C. Collins and D. E. Soper, *Nucl. Phys.* **B193**, 381 (1981); **B213**, 545(E) (1983).  
[3] J. C. Collins and D. E. Soper, *Nucl. Phys.* **B197**, 446 (1982).  
[4] J. Collins, *Foundations of Perturbative QCD*, Vol. 32 (Cambridge University Press, Cambridge, England, 2011).

[5] Y. Li and H. X. Zhu, *Phys. Rev. Lett.* **118**, 022004 (2017).  
[6] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *J. High Energy Phys.* **10** (2017) 041.  
[7] A. A. Vladimirov, *Phys. Rev. Lett.* **118**, 062001 (2017).  
[8] I. Moutl, H. X. Zhu, and Y. J. Zhu, *J. High Energy Phys.* **08** (2022) 280.

- [9] C. Duhr, B. Mistlberger, and G. Vita, *Phys. Rev. Lett.* **129**, 162001 (2022).
- [10] A. Airapetian *et al.* (HERMES Collaboration), *Phys. Rev. Lett.* **103**, 152002 (2009).
- [11] X. Qian *et al.* (Jefferson Lab Hall A Collaboration), *Phys. Rev. Lett.* **107**, 072003 (2011).
- [12] C. Adolph *et al.* (COMPASS Collaboration), *Phys. Lett. B* **717**, 383 (2012).
- [13] A. Airapetian *et al.* (HERMES Collaboration), *Phys. Rev. D* **87**, 074029 (2013).
- [14] M. Anselmino, M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin, *J. High Energy Phys.* **04** (2014) 005.
- [15] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Rev. Lett.* **116**, 132301 (2016).
- [16] C. Adolph *et al.* (COMPASS Collaboration), *Phys. Lett. B* **770**, 138 (2017).
- [17] M. Anselmino, M. Boglione, U. D'Alesio, F. Murgia, and A. Prokudin, *J. High Energy Phys.* **04** (2017) 046.
- [18] A. Airapetian *et al.* (HERMES Collaboration), *J. High Energy Phys.* **12** (2020) 010.
- [19] X. Liu and B.-Q. Ma, *Eur. Phys. J. C* **81**, 635 (2021).
- [20] J. M. Cruz-Martinez, M. Fieg, T. Giani, P. Krack, T. Mäkelä, T. R. Rabemananjara, and J. Rojo, *Eur. Phys. J. C* **84**, 369 (2024).
- [21] G. Aad *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **84**, 315 (2024).
- [22] J. Isaacson, Y. Fu, and C. P. Yuan, *Phys. Rev. D* **110**, 073002 (2024).
- [23] G. D. Alexeev *et al.* (COMPASS Collaboration), *Phys. Rev. Lett.* **133**, 071902 (2024).
- [24] M. Niemiec (COMPASS Collaboration), *Proc. Sci. ICHEP2024* (2025) 557.
- [25] I. Scimemi and A. Vladimirov, *J. High Energy Phys.* **06** (2020) 137.
- [26] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, M. Cerutti, F. Piacenza, M. Radici, and A. Signori (MAP (Multi-dimensional Analyses of Partonic distributions) Collaboration), *J. High Energy Phys.* **10** (2022) 127.
- [27] V. Moos, I. Scimemi, A. Vladimirov, and P. Zurita, *J. High Energy Phys.* **05** (2024) 036.
- [28] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, M. Cerutti, F. Delcarro, M. Radici, L. Rossi, and A. Signori (MAP (Multi-dimensional Analyses of Partonic distributions) Collaboration), *J. High Energy Phys.* **08** (2024) 232.
- [29] Z.-B. Kang, J. Penttala, and C. Zhang, *arXiv:2410.21435*.
- [30] A. B. Martinez, F. Hautmann, L. Keersmaekers, A. Lelek, M. Mendizabal Morentin, S. Taheri Monfared, and A. M. van Kampen, *Phys. Lett. B* **868**, 139762 (2025).
- [31] V. Moos, I. Scimemi, A. Vladimirov, and P. Zurita, *J. High Energy Phys.* **11** (2025) 134.
- [32] A. Bacchetta, V. Bertone, C. Bissolotti, M. Cerutti, M. Radici, S. Rodini, and L. Rossi (MAP (Multi-dimensional Analyses of Partonic distributions) Collaboration), *Phys. Rev. Lett.* **135**, 021904 (2025).
- [33] S. Camarda, G. Ferrera, and L. Rossi, *J. High Energy Phys.* **01** (2026) 150.
- [34] X. Ji, *Phys. Rev. Lett.* **110**, 262002 (2013).
- [35] X. Ji, *Sci. China Phys. Mech. Astron.* **57**, 1407 (2014).
