

RECEIVED: July 28, 2021

REVISED: November 2, 2021

ACCEPTED: November 3, 2021

PUBLISHED: November 19, 2021

# Towards high-precision parton distributions from lattice QCD via distillation

**Colin Egerer,<sup>a,b</sup> Robert G. Edwards,<sup>b</sup> Christos Kallidonis,<sup>b</sup> Kostas Orginos,<sup>a,b</sup>  
Anatoly V. Radyushkin,<sup>b,c</sup> David G. Richards,<sup>b</sup> Eloy Romero<sup>b</sup>  
and Savvas Zafeiropoulos<sup>d</sup> on behalf of the HadStruc collaboration**

<sup>a</sup>*Department of Physics, William and Mary,  
300 Ukrop Way, Williamsburg, Virginia 23187, U.S.A.*

<sup>b</sup>*Thomas Jefferson National Accelerator Facility,  
12000 Jefferson ave., Newport News, VA 23606, U.S.A.*

<sup>c</sup>*Physics Department, Old Dominion University,  
4600 Elkhorn Ave., Norfolk, VA 23529, U.S.A.*

<sup>d</sup>*Aix Marseille Univ, Université de Toulon, CNRS,  
CPT, Marseille, France*

*E-mail:* [cegerer@jlab.org](mailto:cegerer@jlab.org), [edwards@jlab.org](mailto:edwards@jlab.org), [kallidonis@gmail.com](mailto:kallidonis@gmail.com),  
[knorgi@wm.edu](mailto:knorgi@wm.edu), [radyush@jlab.org](mailto:radyush@jlab.org), [dgr@jlab.org](mailto:dgr@jlab.org), [eromero@jlab.org](mailto:eromero@jlab.org),  
[savvas.zafeiropoulos@cpt.univ-mrs.fr](mailto:savvas.zafeiropoulos@cpt.univ-mrs.fr)

**ABSTRACT:** We apply the Distillation spatial smearing program to the extraction of the unpolarized isovector valence PDF of the nucleon. The improved volume sampling and control of excited-states afforded by distillation leads to a dramatically improved determination of the requisite Ioffe-time Pseudo-distribution (pITD). The impact of higher-twist effects is subsequently explored by extending the Wilson line length present in our non-local operators to one half the spatial extent of the lattice ensemble considered. The valence PDF is extracted by analyzing both the matched Ioffe-time Distribution (ITD), as well as a direct matching of the pITD to the PDF. Through development of a novel prescription to obtain the PDF from the pITD, we establish a concerning deviation of the pITD from the expected DGLAP evolution of the pseudo-PDF. The presence of DGLAP evolution is observed once more following introduction of a discretization term into the PDF extractions. Observance and correction of this discrepancy further highlights the utility of distillation in such structure studies.

**KEYWORDS:** Lattice field theory simulation, QCD Phenomenology

**ARXIV EPRINT:** [2107.05199](https://arxiv.org/abs/2107.05199)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>PDFs and Ioffe-time pseudo-distributions</b>	<b>2</b>
<b>3</b>	<b>Numerical methods</b>	<b>4</b>
3.1	Distillation	4
3.2	Matrix element isolation	7
<b>4</b>	<b>Extraction of unpolarized nucleon PDFs</b>	<b>9</b>
4.1	An ill-posed inverse	13
4.2	Direct extraction of PDFs from reduced pseudo-ITDs	15
4.2.1	Data correlation	17
4.3	Classical orthogonal polynomials	19
4.3.1	PDF results with Jacobi polynomials	23
<b>5</b>	<b>On the numerical consistency with DGLAP</b>	<b>30</b>
5.1	Jacobi polynomial corrections — discretization effects	33
<b>6</b>	<b>Conclusions</b>	<b>36</b>
<b>A</b>	<b>Comparison of distillation with conventional smearing</b>	<b>38</b>

---

## 1 Introduction

Elucidation of the dynamical properties of quarks and gluons and their collective emergent phenomena is the principal charge of contemporary hadronic physics and is a central component of the precision phenomenology program at the Large Hadron Collider (LHC). Crucial for the interpretation and prediction of inclusive and semi-inclusive scattering processes, such as deep inelastic scattering (DIS) and semi-inclusive DIS, are the parton distribution functions (PDFs) which emerge in the QCD factorization [1] of such inclusive cross sections. PDFs capture the collinear momentum distribution of partons within a fast moving hadron with  $p^\mu = (p^+, m^2/2p^+, \mathbf{0}_\perp)$ , and offer a probabilistic interpretation featured at leading-twist. Despite complementing the growing hadron tomography efforts both theoretically and at upcoming facilities such as the Electron Ion Collider, determination of PDFs remains a high priority as they are often large sources of hadronic error in collider experiments and thereby affect the precision measurements of an array of Standard Model parameters [2].

The numerical tool of lattice field theory enables the quantitative study of strongly-coupled theories, such as Quantum Chromodynamics (QCD), from first-principles. As PDFs accumulate information on the infrared structure of a hadron, they would seem ideal objects

to target in lattice QCD (LQCD) calculations. However, the Euclidean metric of LQCD precludes direct access to the light-like bilinears required to define PDFs and other light-cone distributions. Attempts to circumvent this preclusion date to early efforts to access the hadronic tensor [3, 4], forward Compton amplitude [5], and light-cone distributions of exclusive processes [6, 7] from suitably constructed Euclidean correlation functions.

The recent cascade of research into the light-cone structure of hadrons from LQCD followed from X. Ji's proposed connection between matrix elements of *space-like* separated parton bilinears and their light-like counterparts [8]. When analyzed with respect to the space-like extent, the resulting distribution of parton longitudinal space-like momenta, or quasi-PDF, can be factorized into the light-cone PDFs  $f(x, \mu^2)$  in the limit of large space-like momenta [9]. Since then, substantial effort has been invested in extracting PDFs from the quasi-distribution formalism [10–20] and the interplay with lattice systematics has been explored [21, 22]. In the forward limit, quasi-distributions have recently been extended to access distribution amplitudes [23–25] and PDFs of explicit higher-twist [26]. The reader is directed to [27, 28] for details on community progress. An alternative coordinate space interpretation, the pseudo-distribution formalism expounded upon in section 2, developed in [29] shares the same matrix elements with quasi-distributions, but its Lorentz-invariant amplitudes factorize in the short-distance space-like regime into PDFs and perturbatively calculable coefficient functions. In this manner, pseudo-distributions are a special case of *Good Lattice Cross Sections* [30, 31], which has offered complementary information on pion structure [32, 33]. Regardless of the methodology adopted, high-momenta is a requirement in order to access the regime of low momentum fraction.

The remainder of this manuscript is organized as follows. In the interest of self-containment, we begin in section 2 with a summary of the pseudo-distribution formalism and how its application to certain space-like matrix elements illuminates the forward lightcone structure of hadrons. After introducing the lattice ensemble employed in this work, we proceed in section 3 with a recapitulation of the Distillation spatial smearing program, a recent extension of the method to high-momentum observables, and argue why its utilization is essential in such structure calculations. Methods for matrix element and subsequent PDF extraction from our lattice data, especially in light of lattice artifacts, are developed subsequently. Results of these protocols are presented in section 4, followed by discussion and concluding remarks in section 6.

## 2 PDFs and Ioffe-time pseudo-distributions

Consider the non-local quark bilinear  $\bar{\psi}(z) \gamma^\alpha \Phi_z^{(f)}(\{z, 0\}) \psi(0)$  connected with a straight  $z$ -separated Wilson line  $\Phi_z^{(f)}(\{z, 0\})$  in the fundamental representation of SU(3). Lorentz invariance dictates the forward helicity-averaged matrix element of this operator decomposes according to

$$M^\alpha(p, z) = \langle h(p) | \underbrace{\bar{\psi}(z) \gamma^\alpha \Phi_z^{(f)}(\{z, 0\}) \psi(0)}_{\hat{\mathcal{O}}_{\text{WL}}^{[\gamma^\mu]}(z)} | h(p) \rangle \quad (2.1)$$

$$= 2p^\alpha \mathcal{M}(\nu, z^2) + 2z^\alpha \mathcal{N}(\nu, z^2), \quad (2.2)$$

with  $\nu \equiv p \cdot z$  and  $z^2$  the Ioffe-time [34, 35] and invariant interval, respectively, of the process. For a fast-moving hadron, the usual unpolarized PDFs are defined with light-cone coordinates where  $\alpha = +$ ,  $p^\alpha = \left(p^+, \frac{m_h^2}{2p^+}, \mathbf{0}_\perp\right)$  and  $z^\alpha = (0, z^-, \mathbf{0}_\perp)$ . In this scenario  $M^+(p, z)$  only receives contributions from  $\mathcal{M}(p^+ z^-, 0)$ . Provided the logarithmic singularity that arises for  $z^2 = 0$  is regularized (typically in  $\overline{\text{MS}}$ ),  $\mathcal{M}(p^+ z^-, 0)$  defines the Ioffe-time distribution (ITD) [35]:

$$\mathcal{M}(p^+ z^-, 0)_{\mu^2} \equiv \mathcal{Q}(\nu, \mu^2) = \int_{-1}^1 dx e^{i\nu x} f_{q/h}(x, \mu^2), \quad (2.3)$$

and obviates the Fourier transform of (2.1) to momentum space which defines the conventional PDFs. Lorentz invariance implies the  $\nu$ -dependence of  $\mathcal{M}(p^+ z^-, 0)_{\mu^2}$  can be computed in any frame, and with any choice of  $\{z, \alpha\}$  that may be convenient. A particular choice amenable to calculation with lattice QCD is  $\alpha = 0$ ,  $p^\alpha = (E, \mathbf{0}_\perp, p_z)$  and  $z^\alpha = (0, \mathbf{0}_\perp, z_3)$ , which excludes the contamination from the pure higher-twist term  $\mathcal{N}(\nu, z^2)$ . The remaining term  $\mathcal{M}(\nu, z^2 \neq 0)$  is deemed the *Ioffe-time Pseudo-distribution* [29] or pseudo-ITD. In addition to the twist-2 contributions, the pseudo-ITD also contains higher-twist contributions  $\mathcal{O}(z^2 \Lambda_{\text{QCD}}^2)$  that vanish only in the light-cone limit. Furthermore, for all relevant Feynman diagrams [36], the Fourier transform of the pseudo-ITD with respect to  $\nu$  has support only on the canonical parton momentum fraction interval  $x \in [-1, 1]$ . The challenge numerically is to extract the leading-twist dependence amongst contributing  $\mathcal{O}(z^2)$  higher-twist terms.

A considerable challenge of using this non-local parton bilinear is the appearance of additional ultraviolet (UV) divergences for space-like separations. Prior to taking the continuum limit, the UV divergences must be regularized and removed. In perturbation theory, such divergences appear first, as power-like  $z/a$  [37] terms in the gauge-link self-energy corrections, with  $a$  being a UV regulator. These exponentiate to all orders [38, 39] producing the factor  $Z_{\text{link}}(z_3, a) \simeq e^{-A|z_3|/a}$ . Second, there are logarithmic  $\ln(-z^2/a^2)$  UV corrections present both in the link self-energy and in the vertex link corrections [40]. As these unwanted UV divergences are multiplicative [40–43], and independent of the Ioffe-time, combining into an overall factor  $Z_{\text{UV}}(z_3, a)$ , we construct the so-called *reduced pseudo-ITD* [29]

$$\mathfrak{M}(\nu, z^2) = \frac{\mathcal{M}(\nu, z^2)}{\mathcal{M}(0, z^2)}. \quad (2.4)$$

Since the rest-frame pseudo-ITD  $\mathcal{M}(0, z^2)$  has the same UV divergences associated with the gauge link, by forming the reduced pseudo-ITD we cancel the divergent factors  $Z_{\text{link}}(z_3, a)$  and  $Z_{\text{UV}}(z_3, a)$ , thereby ensuring a finite continuum limit. The choice to construct the RG invariant reduced pseudo-ITD with  $\mathcal{M}(0, z^2)$  is especially straightforward, as  $\mathcal{M}(0, z^2)$  is simply the bare vector charge  $Z_V^{-1}$  in the light-cone limit thereby leaving the OPE unaltered. Motivation for the reduced pseudo-ITD also extends to the mitigation of higher-twist  $\mathcal{O}(z^2 \Lambda_{\text{QCD}}^2)$  effects [44]. An alternative has recently been proposed that makes use of a vacuum matrix element of the space-like parton bilinear [45] (see also ref. [46]).

Following removal of the UV divergences produced by the space-like Wilson line, the remaining singularities in (2.4) stem from  $\ln(-z^2)$  contributions in QCD. These terms generate the perturbative evolution of the collinear PDFs and complicate the naive  $z^2 \rightarrow 0$

limit. The reduced pseudo-ITD  $\mathfrak{M}(\nu, z^2)$  factorizes in the perturbative small- $z^2$  regime into PDFs with perturbatively calculable hard coefficients. The factorization relationship has been computed to NLO [47–49] and recently NNLO [46, 50]. The NLO relationship that matches the  $\overline{\text{MS}}$  ITD  $\mathcal{Q}(\nu, \mu^2)$  to the reduced pseudo-ITD  $\mathfrak{M}(\nu, z^2)$  reads

$$\mathfrak{M}(\nu, z^2) = \int_0^1 du \mathcal{C}(u, z^2 \mu^2, \alpha_s(\mu)) \mathcal{Q}(u\nu, \mu^2) + \sum_{k=1}^{\infty} \mathcal{B}_k(\nu) (z^2)^k, \quad (2.5)$$

where  $\mathcal{Q}(\nu, \mu^2)$  is the lightcone ITD at a factorization scale  $\mu^2$ . The matching kernel

$$\mathcal{C}(u, z^2 \mu^2, \alpha_s(\mu)) = \delta(1-u) - \frac{\alpha_s}{2\pi} C_F \left[ \ln \left( \frac{e^{2\gamma_E+1}}{4} z^2 \mu^2 \right) B(u) + L(u) \right], \quad (2.6)$$

involves a scale-independent kernel  $L(u) = \left[ 4 \frac{\ln(1-u)}{1-u} - 2(1-u) \right]_+$  that matches the lattice and  $\overline{\text{MS}}$  regularization schemes, and a scale-dependent kernel that relates the  $z^2$  and  $\mu^2$  scales through the flavor non-singlet DGLAP evolution kernel  $B(u) = \left[ \frac{1+u^2}{1-u} \right]_+$  [51–53].

The factorization is valid in so far as the polynomial corrections  $\mathcal{B}_k(\nu) (z^2)^k$  can be mitigated. We follow in this work the plus-prescription defined by  $\int_0^1 du G(u)_+ f(ux) = \int_0^1 du G(u) [f(ux) - f(x)]$ .

### 3 Numerical methods

The pseudo-distribution formalism has been leveraged in several lattice calculations of partonic structure of hadrons, including the valence quark content of the pion [54], and the unpolarized valence quark [44, 55–58] and recently gluon [59] contents of the nucleon. Even though each calculation makes use of standard spatial and momentum smearing techniques, considerable statistical fluctuations are met for Ioffe-times in excess of  $\nu \gtrsim 5$  (and at even smaller values for gluonic matrix elements).

Motivated by the demonstrable success of the union of the distillation paradigm and the momentum smearing idea [60], we apply for the first time the distillation spatial smearing program to the extraction of PDFs from lattice QCD. We employ a 349 configuration isotropic clover ensemble generated by the JLab/W&M/LANL collaboration [61], featuring  $2 \oplus 1$  quark flavors within a  $32^3 \times 64$  lattice volume. The inverse coupling was set to  $\beta = 6.3$ , from which the lattice spacing  $a \simeq 0.094$  fm was obtained from the Wilson-flow scale  $w_0$  [62], which yields a pion mass  $m_\pi = 358$  MeV. The reader is referred to [63, 64] for further details of this ensemble, denoted  $a094m358$  and summarized in table 1. A single application of the stout smearing kernel [65] yields a tadpole-improved tree-level clover coefficient  $c_{\text{sw}}$  near the value determined non-perturbatively from the Schrödinger functional method [61] a posteriori.

#### 3.1 Distillation

The Jacobi smearing kernel  $J_{\sigma, n_\sigma}(T) = (1 + \sigma \nabla^2(T))^{n_\sigma}$  [66] featured in a variety of lattice calculations is but one method with which point-like interpolating fields can be spatially

ID	$a$ (fm)	$m_\pi$ (MeV)	$L^3 \times N_t$	$N_{\text{cfg}}$	$N_{\text{srcs}}$	$R_{\mathcal{D}}$
$a094m358$	0.094(1)	358(3)	$32^3 \times 64$	349	4	64

**Table 1.** Lattice ensemble employed in this work. The number of distinct source positions  $N_{\text{srcs}}$  per configuration and the distillation space rank  $R_{\mathcal{D}}$  are also indicated. We note the resulting number of measurements  $N_{\text{meas}} = N_{\text{cfg}} \times N_{\text{srcs}} = 1396$  is kept fixed for all two- and three-point functions computed in this work, regardless of momentum boost, source-sink separation and Wilson line length.

smeared and made more sensitive to confinement scale physics. This low-mode filter can be made explicit by identifying eigenvectors of the three-dimensional gauge-covariant Laplacian

$$-\nabla^2(T) \xi^{(k)}(T) = \lambda^{(k)}(T) \xi^{(k)}(T) \quad (3.1)$$

and ordering solutions according to the eigenvalue magnitude  $\lambda^{(k)}(T)$ . Forming the outer product of equal-time eigenvectors defines the distillation [67] smearing kernel

$$\square(\vec{x}, \vec{y}; T)_{ab} = \sum_{k=1}^{R_{\mathcal{D}}} \xi_a^{(k)}(\vec{x}, T) \xi_b^{(k)\dagger}(\vec{y}, T), \quad (3.2)$$

where  $R_{\mathcal{D}}$  is the desired distillation space rank and color indices  $\{a, b\}$  are made explicit. Two-point correlation functions formed by Wick-contracting quark fields smeared via (3.2) can be recast into a trace over distinct reusable objects constructed within the distillation space, the so-called *elementals* and *perambulators*. The elementals encode the interpolator construction, which in the case of baryons read

$$\Phi_{\alpha\beta\gamma}^{(i,j,k)}(T) = \epsilon^{abc} \left( \mathcal{D}_1 \xi^{(i)} \right)^a \left( \mathcal{D}_2 \xi^{(j)} \right)^b \left( \mathcal{D}_3 \xi^{(k)} \right)^c (T) S_{\alpha\beta\gamma}, \quad (3.3)$$

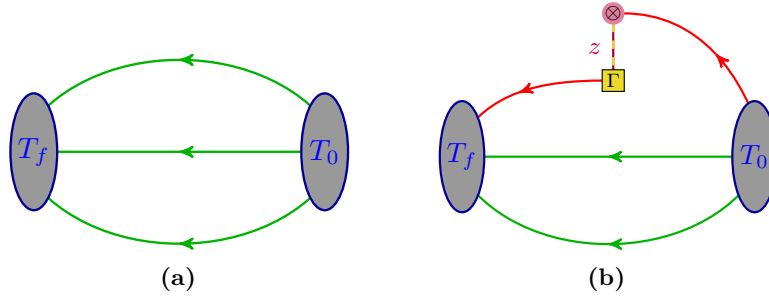
where  $\mathcal{D}_i$  are covariant derivatives, and  $S_{\alpha\beta\gamma}$  encode the patterns of subduction of a continuum interpolator across irreducible representations (irreps) of a hypercubic lattice and its associated little groups. Quark propagation between distillation spaces is captured by the perambulators

$$\tau_{\alpha\beta}^{(l,k)}(T_f, T_0) = \xi^{(l)\dagger}(T_f) M_{\alpha\beta}^{-1}(T_f, T_0) \xi^{(k)}(T_0), \quad (3.4)$$

with  $M$  the Dirac operator. In the case of three-point correlation functions, an additional computational unit appears. The generalized perambulator, or *genprop* for short, carries the same external indices as a standard perambulator, but includes an intermediate operator insertion. In the case of the Wilson-line operator specific to this work, the genprop reads

$$\begin{aligned} \Xi_{\alpha\beta}^{(l,k)}(T_f, T_0; \tau, z_3) &= \sum_{\vec{y}} \xi^{(l)\dagger}(T_f) M_{\alpha\sigma}^{-1}(T_f, \tau; \vec{y} + z_3 \hat{z}) \left[ \gamma^4 \right]_{\sigma\rho} \\ &\quad \times \Phi_{\hat{z}}^{(f)}(\{\vec{y} + z_3 \hat{z}, \vec{y}\}) M_{\rho\beta}^{-1}(\tau, T_0; \vec{y}) \xi^{(k)}(T_0), \end{aligned} \quad (3.5)$$

where the unpolarized PDFs are selected with the Dirac matrix  $\gamma^4$ , and the sum over the initial spatial coordinates of the Wilson line  $\vec{y}$  ensures a zero 3-momentum projection. We



**Figure 1.** (a) Two-point and (b) Wilson-line type three-point correlation functions factorized into distinct elementals (blue), perambulators (green) and generalized perambulators (red).

remark the factorization of correlation functions made manifest by distillation allows for an efficient implementation of the variational method with an extended basis of operators. These factorizations are shown diagrammatically in figure 1.

The viability of the pseudo-distribution formalism hinges on the space-like quark bilinear remaining in a perturbative, or short-distance, regime. To then minimize the impact of polynomial- $z^2$  corrections when mapping the  $\nu$ -dependence of the reduced pseudo-ITD, the matrix elements (2.1) must be isolated in frames of varying external nucleon boosts. Increasing the 3-momentum of a hadronic interpolator not only leads to exponentially worsening statistical fluctuations, but also reduces the efficacy of spatially-smeared interpolators to overlap onto the desired hadronic states. An improvement program capable of increasing interpolator-state overlaps in boosted frames, now known as *momentum smearing* was first established in [68]. Momentum smearing effectively shifts the operator-ground-state overlap peak in momentum space; crucially, this shift also improves the overlap of a boosted interpolator onto neighboring excited-states. In anticipation of an increasingly dense nucleon spectrum in boosted frames, we adopt in this work a modification of the distillation paradigm [60] which simultaneously leverages the momentum smearing heuristic with improved control over excited-states. In the interest of self-containment, we now summarize this procedure.

The momentum smearing idea is incorporated within distillation by applying spatially-varying phases to a pre-computed eigenvector basis

$$\tilde{\xi}^{(k)}(\vec{z}, T) \equiv e^{i\vec{\zeta} \cdot \vec{z}} \xi^{(k)}(\vec{z}, T), \quad (3.6)$$

where we designate the modified eigenvectors as *phased*. As the eigenvector basis already reflects the periodicity of the spatial lattice, the phase factors introduced in this manner are restricted to allowed lattice momenta. This requirement was not found to be limiting, but rather offered broad improvement of momentum space overlaps for a range of nucleon momenta [60]. All distillation components must then be reconstructed on each new modified basis. In the interest of compute cycles and storage, we adopt three eigenvector bases: the pre-computed basis (3.1) for the nucleon at rest and with small  $\hat{z}$ -momenta ( $|ap_z| \leq 3[2\pi/L]$ ), while two additional bases are formed according to

$$\xi_{\pm}^{(k)}(\vec{z}, T) \equiv e^{i\vec{\zeta}_{\pm} \cdot \vec{z}} \xi^{(k)}(\vec{z}, T) \quad (3.7)$$



where  $\vec{\zeta}_{\pm} = \pm 2 \cdot \frac{2\pi}{L} \hat{z}$ . These phased bases were found to be sufficient to resolve the ground-state nucleon in  $\hat{z}$ -boosted frames  $|ap_z| > 3[2\pi/L]$ , respectively [60]. We employ  $R_D = 64$  eigenvectors within each basis.

### 3.2 Matrix element isolation

**Interpolator construction.** The breaking of rotational symmetry by a hypercubic lattice entails baryons appear in lattice calculations according to definite patterns of subduction across the finite number of irreps  $\Lambda$  of the octahedral group double-cover  $O_h^D$ . We elect to use a single spatially local, non-relativistic nucleon interpolating operator constructed according to [69, 70]. The paradigm established in [71] is adopted herein to ensure our interpolator transforms irreducibly under the appropriate little group irreps. A forthcoming calculation will leverage an extended basis of interpolators to extract these same matrix elements.

