Lattice QCD constraints on the parton distribution functions of ³He

William Detmold,¹ Marc Illa,² David J. Murphy,¹ Patrick Oare,¹

Kostas Orginos,^{3,4} Phiala E. Shanahan,¹ Michael L. Wagman,⁵ and Frank Winter⁴

(NPLQCD collaboration)

¹Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Departament de Física Quàntica i Astrofísica and Institut de Ciències del Cosmos,

Universitat de Barcelona, Martí Franquès 1, E08028-Spain

³Department of Physics, College of William and Mary, Williamsburg, VA 23187-8795, USA

⁴ Jefferson Laboratory, 12000 Jefferson Avenue, Newport News, VA 23606, USA

 5Fermi National Accelerator Laboratory, Batavia, IL 60510, USA

The fraction of the longitudinal momentum of ³He that is carried by the isovector combination of u and d quarks is determined using lattice QCD for the first time. The ratio of this combination to that in the constituent nucleons is found to be consistent with unity at the few-percent level from calculations with quark masses corresponding to $m_{\pi} \sim 800$ MeV, extrapolated to the physical quark masses. This constraint is consistent with, and significantly more precise than, determinations from global nuclear parton distribution function fits. Including the lattice QCD determination of the momentum fraction in the nNNPDF global fitting framework results in the uncertainty on the isovector momentum fraction ratio being reduced by a factor of 2.5, and thereby enables a more precise extraction of the u and d parton distributions in ³He.

A central pillar of our understanding of the internal structure of strongly interacting hadronic and nuclear systems is knowledge of their partonic structure as accessed in deep-inelastic scattering (DIS) experiments and other hard processes. Since the 1960s, such experiments have revealed the longitudinal momentum distributions of quarks and gluons in a fast moving proton, known collectively as parton distribution functions (PDFs). The simplest PDFs, $q(x,\mu)$ (and $q(x,\mu)$), describe the probability of a quark of flavor q (or gluon q) carrying a fraction x of the longitudinal momentum of the struck proton at a renormalization scale μ . In 1983, the European Muon Collaboration (EMC) [1] observed that the partonic structure of nuclei differs substantially from that of the constituent protons and neutrons, a landmark in the development of nuclear physics [2-6]. Since the DIS processes observed in the EMC experiments were at very high energy, and the binding energy of a nucleus is small in comparison to its mass, the appearance and size of the EMC effect was surprising at the time. Interest in the EMC effect has been rekindled by recent data from SLAC and JLab [7–11] on EMC ratios for light nuclei. Not only have these data provided precise determinations of the EMC effect for nuclei with small atomic number A, but they have revealed a correlation between the strength of the EMC effect and so called "short range correlations" [12, 13].

In addition to experimental investigations, theoretical calculations of the partonic structure of hadrons and nuclei from the Standard Model can have important impact on our understanding of the structure of matter. For example, Standard Model calculations of nuclear partonic structure would reveal the QCD origin of the EMC effect as well as aid in the flavor-separation of proton PDFs. Parton distributions are inherently rooted in the strong interaction dynamics of QCD and cannot be determined using perturbative methods. Since the seminal works of Refs. [14, 15], lattice Quantum Chromodynamics (LQCD) calculations have addressed the simplest aspects of the parton distributions of the proton, notably determining the first few Mellin moments of the unpolarized, polarized, and transversity quark distributions [16], as well as their gluonic analogues [17–21]. Recently, efforts have been made to extend these studies to the full x-dependence of the proton PDFs [16, 22-24]. More complicated extensions of partonic structure, such as generalized parton distribution functions and transversemomentum dependent parton distribution functions of the proton, have also been studied using LQCD [25–30].