- [36] X. Ji, Y.-S. Liu, Y. Liu, J.-H. Zhang, and Y. Zhao, *Rev. Mod. Phys.* **93**, 035005 (2021).
- [37] X. Ji, *Research* **8**, 0695 (2025).
- [38] M. A. Ebert, I. W. Stewart, and Y. Zhao, *Phys. Rev. D* **99**, 034505 (2019).
- [39] X. Ji, Y. Liu, and Y.-S. Liu, *Nucl. Phys.* **B955**, 115054 (2020).
- [40] P. Shanahan, M. Wagman, and Y. Zhao, *Phys. Rev. D* **102**, 014511 (2020).
- [41] Q.-A. Zhang *et al.* (Lattice Parton Collaboration), *Phys. Rev. Lett.* **125**, 192001 (2020).
- [42] M. Schlemmer, A. Vladimirov, C. Zimmermann, M. Engelhardt, and A. Schäfer, *J. High Energy Phys.* **08** (2021) 004.
- [43] P. Shanahan, M. Wagman, and Y. Zhao, *Phys. Rev. D* **104**, 114502 (2021).
- [44] Y. Li *et al.*, *Phys. Rev. Lett.* **128**, 062002 (2022).
- [45] M.-H. Chu *et al.* (Lattice Parton (LPC) Collaboration), *Phys. Rev. D* **106**, 034509 (2022).
- [46] M.-H. Chu *et al.* (Lattice Parton (LPC) Collaboration), *J. High Energy Phys.* **08** (2023) 172.
- [47] H.-T. Shu, M. Schlemmer, T. Sizmann, A. Vladimirov, L. Walter, M. Engelhardt, A. Schäfer, and Y.-B. Yang, *Phys. Rev. D* **108**, 074519 (2023).
- [48] A. Avkhadiiev, P. E. Shanahan, M. L. Wagman, and Y. Zhao, *Phys. Rev. D* **108**, 114505 (2023).
- [49] W.-Y. Liu, I. Zahed, and Y. Zhao, *Phys. Rev. D* **111**, 074022 (2025).
- [50] A. Avkhadiiev, P. E. Shanahan, M. L. Wagman, and Y. Zhao, *Phys. Rev. Lett.* **132**, 231901 (2024).
- [51] C. Alexandrou, S. Bacchio, K. Cichy, M. Constantinou, A. Sen, G. Spanoudes, F. Steffens, and J. Tarello, *arXiv:2509.26316*.
- [52] X.-d. Ji, J.-p. Ma, and F. Yuan, *Phys. Rev. D* **71**, 034005 (2005).
- [53] J. C. Collins and D. E. Soper, *Nucl. Phys.* **B194**, 445 (1982).
- [54] J. C. Collins, D. E. Soper, and G. F. Sterman, *Nucl. Phys.* **B250**, 199 (1985).
- [55] T. Becher and M. Neubert, *Eur. Phys. J. C* **71**, 1665 (2011).
- [56] X. Ji, Y. Liu, and Y.-S. Liu, *Phys. Lett. B* **811**, 135946 (2020).
- [57] X. Ji and Y. Liu, *Phys. Rev. D* **105**, 076014 (2022).
- [58] J.-C. He, M.-H. Chu, J. Hua, X. Ji, A. Schäfer, Y. Su, W. Wang, Y.-B. Yang, J.-H. Zhang, and Q.-A. Zhang (Lattice Parton (LPC) Collaboration), *Phys. Rev. D* **109**, 114513 (2024).
- [59] L. Walter *et al.* (Lattice Parton (LPC) Collaboration), *Phys. Rev. D* **111**, 094507 (2025).
- [60] L. Ma *et al.* (Lattice Parton Collaboration), *J. High Energy Phys.* **08** (2025) 086.
- [61] Z.-F. Deng, W. Wang, and J. Zeng, *J. High Energy Phys.* **09** (2022) 046.
- [62] Y.-K. Huo *et al.* (Lattice Parton (LPC) Collaboration), *Nucl. Phys.* **B969**, 115443 (2021).
- [63] R. Zhang, A. V. Grebe, D. C. Hackett, M. L. Wagman, and Y. Zhao, *Phys. Rev. D* **112**, L051502 (2025).
- [64] X.-D. Ji and M. J. Musolf, *Phys. Lett. B* **257**, 409 (1991).

- [65] Z.-C. Hu *et al.* (CLQCD Collaboration), *Phys. Rev. D* **109**, 054507 (2024).
- [66] A. Hasenfratz and F. Knechtli, *Phys. Rev. D* **64**, 034504 (2001).
- [67] A. Hasenfratz, R. Hoffmann, and F. Knechtli, *Nucl. Phys. B, Proc. Suppl.* **106–107**, 418 (2002).