The space-like matrix element (2.1) is isolated in a flavor isovector combination according to the kinematics highlighted in section 2, and requires computation of standard two-point

$$C_2(p_z, T) = \langle \mathcal{N}(-p_z, T) \bar{\mathcal{N}}(p_z, 0) \rangle = \sum_n |\mathcal{A}_n|^2 e^{-E_n T} \quad (3.8)$$

and three-point correlation functions featuring the unrenormalized Wilson line operator  $\hat{\mathcal{O}}_{\text{WL}}^{[\gamma_4]}(z_3, \tau)$

$$\begin{aligned} C_3(p_z, T, \tau; z_3) &= V_3 \langle \mathcal{N}(-p_z, T) \hat{\mathcal{O}}_{\text{WL}}^{[\gamma_4]}(z_3, \tau) \bar{\mathcal{N}}(p_z, 0) \rangle \\ &= V_3 \sum_{n, n'} \langle \mathcal{N} | n' \rangle \langle n | \bar{\mathcal{N}} \rangle \langle n' | \hat{\mathcal{O}}_{\text{WL}}^{[\gamma_4]}(z_3, \tau) | n \rangle e^{-E_{n'}(T-\tau)} e^{-E_n T}, \end{aligned} \quad (3.9)$$

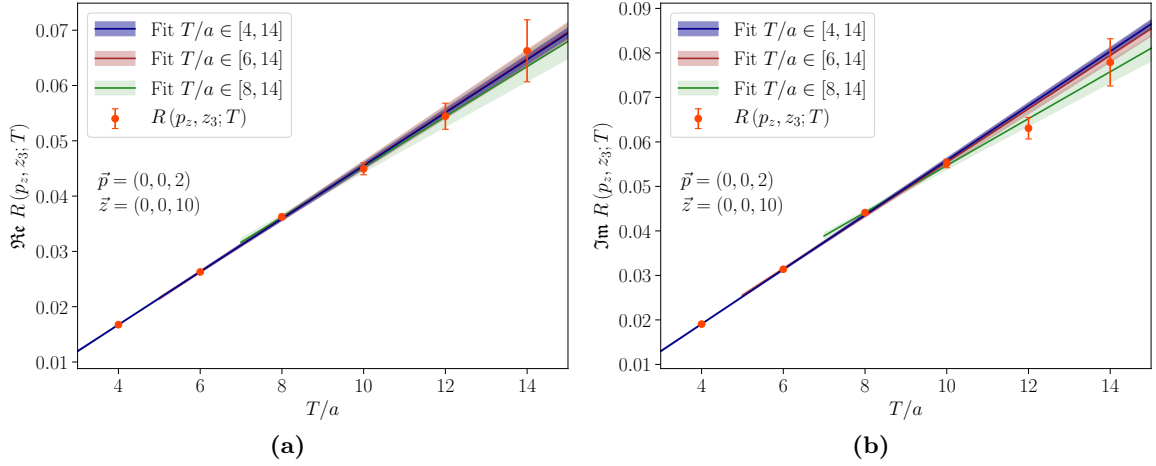
where energy eigenstate contributions are denoted, and the nucleon interpolating fields  $\mathcal{N}$  that are smeared with distillation are separated by a Euclidean time  $T$ . The explicit momentum projection built into the genprop (3.5) leads to an overall spatial volume factor  $V_3$  in the forward case. The Wilson line operator is inserted for  $0 < \tau < T$ . These correlators and the factorization manifest through distillation are illustrated in figure 1.

The spectral representations of (3.8) and (3.9) indicate the desired ground-state matrix element follows from ratios of three-point to two-point correlation functions, which plateau asymptotically for  $0 \ll \tau \ll T$ . The contamination from excited-states is reduced further in this calculation by extracting the matrix elements using the *Summation method* [72, 73], whereby the time slice  $\tau$  on which  $\hat{\mathcal{O}}_{\text{WL}}^{[\gamma_4]}(z_3)$  is introduced is summed over

$$R(p_z, z_3; T) = \sum_{\tau=1}^{T-1} \frac{C_3(p_z, T, \tau; z_3)}{C_2(p_z, T)}. \quad (3.10)$$

Note any contact terms are explicitly excluded. Excited-states in (3.10) scale as  $\exp[-\Delta E T]$ , while in plateau and multi-state fits these effects scale as  $\exp[-\Delta E T/2]$ . As the ground-excited state gap  $\Delta E$  is generally large at low-momenta, the gains afforded by the summation method over plateau/multi-state fits are modest. However, at high-momenta  $\Delta E$  becomes small and the summation method offers considerable suppression of excited-states relative





**Figure 2.** (a) Real and (b) imaginary summation data  $R(p_3, z_3; T)$  for  $p_3 \simeq 0.82$  GeV and  $z/a = 10$ , together with the linear fit (3.11) applied for different time series. The slope of each linear fit yields the bare matrix element, which is seen to be consistently determined for varied fitting windows.

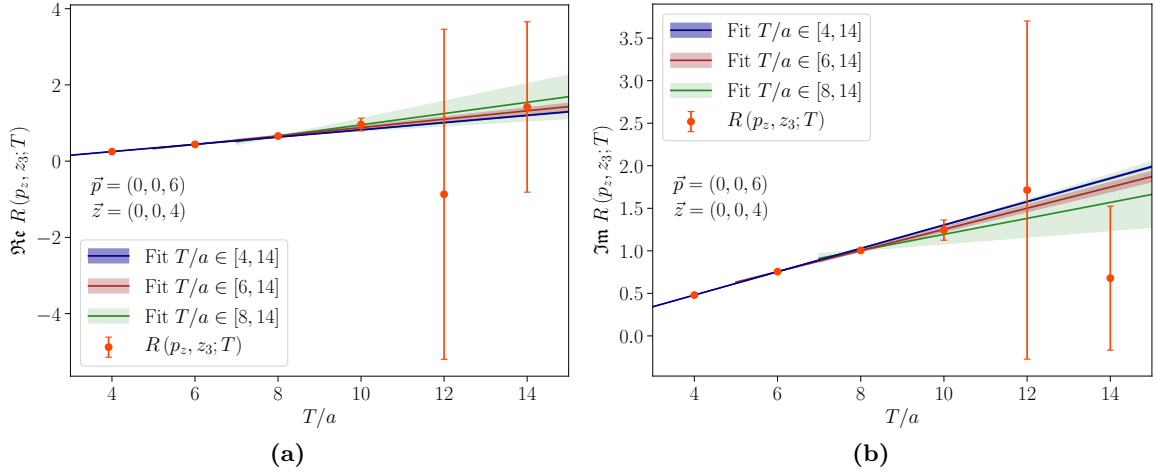
to plateau and multi-state fits. The geometric series resulting from (3.10) depends linearly on the targeted matrix element  $M_4(p_z, z_3)$ , for which we implement the fitting functional

$$R_{\text{fit}}(p_z, z_3; T) = \mathcal{A} + M_4(p_z, z_3)T + \mathcal{O}(e^{-\Delta ET}). \quad (3.11)$$

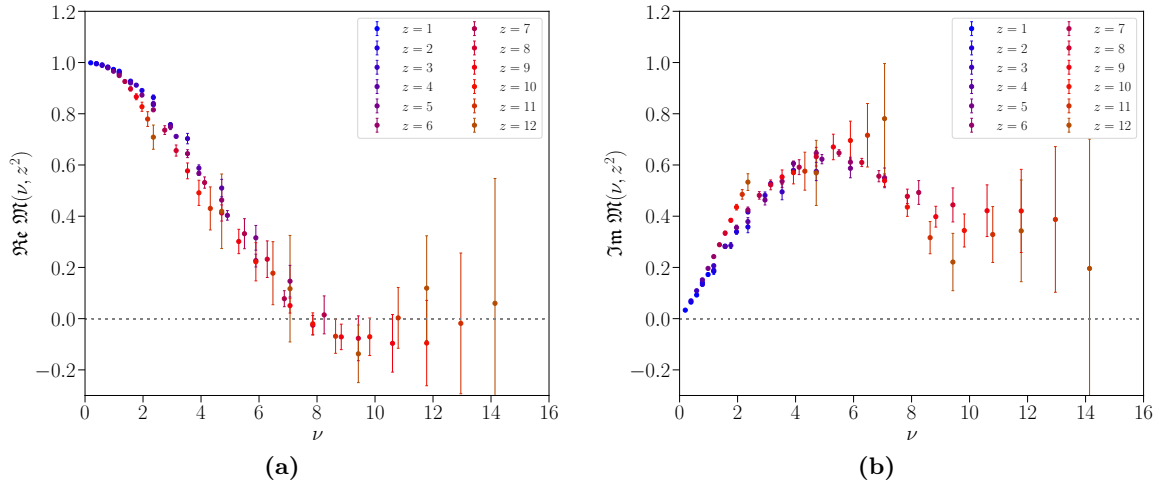
We note in practice, the excited-state term  $\mathcal{O}(e^{-\Delta ET})$  is found to have no impact on our summation fits and is hence omitted from our results.

The two- and three-point functions are computed on four temporal source origins per configuration with  $T/a \in \{4, 6, 8, 10, 12, 14\} \sim 0.38\text{--}1.32$  fm. This number of source-sink separations is chosen to filter out any excited-states that are not captured by the combined effect of distillation and the summation method, and to ensure our linear fits (3.11) do not over fit our data as signal-to-noise problems become unavoidable. We consider nucleon momenta up to  $|p_z| = 6 \times [2\pi/aL] \sim 2.47$  GeV and Wilson line lengths up to  $z_3/a = 16$ , although only  $z_3/a \leq 12$  will be subsequently used in our analysis. As stated in table 1, the number of temporal origins and configurations is held fixed for all computed matrix elements regardless of  $T/a$ ,  $p_z$  and  $z_3$ . A representative set of  $R(p_3, z_3; T)$  and applied linear fits are shown in figures 2 and 3. Repeating the matrix element extraction for all momenta and displacements, we gather the real/imaginary components of the unpolarized reduced pseudo-ITD in figures 4a and 4b.

Our adoption of distillation in this calculation is justified numerically by making a direct comparison in appendix A with an earlier calculation of the nucleon’s unpolarized PDF [58], which instead used spatial Jacobi smearing on the same ensemble. A cost comparison between the two approaches is also presented, making a strong case for distillation in future calculations of hadronic structure.



**Figure 3.** (a) Real and (b) imaginary summation data  $R(p_3, z_3; T)$  for  $p_3 \simeq 2.47$  GeV and  $z/a = 4$ , together with the linear fit (3.11) applied for different time series. The slope of each linear fit yields the bare matrix element. Slight tension in the extracted matrix element is observed as the fitting window is altered. Although minor, this stems jointly from the lack of constraint provided by the  $T/a = 12, 14$  data and the greater flexibility afforded to each fit as the minimum  $T/a$  is increased.



**Figure 4.** (a) Real and (b) imaginary component of the unpolarized reduced pseudo-ITD on the  $a094m358$  ensemble with Wilson line extents  $z/a < 13$ .

## 4 Extraction of unpolarized nucleon PDFs

As PDFs are determined phenomenologically at a factorization scale  $\mu^2$  in  $\overline{\text{MS}}$  to renormalize the associated collinear divergences, the nucleon unpolarized reduced pseudo-ITD shown in figure 4a and 4b must be matched to a common scale in  $\overline{\text{MS}}$  prior to any meaningful comparisons. At one-loop and without loss of generality, negating the sign of the  $\mathcal{O}(\alpha_s)$  correction and interchanging the ITD and reduced pseudo-ITD in (2.5) one obtains the

factorization relationship that matches the reduced pseudo-ITD to the ITD:

$$\mathcal{Q}(\nu, \mu^2) = \mathfrak{M}(\nu, z^2) + \frac{\alpha_s C_F}{2\pi} \int_0^1 du \left[ \ln \left( \frac{e^{2\gamma_E+1}}{4} z^2 \mu^2 \right) B(u) + L(u) \right] \mathfrak{M}(u\nu, z^2). \quad (4.1)$$

This relationship describes the evolution of each distinct set of  $\mathfrak{M}(\nu, z^2)$  data at a given  $z^2$  to a common scale  $\mu^2$  in  $\overline{\text{MS}}$ . Regardless of whether the evolution and matching steps are done separately or in one step, a smooth and continuous description of the reduced pseudo-ITD for each  $z^2$  in the interval  $[0, \nu]$  is required. It is common in the literature to find polynomials in Ioffe-time fit to each set of distinct  $z^2$  data in order to build  $\mathfrak{M}(u\nu, z^2)$  [56–58, 74]. Interpolations are also common, and when used have been found to be consistent with polynomial fits [57, 58].

A polynomial in  $\nu$  is perhaps a dubious choice, as it cannot capture the correct limiting behavior of the ITD at large- $\nu$ . To understand this, consider a simple nucleon valence PDF ansatz

$$f_{q_v/N}(x) = \frac{\Gamma(\alpha + \beta + 2)}{\Gamma(\alpha + 1)\Gamma(\beta + 1)} x^\alpha (1-x)^\beta. \quad (4.2)$$

The cosine transform of this ansatz with respect to Ioffe-time is given by

$$\Re \mathcal{Q}(\nu, \alpha, \beta) = \frac{\pi \Gamma(2 + \alpha + \beta)}{2^{1+\alpha+\beta}} {}_2F_3 \left( \frac{1+\alpha}{2}, \frac{2+\alpha}{2}; \frac{1}{2}, \frac{2+\alpha+\beta}{2}, \frac{3+\alpha+\beta}{2}; -\frac{\nu^2}{4} \right), \quad (4.3)$$

with  ${}_2F_3$  a generalized hypergeometric function and  $\alpha, \beta > -1$ . In the large Ioffe-time regime (4.3) behaves as

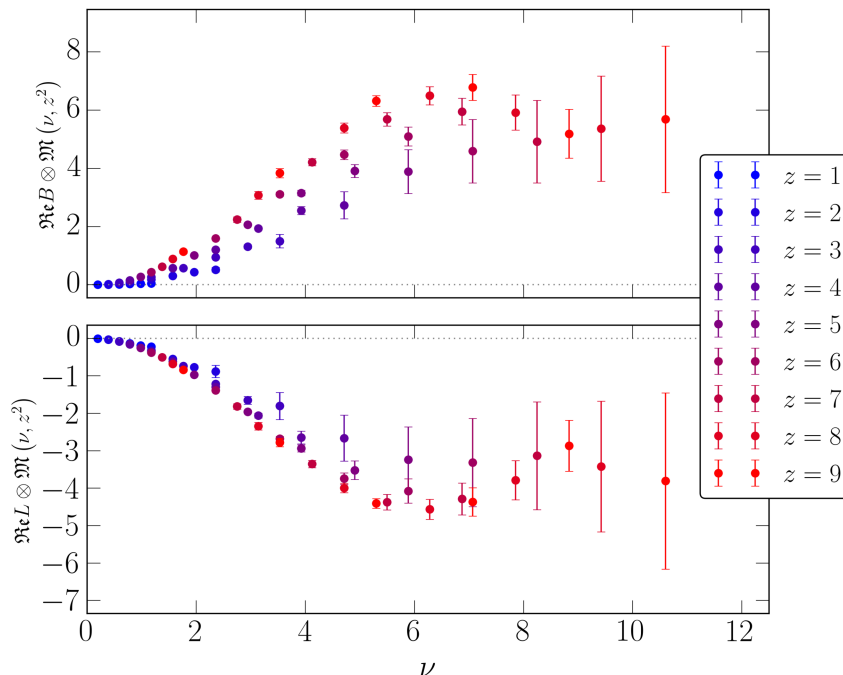
$$\Re \mathcal{Q}(\nu) \simeq \beta \cos \left( \frac{\pi}{2} \alpha \right) \frac{\Gamma(\alpha + 2)}{\nu^{\alpha+2}} - \sin \left( \frac{\pi}{2} \alpha \right) \frac{\Gamma(\alpha + 1)}{\nu^{\alpha+1}}. \quad (4.4)$$

For the real component of the ITD to correspond to a valence PDF with a finite sum rule, the ITD must then vanish for asymptotically large- $\nu$  (i.e.  $\alpha > -1$ ). This suggests the usefulness of a smooth polynomial in  $\nu$  extends only so far as interpolating the discrete pseudo-ITD data, and should not be used as a measure of the moments of the pseudo-PDFs given their divergent behavior at large- $\nu$ . This motivates methods to directly extract the PDFs from the reduced pseudo-ITD, thereby obviating the need for a continuous description of  $\mathfrak{M}(\nu, z^2)$  in order to perform the evolution/matching steps. This will be developed in sections 4.2 and 4.3.

To get a handle on the scale dependence of our data and ground the ensuing discussion, we nonetheless start with a provisional sixth order polynomial fit in Ioffe-time to the reduced pseudo-ITD for constant  $z^2$ :

$$\mathfrak{M}(\nu, z^2) = 1 + \sum_{n=1}^3 \left( c_{2n} \nu^{2n} + i c_{2n-1} \nu^{2n-1} \right). \quad (4.5)$$

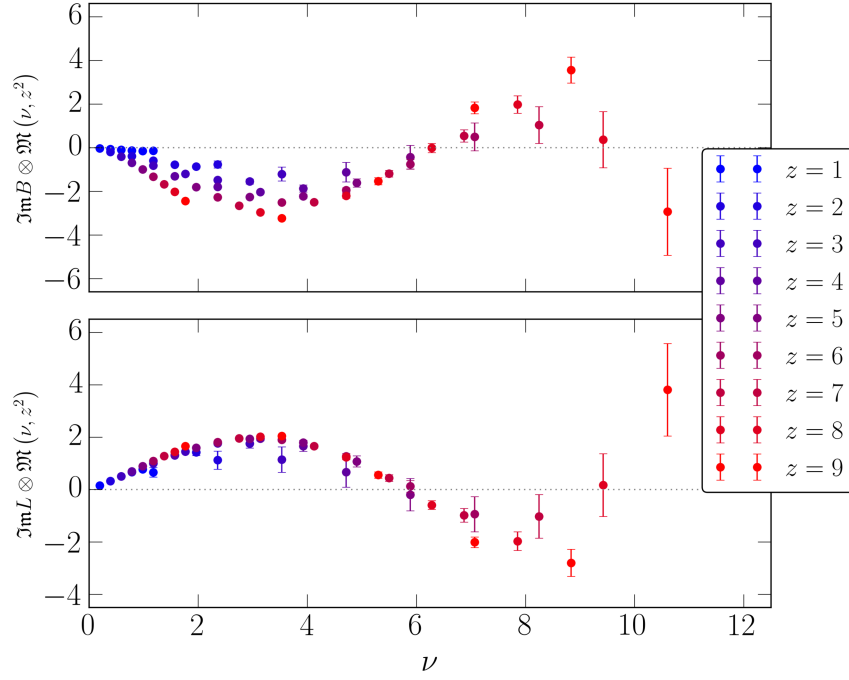
The even (odd) powers of the polynomial are applied to jackknife samples of the real (imaginary) component of  $\mathfrak{M}(\nu, z^2)$  given in figures 4a and 4b. Higher order polynomials were considered, but were found to be unconstrained by the data. With the polynomial fits



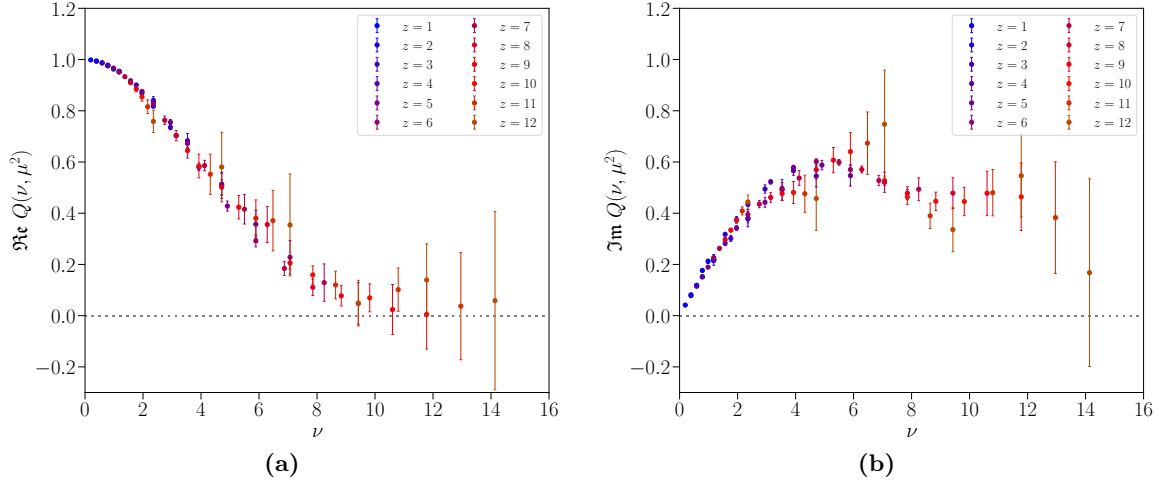
**Figure 5.** Convolutions needed to evolve (upper) and match (lower) the real component of the reduced pseudo-ITD to a common scale of 2 GeV in  $\overline{\text{MS}}$ . The NLO prefactor  $\alpha_s C_F/2\pi$  is included in these data, but omitted from the labels for clarity. The convolutions were performed up to  $z/a = 16$ , but data for  $z/a > 9$  are generally noisy and not shown.

in hand, we perform the evolution and scheme conversion convolutions (4.1) in a single step. The matched  $\overline{\text{MS}}$  scale  $\mu = 2 \text{ GeV}$  was chosen, and the strong coupling  $\alpha_s(2 \text{ GeV}) \simeq 0.303$  was adopted from LHAPDF6 [75]. The scale  $\mu = 2 \text{ GeV}$  corresponds to the reduced pseudo-ITD being evolved to the common scale  $z_0^2 = 4e^{-2\gamma_E-1}(2 \text{ GeV})^{-2} \simeq 0.12 \text{ GeV}^{-2}$  or  $z_0^{-1} \simeq 2.94 \text{ GeV}$ . On this ensemble *a094m358*, this common scale then equates to  $z_0^2/a^2 \simeq 0.511$ . The computed evolution and scheme matching convolutions are depicted in figure 5 and figure 6 for the real and imaginary components, respectively. It is curious the evolution and matching convolutions appear to be nearly equal in magnitude but opposite in sign. This feature of the pseudo-distributions has been observed in independent calculations [56, 57] and hints an NNLO matching relation may not be needed. Nonetheless, a future study will explore to what effect the matching relation (2.5), truncated at NLO in this work, can be improved at NNLO [46].

When the scale and scheme conversion are incorporated, we observe in figure 7a and figure 7b a dramatic collapse of the reduced pseudo-ITD onto a common curve for  $z/a \lesssim 10$ . The lack of residual  $z^2$ -dependence is particularly striking in the real component of the ITD, but less so in the imaginary component. This confirms the formation of the reduced ratio (2.4) indeed cancels much of the  $z^2$ -dependence in the pseudo-ITD, with any remaining at small- $z^2$  ideally described by the coordinate-space DGLAP evolution [29].



**Figure 6.** Convolutions needed to (upper) evolve and (lower) match the imaginary component of the reduced pseudo-ITD to a common scale of 2 GeV in  $\overline{\text{MS}}$ . The NLO prefactor  $\alpha_s C_F/2\pi$  is included in these data, but omitted from the labels for clarity. The convolutions were performed up to  $z/a = 16$ , but data for  $z/a > 9$  are generally noisy and not shown.



**Figure 7.** (a) Real and (b) imaginary component of the unpolarized ITD at a scale of 2 GeV in  $\overline{\text{MS}}$  obtained from the matching relation (4.1), where a smooth and continuous description of the  $\mathfrak{M}(\nu, z^2)$  data is provided by a polynomial fit for each  $z^2$ . Data are shown for Wilson line extents  $z/a \leq 12$ ; extents  $z/a > 12$  are considerably uncertain, and thus excluded from our ensuing analysis.

#### 4.1 An ill-posed inverse

The Fourier transform relating the  $x$ -dependence of the PDF to the  $\nu$ -dependence of the ITD

$$\mathcal{Q}(\nu, \mu^2) = \int_{-1}^1 dx e^{i\nu x} f_{a/h}(x, \mu^2) \quad (4.6)$$

is an ill-posed inverse problem, as the ITD is computed in a discrete and limited range of Ioffe-time. Regularization at this stage is however numerically cheap and more stable relative to a direct matching of the  $\{x, \mu^2\}$  dependencies of the PDF to the  $\{\nu, z^2\}$  dependencies of the reduced pseudo-ITD. A direct inversion of (4.6) is satisfied by an infinite number of solutions, each of which having little predictive or postdictive credibility. The futility of direct inversions has been demonstrated in a few LCS calculations [57, 58], wherein each inversion yielded unstable PDFs with spurious oscillations. The limited range of Ioffe-time accessible to present PDF calculations only compounds the need for refined extraction methods.