In this letter, the partonic structure of light nuclei is studied in LQCD for the first time through an investigation of the isovector quark momentum fractions (the first moments of the corresponding isovector PDFs) of the proton, diproton, and ³He. At the heavier-than-physical quark masses used in this LQCD study, percent-level nuclear effects are resolved in the momentum fraction of ³He. After an extrapolation to the physical quark masses, these calculations provide a constraint on the isovector momentum fraction that is used as an additional input into the nNNPDF2.0 [31] global nuclear PDF analysis framework. Since the isovector combination of nuclear PDFs is poorly determined from experiment, this LQCD constraint significantly reduces the uncertainties on the ³He PDFs and thereby improves knowledge of nuclear structure.

LQCD methodology — The existence of strong interactions between quarks and gluons necessitates the use of LQCD for calculations of the partonic structure of nuclei. The calculations presented here are performed using a single ensemble of gauge-field configurations generated with a Lüscher-Weisz gauge action [32] with $N_f = 3$ degenerate light-quark flavors with the clover-improved Wilson fermion action [33], and quark masses tuned to produce a pion mass of $m_{\pi} = 806$ MeV. The lattice geometry is $L^3 \times T = 32^3 \times 48$, and the lattice spacing is determined to be $a \sim 0.145$ fm from Υ spectroscopy [34]. This ensemble, and two others with different spacetime volumes have previously been used to study the spectrum [34, 35] and properties [36–46] of light nuclei up to atomic number A = 4. The multi-volume spectroscopy studies show that the pp and ³He states that are investigated here are bound systems with infinite volume energies below threshold. Consequently, matrix elements in these states are expected to receive only exponentially small finite volume effects, $\mathcal{O}(e^{-\kappa L}, e^{-m_{\pi}L})$, that will be neglected in this work [47-53].

The Mellin moments of the unpolarized isovector quark PDFs, $q_3^{(h)}(x,\mu) = u^{(h)}(x,\mu) - d^{(h)}(x,\mu)$, in a hadronic or nuclear state h, defined as $\langle x^n \rangle_{u-d}^{(h)}(\mu) \equiv \int_{-1}^1 dx \, x^n q_3^{(h)}(x,\mu)$, are determined from matrix elements of twist-two operators as

$$\langle h | \mathcal{O}_{\mu_0 \dots \mu_n} | h \rangle \equiv \langle h | \overline{q} \tau_3 \gamma_{\{\mu_0} (i \overleftrightarrow{D}_{\mu_1}) \dots (i \overleftrightarrow{D}_{\mu_n\}}) q | h \rangle$$

= $\langle x^n \rangle_{u-d}^{(h)} (\mu) p_{\{\mu_0} \dots p_{\mu_n\}},$ (1)

where p is the momentum of the state h, τ_3 is a Pauli matrix in flavor space, $\overrightarrow{D}_{\mu} = (\overrightarrow{D}_{\mu} - \overleftarrow{D}_{\mu})/2$ where D_{μ} is the gauge covariant derivative, and $\{\ldots\}$ indicates symmetrization and trace-subtraction of the enclosed indices. The above operators are constructed to transform irreducibly under the Lorentz group, but the hypercubic spacetime lattice used in the LQCD calculations reduces these symmetries, in general inducing mixing between operators of different Lorentz spin. In particular, the two-index operators that determine the isovector quark momentum fraction, $\langle x \rangle_{u-d}^{(h)}$, subduce to operators in two different irreducible representations of the hypercubic group. In this work, matrix elements of an Euclidean operator in the $\tau_1^{(3)}$ representation [54] are computed, namely

$$\mathcal{T} = \frac{1}{\sqrt{2}} \left(\mathcal{T}_{33} - \mathcal{T}_{44} \right), \text{ with } \mathcal{T}_{\mu\nu} = \overline{q} \tau_3 \gamma_{\{\mu} \overleftrightarrow{D}_{\nu\}} q, \quad (2)$$

where γ_{ν} is also Euclidean. With a lattice regulator, this operator is discretized as a covariant finite difference whose form is given in the Supplementary Material. For both spin-zero and spin-half systems, spin-averaged in the latter case, matrix elements in states with zero three-momentum determine the momentum fraction as $\langle h | \mathcal{T} | h \rangle = \langle x \rangle_{u-d}^{(h)} M_h / \sqrt{2}.$