- [68] J. Green, K. Jansen, and F. Steffens, *Phys. Rev. Lett.* **121**, 022004 (2018).
- [69] J.-W. Chen, X. Ji, and J.-H. Zhang, *Nucl. Phys.* **B915**, 1 (2017).
- [70] J.-H. Zhang, J.-W. Chen, X. Ji, L. Jin, and H.-W. Lin, *Phys. Rev. D* **95**, 094514 (2017).
- [71] B. U. Musch, P. Hagler, J. W. Negele, and A. Schafer, *Phys. Rev. D* **83**, 094507 (2011).
- [72] J.-H. Zhang, L. Jin, H.-W. Lin, A. Schäfer, P. Sun, Y.-B. Yang, R. Zhang, Y. Zhao, and J.-W. Chen (LP3 Collaboration), *Nucl. Phys.* **B939**, 429 (2019).
- [73] M. Luscher and P. Weisz, *J. High Energy Phys.* **09** (2001) 010.
- [74] J.-W. Chen *et al.*, [arXiv:2505.14619](https://arxiv.org/abs/2505.14619).
- [75] A.-S. Xiong, J. Hua, T. Wei, F.-S. Yu, Q.-A. Zhang, and Y. Zheng, *Eur. Phys. J. C* **85**, 1409 (2025).
- [76] Y.-F. Ling, M.-H. Chu, J. Liang, J. Hua, A.-S. Xiong, and Q.-A. Zhang, [arXiv:2511.03593](https://arxiv.org/abs/2511.03593).
- [77] X. Ji, Y. Liu, A. Schäfer, W. Wang, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, *Nucl. Phys.* **B964**, 115311 (2021).
- [78] T. Regge, *Nuovo Cimento* **14**, 951 (1959).
- [79] D. Bollweg, X. Gao, S. Mukherjee, and Y. Zhao, *Phys. Lett. B* **852**, 138617 (2024).
- [80] D. Bollweg, X. Gao, J. He, S. Mukherjee, and Y. Zhao, *Phys. Rev. D* **112**, 034501 (2025).
- [81] J. Collins and T. Rogers, *Phys. Rev. D* **91**, 074020 (2015).
- [82] I. Scimemi and A. Vladimirov, *J. High Energy Phys.* **08** (2018) 003.
- [83] M. G. Echevarria, A. Idilbi, A. Schäfer, and I. Scimemi, *Eur. Phys. J. C* **73**, 2636 (2013).
- [84] J. M. Henn, G. P. Korchemsky, and B. Mistlberger, *J. High Energy Phys.* **04** (2020) 018.
- [85] S. Moch, J. A. M. Vermaseren, and A. Vogt, *Nucl. Phys.* **B688**, 101 (2004).
- [86] A. Vogt, S. Moch, and J. A. M. Vermaseren, *Nucl. Phys.* **B691**, 129 (2004).
- [87] R. von Manteuffel, E. Panzer, and R. M. Schabinger, *Phys. Rev. Lett.* **124**, 162001 (2020).
- [88] A. Avkhadiev, V. Bertone, C. Bissolotti, M. Cerutti, Y. Fu, S. Rodini, P. Shanahan, M. Wagman, and Y. Zhao, [arXiv:2510.26489](https://arxiv.org/abs/2510.26489).
- [89] X. Jiang, C. Shi, Y. Chen, M. Gong, and Y.-B. Yang, [arXiv:2411.08461](https://arxiv.org/abs/2411.08461).
- [90] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi (QUADA Collaboration), *Comput. Phys. Commun.* **181**, 1517 (2010).
- [91] R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower, and S. Gottlieb (QUADA Collaboration), in *International Conference for High Performance Computing, Networking, Storage and Analysis* (2011), [10.1145/2063384.2063478](https://arxiv.org/abs/10.1145/2063384.2063478).
- [92] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. C. Brower (QUADA Collaboration), in *International Conference for High Performance Computing, Networking, Storage and Analysis* (2016), [10.5555/3014904.3014995](https://arxiv.org/abs/10.5555/3014904.3014995).
- [93] Y.-J. Bi, Y. Xiao, W.-Y. Guo, M. Gong, P. Sun, S. Xu, and Y.-B. Yang, *Proc. Sci. LATTICE2019* (2020) 286 [[arXiv:2001.05706](https://arxiv.org/abs/2001.05706)].
- [94] X.-D. Ji, [arXiv:hep-lat/9506034](https://arxiv.org/abs/hep-lat/9506034).
- [95] K. Zhang, X. Ji, Y.-B. Yang, F. Yao, and J.-H. Zhang (Lattice Parton (LPC) Collaboration), *Phys. Rev. Lett.* **129**, 082002 (2022).