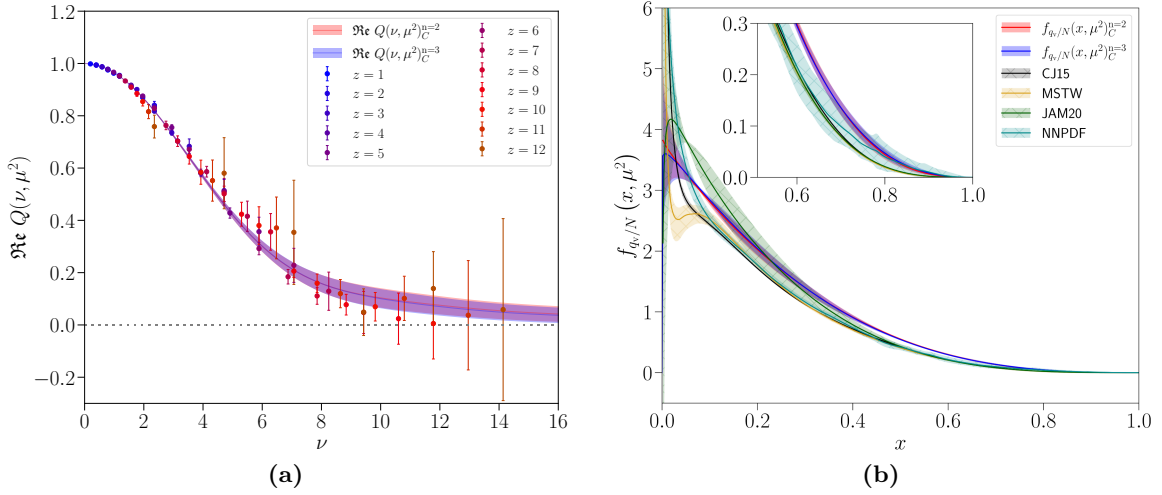
This inversion problem is shared with other lattice formalisms that rely on QCD factorization, and indeed the global analysis of inclusive/semi-inclusive processes. Although different in character, this problem pervades the quantitative sciences and even impacts the image reconstruction of black holes [76]. Arguably the most serious systematic that must then be confronted in LCS studies is how to reliably extract a targeted distribution, while minimizing numerical artifacts and bias. Numerous sophisticated methods, such as the Backus-Gilbert [77], maximum entropy [78], and Bayesian reconstruction methods have been explored as tools to aid in PDF extractions [57, 79] and other observables more generally [80, 81].

A common heuristic to regularize the inverse problem at hand, both in the global fitting of inclusive scattering data [82–87] and analogous lattice calculations, is to supply additional information in the form of physically motivated PDF parameterizations. In particular, a parametric form can incorporate the known divergence/convergence of PDFs at small-/large- $x$  and enforce any parton sum rules explicitly. We note parametric frameworks in the literature generally differ in what functional form is used to smoothly connect the two limiting  $x$ -space regimes. Inspired by the phenomenological forms of the global fitting community and to establish a benchmark for the alternate extraction methods that follow, we opt to first regularize the inverse problem by parameterizing the valence/plus quark distributions according to

$$f_{q_v/N}(x, \mu^2) = f_{q/N}(x, \mu^2) - f_{\bar{q}/N}(x, \mu^2) = N_v x^\alpha (1-x)^\beta P(x) \quad (4.7)$$

$$f_{q_+/N}(x, \mu^2) = f_{q_v/N}(x, \mu^2) + 2f_{\bar{q}/N}(x, \mu^2) = N_+ x^{\alpha_+} (1-x)^{\beta_+} P(x), \quad (4.8)$$

where  $P(x)$  is a smooth interpolating polynomial and  $N_v^{-1} = \left[ B(\alpha+1, \beta+1) + \sum_k \lambda_k B(\alpha + 1 + \frac{k+1}{2}, \beta+1) \right]$  ensures the valence quark sum rule  $\int_0^1 dx f_{q_v/N}(x, \mu^2) = 1$  is satisfied; the normalization of  $f_{q_+/N}(x, \mu^2)$  is not fixed by a sum rule, and is left to float in our fits. Given the limited range of Ioffe-time in our results, we will find the simplest 2-parameter ansatz with  $P(x) = 1$  cannot be avoided. Where possible, the bias introduced by this



**Figure 8.** (a) The real component of the matched ITD at  $\mu = 2$  GeV in  $\overline{\text{MS}}$  fit by cosine transforms of two- and three-parameter model PDFs (4.9). Data has been fit for  $z/a \leq 12$ , and correlations have been neglected. The resulting PDF parameters and figure of merit are gathered in table 2. (b) The nucleon unpolarized valence quark PDF at 2 GeV in  $\overline{\text{MS}}$  determined from the uncorrelated cosine transform fits (4.9) applied to real component of the matched ITD. Comparisons are made with the NLO global analyses of CJ15 [82] and JAM20 [88], and the NNLO analyses of MSTW [89] and NNPDF [87] at the same scale.

highly-constraining choice will be studied by supplementing  $P(x)$  with additional half-integer powers of  $x$ :  $P(x) = 1 + \sum_k \lambda_k x^{(k+1)/2}$ , thereby increasing the flexibility of our parameterizations beyond the nominal PDF behavior  $x^\alpha (1-x)^\beta$ .

We start with two- and three-parameter PDF parameterizations, where in the latter we take  $P(x) = 1 + \delta x$ . The cosine/sine transforms of the PDF forms (4.7) and (4.8)

$$\Re Q(\nu, \mu^2) = \int_0^1 dx \cos(\nu x) f_{q_v/N}(x, \mu^2) \quad (4.9)$$

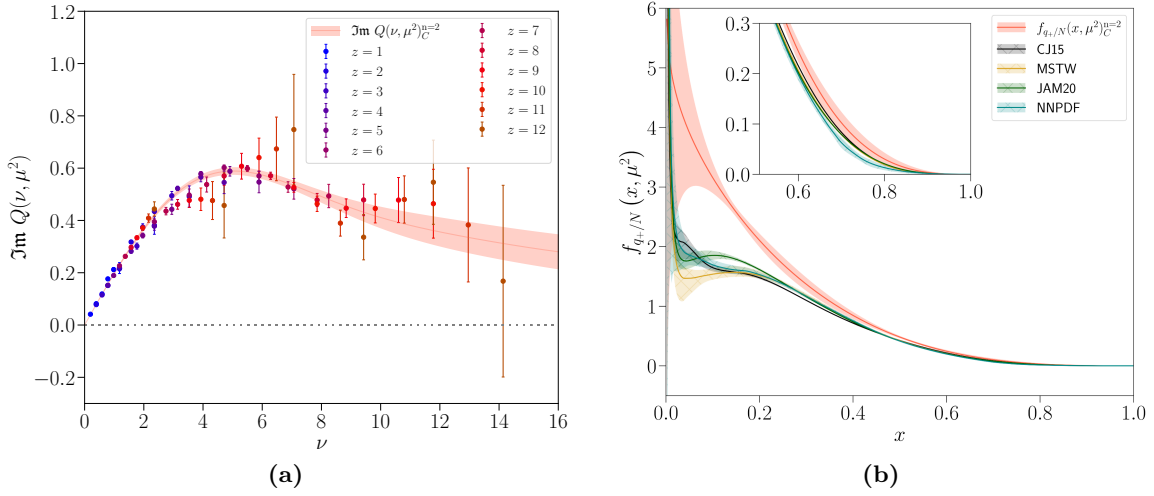
$$\Im Q(\nu, \mu^2) = \int_0^1 dx \sin(\nu x) f_{q_+/N}(x, \mu^2) \quad (4.10)$$

are fit to the real/imaginary ITD data using first an uncorrelated least-squares regression

$$\chi^2 = \sum_{\nu_{\min}}^{\nu_{\max}} \frac{[Q(\nu, \mu^2) - Q_{\text{fit}}(\nu, \mu^2)]^2}{\sigma_Q^2}, \quad (4.11)$$

with  $\sigma_Q^2$  the variance of the ITD, and  $\{\nu_{\min}, \nu_{\max}\}$  representing potential cuts on the data. These uncorrelated fits include all  $z/a \in \{1, \dots, 12\}$  and  $ap_z \in \{1, \dots, 6\} \times 2\pi/L$ . For ease of later reference, this method of extraction is denoted type-*C*. The fits to the real and imaginary components of the ITD are shown in figure 8a and figure 9a. The resulting valence and plus quark PDFs are juxtaposed with phenomenological determinations in figure 8b and figure 9b. The phenomenological PDFs are three flavor NLO determinations





**Figure 9.** (a) The imaginary component of the matched ITD at  $\mu = 2 \text{ GeV}$  in  $\overline{\text{MS}}$  fit by the sine transform of a two-parameter model PDF (4.10). Data has been fit for  $z/a \leq 12$ , and correlations have been neglected. The resulting PDF parameters and figure of merit are gathered in table 2. (b) The nucleon unpolarized plus quark PDF at  $2 \text{ GeV}$  in  $\overline{\text{MS}}$  determined from the uncorrelated sine transform fits (4.10) applied to the imaginary component of the matched ITD. Comparisons are made with the NLO global analyses of CJ15 [82] and JAM20 [88], and the NNLO analyses of MSTW [89] and NNPDF [87] at the same scale.

by the CJ [82] and JAM [88] collaborations, and three flavor NNLO determinations of MSTW [89] and NNPDF [87].<sup>1</sup>

Apart from the  $z/a \geq 9$  data, such an uncorrelated fit would seem to well describe the  $\Re Q(\nu, \mu^2)$  data and lead to valence PDFs that feature many structural similarities with phenomenological determinations at the same scale. The statistically consistent figure of merit for the two- and three-parameter fits tabulated in table 2, however indicates the data cannot distinguish between these models. The two-parameter fit to  $\Im Q(\nu, \mu^2)$  is clearly more heavily constrained by the  $z/a \lesssim 7$  data, and all but avoids points of the ITD originating from larger separations. Above  $x \sim 0.4$  the extracted  $f_{q+/N}(x, \mu^2)_C^{n=2}$  result likewise shares structural similarities with the shown phenomenological results. The lack of any large- $\nu$  constraint provided by  $\Im Q(\nu, \mu^2)$  entails a generally unconstrained fitted PDF in the small- $x$  regime, although this relation is not bijective.

## 4.2 Direct extraction of PDFs from reduced pseudo-ITDs

A separate, though in principle equivalent, route to extract PDFs from these data is to directly apply the factorized relationship (2.5) having substituted the definition of the ITD (4.6):

$$\mathfrak{M}(\nu, z^2) = \int_{-1}^1 dx \int_0^1 du \, C(u, z^2 \mu^2, \alpha_s(\mu^2)) e^{ix\nu} f_{q/N}(x, \mu^2) + \sum_{k=1}^{\infty} \mathcal{B}_k(\nu) (z^2)^k. \quad (4.12)$$

<sup>1</sup>LHAPDF set names: CJ15 — CJ15nlo, JAM20 — JAM20-SIDIS\_PDF\_proton\_nlo, MSTW — MSTW2008nnlo68c1\_nf4, NNPDF — NNPDF31\_nnlo\_pch\_as\_0118\_mc\_164.

$N_{\text{param}}$	$N_{v/+}$	$\alpha$	$\beta$	$\delta$	$\chi_r^2$
2	—	−0.006(98)	2.754(285)	—	2.183(483)
3	—	0.019(98)	2.212(291)	−0.737(12)	2.192(490)
2	3.616(2.260)	−0.077(275)	2.983(606)	—	2.780(806)

**Table 2.** Unpolarized nucleon valence and plus quark PDF parameters obtained from performing uncorrelated cosine/sine transform fits to the real/imaginary component of the matched ITD at 2 GeV in  $\overline{\text{MS}}$ . Results for the plus quark PDF are only shown for  $N_{\text{param}} = 2$ , where the smooth polynomial  $P(x) = 1$ , as higher numbers of parameters led to uncontrolled fits. The uncorrelated figure of merit is also shown.

By assuming a PDF parameterization and performing a maximum likelihood regression of the double convolution and  $\mathfrak{M}(\nu, z^2)$ , the introduction of additional systematic errors from the evolution/matching steps and a potentially incorrect functional description of  $\mathfrak{M}(\nu, z^2)$  when interpolating its  $\nu$ -dependence (e.g. eq. (4.5)) can all be avoided. The direct matching relationship between the PDFs and the reduced pseudo-ITD is given by

$$\left\{ \begin{array}{l} \Re \\ \Im \end{array} \right\} \mathfrak{M}(\nu, z^2) = \int_0^1 dx \left\{ \begin{array}{l} \mathcal{K}_v(x\nu, z^2\mu^2) f_{q_v/N}(x, \mu^2) \\ \mathcal{K}_+(x\nu, z^2\mu^2) f_{q_+/N}(x, \mu^2) \end{array} \right\} + \sum_{k=1}^{\infty} \mathcal{B}_k(\nu) (z^2)^k, \quad (4.13)$$

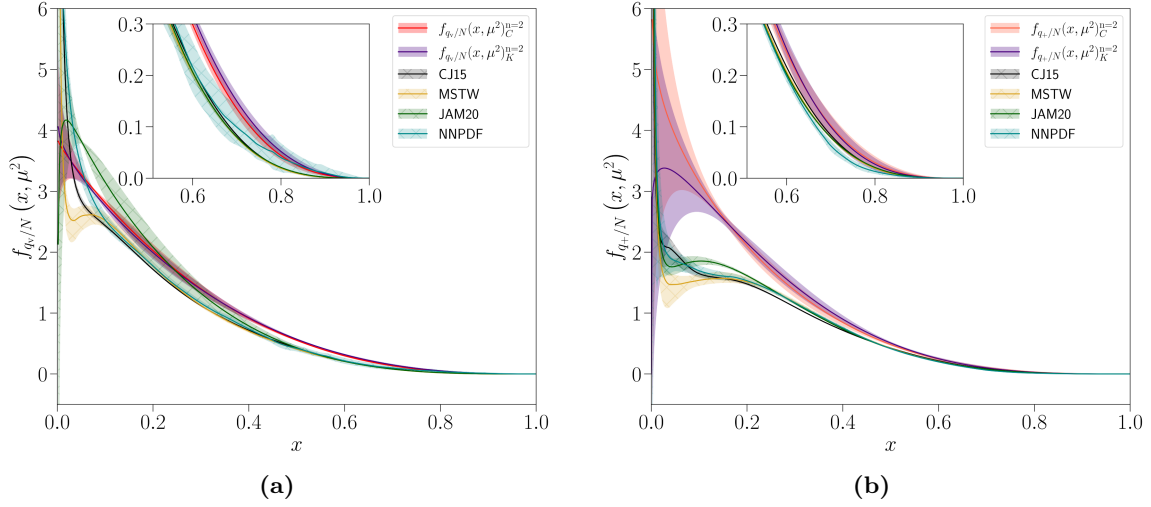
where the one-loop kernels that match the  $\{x, \mu^2\}$ -dependencies of the valence/plus quark PDFs to the reduced pseudo-ITD are given by

$$\begin{aligned} \mathcal{K}_v(x\nu, z^2\mu^2) &= \cos(x\nu) - \frac{\alpha_s C_F}{2\pi} \left[ \ln \left( \frac{e^{2\gamma_E+1}}{4} z^2\mu^2 \right) \tilde{B}_v(x\nu) + \tilde{D}_v(x\nu) \right] \\ \mathcal{K}_+(x\nu, z^2\mu^2) &= \sin(x\nu) - \frac{\alpha_s C_F}{2\pi} \left[ \ln \left( \frac{e^{2\gamma_E+1}}{4} z^2\mu^2 \right) \tilde{B}_+(x\nu) + \tilde{D}_+(x\nu) \right], \end{aligned} \quad (4.14)$$

with the Altarelli-Parisi and scheme matching kernels modified to

$$\begin{aligned} \tilde{B}_v(y) &= \frac{1 - \cos(y)}{y^2} + \frac{3 - 4\gamma_E}{2} \cos(y) + 2 \sin(y) \frac{y \text{Si}(y) - 1}{y} + 2 \cos(y) [\text{Ci}(y) - \ln(y)] \\ \tilde{B}_+(y) &= -\frac{\sin(y) + y}{y^2} + \frac{3 - 4\gamma_E}{2} \sin(y) - 2 \cos(y) \frac{y \text{Si}(y) - 1}{y} + 2 \sin(y) [\text{Ci}(y) - \ln(y)] \\ \tilde{D}_v(y) &= -4y \Im \left[ e^{ix} {}_3F_3(111; 222; -iy) \right] + \left[ \cos(y) \left( 1 + \frac{2}{y^2} \right) - \frac{2}{y^2} \right] \\ \tilde{D}_+(y) &= 4y \Re \left[ e^{ix} {}_3F_3(111; 222; -iy) \right] + \left[ \sin(y) \left( 1 + \frac{2}{y^2} \right) - \frac{2}{y} \right], \end{aligned}$$

and  $\text{Si}(y)/\text{Ci}(y)$  are the integral sine/cosine functions and  ${}_3F_3(111; 222; -iy)$  is a generalized hypergeometric function [47, 90]. A notable challenge of this direct approach is that accurate computation of the generalized hypergeometric function requires multi-precision arithmetic, which is computationally inefficient and invariably slows the parameter optimization. PDFs extracted by directly performing parametric fits (4.13) to the reduced pseudo-ITD data are denoted type-*K*.



**Figure 10.** Two-parameter valence (a) and plus (b) quark PDF resulting from type- $C$  (red) and type- $K$  (indigo) fits to the unpolarized nucleon ITD and reduced pseudo-ITD, respectively. The direct matching fits are consistent with the cosine/sine transform of the model PDF fit to the ITD.

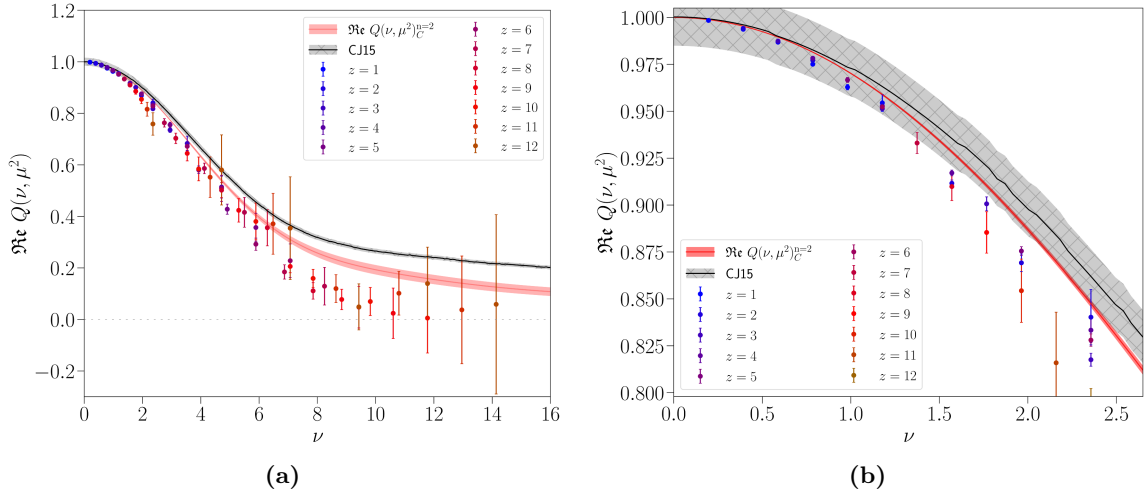
$N = 2_{v/+}$	$N_{v/+}$	$\alpha$	$\beta$	$\chi_r^2$
$2_v$	—	$-0.030(96)$	$2.601(277)$	$7.364(761)$
$2_+$	$5.131(3.405)$	$0.091(299)$	$3.244(638)$	$4.536(902)$

**Table 3.** Unpolarized nucleon valence and plus quark PDF parameters obtained from type- $K$  fits to the real/imaginary component of  $\mathfrak{M}(\nu, z^2)$ .

Simple two-parameter PDFs obtained from uncorrelated type- $K$  fits are shown in figure 10a and figure 10b, together with the same phenomenological determinations and the uncorrelated type- $C$  two-parameter PDF fits. The type- $C$  and type- $K$  fits are statistically consistent. This is confirmed by comparing the type- $K$  fit results in table 3 to the type- $C$  results in table 2. However, the central values of the type- $K$  fits suggest that at small- $x$  the  $f_{q_v/N}(x)$  is more divergent and the  $f_{q_+/N}(x)$  is instead convergent for small- $x$  at the scale  $\mu = 2 \text{ GeV}$ . The factor of two or three increase in the figure of merit when switching from type- $C$  to type- $K$  fits is the first indication of puzzling behavior in  $\mathfrak{M}(\nu, z^2)$ . We reiterate the naive two-parameter fits capture the known limiting regimes of the PDFs. The poor figures of merit in the type- $K$  fits hint that  $\mathfrak{M}(\nu, z^2)$  at this stage apparently does not align well with expectations from the direct matching (4.13). This is a potentially disastrous conclusion. To gain some insight, we now consider the data correlations.

#### 4.2.1 Data correlation

The data featured in this work, and indeed any lattice calculation, naturally are correlated. By beginning this subsection with uncorrelated fits, we highlight that without knowledge or through the simple neglect of data correlations, which appears to be common in the literature, one might incorrectly assume an adequate description of the data has been



**Figure 11.** (a) Real component of the matched ITD at  $\mu = 2$  GeV in  $\overline{\text{MS}}$  fit by the cosine transform of a two-parameter model PDF (4.7). Data have been fit for  $z/a \leq 12$  and *data correlations have been incorporated*. The fit clearly misses each point of the ITD. The derived CJ15 ITD at the same scale is shown for reference. (b) Focus is given to the small- $\nu$  region. The correlated two-parameter fit is seen to deviate appreciably from the precise small- $\nu$  data.

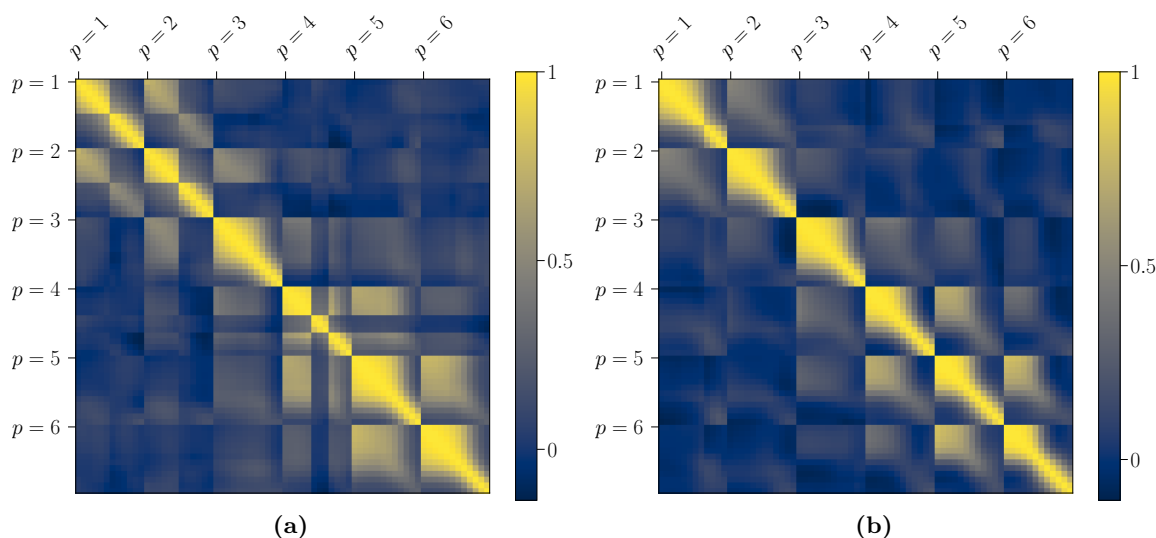
achieved. These correlations must be taken into account in order to provide a rigorous accounting of mutual fluctuations in the data and thus an agnostic PDF determination.

Simply repeating the two-parameter fit to  $\Re Q(\nu, \mu^2)$ , only this time accounting for the data covariance **Cov**

$$\chi_r^2 = \sum_{i,j=\nu_{\min}}^{\nu_{\max}} q_i^\top \mathbf{Cov}_{ij}^{-1} q_j, \quad (4.15)$$

with  $q_k = \mathcal{Q}(\nu, \mu^2)_k - \mathcal{Q}_{\text{fit}}(\nu, \mu^2)_k$ , we arrive at a much different conclusion shown in figure 11a. The visual discrepancy between the ITD and two-parameter fit is stark, and leads to a correlated figure of merit of  $\mathcal{O}(40)$ . Although the fit misses nearly all of the moderate to large- $z$  points, figure 11b illustrates the large increase in the figure of merit is primarily due to the slight deviation from the very precise  $z/a \lesssim 4$  data.