The renormalized operator in the modified minimal subtraction scheme $(\overline{\text{MS}})$ is related to the bare lattice operator in Eq. (2) as

$$\mathcal{T}^{(\overline{\mathrm{MS}})}(\mu) = \mathcal{R}^{\overline{\mathrm{MS}}/\mathrm{RI'MOM}}(\mu,\mu_0)\mathcal{Z}^{\mathrm{RI'MOM}}(\mu_0,a)\mathcal{T}(a),$$
(3)

where the renormalization coefficient $\mathcal{Z}^{\text{RI'MOM}}(\mu_0, a)$ is defined non-perturbatively in a regularizationindependent momentum-subtraction scheme [55] at a scale μ_0 and then matched to $\overline{\text{MS}}$ through the three-loop perturbative coefficient $\mathcal{R}^{\overline{\text{MS}}/\text{RI'MOM}}(\mu, \mu_0)$ [56, 57], as detailed in the Supplementary Material. For $\mu = 2$ GeV, $\mathcal{R}^{\overline{\text{MS}}/\text{RI'MOM}}(\mu, \mu_0)\mathcal{Z}^{\text{RI'MOM}}(\mu_0, a) = 0.89(4).$

The techniques needed to compute matrix elements of this operator are simple generalizations of those used for calculations of isovector matrix elements of quark currents using the compound-propagator background-field method introduced in Ref. [42] and further detailed in Refs. [43, 44, 58] and the Supplementary Material. Quark propagators and \mathcal{T} -compound propagators are computed from an average of $N_{\rm src} = 24$ source points randomly distributed on $N_{\rm cfg} = 2290$ gauge-field configurations for $N_B = 5$ different background field strengths. These compound propagators are then used to construct baryon two-point correlation functions,

$$G_h(t;\lambda) = \sum_{\mathbf{x}} \left\langle 0 \left| \chi_h(\mathbf{x},t) \chi_h^{\dagger}(0) \right| 0 \right\rangle_{\lambda} , \qquad (4)$$

where λ is the \mathcal{T} -background field strength, χ_h is an interpolating field for states with the quantum numbers of the hadron or nucleus h, and spinor indices on the interpolating operators are suppressed. Correlation functions are constructed from Gaussian-smeared source interpolating operators [59], while the sink interpolating operators are either smeared or point-like and the multibaryon contractions are performed using the techniques of Refs. [60]. This quantity contains responses to the field up to $\mathcal{O}(\lambda^{N_Q})$, with N_Q being the number of valence quarks in the state. The linear response of this background-field two-point function, $G_h(t;\lambda)|_{\mathcal{O}(\lambda)}$, is determined by the matrix element of \mathcal{T} . This term can be extracted exactly from the computed set of fixedorder background field correlation functions with N_Q field strengths [42, 43].

Combining the linear response of the two-point correlation function with the zero-field correlation function, it is straightforward to show that the ratio

$$\mathcal{R}_h(t) = \frac{G_h(t;\lambda)|_{\mathcal{O}(\lambda)}}{G_h(t;0)} - \frac{G_h(t-a;\lambda)|_{\mathcal{O}(\lambda)}}{G_h(t-a;0)}$$
(5)

is related to matrix elements of \mathcal{T} through the spectral representation of each term in Eq. (5), in particular



FIG. 1. The effective matrix element, Eq. (5), associated with the isovector quark momentum fractions of the proton, pp and ³He. Blue (orange) points, labelled SS (SP), show results for interpolating operators with smeared sources and smeared (point-like) sinks. For each effective matrix element, points are shown for $t \leq t_{\max}$, where t_{\max} is the minimum twhere the signal-to-noise ratio of $G_h(t+a;\lambda)|_{\mathcal{O}(\lambda)}$ is less than 0.5. Colored bands show the highest weight fit to the combined dataset and the shaded gray bands show the weighted average of all accepted fits and the total statistical plus fitting systematic uncertainties.

asymptoting as

$$\mathcal{R}_{h}(t) \stackrel{t \to \infty}{\longrightarrow} \langle h | \mathcal{T} | h \rangle, \qquad (6)$$

with exponentially vanishing contamination at early times that involves excited-state overlap factors and transition matrix elements.