Visualizing the data covariance in the real component in figure 12a, it is clear the low-momentum data  $ap_z \leq 4\pi/L$  are strongly correlated amongst each other and correlate weakly with the  $ap_z \geq 8\pi/L$  data; some mild correlation is visible with the  $ap_z = 6\pi/L$  data with  $z/a \leq 6$ . Within the  $ap_z \leq 4\pi/L$  channels the strongest correlation can be found in the shortest Wilson line data. These observations provide an explanation for the poor correlated two-parameter fit in figure 11a — the strongest correlation is with the most precise data in our calculation causing any correlated fit to favor the small- $\nu$  data. Indeed strong correlation is also observed amongst the momentum channels  $ap_z = \{4, 5, 6\} \times 2\pi/L$ , but the signal-to-noise degradation for these high-momentum data minimizes their effect on any fit. It is interesting this delineation corresponds to the transition from an unphased to phased eigenvector basis. The data covariance in the imaginary component, shown in figure 12b, shows the strongest correlations within each momentum channel and between



**Figure 12.** Data covariance in the real (a) and imaginary (b) components of the matched ITD at 2 GeV, normalized according to  $\text{Cov}_{ij}/\sqrt{\text{Cov}_{ii}\text{Cov}_{jj}}$ . Within each lattice momentum block, entries are ordered in ascending Wilson line lengths.

adjacent Wilson line lengths (e.g.  $z/a = 4$  and  $z/a = 5$ ). It is then no surprise that a correlated two-parameter PDF parameterization of  $\Im \mathcal{Q}(\nu, \mu^2)$  is also met with a poor figure of merit. The non-trivial structures of correlation evident in these data are indicative of our simple PDF parameterizations (4.7) and (4.8) being inappropriate for these data. The above puzzling, and indeed worrisome, conclusions are given a deeper quantitative understanding in the following sections.

### 4.3 Classical orthogonal polynomials

The phenomenological parameterizations we have considered thus far are but one way to regulate the ill-posed inverse relation between the ITD/reduced pseudo-ITD and the corresponding PDF. These parameterizations nevertheless introduce a model dependence into the extracted PDF. Any PDF faithfully reported from a lattice calculation should take into account the space of functions that smoothly connects the  $x \rightarrow 0$  and  $x \rightarrow 1$  limits. As an alternative means to describe the valence/plus quark sectors and minimize model bias, we propose to parameterize the PDFs by a complete basis of classical orthogonal polynomials [55]. The leveraging of orthogonal polynomials to obtain an unknown distribution is not unique to this work. The approach we adopt parallels efforts to extract PDFs from phenomenological fits of inclusive processes [84, 91], as well as distribution amplitudes [92–94] and inelastic scattering cross sections [95] from matrix elements calculated in lattice QCD.

Consider the Jacobi (hypergeometric) polynomials

$$P_n^{(\alpha, \beta)}(z) = \frac{\Gamma(\alpha + n + 1)}{n! \Gamma(\alpha + \beta + n + 1)} \times \sum_{j=0}^n \binom{n}{j} \frac{\Gamma(\alpha + \beta + n + j + 1)}{\Gamma(\alpha + j + 1)} \left(\frac{z-1}{2}\right)^j, \quad (4.16)$$

which for  $\alpha, \beta > -1$  form a basis of orthogonal polynomials on the interval  $[-1, 1]$  with respect to the metric  $(1-z)^\alpha (1+z)^\beta$ . Under the mapping  $z \mapsto 1-2x$  the shifted Jacobi polynomials

$$\Omega_n^{(\alpha, \beta)}(x) = \sum_{j=0}^n \omega_{n,j}^{(\alpha, \beta)} x^j \quad (4.17)$$

form a complete basis of orthogonal polynomials on the interval  $[0, 1]$  with respect to the metric  $x^\alpha (1-x)^\beta$ . In eq. (4.17) we have defined

$$\omega_{n,j}^{(\alpha, \beta)} = \frac{\Gamma(\alpha + n + 1)}{n! \Gamma(\alpha + \beta + n + 1)} \binom{n}{j} \frac{(-1)^j \Gamma(\alpha + \beta + n + j + 1)}{\Gamma(\alpha + j + 1)}. \quad (4.18)$$

As the set of polynomials  $\{\Omega_n^{(\alpha, \beta)}\}$  span  $x \in [0, 1]$ , a PDF can be expressed generically as

$$f_{a/h}(x) = x^\alpha (1-x)^\beta \sum_{n=0}^{\infty} C_{a,n}^{(\alpha, \beta)} \Omega_n^{(\alpha, \beta)}(x), \quad (4.19)$$

with expansion coefficients  $C_{a,n}^{(\alpha, \beta)}$ . The parameters  $\{\alpha, \beta\}$  lose their familiar characterization of the  $x \rightarrow 0/x \rightarrow 1$  PDF behaviors in place of delineating between different choices of bases. The expansion in Jacobi polynomials in eq. (4.19) is thus entirely generic and model-independent. However, the series of Jacobi polynomials must in practice be truncated at some finite order,  $n$ . The bias then introduced may be studied by fixing the order of truncation and determining the optimum  $\{\alpha, \beta\}$ , or tuning  $\{\alpha, \beta\}$  to capture generic properties of a PDF and subsequently optimize the order of truncation — we will adopt the former.

Our strategy to parameterize the reduced pseudo-ITD using a set of  $\Omega_n^{(\alpha, \beta)}(x)$  will be met by similar numerical difficulties as the type-K fits discussed above. The numerical effort is lessened by considering a Taylor series expansion in  $\nu$  for fixed separations  $z^2$  of the direct matching kernels  $\mathcal{K}_{v/+}(x\nu, z^2\mu^2)$  and eq. (4.19). The contribution of an  $n^{\text{th}}$ -order Jacobi polynomial  $\Omega_n^{(\alpha, \beta)}(x)$  to  $\Re \mathfrak{M}(\nu, z^2)$  and  $\Im \mathfrak{M}(\nu, z^2)$  is given by

$$\begin{aligned} \sigma_n^{(\alpha, \beta)}(\nu, z^2\mu^2) &= \Re \int_0^1 dx \mathcal{K}_v(x\nu, z^2\mu^2) x^\alpha (1-x)^\beta \Omega_n^{(\alpha, \beta)}(x) \\ \eta_n^{(\alpha, \beta)}(\nu, z^2\mu^2) &= \Im \int_0^1 dx \mathcal{K}_+(x\nu, z^2\mu^2) x^\alpha (1-x)^\beta \Omega_n^{(\alpha, \beta)}(x). \end{aligned}$$

Expanding the direct matching kernels  $\mathcal{K}_{v/+}(x\nu, z^2\mu^2)$  in even/odd powers of  $\nu$  one finds

$$\sigma_n^{(\alpha, \beta)}(\nu, z^2\mu^2) = \sum_{j=0}^n \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} c_{2k}(z^2\mu^2) \omega_{n,j}^{(\alpha, \beta)} B(\alpha + 2k + j + 1, \beta + 1) \nu^{2k} \quad (4.20)$$

$$\eta_n^{(\alpha, \beta)}(\nu, z^2\mu^2) = \sum_{j=0}^n \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} c_{2k+1}(z^2\mu^2) \omega_{n,j}^{(\alpha, \beta)} B(\alpha + 2k + j + 2, \beta + 1) \nu^{2k+1}, \quad (4.21)$$

where

$$c_n(z^2\mu^2) = 1 - \frac{\alpha_s C_F}{2\pi} \left[ \gamma_n \ln \left( \frac{e^{2\gamma_E+1}}{4} z^2\mu^2 \right) + d_n \right] \quad (4.22)$$

and the constants  $\gamma_n$  and  $d_n$  are the leading order moments of the Altarelli-Parisi and scheme matching kernels derived in [96], respectively. The sum over  $k$  is to be performed to assure convergence for a given value of  $\nu$  — we have identified  $k_{\max} = 75$  as providing more than adequate numerical precision, in reasonable computation time. With the above definitions, the leading-twist valence and plus quark PDFs describe the reduced pseudo-ITD components according to

$$\Re \mathcal{M}^{lt}(\nu, z^2) = \sum_{n=0}^{\infty} \sigma_n^{(\alpha, \beta)}(\nu, z^2 \mu^2) C_{v,n}^{lt(\alpha, \beta)} \quad (4.23)$$

$$\Im \mathcal{M}^{lt}(\nu, z^2) = \sum_{n=0}^{\infty} \eta_n^{(\alpha, \beta)}(\nu, z^2 \mu^2) C_{+,n}^{lt(\alpha, \beta)}, \quad (4.24)$$

where the  $C_{v/+,n}^{lt(\alpha, \beta)}$  are the Jacobi polynomial expansion coefficients.

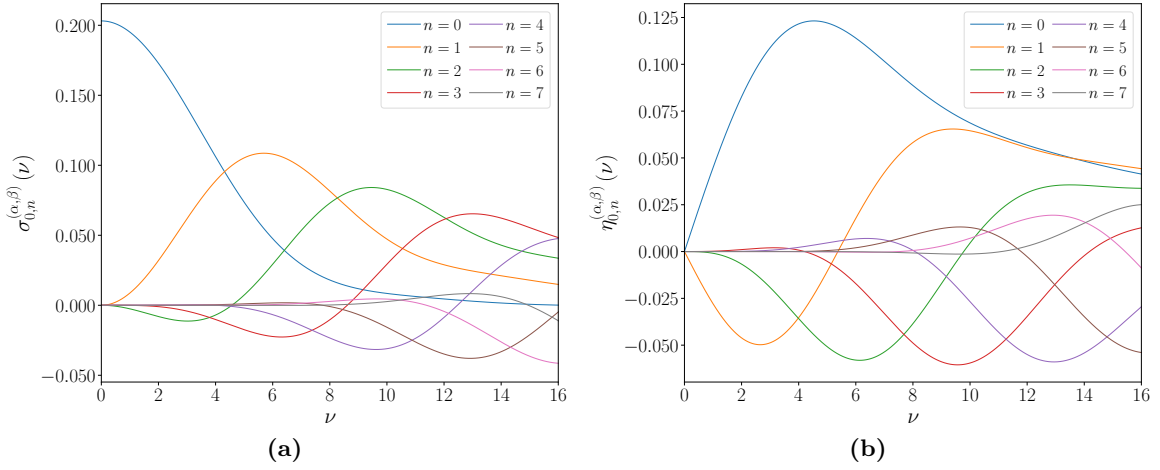
Since the space-like matrix element (2.1) is on-shell and given our choice of an  $\mathcal{O}(a)$  improved lattice QCD formulation, any contaminating effects of this calculation must scale as  $a^2 \Lambda_{\text{QCD}}^2$ ,  $z^2 \Lambda_{\text{QCD}}^2$ , and as some ratio of the lattice spacing  $a$  and Wilson line length  $z$ . Apart from the Ioffe-time  $\nu = p \cdot z$ , these terms are the only possible dimensionless combinations of dimensionful parameters in this calculation, embodying discretization errors of the reduced pseudo-ITD that vanish in the continuum limit and higher-twist errors that survive the continuum limit. Since this calculation has been performed at only a single lattice spacing, the  $a^2 \Lambda_{\text{QCD}}^2$  effect cannot be quantified, leaving the nominal polynomial corrections to the short-distance factorization of the reduced pseudo-ITD (2.5) and the short-distance error arising from the bilocal operator  $\hat{\mathcal{O}}_{\text{WL}}^{[\gamma_4]}(z)$  as the only systematic effects these data are sensitive to. A reliable determination of the leading-twist PDF then depends on parameterization and removal of these effects. The precise form of the short-distance discretization effect can be deduced by deferring to the equalities

$$\mathfrak{M}(p, z) = \mathfrak{M}(p, -z)^* = \mathfrak{M}(-p, z)^* = \mathfrak{M}(-p, -z). \quad (4.25)$$

The leading discretization effect to the real component of  $\mathfrak{M}(\nu, z^2)$  must then scale as either  $\mathcal{O}(a^2/z^2)$  or  $\mathcal{O}(a/|z|)$ , while in the imaginary component the effect must arise at  $\mathcal{O}(a/|z|)$ . Rather than make a selection for the form of a potential discretization effect, we allow our data to dictate the precise form. This shall be discussed in the ensuing paragraphs.

As the Fourier transform in  $\nu$  of the reduced pseudo-ITD only has support on the momentum fraction interval  $x \in [-1, 1]$  [36], the contaminating effects we have discussed must also only have support in this interval. By construction, these corrections must be functions of  $\nu^2$  and  $\nu$  in the real and imaginary components of  $\mathfrak{M}(\nu, z^2)$ , respectively. The same basis of Jacobi polynomials in (4.23) and (4.24) can then be used to parameterize the Ioffe-time dependence of these contaminating effects, with  $x$ -space distributions of the same form as (4.19) but with distinct expansion coefficients. We denote the coefficients of the corrections as  $C_{\tau,n}^{\text{corr}(\alpha, \beta)}$ , with  $\tau = \{v, +\}$  indicating whether the effect arises in the valence/plus quark PDFs. Supposing, for simplicity, these effects enter at tree-level, their contributions to the reduced pseudo-ITD signals with  $\sigma_{0,n}^{(\alpha, \beta)} \equiv \sigma_n^{(\alpha, \beta)}(\nu, z^2 \mu^2)|_{\alpha_s=0}$  and





**Figure 13.** The eight lowest-order  $\sigma_{0,n}^{(\alpha,\beta)}$  (a) and  $\eta_{0,n}^{(\alpha,\beta)}$  (b) polynomials for an arbitrarily chosen basis  $\alpha = 0.125$  and  $\beta = 2.85$ . Each polynomial features an extremum in a range of Ioffe-time accessible in our lattice calculation, and asymptotically approaches zero for  $\nu \rightarrow \infty$ .

$\eta_{0,n}^{(\alpha,\beta)} \equiv \eta_n^{(\alpha,\beta)}(\nu, z^2 \mu^2) |_{\alpha_s=0}$  are then

$$\Re \mathcal{M}^{\text{corr}}(\nu, z^2) = \kappa_{\text{corr}} \sum_{n=1}^{\infty} \sigma_{0,n}^{(\alpha,\beta)} C_{v,n}^{\text{corr}(\alpha,\beta)} \quad (4.26)$$

$$\Im \mathcal{M}^{\text{corr}}(\nu, z^2) = \kappa_{\text{corr}} \sum_{n=0}^{\infty} \eta_{0,n}^{(\alpha,\beta)} C_{+,n}^{\text{corr}(\alpha,\beta)}, \quad (4.27)$$

where  $\kappa_{\text{corr}}$  is the dimensionless parameter which describes the scaling of each correction (e.g.  $\kappa_{\text{corr}} = z^2 \Lambda_{\text{QCD}}^2$ ). Visualizing the Taylor-expanded matching kernels (4.20) and (4.21) at tree-level across a range of Ioffe-times in figure 13a and figure 13b, it is seen the polynomials  $\sigma_{0,n}^{(\alpha,\beta)}, \eta_{0,n}^{(\alpha,\beta)}$  reach an extremum in Ioffe-time commensurate with the polynomial order and asymptotically approach zero. This conveniently reflects the correct large- $\nu$  behavior of the ITD in the same limit (see eq. (4.4)). Since  $\mathcal{M}(0, z^2) = 1$  by construction, all corrections must vanish at zero Ioffe-time. Of the Jacobi polynomials used to describe the corrections at tree-level, only  $\sigma_{0,0}^{(\alpha,\beta)}(0) \neq 0$  (blue curve of figure 13a). Hence the corrections to the real component of  $\mathcal{M}(\nu, z^2)$  shown in eq. (4.26) are restricted to order  $n \geq 1$ .

The complete functional forms we apply to each component of  $\mathcal{M}(\nu, z^2)$  are:

$$\begin{aligned} \Re \mathcal{M}_{\text{fit}}(\nu, z^2) &= \sum_{n=0}^{\infty} \sigma_n^{(\alpha,\beta)}(\nu, z^2 \mu^2) C_{v,n}^{lt(\alpha,\beta)} + \frac{a}{|z|} \sum_{n=1}^{\infty} \sigma_{0,n}^{(\alpha,\beta)}(\nu) C_{v,n}^{az(\alpha,\beta)} \\ &\quad + z^2 \Lambda_{\text{QCD}}^2 \sum_{n=1}^{\infty} \sigma_{0,n}^{(\alpha,\beta)}(\nu) C_{v,n}^{t4(\alpha,\beta)} + z^4 \Lambda_{\text{QCD}}^4 \sum_{n=1}^{\infty} \sigma_{0,n}^{(\alpha,\beta)}(\nu) C_{v,n}^{t6(\alpha,\beta)} \end{aligned} \quad (4.28)$$

$$\begin{aligned} \Im \mathcal{M}_{\text{fit}}(\nu, z^2) &= \sum_{n=0}^{\infty} \eta_n^{(\alpha,\beta)}(\nu, z^2 \mu^2) C_{+,n}^{lt(\alpha,\beta)} + \frac{a}{|z|} \sum_{n=0}^{\infty} \eta_{0,n}^{(\alpha,\beta)}(\nu) C_{+,n}^{az(\alpha,\beta)} \\ &\quad + z^2 \Lambda_{\text{QCD}}^2 \sum_{n=0}^{\infty} \eta_{0,n}^{(\alpha,\beta)}(\nu) C_{+,n}^{t4(\alpha,\beta)} + z^4 \Lambda_{\text{QCD}}^4 \sum_{n=0}^{\infty} \eta_{0,n}^{(\alpha,\beta)}(\nu) C_{+,n}^{t6(\alpha,\beta)}. \end{aligned} \quad (4.29)$$

The leading-twist ( $C_{\tau,n}^{lt}$ ) and discretization ( $C_{\tau,n}^{az}$ ) corrections are accompanied by twist-4 ( $C_{\tau,n}^{t4}$ ) and twist-6 ( $C_{\tau,n}^{t6}$ ) corrections, where we note each higher-twist correction is a hypercubic invariant. The twist-6 corrections have been included for exploration purposes despite the likelihood of their effect being minimal in the range of Ioffe-time for which we have statistically clean data ( $\nu \sim 10$ ). We point out the discretization effect in (4.28) we adopt herein scales as  $a/|z|$ ; the correlated figure-of-merit of this fit to  $\Re \mathfrak{M}(\nu, z^2)$  was found to be several times smaller when  $a/|z|$  was used in place of  $a^2/z^2$ . In this sense, the reduced pseudo-ITD data was found to favor a short-distance discretization effect of the form  $a/|z|$  in both real and imaginary components. Indeed other correlations of the lattice spacing with  $z$  (e.g.  $a^2/z^2$ ) may be present in our data, but are evidently sub-leading to  $a/|z|$  which we consider for the remainder of this work.

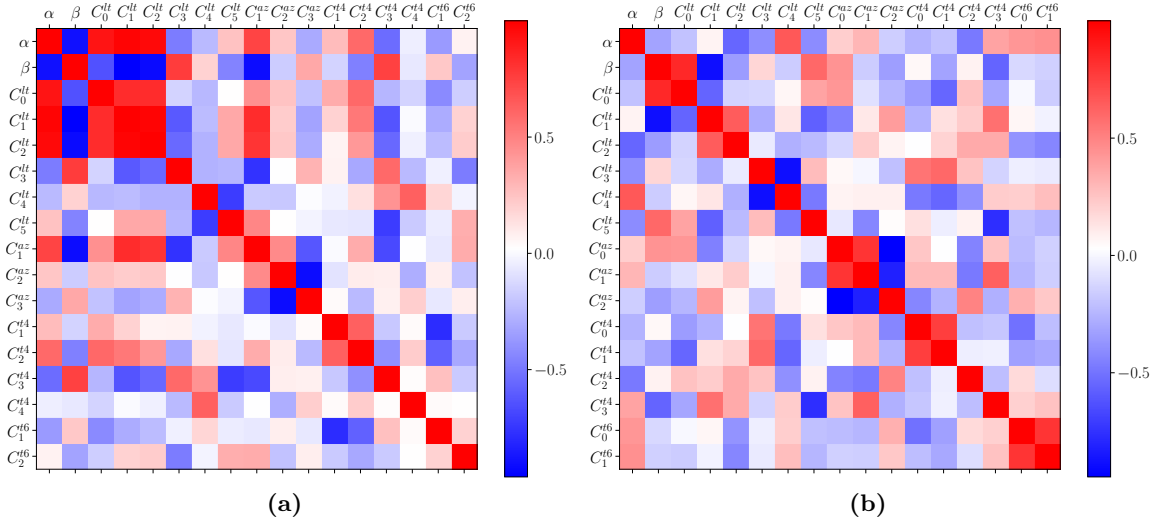
There are further sources of systematic error in  $\mathfrak{M}(\nu, z^2)$  that have been neglected by our parameterizations (4.28) and (4.29). Were the full NLO coordinate space matching kernel considered in the expansion of each correction in Jacobi polynomials (cf. eq. (4.26) & (4.27)), it becomes clear the discretization and higher-twist corrections we consider would enter at NLO with a more complicated dependence on  $z$ , namely an additional logarithmic dependence on  $z^2$ . These additional terms, however, each receive a further suppression by a factor of  $\mathcal{O}(\alpha_s/\pi) \sim 0.1$ , and are on the order of a few percent in the range of  $\nu$  for which we have computed  $\mathfrak{M}(\nu, z^2)$ . Only for considerably large- $z$  will the  $\mathcal{O}(\alpha_s)$  components of  $\sigma_n^{(\alpha,\beta)}$  and  $\eta_n^{(\alpha,\beta)}$ , that is  $\sigma_n^{(\alpha,\beta)} - \sigma_{0,n}^{(\alpha,\beta)}$  and  $\eta_n^{(\alpha,\beta)} - \eta_{0,n}^{(\alpha,\beta)}$ , become relevant. These additional sources of  $z^2$  contamination are avoided, and hence excluded from (4.28) and (4.29), by noting the short-distance factorization (2.5) would be invalidated by such large separations.

### 4.3.1 PDF results with Jacobi polynomials

In the fits we perform according to (4.28) and (4.29), we elect to fix the order of truncation for the leading-twist and each type of correction, and numerically search for the optimal  $\{\alpha, \beta, C_{\tau,n}^{\text{corr}}\}$ . Treating each fitted parameter as non-linear in a maximum likelihood fit leads to wildly unstable results. The way forward is to recognize  $\alpha, \beta$  are the only fitted parameters that are truly non-linear; the correction coefficients  $C_{\tau,n}^{\text{corr}}$  are all linear. A maximum likelihood fit of the posterior distribution of the linear terms is then Gaussian and cheap to obtain.

The reduced pseudo-ITD fits of eq. (4.28) and eq. (4.29) are implemented with the help of the Variable Projection (VarPro) algorithm [97] for separable non-linear optimization problems. This reduces the dimension of the non-linear optimization from  $d = N_{\text{corr}} + 2$  to  $d = 2$ , where  $N_{\text{corr}}$  are the number of linear correction coefficients. In our case, minimization is performed in the  $d = 2$  Jacobi polynomial basis  $\{\alpha, \beta\}$ , and any correction terms  $C_{\tau,n}^{\text{corr}(\alpha,\beta)}$  are solved for exactly in terms of the non-linear basis functions  $\{\sigma_n^{(\alpha,\beta)}, \sigma_{0,n}^{(\alpha,\beta)}, \eta_n^{(\alpha,\beta)}, \eta_{0,n}^{(\alpha,\beta)}\}$ . We note without VarPro, optimizations with  $N_{\text{corr}} \geq 4$  are numerically unstable, regardless of the type of correction included in the model.

Care needs to be taken as correction terms are included in eqs. (4.28) and (4.29), as physical insight can quickly be replaced with over fitting. The first sensible restriction to impose is for all  $x$ -space corrections  $\mathcal{O}(a/|z|)$ ,  $\mathcal{O}(z^2 \Lambda_{\text{QCD}}^2)$  and  $\mathcal{O}(z^4 \Lambda_{\text{QCD}}^4)$  to be sub-leading relative to the leading-twist PDF. It would be alarming to obtain, say, a twist-4 contribution



**Figure 14.** Parameter covariances in Jacobi polynomial fits with  $[n_{lt}, n_{az}, n_{t4}, n_{t6}] = [6342]$  to the real (a) and imaginary (b) components of the unpolarized reduced pseudo-ITD for  $z/a \leq 12$ . Entries are normalized according to  $\text{Cov}_{ij} / \sqrt{\text{Cov}_{ii} \text{Cov}_{jj}}$ .

that is larger than the leading-twist PDF, given that the short-distance factorization of the pseudo-distributions implies leading-twist dominance. Such disastrous scenarios are avoided with several Bayesian constraints of a Gaussian form. So as to allow the reduced pseudo-ITD to dictate the best fit results, all Bayesian priors are fixed to zero. The hierarchy we desire is realized with the following prior widths:

- Leading-twist:

$$\begin{aligned}
 \delta C_{\tau,0}^{lt(\alpha,\beta)} &= 1.1, & \delta C_{\tau,1}^{lt(\alpha,\beta)} &= 0.75, & \delta C_{\tau,2}^{lt(\alpha,\beta)} &= 0.5, \\
 \delta C_{\tau,3}^{lt(\alpha,\beta)} &= 0.25, & \delta C_{\tau,4}^{lt(\alpha,\beta)} &= 0.125, & \delta C_{\tau,5}^{lt(\alpha,\beta)} &= 0.1, \\
 \delta C_{\tau,6}^{lt(\alpha,\beta)} &= 0.05, & \delta C_{\tau,7}^{lt(\alpha,\beta)} &= 0.025
 \end{aligned}$$

- Corrections:

$$\delta C_{\tau,n \in \mathbb{Z}_3}^{\text{corr}(\alpha,\beta)} = 0.25, \quad \delta C_{\tau,n=3,4,5}^{\text{corr}(\alpha,\beta)} = 0.125, \quad \delta C_{\tau,n=6,7}^{\text{corr}(\alpha,\beta)} = 0.1.$$

The validity of the entire Jacobi polynomial parameterization is guaranteed using shifted log-normally distributed priors to ensure  $\alpha, \beta > -1$ . In practice, the log-normal prior on beta is shifted to  $\beta = 0$  to secure  $\beta > 0$  and hence convergent PDFs at  $x = 1$ .

As leading-twist and correction terms are added, the question becomes at which order each series of Jacobi polynomials should be truncated. We address this by scanning over all possible combinations of truncation orders for  $n_{lt} \in \{3, 4, 5, 6\}$ ,  $n_{az} \in \mathbb{Z}_4$ ,  $n_{t4} \in \mathbb{Z}_5$ ,  $n_{t6} \in \mathbb{Z}_3$ , where  $n_*$  are the orders of truncation in the fits (4.28) and (4.29). Figure 14 illustrates, for a rather large number of Jacobi polynomials  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\} = \{6, 3, 4, 2\}$ , the covariances of  $\alpha, \beta$  and each linear correction term  $C_{\tau,n}^{\text{corr}(\alpha,\beta)}$  in fits to the real (4.28)

and imaginary (4.29) reduced pseudo-ITD components for Wilson line lengths  $z/a \leq 12$ . The covariance of each pair of fitted parameters is estimated via jackknife resampling

$$\text{Cov}_{ij} \simeq \frac{N-1}{N} \sum_{n=1}^N \left( \mathbf{f}_{n,i} - \bar{\mathbf{f}}_i \right) \left( \mathbf{f}_{n,j} - \bar{\mathbf{f}}_j \right), \quad (4.30)$$

where fit parameters associated with jackknife sample  $n$  are denoted by  $\mathbf{f}_{n,k}$ , with jackknife average  $\bar{\mathbf{f}}_k$ . Without observing the quality of agreement between each fit and the reduced pseudo-ITD, it is clear several parameters correlate weakly or not at all with other parameters in the fit. This implies these weakly correlated parameters are not well-constrained by the data, and their removal will not affect the information content of the fit. For instance, the real component fit parameter covariances, shown in figure 14a, suggest the leading-twist expansion coefficients  $C_{n_{lt}}^{lt}$  are constrained by the data for  $n_{lt} \leq 3$ , while  $C_4^{lt}, C_5^{lt}$  weakly correlate with the remaining parameters. The discretization, twist-4 and twist-6 corrections exhibit mild correlation for  $C_1^{az}, C_2^{t4}, C_3^{t4}, C_1^{t6}$ , with the remaining correction parameters largely unconstrained. In the imaginary component, the fit parameter covariances shown in figure 14b suggest a more nuanced pattern of correlation. Several leading-twist Jacobi polynomials appear to be well-constrained by the data, while the relative correlation between the  $C^{az}$  and  $C^{t4}, C^{t6}$  parameters is increased relative to the corresponding entries in figure 14a.

This exercise demonstrates an important point. Although the VarPro implementation of the Jacobi polynomial fits allows for arbitrarily many leading-twist and correction coefficients, the reduced pseudo-ITD data simply do not contain enough information to constrain so many parameters. One should then expect the likelihood function is maximized for the real component of the reduced pseudo-ITD with truncation orders  $n_{lt} \sim 3$  and  $n_{az} \sim 1-2$ , and  $n_{lt} \sim 3$  and  $n_{az} \sim 3$  for the imaginary component.

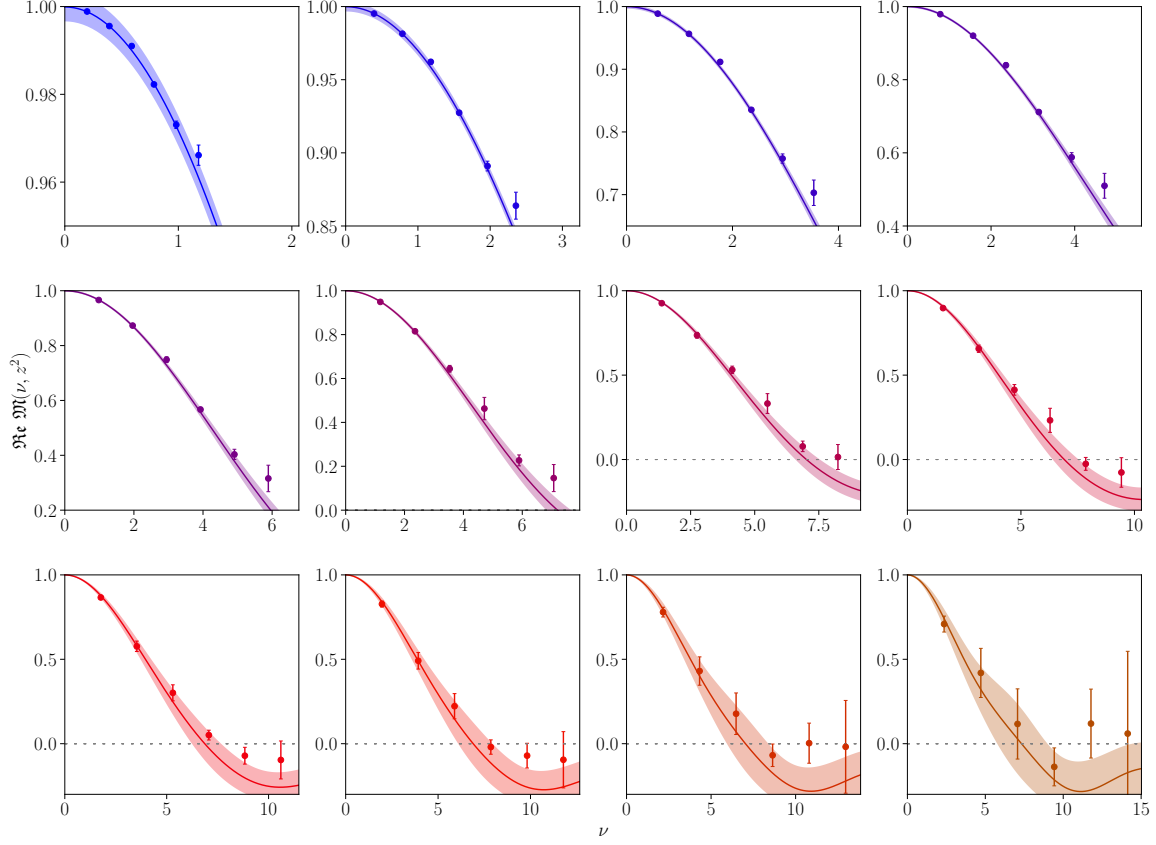
By scanning over the order of truncation for the leading-twist and correction terms parameterized by Jacobi polynomials, we find the likelihood of the functional (4.28) to describe  $\Re \mathcal{M}(\nu, z^2)$  with  $z/a \leq 12$  to be maximized for  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$ . Likewise, the likelihood of the functional (4.29) to describe  $\Im \mathcal{M}(\nu, z^2)$  with  $z/a \leq 12$  is maximized for  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$ . The fit results for each and their respective figures-of-merit are given in table 4. The Jacobi polynomial fits of orders  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  and  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  applied to the real and imaginary components of  $\mathcal{M}(\nu, z^2)$  are presented in figure 15 and figure 16, respectively.

Considering first the real component fit, each set of  $\Re \mathcal{M}(\nu, z^2)$  for  $z/a \leq 8$  are well represented by the expansion in Jacobi polynomials. The main exception is the highest momentum point  $p_z = 6 \times (2\pi/aL) \sim 2.47 \text{ GeV}$ . The  $\Re \mathcal{M}(\nu, z^2)$  data for  $z/a > 8$  are also reasonably well described, however the highest two momenta are seen to deviate. This behavior is not surprising despite the twist-4 and twist-6 corrections, which capture large- $z^2$  deviations, as the highest momentum data are subject to loss of signal in both the two- and three-point functions. The associated fit parameter covariances shown in figure 17a demonstrate the leading-twist, discretization and twist-4 corrections are well constrained by the  $\Re \mathcal{M}(\nu, z^2)$  data; as expected, the twist-6 corrections are only weakly constrained.

$\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_{v/+}$	$\{4, 1, 3, 2\}_v$	$\{4, 0, 3, 2\}_v$	$\{3, 3, 1, 0\}_+$	$\{3, 0, 1, 0\}_+$	$\{6, 3, 4, 2\}_v$	$\{6, 3, 4, 2\}_+$
$\alpha$	-0.209(147)	-0.376(37)	-0.328(20)	-0.331(31)	-0.264(117)	-0.326(20)
$\beta$	1.330(415)	2.032(496)	2.361(243)	3.227(297)	1.438(404)	2.051(260)
$C_{\tau,0}^{lt}$	1.606(257)	1.340(165)	2.041(108)	1.156(83)	1.489(213)	1.954(107)
$C_{\tau,1}^{lt}$	0.427(752)	0.335(261)	0.123(248)	0.161(243)	0.174(620)	0.404(213)
$C_{\tau,2}^{lt}$	-0.880(409)	-0.125(100)	-0.464(121)	0.700(98)	-1.002(301)	-0.241(118)
$C_{\tau,3}^{lt}$	-0.675(122)	-0.651(140)	—	—	-0.568(118)	-0.018(79)
$C_{\tau,4}^{lt}$	—	—	—	—	0.089(28)	0.020(27)
$C_{\tau,5}^{lt}$	—	—	—	—	0.020(12)	-0.023(10)
$C_{\tau,0}^{az}$	—	—	-0.001(43)	—	—	0.054(35)
$C_{\tau,1}^{az}$	-0.279(48)	—	-0.338(39)	—	-0.226(53)	-0.219(46)
$C_{\tau,2}^{az}$	—	—	0.434(74)	—	0.209(67)	0.283(67)
$C_{\tau,3}^{az}$	—	—	—	—	-0.164(48)	—
$C_{\tau,0}^{t4}$	—	—	0.170(28)	0.391(46)	—	0.185(47)
$C_{\tau,1}^{t4}$	0.052(53)	-0.090(52)	—	—	0.060(50)	0.032(68)
$C_{\tau,2}^{t4}$	-0.371(106)	-0.112(77)	—	—	-0.341(93)	-0.200(79)
$C_{\tau,3}^{t4}$	-0.407(122)	0.274(99)	—	—	-0.397(131)	0.076(29)
$C_{\tau,4}^{t4}$	—	—	—	—	0.088(30)	—
$C_{\tau,0}^{t6}$	—	—	—	—	—	-0.067(34)
$C_{\tau,1}^{t6}$	-0.045(37)	0.011(39)	—	—	-0.045(36)	-0.079(53)
$C_{\tau,2}^{t6}$	0.228(52)	0.397(84)	—	—	0.227(53)	—
$\chi_r^2$	2.620(345)	45.68(1.72)	2.845(387)	123.16(2.73)	2.809(374)	3.110(431)

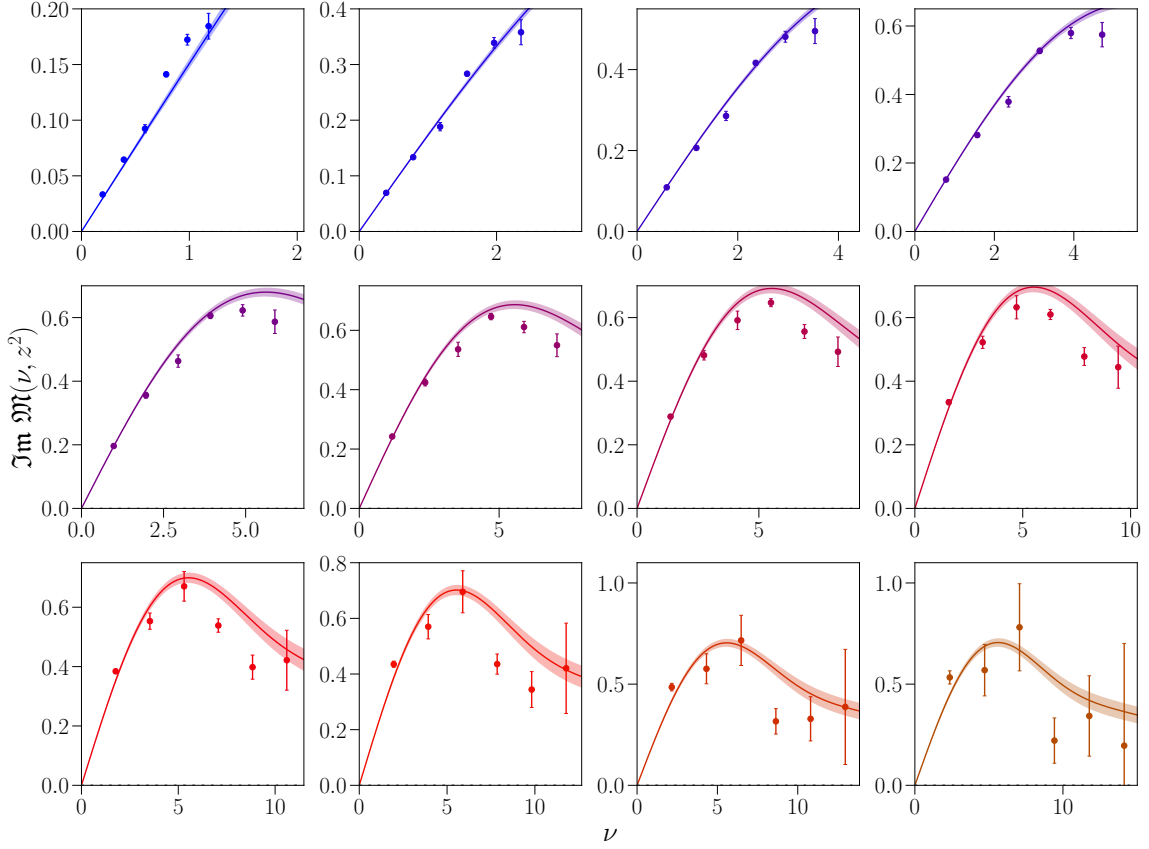
**Table 4.** Various Jacobi polynomial fits to the real and imaginary components of the unpolarized reduced pseudo-ITD for  $z/a \leq 12$ . Each column represents distinct orders of truncation in the Jacobi polynomial expansions to the leading-twist, discretization, twist-4 and twist-6 corrections. The real and imaginary component fits were found to have the highest likelihoods of describing the data with truncation orders  $\{4, 1, 3, 2\}_v$  and  $\{3, 3, 1, 0\}_+$ , respectively. The dramatic effect even a single discretization term has on each fit is shown in the columns  $\{4, 0, 3, 2\}_v$  and  $\{3, 0, 1, 0\}_+$ .

The resultant leading-twist PDF  $f_{q/N}(x)$  and  $x$ -space distributions corresponding to the  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  Jacobi polynomial fit are gathered in figure 18b. As expected, the corrections in  $x$ -space are sub-leading to the leading-twist PDF. The Jacobi-parameterized leading-twist PDF, however, features many structural differences with the included phenomenological PDFs and the uncorrelated two-parameter PDF fit. Most evident is the softer approach to  $x = 1$ . Due to the valence quark sum rule, this enhances the low- to moderate- $x$  region and leads to further tension with the phenomenological results. By evaluating the cosine transform of the pure leading-twist component, we see in figure 18a that the  $z/a \gtrsim 7$  ITD data deviate successively further from the derived leading-twist ITD shown in purple. Whereas the uncorrelated two-parameter ITD fit shown in red attempts to capture all the  $z/a \leq 12$  data and indeed the unwanted impact from higher-twist effects, the Jacobi polynomial parameterization has effectively isolated and removed these polynomial- $z^2$  effects, leaving the pure leading-twist contribution.



**Figure 15.** Fit to the real component of the unpolarized reduced pseudo-ITD where the leading-twist, discretization, twist-4, and twist-6 corrections have been expanded in Jacobi polynomials up to order  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$ . Starting from the upper left panel and traversing horizontally, the leading-twist plus corrections are shown for each  $z/a \leq 12$ .

The quality of the Jacobi polynomial fit to the imaginary component of the reduced pseudo-ITD shown in figure 16 is more puzzling. The  $z/a \leq 4$  appear reasonably well represented by the expansion in Jacobi polynomials, but by  $z/a = 5$  it is evident the data for a given  $z^2$  segregate into two distinct groups — one for lattice momenta  $p_{\text{latt}} \in \{1, 2, 3\}$  and another for  $p_{\text{latt}} \in \{4, 5, 6\}$ . This distinction coincides with the switch from an unphased eigenvector basis to the phased bases  $\vec{\zeta}_{\pm}$  defined in eq. (3.7). The fit parameter covariances shown in figure 17b demonstrate a milder constraint of the first and second order leading-twist Jacobi polynomials compared to the best fit of the real component. The discretization and twist-4 corrections are also seen to be well constrained by the data. The resultant leading-twist plus quark PDF  $f_{q+/N}(x)$  and  $x$ -space distributions corresponding to the  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  Jacobi polynomial fit are illustrated in figure 19b. As in the real component fit, the corrections are sub-leading to the leading-twist PDF, which in this case is in agreement with the NNPDF result [87] for  $x \geq 0.5$ . At small values of  $x$ , the leading-twist PDF parameterized by Jacobi polynomials is generally consistent with the two-parameter uncorrelated PDF fit. The sine transform of the pure leading-twist component is shown in figure 19a together with the  $\Im Q(\nu, \mu^2)$  data at 2 GeV.

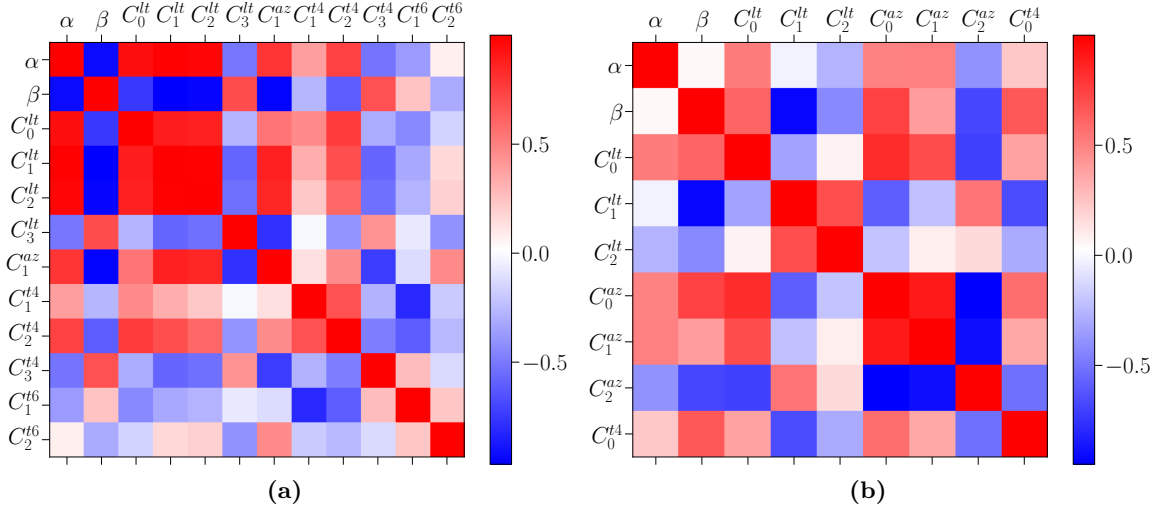


**Figure 16.** Fit to the imaginary component of the unpolarized reduced pseudo-ITD where the leading-twist, discretization, twist-4, and twist-6 corrections have been expanded in Jacobi polynomials up to order  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$ . Starting from the upper left panel and traversing horizontally, the leading-twist plus corrections are shown for each  $z/a \leq 12$ .

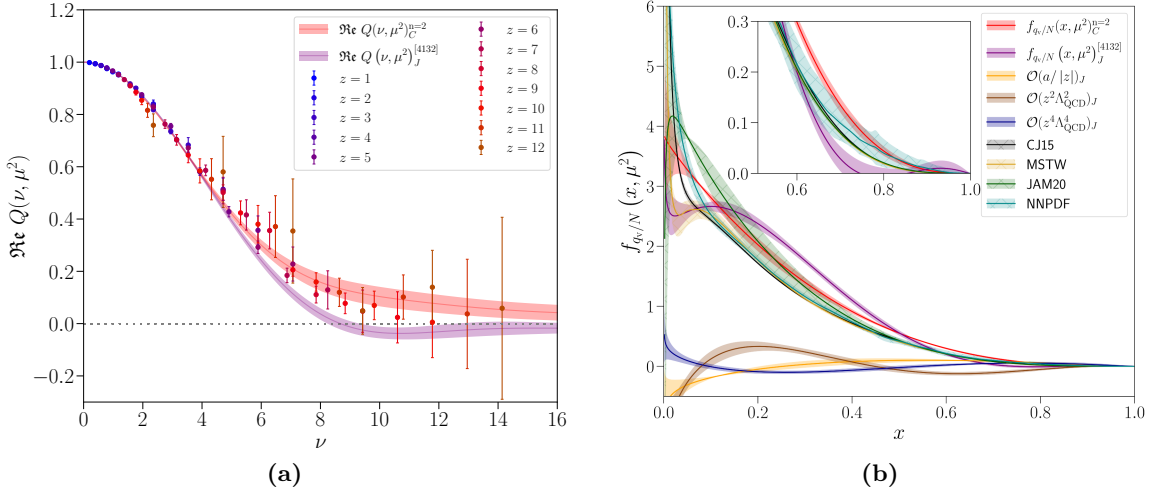
Unlike the real component of the derived leading-twist ITD in figure 18a, the derived imaginary component of the leading-twist ITD does not agree with the  $\text{Im } \mathcal{Q}(\nu, \mu^2)$  data for any of the  $ap_z \gtrsim 4\pi/L$  data with  $z/a \gtrsim 7$ . As the imaginary component of  $\mathfrak{M}(\nu, z^2)$  is optimally fit with three discretization corrections and only one higher-twist term, the  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  fit would suggest the imaginary component of the ITD is susceptible to less higher-twist effects in exchange for greater discretization effects. This is a tenuous conclusion, however, in light of the segregation of the  $\text{Im } \mathfrak{M}(\nu, z^2)$  data into two distinct clusters, a low- and high-momentum set, for large Wilson line lengths. A future study exploring the side effects of phased distillation is warranted.

By far the biggest indicator of a reasonable description of the reduced pseudo-ITD data via Jacobi polynomials is a discretization term. Repeating the above Jacobi polynomial fits but leaving out any discretization corrections, namely  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 0, 3, 2\}_v$  and  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 0, 1, 0\}_+$ , the correlated figures of merit increase considerably to unacceptable values (see table 4). This same conclusion is reached when cuts on momentum and Wilson line lengths are made. Since the discretization term we have included is of  $\mathcal{O}(a/|z|)$ , its effect is most pronounced at short distances. This is precisely the regime

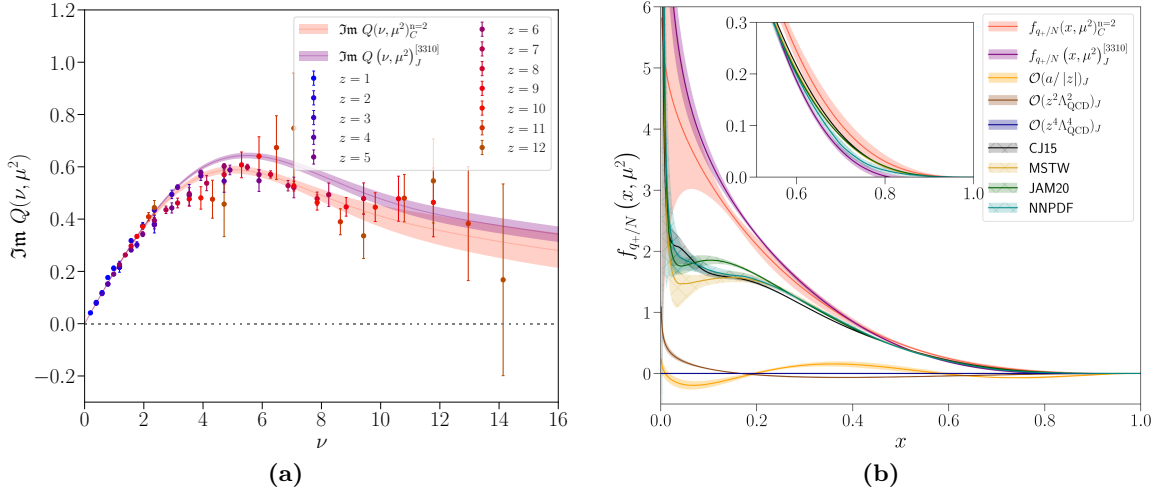




**Figure 17.** Parameter covariances of the optimal Jacobi polynomial fit to the real (a) and the imaginary (b) component of the unpolarized reduced pseudo-ITD for  $z/a \leq 12$  with truncation orders  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  and  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$ , respectively. Entries are normalized according to  $\text{Cov}_{ij}/\sqrt{\text{Cov}_{ii}\text{Cov}_{jj}}$ .