Ground-state matrix elements are extracted from $\mathcal{R}_h(t)$, and systematic fitting uncertainties are estimated, using a procedure for sampling from all possible fit ranges and models analogous to the procedure described for twopoint correlation functions in Ref. [61]. In summary, in analyzing $\mathcal{R}_h(t)$ to extract the momentum fractions, the full t dependence that results from the spectral decomposition of each term in Eq. (5) is fit, and combined fits to two- and three-point correlation functions are used to constrain the relevant energies, overlap factors, and matrix elements. All possible choices of fit ranges and up



FIG. 2. Left: Renormalized isovector momentum fractions for $h \in \{p, pp, {}^{3}\text{He}\}$ at a scale of $\mu = 2$ GeV. Right: Ratios of the isovector nuclear momentum fractions to that of the constituent nucleons.

	p	pp	$^{3}\mathrm{He}$
$\langle x \rangle_{u-d}^{(h)}$	0.191(1)(9)	0.194(2)(9)	0.066(1)(3)
$\left(\frac{A}{Z-N}\right)\langle x\rangle_{u-d}^{(h)}/\langle x\rangle_{u-d}^{(p)}$		1.007(14)	1.028(15)

TABLE I. The isovector quark momentum fractions in p, pp and ³He, calculated at $m_{\pi} = 806$ MeV in $\overline{\text{MS}}$ -scheme at $\mu = 2$ GeV. The first uncertainty combines LQCD statistical and systematic uncertainties and the second uncertainty is from operator renormalization. The correlated ratios of the isovector momentum fraction in nuclei to those in the constituent nucleons, in which the renormalization constants and their uncertainties cancel, are also given.

to 4 states contributing to the spectral decompositions are considered using a model selection process described in the Supplementary Material. A weighted average over fits from all acceptable fit ranges is used to define ground-state energy results, including systematic uncertainties from fit range and model variation. Results are shown in Fig. 1 for the proton, diproton and ³He.

Results and Discussion — The extracted values of the isovector quark momentum fractions for $p, pp, {}^{3}\text{He}$ at quark masses corresponding to $m_{\pi} = m_K = 806$ MeV are shown in Tab. I and displayed graphically in Fig. 2. The uncertainties are separated into those from the LQCD calculation of the bare matrix elements, and the (larger) uncertainty from the renormalization and matching to the MS scheme. The proton isovector momentum fraction is consistent with other LQCD extractions at similar values of the quark masses [62] given the different renormalization procedures and lattice spacings. The pp and ³He momentum fractions are determined with $\mathcal{O}(5\%)$ uncertainties and are found to be approximately consistent with those of the constituent nucleons. The ratios of the nuclear momentum fractions to that of the proton are independent of operator renormalization to $\mathcal{O}(\alpha_s)$, and are determined at few-percent precision even for ³He.

In Refs. [63–65], nuclear effective field theory (EFT)

was used to study nuclear effects in PDF moments. In particular, it was shown that the leading source of such effects is the two-nucleon correlations that couple to the twist-two operators defining the PDF moments. In terms of the parameters defined in that work, nuclear effects in the isovector momentum fraction are encapsulated in the low energy constant (LEC) $\alpha_{3,2}$ and nuclear factor $\mathcal{G}_3({}^{3}\text{He})$; their product is bounded as $\alpha_{3,2}\mathcal{G}_3({}^{3}\text{He}) =$ 0.0018(14) at $\mu = 2$ GeV from the numerical calculations presented here (see the Supplementary Material for details). While the quark momentum fractions themselves have nonanalytic dependence on the quark masses [66-68], this two-body LEC is expected to be relatively insensitive to variation of the quark masses, as seen for the the analogous two-body contribution in the $np \rightarrow d\gamma$ [39] and $pp \rightarrow de^+\nu_e$ [42, 69] processes. This relative mass-independence assumption allows an extrapolation to the physical quark masses: a naive estimate is given by taking the central value determined at $m_{\pi} = 806$ MeV and inflating the uncertainty by 50% to account for possible quark-mass dependence as well as the effects of the nonzero lattice spacing and finite volume (this uncertainty is estimated based on the mass dependence seen for the analogous two-body LECs in Refs. [39, 42, 69]). This extrapolated value can be combined with the physical value of the nucleon momentum fraction, $\langle x \rangle_{u-d}^{(p)} = 0.160(7)$ at $\mu = 2$ GeV from the nNNPDF2.0 analysis [31], to determine the isovector momentum fraction ratio $3\langle x \rangle_{u-d}^{(^{3}\text{He})}/\langle x \rangle_{u-d}^{(p)}|_{\text{LQCD}} = 1.035(26)$ at the physical quark masses (see the Supplementary Material for more details).