**Figure 18.** (a) The real component of the leading-twist ITD (purple) at 2 GeV derived from the Jacobi polynomial expansion of the reduced pseudo-ITD for  $z/a \leq 12$  with  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$ . The result is compared with the uncorrelated 2-parameter phenomenological form of eq. (4.7) shown in red. (b) The valence quark leading-twist PDF (purple) obtained from the  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  Jacobi polynomial expansion of the reduced pseudo-ITD. The  $a/|z|$  (orange), twist-4 (brown), and twist-6 (navy)  $x$ -space distributions are also shown and seen to be sub-leading. The distributions are compared with the uncorrelated 2-parameter phenomenological fit of eq. (4.7) (red), as well as the NLO global analyses of CJ15 [82] and JAM20 [88], and the NNLO analyses of MSTW [89] and NNPDF [87] at the same scale.



**Figure 19.** (a) The imaginary component of the leading-twist ITD (purple) at 2 GeV derived from the Jacobi polynomial expansion of the reduced pseudo-ITD for  $z/a \leq 12$  with  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$ . The result is compared with the uncorrelated 2-parameter phenomenological form of eq. (4.8) shown in red. (b) The plus quark leading-twist PDF (purple) obtained from the  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  Jacobi polynomial expansion of the reduced pseudo-ITD. The  $a/|z|$  (orange) and twist-4 (brown)  $x$ -space distributions are also shown and seen to be sub-leading. The distributions are compared with the uncorrelated 2-parameter phenomenological fit of eq. (4.8) (red), as well as the NLO global analyses of CJ15 [82] and JAM20 [88], and the NNLO analyses of MSTW [89] and NNPDF [87] at the same scale.

wherein the short distance factorization (4.1), or equivalently (4.12), is applicable. This motivates a more detailed look at the short-distance behavior of the computed reduced pseudo-ITD.

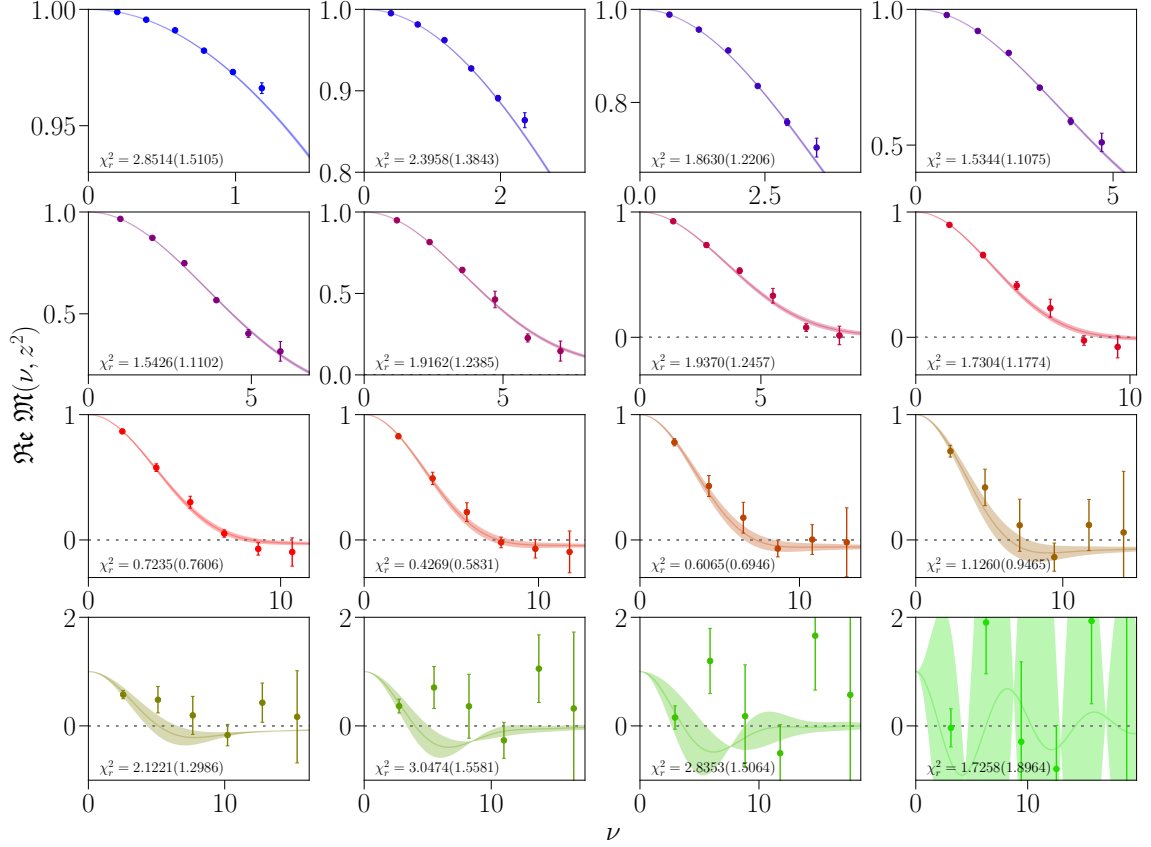
## 5 On the numerical consistency with DGLAP

The one-loop matching relationship between the ITD and the reduced pseudo-ITD (4.1) implies that  $\mathcal{Q}(\nu, \mu^2) = \mathfrak{M}(\nu, z^2)$  at tree-level. The scatter that exists for a given  $z^2$  should ideally be compensated at  $\mathcal{O}(\alpha_s)$  by the  $\ln z^2$ -dependence produced by the DGLAP evolution, up to large- $z^2$  higher-twist corrections. In this section we study the  $z^2$ -dependence of  $\mathfrak{M}(\nu, z^2)$  more closely, and investigate whether the observed dependence is numerically consistent with DGLAP, thus yielding a truly  $z^2$ -independent ITD.

We begin by focusing on the real component of the reduced pseudo-ITD. The dependence of  $\Re \mathfrak{M}(\nu, z^2)$  on the invariant space-like interval  $z^2$  can be most easily visualized by parameterizing the valence pseudo-PDF  $\mathcal{P}_v(x, z^2; \alpha, \beta)$  by a simple two-parameter phenomenological form

$$\mathcal{P}_v(x, z^2; \alpha, \beta) = \frac{\Gamma(2 + \alpha + \beta)}{\Gamma(1 + \alpha)\Gamma(1 + \beta)} x^\alpha (1 - x)^\beta, \quad (5.1)$$

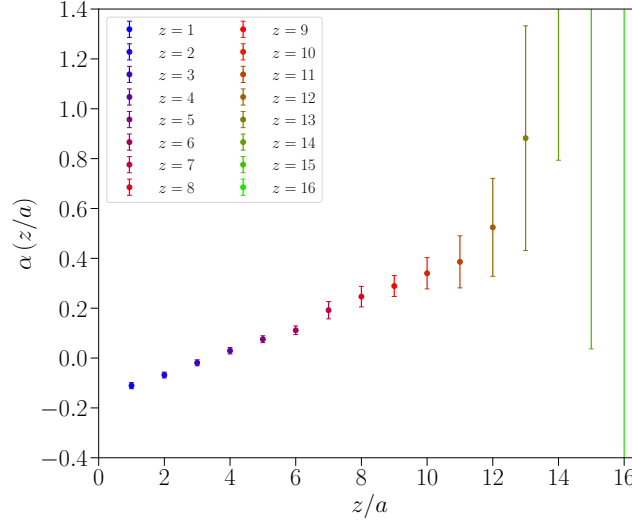
and fitting its cosine-transform to  $\Re \mathfrak{M}(\nu, z^2)$  separately for each  $z^2$ . In order to more readily expose the role of the Altarelli-Parisi kernel, we impose the added restriction  $\beta = 3$ .



**Figure 20.** Cosine transform of the model pseudo-PDF in eq. (5.1) fit separately to  $\Re \mathfrak{M}(\nu, z^2)$  for distinct  $z^2$ ; data correlations have been included in each fit. Starting from the upper left panel and traversing horizontally,  $z/a$  increases from unity. The correlated figure of merit for each separate fit is also indicated.

This choice not only captures the naive  $x \rightarrow 1$  behavior of the nucleon’s valence quark PDF [98], but also forces  $\alpha$  to reflect any  $z$ -dependence in the reduced distribution; further, this value of  $\beta$  is in statistical agreement with those obtained from the uncorrelated ITD fits (see table 2).

Figure 20 illustrates the cosine-transform of the model valence pseudo-PDF (5.1) fit separately to each  $z^2$  of the real component of the reduced pseudo-ITD. The cosine-transforms of  $\mathcal{P}_v(x, z^2; \alpha, \beta = 3)$  are seen to describe  $\Re \mathfrak{M}(\nu, z^2)$  quite well for  $z/a \lesssim 10$ , with the greatest tension seen for the highest momentum point for each separation. The fits for  $z/a \geq 13$  are also shown for completeness, but are clearly noise dominated. Also noteworthy, the highest figures-of-merit are observed for the smallest separations, with a somewhat monotonic reduction until  $z/a \simeq 11$ . The dependence of the fitted values of  $\alpha$  on the separation  $z/a$  is visualized for  $\Re \mathfrak{M}(\nu, z^2)$  in figure 21. As a function of  $z/a$ ,  $\alpha$  decreases with the Wilson line length, matching expectations from the Altarelli-Parisi evolution of the pseudo-PDF. However, it is clear  $\Re \mathfrak{M}(\nu, z^2)$  depends linearly on  $z/a$  for  $z/a \lesssim 12$ , most notably for small- $z$ .



**Figure 21.** The fitted value of  $\alpha$  as a function of  $z/a$  resulting from the cosine-transform of the model pseudo-PDF in eq. (5.1) fit to  $\Re \mathfrak{M}(\nu, z^2)$ . The decrease of  $\alpha$  with  $z/a$  is in agreement with expectations from the Altarelli-Parisi evolution of the pseudo-PDF. This dependence is however clearly linear.

This manifest lack of  $\ln z^2$  behavior of  $\Re \mathfrak{M}(\nu, z^2)$  at short distances immediately suggests tension with the presumed DGLAP evolution of the pseudo-PDF. To determine if this  $z^2$ -dependence in  $\Re \mathfrak{M}(\nu, z^2)$  is nevertheless numerically consistent with DGLAP, the one-loop matching relationship between the reduced pseudo-ITD and ITD is applied. In the ideal scenario where the  $z^2$ -dependence of  $\Re \mathfrak{M}(\nu, z^2)$  is exactly described by DGLAP, the matched ITD will be independent of the interval  $z^2$  up to polynomial corrections for large- $z^2$ . Rather than perform the matching step to a common scale in  $\overline{\text{MS}}$  using a smooth polynomial in Ioffe-time (e.g. eq. (4.5)) as was done in section 4, we leverage the cosine-transform of the model pseudo-PDF (5.1) as the smooth and continuous description of the reduced pseudo-ITD data. That is, we perform the matching of  $\Re \mathfrak{M}(\nu, z^2)$  to a common scale in  $\overline{\text{MS}}$  according to

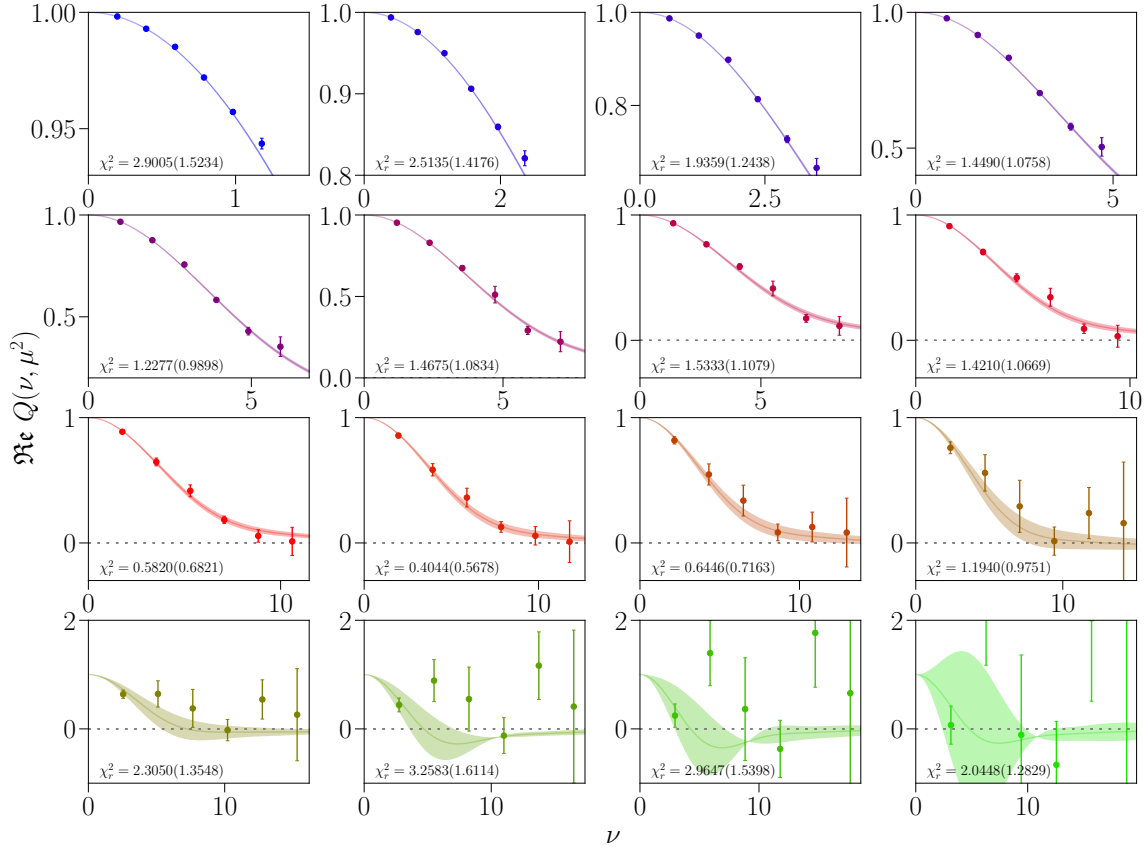
$$\Re \mathfrak{Q}(\nu, \mu^2) = \Re \mathfrak{M}(\nu, z^2) + \frac{\alpha_s C_F}{2\pi} \int_0^1 du \, \mathfrak{P}(u\nu, z^2; \alpha, \beta=3) \left[ \ln \left( \frac{z^2 \mu^2 e^{2\gamma_E+1}}{4} \right) B(u) + L(u) \right], \quad (5.2)$$

where  $\mathfrak{P}(u\nu, z^2; \alpha, \beta=3)$  is the cosine-transform of the model pseudo-PDF  $\mathcal{P}_v(x, z^2; \alpha, \beta=3)$  expressed in a closed form by a generalized hypergeometric function

$$\mathfrak{P}(\nu, z^2; \alpha, \beta=3) = {}_2F_3 \left( \frac{1+\alpha}{2}, \frac{2+\alpha}{2}; \frac{1}{2}, \frac{5+\alpha}{2}, \frac{6+\alpha}{2}; -\frac{\nu^2}{4} \right). \quad (5.3)$$

For an explicit, albeit crude, conversion to  $\overline{\text{MS}}$ , we set  $\alpha = 0.2$  in eq. (5.3).

Our strategy to expose any  $z$ -dependence in the ITD  $\mathfrak{Q}(\nu, \mu^2)$  remains identical to the reduced distribution above. The resultant matched ITD at 2 GeV in  $\overline{\text{MS}}$  is once more fit using the two-parameter form in eq. (5.1) independently for each  $z^2$  and with  $\beta = 3$ . The parameterized distribution in this case is of course no longer the valence pseudo-PDF,

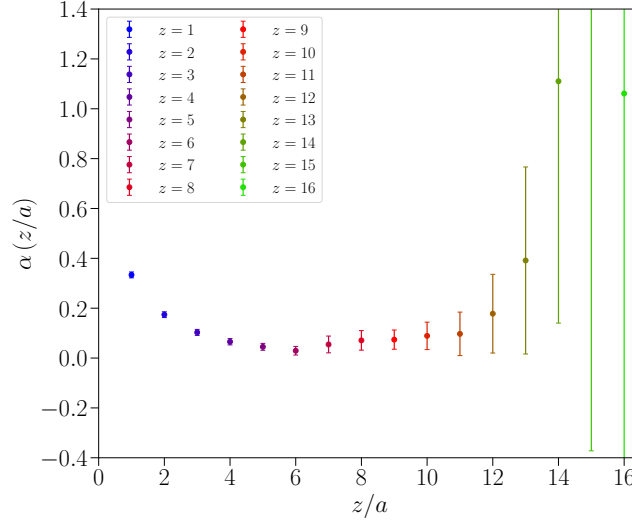


**Figure 22.** Cosine transform of the two-parameter model PDF, with the same functional form as (5.1), fit separately to each  $z^2$  of the matched ITD. The ITD was obtained using (5.3) for the evolution/matching step. Data correlations have been included in each fit. Starting from the upper left panel and traversing horizontally,  $z/a$  increases from unity. The correlated figure of merit for each separate fit is also indicated.

but rather the valence PDF itself. As illustrated in figure 22, each  $z^2$  of the matched ITD is well described by the simple two-parameter form. The poorest figures-of-merit are again observed for the smallest ( $z/a \lesssim 3$ ) and largest ( $z/a \gtrsim 13$ ) separations. The dependence of the fitted values of  $\alpha$  on the separation  $z/a$  for  $\Re Q(\nu, \mu^2)$  is illustrated in figure 23. For  $4 \lesssim z/a \lesssim 11$  the fitted value of  $\alpha$  is observed to be independent of  $z/a$  and hence numerically consistent with DGLAP in said interval. Remarkably, however, the values of  $\alpha$  for the shortest separations, namely  $z/a \lesssim 4$ , deviate increasingly from this constancy as  $z/a \rightarrow 1$ . A subsequent analysis of the imaginary component of both the reduced pseudo-ITD and matched ITD arrived at a similar conclusion, but has been omitted for brevity.

### 5.1 Jacobi polynomial corrections — discretization effects

The findings above rigorously demonstrate the reduced pseudo-ITD is numerically inconsistent with DGLAP in the small- $z$  regime. Whether matching the reduced pseudo-ITD to the light-cone ITD or directly to the light-cone PDF, the presence of the Altarelli-Parisi



**Figure 23.** The fitted values of  $\alpha$  from the cosine-transform of the two-parameter PDF functional form (5.1) fit to each  $z^2$  of the matched ITD. The latter was obtained using (5.3) for the evolution/matching step. The values of  $\alpha$  are statistically constant for  $4 \lesssim z/a \lesssim 11$ , with sharp deviations for small- $z/a$ .

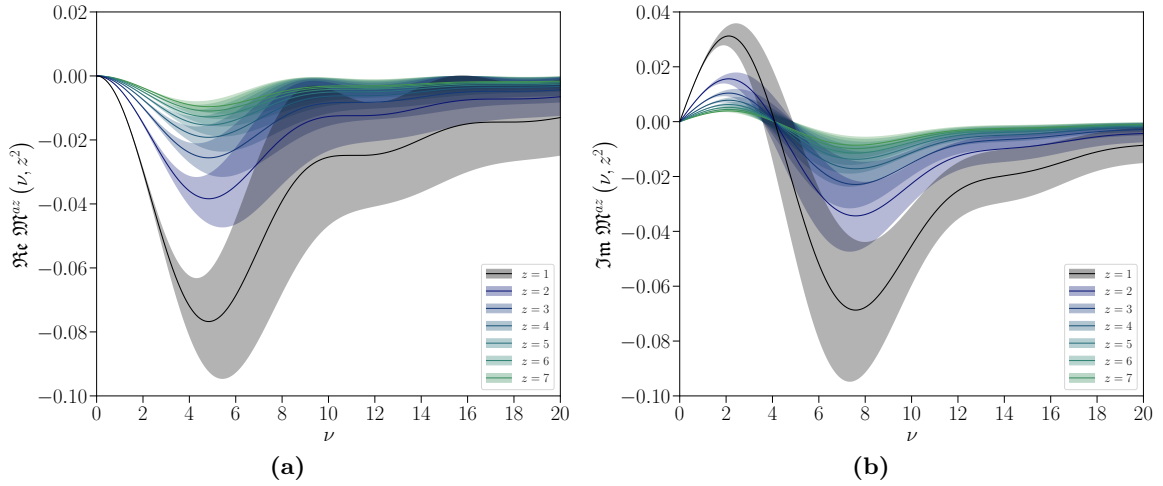
evolution kernel should in principle capture and remove the  $\ln z^2$  scatter that theoretically exists in  $\mathfrak{M}(\nu, z^2)$  for small- $z$ . As  $\mathfrak{M}(\nu, z^2)$  was found to depend only linearly on the separation  $z$  (figure 21), the Altarelli-Parisi kernel effectively introduces a  $\ln z^2$ -dependence into the small- $z$  ITD and thus explains the dependence of  $\alpha$  on  $z/a$  in figure 23. Despite this concerning conclusion, a broad subset of  $\mathfrak{M}(\nu, z^2)$  remains consistent with DGLAP; the statistically constant value of  $\alpha(z/a)$  observed in the ITD fits in the interval  $4 \lesssim z/a \lesssim 11$  (e.g. figure 23) validates the nice collapse of the  $\mathfrak{M}(\nu, z^2)$  data onto a common curve (figure 7a) when matched to a common scale in  $\overline{\text{MS}}$ .

To gain further insight into the regions wherein DGLAP is not respected, we return to the optimal Jacobi polynomial fits  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  and  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$ . The reader is reminded section 4 concluded with the realization that a suitable description of  $\mathfrak{M}(\nu, z^2)$  was only possible with the nominal inclusion of an  $\mathcal{O}(a/|z|)$  correction in eq. (4.28) and eq. (4.29). The discretization effect parameterized by each of these fits is given by

$$\mathfrak{M}^{az}(\nu, z^2) = \frac{a}{|z|} \times \begin{cases} C_{v,1}^{az(\alpha,\beta)} \sigma_{0,1}^{(\alpha,\beta)}(\nu) & \text{for } \{4, 1, 3, 2\}_v \\ \sum_{n=0}^2 C_{+,n}^{az(\alpha,\beta)} \eta_{0,n}^{(\alpha,\beta)}(\nu) & \text{for } \{3, 3, 1, 0\}_+ \end{cases}, \quad (5.4)$$

and visualized in figure 24. The discretization effect  $\Re \mathfrak{M}^{az}(\nu, z^2)$  is seen to be strictly negative in the interval of Ioffe-time in which  $\mathfrak{M}(\nu, z^2)$  has been computed. By comparison, the discretization effect  $\Im \mathfrak{M}^{az}(\nu, z^2)$  involves three Jacobi polynomials and suggests the  $\Im \mathfrak{M}(\nu, z^2)$  data are subject to a discretization effect that is opposite in sign at small and large values of Ioffe-time.

We now justify the necessity of the  $\mathcal{O}(a/|z|)$  discretization correction by considering the removal of the  $\Re \mathfrak{M}^{az}(\nu, z^2)$  effect from the computed  $\Re \mathfrak{M}(\nu, z^2)$  data, which we

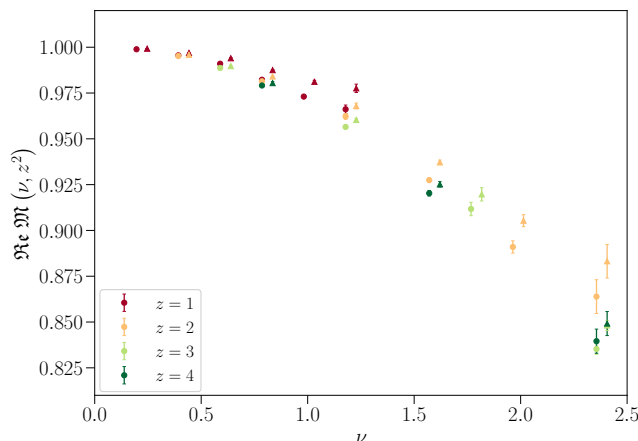


**Figure 24.** Visualization of the discretization effects determined by the optimal Jacobi polynomial fits  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  (a) and  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  (b) for  $z/a \leq 7$ .

denote  $\Re \mathcal{M}'(\nu, z^2) \equiv \Re \mathcal{M}(\nu, z^2) - \Re \mathcal{M}^{az}(\nu, z^2)$ . Based on figure 24a, the removal should shift the small- $z$  points of  $\Re \mathcal{M}(\nu, z^2)$  to larger values, with the largest impact for  $\nu \sim 4.5$ . Figure 25 juxtaposes the original  $\Re \mathcal{M}(\nu, z^2)$  and discretization corrected  $\Re \mathcal{M}'(\nu, z^2)$  in the interval  $\nu \in [0, 2.5]$ . Although the differences are numerically small, at small Ioffe-times  $\Re \mathcal{M}'(\nu, z^2)$  is noticeably larger than the uncorrected reduced pseudo-ITD. The importance of removing this discretization effect is quantitatively discerned by repeating the DGLAP investigation for  $\Re \mathcal{M}'(\nu, z^2)$ .

Parameterizing the discretization corrected valence pseudo-PDF  $\mathcal{P}'_v(x, z^2)$  with the two-parameter form in eq. (5.1) and fitting its cosine-transform to  $\Re \mathcal{M}'(\nu, z^2)$  with  $\beta = 3$ , the  $z$ -dependence of  $\Re \mathcal{M}'(\nu, z^2)$  is once more reflected in the variation of  $\alpha$  with  $z/a$ . As illustrated in figure 26a,  $\alpha$  now varies non-linearly with  $z/a$  for  $z/a \lesssim 4$  and linearly for  $4 \lesssim z/a \lesssim 11$ . Whether this markedly distinct  $z$ -dependence (cf. figure 21) is numerically consistent with DGLAP is once again checked by performing the matching to a common scale in  $\overline{\text{MS}}$  using eq. (5.3), and repeating the two-parameter fits to the discretization corrected ITD  $\Re \mathcal{Q}'(\nu, \mu^2)$  for each  $z^2$  and with  $\beta = 3$ . The resulting fitted values of  $\alpha$  are presented in figure 26b. Relative to the  $z$ -dependence of the uncorrected ITD shown in figure 23, the variation of  $\alpha$  with  $z/a$  is considerably more constant for  $z/a \lesssim 11$ . In other words, the ITD is seen to fall into better agreement with DGLAP in the short-distance regime following removal of the  $\mathcal{O}(a/|z|)$  effect. That the optimal Jacobi polynomial fits  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  and  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  provide the best description of  $\mathcal{M}(\nu, z^2)$  can now be quantitatively explained by the compensating effect the  $\mathcal{O}(a/|z|)$  term provides. The poor quality of the correlated phenomenological fits to the matched ITD, as well as the correlated Jacobi polynomial fits to  $\mathcal{M}(\nu, z^2)$  without any corrections, are a direct result of attempting to fit a singular function in  $z$  to data that do not exhibit singular behavior. By excluding  $z/a \lesssim 4$  and  $z/a \gtrsim 11$ , the short-distance tension and any large- $z$  polynomial effects can be removed yielding reduced pseudo-ITD





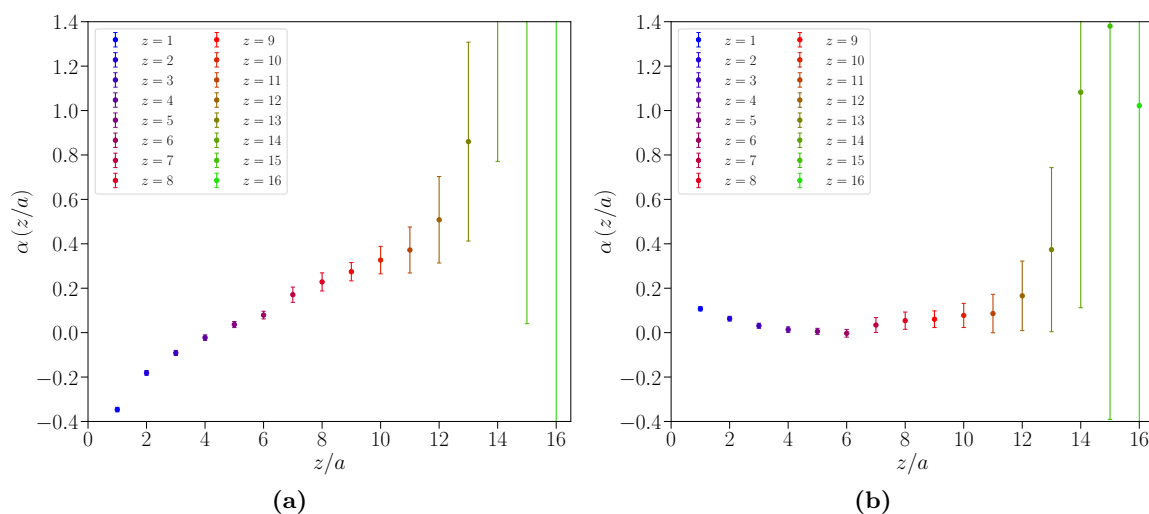
**Figure 25.** Juxtaposition of the raw  $\Re \mathcal{M}(\nu, z^2)$  and discretization corrected  $\Re \mathcal{M}'(\nu, z^2)$  distributions, represented by circles and wedges respectively. The  $\Re \mathcal{M}'(\nu, z^2)$  data is shifted horizontally for legibility.

or matched ITD data that are well in line with theoretical expectations. Such cuts are common in the literature, however their nominal effect is to neglect deviating behavior.

Although the DGLAP investigation has been shown for the real component of the reduced pseudo-ITD, the considerable reduction in the correlated figure-of-merit when discretization effects are included in fits to  $\Im \mathcal{M}(\nu, z^2)$  (e.g.  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 3, 1, 0\}_+$  versus  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_+ = \{3, 0, 1, 0\}_+$  in table 4) indicates the imaginary component of the raw reduced pseudo-ITD likewise deviates from expectations of DGLAP at short-distances. The central question left for future research is the origin of this discretization effect.

## 6 Conclusions

In this work we presented in detail the first lattice calculation of the unpolarized nucleon PDF in the distillation framework, showing its efficacy and emphasizing in detail the advantages of this smearing technique. We have observed that distillation can yield a pseudo-distribution of superior quality, addressing one of the principal components limiting the competitiveness of contemporary lattice QCD calculations of PDFs with respect to precision phenomenological extractions. For the unpolarized PDF, the impact of this lattice calculation on global fits was found to be rather marginal. The quality of data, however, hints that a future calculation of a distribution less constrained by experiment may benefit from the use of distillation. By validating the distillation method in the unpolarized nucleon PDF, our study opens new avenues of synergy between lattice and phenomenology in the spirit of [11, 99, 100]. We performed a careful study of the correct extraction of the matrix elements that yield the pseudo-ITD and we analyzed in detail its real and imaginary components. This was followed by a discussion of correlations of lattice data and how the commonly used uncorrelated fits can yield erroneous results and can severely flaw the extraction of the PDFs. We regulate the ill-posed inverse problem that relates the ITD to the PDF with the use of Jacobi polynomials, and stress the necessity of removing discretization



**Figure 26.** (a) The variation of  $\alpha$  with  $z/a$  resulting from the cosine-transform of the model pseudo-PDF in eq. (5.1) fit to  $\Re \mathcal{M}'(\nu, z^2)$  for each  $z^2$ . The discretization effect captured by the optimal Jacobi polynomial expansion  $\{n_{lt}, n_{az}, n_{t4}, n_{t6}\}_v = \{4, 1, 3, 2\}_v$  is subtracted from  $\Re \mathcal{M}(\nu, z^2)$  prior to performing each fit. (b) The variation of  $\alpha$  with  $z/a$  resulting from the cosine-transform of the two-parameter PDF form in eq. (5.1) fit to the discretization corrected and matched ITD  $\Re \mathcal{Q}'(\nu, z^2)$  for each  $z^2$ . The discretization corrected ITD is considerably more independent of the interval  $z^2$ .

errors as well as higher twist effects in order to ensure a trustworthy extraction of the PDF. The quality of the results herein bolsters confidence that distillation is the natural path to be pursued in these types of hadron structure calculations. We soon plan to extend these analyses to include the extrapolation to the physical pion mass and the continuum limit.

## Acknowledgments

We would like to thank all the members of the HadStruc collaboration for fruitful and stimulating exchanges. This work is supported by Jefferson Science Associates, LLC under U.S. DOE Contract #DE-AC05-06OR23177. KO was supported in part by U.S. DOE grant #DE-FG02-04ER41302 and in part by the Center for Nuclear Femtography grants #C2-2020-FEMT-006, #C2019-FEMT-002-05. AR was supported in part by U.S. DOE Grant #DE-FG02-97ER41028. CE was supported in part by the U.S. Department of Energy under Contract No. DEFG02-04ER41302, a Department of Energy Office of Science Graduate Student Research fellowship, through the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program, and a Jefferson Science Associates graduate fellowship. The SCGSR program is administered by the Oak Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by ORAU under Contract No. DE-SC0014664. We would also like to thank the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing HPC resources on

Frontera [101] that have contributed to the results in this paper. We acknowledge the facilities of the USQCD Collaboration used for this research in part, which are funded by the Office of Science of the U.S. Department of Energy. This work was performed in part using computing facilities at the College of William and Mary which were provided by contributions from the National Science Foundation (MRI grant PHY-1626177), and the Commonwealth of Virginia Equipment Trust Fund. The authors acknowledge William & Mary Research Computing for providing computational resources and/or technical support that have contributed to the results reported within this paper. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 [102]. In addition, this work used resources at NERSC, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract #DE-AC02-05CH11231, as well as resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. #DE-AC05-00OR22725. In addition, this work was made possible using results obtained at NERSC, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract #DE-AC02-05CH11231, as well as resources of the Oak Ridge Leadership Computing Facility (ALCC and INCITE) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. #DE-AC05-00OR22725. The software libraries used on these machines were Chroma [103], QUDA [104, 105], QDP-JIT [106] and QPhiX [107, 108] developed with support from the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research and Office of Nuclear Physics, Scientific Discovery through Advanced Computing (SciDAC) program, and of the U.S. Department of Energy Exascale Computing Project.

We acknowledge PRACE (Partnership for Advanced Computing in Europe) for awarding us access to the high performance computing system Marconi100 at CINECA (Consorzio Interuniversitario per il Calcolo Automatico dell'Italia Nord-orientale) under the grant Pra21-5389. Results were obtained also by using Piz Daint at Centro Svizzero di Calcolo Scientifico (CSCS), via the project with id s994. We thank the staff of CSCS for access to the computational resources and for their constant support. This work also benefited from access to the Jean Zay supercomputer at the Institute for Development and Resources in Intensive Scientific Computing (IDRIS) in Orsay, France under project A0080511504.

## A Comparison of distillation with conventional smearing

The results of this work clearly demonstrate the utility of distillation in the calculation of the collinear structure of hadrons, and how dramatically improved data quality facilitated by distillation has led to invaluable insight into the pseudo-distribution formalism. In this appendix, we touch on the numerical cost of distillation compared to conventional spatial smearing schemes, making direct comparisons to an earlier calculation by some of the authors herein that made use of Jacobi smearing.

The inversion cost present in lattice calculations that implement a standard spatial smearing kernel must solve for the Green's function represented by

$$S(\vec{x}, \vec{z})_{ba}^{\beta\alpha} = \sum_{\vec{y}, \gamma, c} D^{-1}(\vec{x}, \vec{y})_{bc}^{\beta\gamma} \delta(\vec{y} - \vec{z}) \delta_{\gamma\alpha} \delta_{ca}, \quad (\text{A.1})$$

where  $D$  is the chosen discretization of the Dirac operator, and the  $N_c$  color and  $N_s$  spinor components are denoted by Latin and Greek indices. Such point-to-all propagators stand in contrast to the so-called solution vectors required of a distillation based calculation:

$$S^{(k)}(\vec{x}, t'; t)_c^{\beta\alpha} = \sum_{\vec{y}, b} D^{-1}(\vec{x}, t'; \vec{y}, t)_{cb}^{\beta\alpha} \xi^{(k)}(\vec{y}, t)_b, \quad (\text{A.2})$$

where  $\xi^{(k)}$  is an eigenvector of the gauge covariant Laplacian with  $k^{\text{th}}$  largest eigenvalue. The separation of the elemental (interpolator) construction from the solution vector computation with distillation implies the total inversion cost per configuration needed to realize matrix elements sensitive to the Ioffe-time pseudo-distribution is given by

$$N_{\text{inv}}^{\text{dist}}/\text{cfg} = (1 + N_{\text{tsep}}) N_{\text{eigs}} N_{\zeta} N_{\text{src}} N_c, \quad (\text{A.3})$$

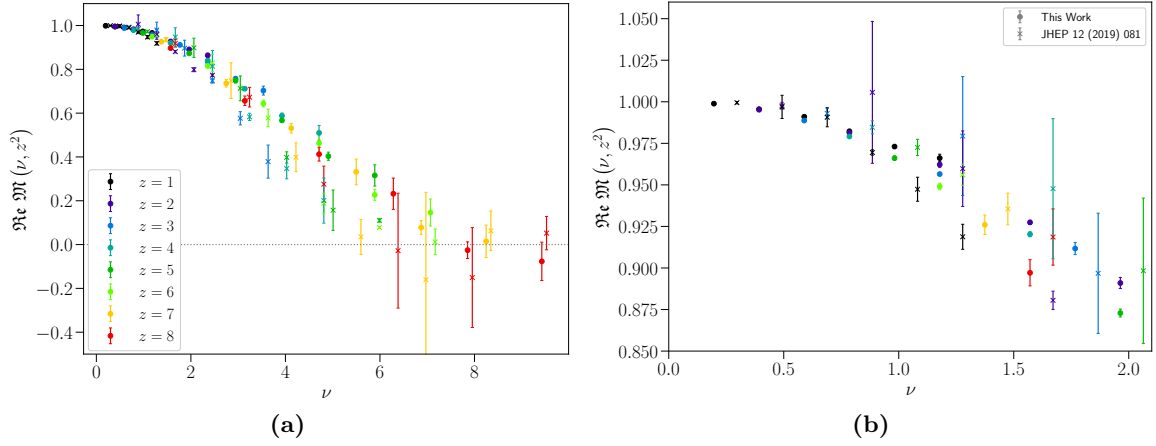
where  $N_{\text{eigs}}$  is the rank of the distillation space,  $N_{\text{tsep}}$  the number of three-point source-sink separations,  $N_{\zeta}$  the number of phased eigenvector bases (including unphased), and  $N_{\text{src}}$  the number of temporal source origins. In this calculation we have selected  $N_{\text{tsep}} = 6$ ,  $N_{\text{eigs}} = 64$ ,  $N_{\zeta} = 3$ , and  $N_{\text{src}} = 4$ , for which  $N_{\text{inv}}^{\text{dist}}/\text{cfg} = 16128$ . We invite the reader to compare this inversion cost with that needed to isolate matrix elements when using sequential source methods (discussed, for example, in ref. [109]), which are common cost saving measures when exploiting standard spatial smearing kernels.

One aim of this body of work was to demonstrate the benefits distillation can provide in calculations of hadronic structure. An earlier calculation of the isovector unpolarized valence quark content of the nucleon using the pseudo-distribution formalism on the  $a094m358$  ensemble used in this work was presented in ref. [58]; this previous calculation instead used Jacobi smearing. Figures 27 and 28 overlay the real and imaginary components, respectively, of  $\mathfrak{M}(\nu, z^2)$  reported in ref. [58] with our results first presented in figures 4a and 4b. Evidently the pseudo-ITD obtained with distillation is considerably more precise than the Jacobi smeared data of ref. [58], with improvement factors at times exceeding an order of magnitude.

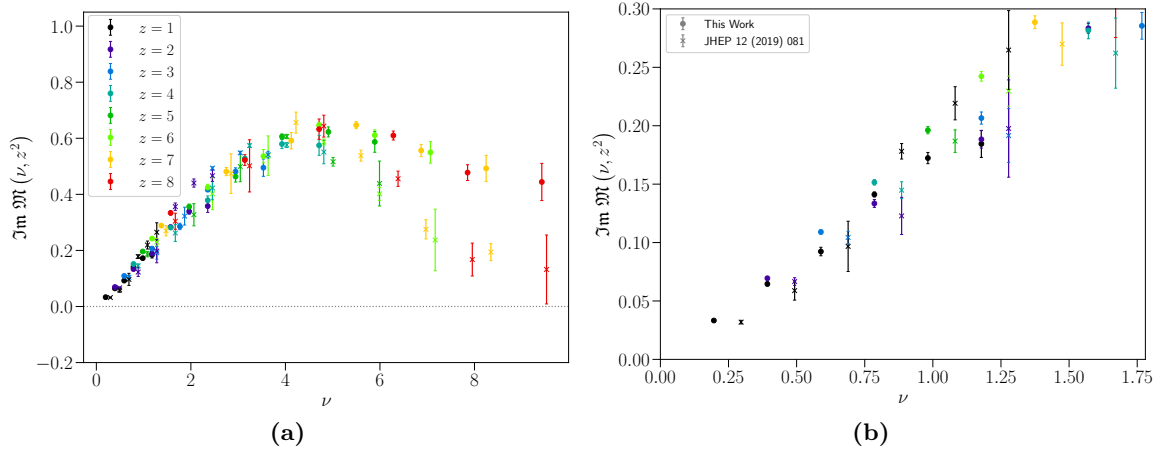
The data from ref. [58] was obtained using the sequential operator technique, for which the number of inversions needed per configuration is given by

$$N_{\text{inv}}^{\text{Seq.Op}}/\text{cfg} = (1 + N_{\text{zsep}}) N_{\zeta} N_{\text{src}} N_c N_s, \quad (\text{A.4})$$

with  $N_{\text{zsep}}$  the number of Wilson line lengths. With  $N_{\text{zsep}} = 17$  and  $N_{\zeta} = 5$  in ref. [58], it follows  $N_{\text{inv}}^{\text{Seq.Op}}/\text{cfg} = 8640$ . In other words, by roughly doubling the number of inversions per configuration, a step quite feasible with contemporary computing resources, we have achieved with distillation a reduction in error far greater than the naive  $1/\sqrt{2}$  reduction factor. Furthermore, we have relied on  $N_{\text{cfg}} \times N_{\text{srcs}} = 349 \times 4 = 1396$  measurements, while  $N_{\text{cfg}} \times N_{\text{srcs}} = 417 \times 8 = 3336$  measurements were used in ref. [58]. The reader is directed to



**Figure 27.** (a) The real component of the Ioffe-time pseudo-distribution computed in this work (circles) compared with the results of ref. [58] using Jacobi smearing (crosses). (b) Focusing on the small- $\nu$  region, the improvements afforded by distillation are seen to be considerable. Data from ref. [58] has been shifted horizontally for legibility.



**Figure 28.** (a) The imaginary component of the Ioffe-time pseudo-distribution computed in this work (circles) compared with the results of ref. [58] using Jacobi smearing (crosses). (b) Focusing on the small- $\nu$  region, the improvements afforded by distillation are seen to be considerable. Data from ref. [58] has been shifted horizontally for legibility.

ref. [109] for further comparisons between distillation and Jacobi smearing in the calculation of nucleon charges.

An alternate inversion strategy, but no less ubiquitous, is that of the sequential propagator approach. In this case the source and sink interpolating operators are held fixed, with the series of inversions beginning at the source and running through the sink. In this manner, matrix elements of a variety of operator insertions and momentum transfers can be computed without additional inversions. The number of inversions needed per configuration is given by

$$N_{\text{inv}}^{\text{Seq. Inv.}} / \text{cfg} = (1 + 2N_{\text{tsep}}) N_c N'_s N_{\text{mom}} N_{\text{src}} N_{\text{proj}} N_{\text{op}}, \quad (\text{A.5})$$

where  $N_{\text{tsep}}$  are the number of source-sink separations,  $N_{\text{mom}}$  the number of momentum projections at the sink,  $N_{\text{proj}}$  the number of spin projectors applied to the nucleon interpolating field,  $N'_s$  the number of open spinor indices after projection of the nucleon interpolator, and  $N_{\text{op}}$  the number of distinct interpolators at the sink. As standard three-point functions enforce three-momentum conservation on an ensemble average by performing explicit momentum projections only at the operator insertion and sink interpolator, the sequential propagator cost (A.5) only scales linearly with  $N_{\text{mom}}$  and  $N_{\text{op}}$  rather than also with the number of interpolators and momentum projections at the source. The factor of two in eq. (A.5) designates one sequential propagator for each of the two quark flavors in the nucleon. Translating these factors into our calculation with distillation using a single unpolarized non-relativistic nucleon interpolator, in the sequential propagator framework one would have  $N_{\text{op}} = 1$ ,  $N_{\text{proj}} = 1$  and  $N'_s = 2$  for a total of  $N_{\text{inv}}^{\text{Seq.Inv.}}/\text{cfg} = 4056$ . Indeed the sequential propagator strategy is often favored in place of the sequential operator framework discussed above and featured in ref. [58], as the inversion cost is roughly half and should produce identical correlation functions on an ensemble average.

Implicit in the sequential operator cost (A.4) is the added cost per operator insertion and momentum transfer:  $(1 + N_{\text{zsep}}N_{\text{ins}}N_{\vec{q}})$  which reduces to  $(1 + N_{\text{zsep}})$  in the case of the unpolarized PDF. The fixed inversion overhead for distillation represented by eq. (A.3) means the unpolarized, helicity and transversity PDFs are simultaneously accessible with the same set of inversions, while the inversion cost with the sequential operator method would increase by a factor of  $\frac{1+3N_{\text{zsep}}}{1+N_{\text{zsep}}}$ . Conventional smearing and the associated sequential inversion methods are then far outpaced by distillation when one is interested in off-forward observables — the sequential operator approach would require an additional  $\frac{1+N_{\text{zsep}}N_{\text{ins}}N_{\vec{q}}}{1+N_{\text{zsep}}}$  inversions from the unpolarized PDF case, while the sequential propagator method would require a prohibitive multiplicative increase by a factor  $N_{\text{mom}}$  for the number of distinct sink momenta. In an analogous fashion, were one interested in a variational analysis of a matrix of three point correlators, the sequential propagator method would require a further multiplicative increase by a factor  $N_{\text{ops}}$ , while the cost increase with the sequential operator method would manifest in additional contractions rather than inversions. This paradigm bears considerable weight on the few calculations of Generalized Parton Distributions (GPDs) from lattice QCD (e.g. [110–112]) present in the literature, wherein a large set of momenta and conceivably operator bases are required to map out the multi-dimensional structure of GPDs while effectively controlling excited-states. The cost of sequential inversion methods used in tandem with conventional smearing clearly grows rapidly in the off-forward regime, while the inversion cost with distillation remains fixed at (A.3).

The numerical benefits presented herein and the cost saving prospects possible with distillation, especially in GPD calculations, strongly motivates its continued use in future calculations of hadronic structure.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] J.C. Collins, D.E. Soper and G.F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [[hep-ph/0409313](#)] [[INSPIRE](#)].
- [2] J. Gao, L. Harland-Lang and J. Rojo, *The Structure of the Proton in the LHC Precision Era*, *Phys. Rept.* **742** (2018) 1 [[arXiv:1709.04922](#)] [[INSPIRE](#)].
- [3] K.-F. Liu and S.-J. Dong, *Origin of difference between  $\bar{d}$  and  $\bar{u}$  partons in the nucleon*, *Phys. Rev. Lett.* **72** (1994) 1790 [[hep-ph/9306299](#)] [[INSPIRE](#)].
- [4] K.-F. Liu, *Parton degrees of freedom from the path integral formalism*, *Phys. Rev. D* **62** (2000) 074501 [[hep-ph/9910306](#)] [[INSPIRE](#)].
- [5] W. Detmold and C.J.D. Lin, *Deep-inelastic scattering and the operator product expansion in lattice QCD*, *Phys. Rev. D* **73** (2006) 014501 [[hep-lat/0507007](#)] [[INSPIRE](#)].
- [6] U. Aglietti, M. Ciuchini, G. Corbo, E. Franco, G. Martinelli and L. Silvestrini, *Model independent determination of the light cone wave functions for exclusive processes*, *Phys. Lett. B* **441** (1998) 371 [[hep-ph/9806277](#)] [[INSPIRE](#)].
- [7] V. Braun and D. Müller, *Exclusive processes in position space and the pion distribution amplitude*, *Eur. Phys. J. C* **55** (2008) 349 [[arXiv:0709.1348](#)] [[INSPIRE](#)].
- [8] X. Ji, *Parton Physics on a Euclidean Lattice*, *Phys. Rev. Lett.* **110** (2013) 262002 [[arXiv:1305.1539](#)] [[INSPIRE](#)].
- [9] X. Ji, *Parton Physics from Large-Momentum Effective Field Theory*, *Sci. China Phys. Mech. Astron.* **57** (2014) 1407 [[arXiv:1404.6680](#)] [[INSPIRE](#)].
- [10] C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen and F. Manigrasso, *Flavor decomposition for the proton helicity parton distribution functions*, *Phys. Rev. Lett.* **126** (2021) 102003 [[arXiv:2009.13061](#)] [[INSPIRE](#)].
- [11] K. Cichy, L. Del Debbio and T. Giani, *Parton distributions from lattice data: the nonsinglet case*, *JHEP* **10** (2019) 137 [[arXiv:1907.06037](#)] [[INSPIRE](#)].
- [12] T. Izubuchi et al., *Valence parton distribution function of pion from fine lattice*, *Phys. Rev. D* **100** (2019) 034516 [[arXiv:1905.06349](#)] [[INSPIRE](#)].
- [13] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato and F. Steffens, *Transversity parton distribution functions from lattice QCD*, *Phys. Rev. D* **98** (2018) 091503 [[arXiv:1807.00232](#)] [[INSPIRE](#)].
- [14] J.-W. Chen, S.D. Cohen, X. Ji, H.-W. Lin and J.-H. Zhang, *Nucleon Helicity and Transversity Parton Distributions from Lattice QCD*, *Nucl. Phys. B* **911** (2016) 246 [[arXiv:1603.06664](#)] [[INSPIRE](#)].
- [15] C. Alexandrou et al., *Computation of parton distributions from the quasi-PDF approach at the physical point*, *EPJ Web Conf.* **175** (2018) 14008 [[arXiv:1710.06408](#)] [[INSPIRE](#)].
- [16] LP3 collaboration, *Improved parton distribution functions at the physical pion mass*, *Phys. Rev. D* **98** (2018) 054504 [[arXiv:1708.05301](#)] [[INSPIRE](#)].
- [17] Z.-Y. Fan, Y.-B. Yang, A. Anthony, H.-W. Lin and K.-F. Liu, *Gluon Quasi-Parton-Distribution Functions from Lattice QCD*, *Phys. Rev. Lett.* **121** (2018) 242001 [[arXiv:1808.02077](#)] [[INSPIRE](#)].



- [18] J.-H. Zhang, X. Ji, A. Schäfer, W. Wang and S. Zhao, *Accessing Gluon Parton Distributions in Large Momentum Effective Theory*, *Phys. Rev. Lett.* **122** (2019) 142001 [[arXiv:1808.10824](#)] [[INSPIRE](#)].
- [19] C. Alexandrou et al., *A complete non-perturbative renormalization prescription for quasi-PDFs*, *Nucl. Phys. B* **923** (2017) 394 [[arXiv:1706.00265](#)] [[INSPIRE](#)].
- [20] J.-W. Chen et al., *Parton distribution function with nonperturbative renormalization from lattice QCD*, *Phys. Rev. D* **97** (2018) 014505 [[arXiv:1706.01295](#)] [[INSPIRE](#)].
- [21] C. Alexandrou et al., *Lattice continuum-limit study of nucleon quasi-PDFs*, *Phys. Rev. D* **103** (2021) 094512 [[arXiv:2011.00964](#)] [[INSPIRE](#)].
- [22] C. Alexandrou et al., *Systematic uncertainties in parton distribution functions from lattice QCD simulations at the physical point*, *Phys. Rev. D* **99** (2019) 114504 [[arXiv:1902.00587](#)] [[INSPIRE](#)].
- [23] J.-H. Zhang, J.-W. Chen, X. Ji, L. Jin and H.-W. Lin, *Pion Distribution Amplitude from Lattice QCD*, *Phys. Rev. D* **95** (2017) 094514 [[arXiv:1702.00008](#)] [[INSPIRE](#)].
- [24] LATTICE PARTON collaboration, *Distribution Amplitudes of  $K^*$  and  $\phi$  at the Physical Pion Mass from Lattice QCD*, *Phys. Rev. Lett.* **127** (2021) 062002 [[arXiv:2011.09788](#)] [[INSPIRE](#)].
- [25] W. Wang, Y.-M. Wang, J. Xu and S. Zhao, *B-meson light-cone distribution amplitude from Euclidean quantities*, *Phys. Rev. D* **102** (2020) 011502 [[arXiv:1908.09933](#)] [[INSPIRE](#)].
- [26] S. Bhattacharya, K. Cichy, M. Constantinou, A. Metz, A. Scapellato and F. Steffens, *Insights on proton structure from lattice QCD: The twist-3 parton distribution function  $g_T(x)$* , *Phys. Rev. D* **102** (2020) 111501 [[arXiv:2004.04130](#)] [[INSPIRE](#)].
- [27] M. Constantinou, *The  $x$ -dependence of hadronic parton distributions: A review on the progress of lattice QCD*, *Eur. Phys. J. A* **57** (2021) 77 [[arXiv:2010.02445](#)] [[INSPIRE](#)].
- [28] K. Cichy and M. Constantinou, *A guide to light-cone PDFs from Lattice QCD: an overview of approaches, techniques and results*, *Adv. High Energy Phys.* **2019** (2019) 3036904 [[arXiv:1811.07248](#)] [[INSPIRE](#)].
- [29] A.V. Radyushkin, *Quasi-parton distribution functions, momentum distributions, and pseudo-parton distribution functions*, *Phys. Rev. D* **96** (2017) 034025 [[arXiv:1705.01488](#)] [[INSPIRE](#)].
- [30] Y.-Q. Ma and J.-W. Qiu, *Extracting Parton Distribution Functions from Lattice QCD Calculations*, *Phys. Rev. D* **98** (2018) 074021 [[arXiv:1404.6860](#)] [[INSPIRE](#)].
- [31] Y.-Q. Ma and J.-W. Qiu, *Exploring Partonic Structure of Hadrons Using *ab initio* Lattice QCD Calculations*, *Phys. Rev. Lett.* **120** (2018) 022003 [[arXiv:1709.03018](#)] [[INSPIRE](#)].
- [32] R.S. Sufian, J. Karpie, C. Egerer, K. Orginos, J.-W. Qiu and D.G. Richards, *Pion Valence Quark Distribution from Matrix Element Calculated in Lattice QCD*, *Phys. Rev. D* **99** (2019) 074507 [[arXiv:1901.03921](#)] [[INSPIRE](#)].
- [33] R.S. Sufian et al., *Pion Valence Quark Distribution from Current-Current Correlation in Lattice QCD*, *Phys. Rev. D* **102** (2020) 054508 [[arXiv:2001.04960](#)] [[INSPIRE](#)].
- [34] B.L. Ioffe, *Space-time picture of photon and neutrino scattering and electroproduction cross-section asymptotics*, *Phys. Lett. B* **30** (1969) 123 [[INSPIRE](#)].
- [35] V. Braun, P. Gornicki and L. Mankiewicz, *Ioffe-time distributions instead of parton momentum distributions in description of deep inelastic scattering*, *Phys. Rev. D* **51** (1995) 6036 [[hep-ph/9410318](#)] [[INSPIRE](#)].



- [36] A. Radyushkin, *Nonperturbative Evolution of Parton Quasi-Distributions*, *Phys. Lett. B* **767** (2017) 314 [[arXiv:1612.05170](#)] [[INSPIRE](#)].
- [37] A.M. Polyakov, *Gauge Fields as Rings of Glue*, *Nucl. Phys. B* **164** (1980) 171 [[INSPIRE](#)].
- [38] V.S. Dotsenko and S.N. Vergeles, *Renormalizability of Phase Factors in the Nonabelian Gauge Theory*, *Nucl. Phys. B* **169** (1980) 527 [[INSPIRE](#)].
- [39] R.A. Brandt, F. Neri and M.-a. Sato, *Renormalization of Loop Functions for All Loops*, *Phys. Rev. D* **24** (1981) 879 [[INSPIRE](#)].
- [40] N.S. Craigie and H. Dorn, *On the Renormalization and Short Distance Properties of Hadronic Operators in QCD*, *Nucl. Phys. B* **185** (1981) 204 [[INSPIRE](#)].
- [41] T. Ishikawa, Y.-Q. Ma, J.-W. Qiu and S. Yoshida, *Renormalizability of quasiparton distribution functions*, *Phys. Rev. D* **96** (2017) 094019 [[arXiv:1707.03107](#)] [[INSPIRE](#)].
- [42] X. Ji, J.-H. Zhang and Y. Zhao, *Renormalization in Large Momentum Effective Theory of Parton Physics*, *Phys. Rev. Lett.* **120** (2018) 112001 [[arXiv:1706.08962](#)] [[INSPIRE](#)].
- [43] J. Green, K. Jansen and F. Steffens, *Nonperturbative Renormalization of Nonlocal Quark Bilinears for Parton Quasidistribution Functions on the Lattice Using an Auxiliary Field*, *Phys. Rev. Lett.* **121** (2018) 022004 [[arXiv:1707.07152](#)] [[INSPIRE](#)].
- [44] K. Orginos, A. Radyushkin, J. Karpie and S. Zafeiropoulos, *Lattice QCD exploration of parton pseudo-distribution functions*, *Phys. Rev. D* **96** (2017) 094503 [[arXiv:1706.05373](#)] [[INSPIRE](#)].
- [45] V.M. Braun, A. Vladimirov and J.-H. Zhang, *Power corrections and renormalons in parton quasidistributions*, *Phys. Rev. D* **99** (2019) 014013 [[arXiv:1810.00048](#)] [[INSPIRE](#)].
- [46] Z.-Y. Li, Y.-Q. Ma and J.-W. Qiu, *Extraction of Next-to-Next-to-Leading-Order Parton Distribution Functions from Lattice QCD Calculations*, *Phys. Rev. Lett.* **126** (2021) 072001 [[arXiv:2006.12370](#)] [[INSPIRE](#)].
- [47] T. Izubuchi, X. Ji, L. Jin, I.W. Stewart and Y. Zhao, *Factorization Theorem Relating Euclidean and Light-Cone Parton Distributions*, *Phys. Rev. D* **98** (2018) 056004 [[arXiv:1801.03917](#)] [[INSPIRE](#)].
- [48] A. Radyushkin, *One-loop evolution of parton pseudo-distribution functions on the lattice*, *Phys. Rev. D* **98** (2018) 014019 [[arXiv:1801.02427](#)] [[INSPIRE](#)].
- [49] J.-H. Zhang, J.-W. Chen and C. Monahan, *Parton distribution functions from reduced Ioffe-time distributions*, *Phys. Rev. D* **97** (2018) 074508 [[arXiv:1801.03023](#)] [[INSPIRE](#)].
- [50] L.-B. Chen, W. Wang and R. Zhu, *Next-to-Next-to-Leading Order Calculation of Quasiparton Distribution Functions*, *Phys. Rev. Lett.* **126** (2021) 072002 [[arXiv:2006.14825](#)] [[INSPIRE](#)].
- [51] Y.L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics*, *Sov. Phys. JETP* **46** (1977) 641 [[INSPIRE](#)].
- [52] V.N. Gribov and L.N. Lipatov,  *$e^+e^-$  pair annihilation and deep inelastic  $e p$  scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 675 [[INSPIRE](#)].
- [53] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298 [[INSPIRE](#)].
- [54] B. Joó et al., *Pion valence structure from Ioffe-time parton pseudodistribution functions*, *Phys. Rev. D* **100** (2019) 114512 [[arXiv:1909.08517](#)] [[INSPIRE](#)].

- [55] HADSTRUC collaboration, *The continuum and leading twist limits of parton distribution functions in lattice QCD*, *JHEP* **11** (2021) 024 [[arXiv:2105.13313](#)] [[INSPIRE](#)].
- [56] B. Joó, J. Karpie, K. Orginos, A.V. Radyushkin, D.G. Richards and S. Zafeiropoulos, *Parton Distribution Functions from Ioffe Time Pseudodistributions from Lattice Calculations: Approaching the Physical Point*, *Phys. Rev. Lett.* **125** (2020) 232003 [[arXiv:2004.01687](#)] [[INSPIRE](#)].
- [57] M. Bhat, K. Cichy, M. Constantinou and A. Scapellato, *Flavor nonsinglet parton distribution functions from lattice QCD at physical quark masses via the pseudodistribution approach*, *Phys. Rev. D* **103** (2021) 034510 [[arXiv:2005.02102](#)] [[INSPIRE](#)].
- [58] B. Joó, J. Karpie, K. Orginos, A. Radyushkin, D. Richards and S. Zafeiropoulos, *Parton Distribution Functions from Ioffe time pseudo-distributions*, *JHEP* **12** (2019) 081 [[arXiv:1908.09771](#)] [[INSPIRE](#)].
- [59] Z. Fan, R. Zhang and H.-W. Lin, *Nucleon gluon distribution function from 2 + 1 + 1-flavor lattice QCD*, *Int. J. Mod. Phys. A* **36** (2021) 2150080 [[arXiv:2007.16113](#)] [[INSPIRE](#)].
- [60] C. Egerer, R.G. Edwards, K. Orginos and D.G. Richards, *Distillation at High-Momentum*, *Phys. Rev. D* **103** (2021) 034502 [[arXiv:2009.10691](#)] [[INSPIRE](#)].
- [61] R. Edwards, B. Joó, K. Orginos, D. Richards and F. Winter, *U.S. 2 + 1 flavor clover lattice generation program*, unpublished (2016).
- [62] S. Borsányi et al., *High-precision scale setting in lattice QCD*, *JHEP* **09** (2012) 010 [[arXiv:1203.4469](#)] [[INSPIRE](#)].
- [63] B. Yoon et al., *Isovector charges of the nucleon from 2 + 1-flavor QCD with clover fermions*, *Phys. Rev. D* **95** (2017) 074508 [[arXiv:1611.07452](#)] [[INSPIRE](#)].
- [64] B. Yoon et al., *Controlling Excited-State Contamination in Nucleon Matrix Elements*, *Phys. Rev. D* **93** (2016) 114506 [[arXiv:1602.07737](#)] [[INSPIRE](#)].
- [65] C. Morningstar and M.J. Peardon, *Analytic smearing of SU(3) link variables in lattice QCD*, *Phys. Rev. D* **69** (2004) 054501 [[hep-lat/0311018](#)] [[INSPIRE](#)].
- [66] UKQCD collaboration, *Gauge invariant smearing and matrix correlators using Wilson fermions at  $\beta = 6.2$* , *Phys. Rev. D* **47** (1993) 5128 [[hep-lat/9303009](#)] [[INSPIRE](#)].
- [67] HADRON SPECTRUM collaboration, *A Novel quark-field creation operator construction for hadronic physics in lattice QCD*, *Phys. Rev. D* **80** (2009) 054506 [[arXiv:0905.2160](#)] [[INSPIRE](#)].
- [68] G.S. Bali, B. Lang, B.U. Musch and A. Schäfer, *Novel quark smearing for hadrons with high momenta in lattice QCD*, *Phys. Rev. D* **93** (2016) 094515 [[arXiv:1602.05525](#)] [[INSPIRE](#)].
- [69] R.G. Edwards, J.J. Dudek, D.G. Richards and S.J. Wallace, *Excited state baryon spectroscopy from lattice QCD*, *Phys. Rev. D* **84** (2011) 074508 [[arXiv:1104.5152](#)] [[INSPIRE](#)].
- [70] J.J. Dudek and R.G. Edwards, *Hybrid Baryons in QCD*, *Phys. Rev. D* **85** (2012) 054016 [[arXiv:1201.2349](#)] [[INSPIRE](#)].
- [71] C.E. Thomas, R.G. Edwards and J.J. Dudek, *Helicity operators for mesons in flight on the lattice*, *Phys. Rev. D* **85** (2012) 014507 [[arXiv:1107.1930](#)] [[INSPIRE](#)].
- [72] L. Maiani, G. Martinelli, M.L. Paciello and B. Taglienti, *Scalar Densities and Baryon Mass Differences in Lattice QCD With Wilson Fermions*, *Nucl. Phys. B* **293** (1987) 420 [[INSPIRE](#)].

- [73] S. Capitani et al., *The nucleon axial charge from lattice QCD with controlled errors*, *Phys. Rev. D* **86** (2012) 074502 [[arXiv:1205.0180](#)] [[INSPIRE](#)].
- [74] X. Gao et al., *Valence parton distribution of the pion from lattice QCD: Approaching the continuum limit*, *Phys. Rev. D* **102** (2020) 094513 [[arXiv:2007.06590](#)] [[INSPIRE](#)].
- [75] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [[arXiv:1412.7420](#)] [[INSPIRE](#)].
- [76] EVENT HORIZON TELESCOPE collaboration, *First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole*, *Astrophys. J. Lett.* **875** (2019) L4 [[arXiv:1906.11241](#)] [[INSPIRE](#)].
- [77] G. Backus and F. Gilbert, *The Resolving Power of Gross Earth Data*, *Geophys. J. Int.* **16** (1968) 169.
- [78] M. Asakawa, T. Hatsuda and Y. Nakahara, *Maximum entropy analysis of the spectral functions in lattice QCD*, *Prog. Part. Nucl. Phys.* **46** (2001) 459 [[hep-lat/0011040](#)] [[INSPIRE](#)].
- [79] J. Karpie, K. Orginos, A. Rothkopf and S. Zafeiropoulos, *Reconstructing parton distribution functions from Ioffe time data: from Bayesian methods to Neural Networks*, *JHEP* **04** (2019) 057 [[arXiv:1901.05408](#)] [[INSPIRE](#)].
- [80] XQCD collaboration, *Towards the nucleon hadronic tensor from lattice QCD*, *Phys. Rev. D* **101** (2020) 114503 [[arXiv:1906.05312](#)] [[INSPIRE](#)].
- [81] M.T. Hansen, H.B. Meyer and D. Robaina, *From deep inelastic scattering to heavy-flavor semileptonic decays: Total rates into multihadron final states from lattice QCD*, *Phys. Rev. D* **96** (2017) 094513 [[arXiv:1704.08993](#)] [[INSPIRE](#)].
- [82] A. Accardi, L.T. Brady, W. Melnitchouk, J.F. Owens and N. Sato, *Constraints on large- $x$  parton distributions from new weak boson production and deep-inelastic scattering data*, *Phys. Rev. D* **93** (2016) 114017 [[arXiv:1602.03154](#)] [[INSPIRE](#)].
- [83] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189 [[arXiv:0901.0002](#)] [[INSPIRE](#)].
- [84] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204 [[arXiv:1412.3989](#)] [[INSPIRE](#)].
- [85] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) 014013 [[arXiv:1912.10053](#)] [[INSPIRE](#)].
- [86] NNPDF collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [[arXiv:1410.8849](#)] [[INSPIRE](#)].
- [87] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [[arXiv:1706.00428](#)] [[INSPIRE](#)].
- [88] JEFFERSON LAB ANGULAR MOMENTUM (JAM) collaboration, *Simultaneous Monte Carlo analysis of parton densities and fragmentation functions*, *Phys. Rev. D* **104** (2021) 016015 [[arXiv:2101.04664](#)] [[INSPIRE](#)].
- [89] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions*, *Eur. Phys. J. C* **70** (2010) 51 [[arXiv:1007.2624](#)] [[INSPIRE](#)].
- [90] A.V. Radyushkin, *Generalized parton distributions and pseudodistributions*, *Phys. Rev. D* **100** (2019) 116011 [[arXiv:1909.08474](#)] [[INSPIRE](#)].

- [91] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) 033006 [[arXiv:1506.07443](#)] [[INSPIRE](#)].
- [92] G.S. Bali et al., *Pion distribution amplitude from Euclidean correlation functions: Exploring universality and higher-twist effects*, *Phys. Rev. D* **98** (2018) 094507 [[arXiv:1807.06671](#)] [[INSPIRE](#)].
- [93] RQCD collaboration, *Light-cone distribution amplitudes of pseudoscalar mesons from lattice QCD*, *JHEP* **08** (2019) 065 [Addendum *ibid.* **11** (2020) 037] [[arXiv:1903.08038](#)] [[INSPIRE](#)].
- [94] J. Segovia, L. Chang, I.C. Cloët, C.D. Roberts, S.M. Schmidt and H.-s. Zong, *Distribution amplitudes of light-quark mesons from lattice QCD*, *Phys. Lett. B* **731** (2014) 13 [[arXiv:1311.1390](#)] [[INSPIRE](#)].
- [95] H. Fukaya, S. Hashimoto, T. Kaneko and H. Ohki, *Towards fully nonperturbative computations of inelastic  $\ell N$  scattering cross sections from lattice QCD*, *Phys. Rev. D* **102** (2020) 114516 [[arXiv:2010.01253](#)] [[INSPIRE](#)].
- [96] J. Karpie, K. Orginos and S. Zafeiropoulos, *Moments of Ioffe time parton distribution functions from non-local matrix elements*, *JHEP* **11** (2018) 178 [[arXiv:1807.10933](#)] [[INSPIRE](#)].
- [97] G.H. Golub and V. Pereyra, *The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate*, *SIAM J. Numer. Anal.* **10** (1973) 413.
- [98] S.J. Brodsky and G.R. Farrar, *Scaling Laws at Large Transverse Momentum*, *Phys. Rev. Lett.* **31** (1973) 1153 [[INSPIRE](#)].
- [99] L. Del Debbio, T. Giani, J. Karpie, K. Orginos, A. Radyushkin and S. Zafeiropoulos, *Neural-network analysis of Parton Distribution Functions from Ioffe-time pseudodistributions*, *JHEP* **02** (2021) 138 [[arXiv:2010.03996](#)] [[INSPIRE](#)].
- [100] J. Bringewatt, N. Sato, W. Melnitchouk, J.-W. Qiu, F. Steffens and M. Constantinou, *Confronting lattice parton distributions with global QCD analysis*, *Phys. Rev. D* **103** (2021) 016003 [[arXiv:2010.00548](#)] [[INSPIRE](#)].
- [101] D. Stanzione, J. West, R.T. Evans, T. Minyard, O. Ghattas and D.K. Panda, *Frontera: The evolution of leadership computing at the national science foundation*, in *Practice and Experience in Advanced Research Computing*, PEARC '20, New York, NY, U.S.A., pp. 106–111, Association for Computing Machinery (2020) [[DOI](#)].
- [102] J. Towns et al., *XSEDE: Accelerating Scientific Discovery*, *Comput. Sci. Eng.* **16** (2014) 62 [[INSPIRE](#)].
- [103] SciDAC, LHPC and UKQCD collaborations, *The Chroma software system for lattice QCD*, *Nucl. Phys. B Proc. Suppl.* **140** (2005) 832 [[hep-lat/0409003](#)] [[INSPIRE](#)].
- [104] M.A. Clark, R. Babich, K. Barros, R.C. Brower and C. Rebbi, *Solving Lattice QCD systems of equations using mixed precision solvers on GPUs*, *Comput. Phys. Commun.* **181** (2010) 1517 [[arXiv:0911.3191](#)] [[INSPIRE](#)].
- [105] R. Babich, M.A. Clark and B. Joo, *Parallelizing the QUDA Library for Multi-GPU Calculations in Lattice Quantum Chromodynamics*, in *SC 10 (Supercomputing 2010)*, (2010) [[arXiv:1011.0024](#)] [[INSPIRE](#)].
- [106] F.T. Winter, M.A. Clark, R.G. Edwards and B. Joó, *A Framework for Lattice QCD Calculations on GPUs*, in *28th IEEE International Parallel and Distributed Processing Symposium*, (2014) [[DOI](#)] [[arXiv:1408.5925](#)] [[INSPIRE](#)].

- [107] B. Joó et al., *Lattice QCD on Intel® Xeon Phi™ Coprocessors*, *Lect. Notes Comput. Sci.* **7905** (2013) 40.
- [108] B. Joó, D.D. Kalamkar, T. Kurth, K. Vaidyanathan and A. Walden, *Optimizing Wilson-Dirac operator and linear solvers for Intel® KNL*, in *High Performance Computing*, M. Taufer, B. Mohr and J.M. Kunkel eds., Cham, pp. 415–427, Springer International Publishing (2016) [DOI].
- [109] C. Egerer, D. Richards and F. Winter, *Controlling excited-state contributions with distillation in lattice QCD calculations of nucleon isovector charges  $g_S^{u-d}$ ,  $g_A^{u-d}$ ,  $g_T^{u-d}$* , *Phys. Rev. D* **99** (2019) 034506 [arXiv:1810.09991] [INSPIRE].
- [110] J.-W. Chen, H.-W. Lin and J.-H. Zhang, *Pion generalized parton distribution from lattice QCD*, *Nucl. Phys. B* **952** (2020) 114940 [arXiv:1904.12376] [INSPIRE].
- [111] C. Alexandrou et al., *Unpolarized and helicity generalized parton distributions of the proton within lattice QCD*, *Phys. Rev. Lett.* **125** (2020) 262001 [arXiv:2008.10573] [INSPIRE].
- [112] C. Alexandrou et al., *Transversity GPDs of the proton from lattice QCD*, arXiv:2108.10789 [INSPIRE].