It is interesting to compare the LQCD results for the momentum fractions and their ratios to phenomenology. In particular, the isovector momentum fractions determined here provide valuable information that is complementary to experimental constraints on the nuclear modification of PDFs; almost all information on the nuclear modification of partonic structure has been obtained for the ratio of isoscalar-corrected F_2 structure functions of nuclei to that of the deuteron [3, 5, 6]. Additional constraints are especially valuable in the context of the intriguing question as to whether there is flavordependence to the EMC effect. Such flavor dependence has been conjectured in models of QCD [70-75] and in EFT [63–65] and is included in recent data-driven analyses of experimental results [76, 77] and provides a potential explanation of the NuTeV anomaly in $\sin^2 \theta_W$ [78].

Fig. 3 shows the constraint on the isovector momentum fraction ratio for ³He obtained from the results presented here, compared with the constraints on the isovector and isoscalar momentum fraction ratios from the recent nNNPDF2.0 [31] global nuclear PDF fits. The nNNPDF2.0 ellipse is generated by combining the Monte Carlo replica sets for the bound proton PDFs in ⁴He appropriately to form the PDFs of ³He (under the as-



FIG. 3. The ratio of the isovector momentum fractions of ³He and *p* determined in this work compared to constraints on the isovector and isoscalar momentum fraction ratios from the nNNPDF2.0 [31] global analysis before and after the LQCD constraint is imposed. Both axes are normalized to unity in the absence of nuclear effects. The LQCD constraint on the isovector ratio at $m_{\pi} = 806$ MeV is also displayed. In all cases, 68% confidence intervals are shown.

sumption that the nuclear effects vary slowly with A). In this way, correlations between the ³He and proton PDFs are accounted for. For the isovector combination, the 68% confidence interval is $3\langle x \rangle_{u-d}^{(^{3}\text{He})} / \langle x \rangle_{u-d}^{(p)}|_{\text{nNNPDF2.0}} = 1.007(63)$. In the nNNPDF approach, it is also straightforward to impose the LQCD constraint on the nuclear PDFs by reweighting the Monte Carlo replicas as discussed in Ref. [79]; the combined confidence region is shown in Fig. 3. The 68% confidence interval reduces to $3\langle x \rangle_{u-d}^{(^{3}\text{He})} / \langle x \rangle_{u-d}^{(p)} |_{\text{nNNPDF2.0+LQCD}} = 1.028(25).$ Fig. 4 compares the ratio of the isovector PDF for ³He to that of the constituent nucleons, with and without the imposition of the LQCD constraint. As can be seen from the reduced uncertainties in Figs. 3 and 4, LQCD calculations such as those presented here, as well as new experimental constraints [80, 81], can significantly improve our knowledge of the flavor dependence of nuclear PDFs.

Summary — In this work, the isovector momentum fractions of the proton, diproton and ³He systems have been determined using LQCD, complementing a previous study of the gluon momentum fraction on the same ensemble [45]. These calculations were performed at a single set of unphysical SU(3)-symmetric values for the quark masses corresponding to $m_{\pi} = 806$ MeV, and in a single lattice volume and at a single lattice spacing. Bearing these caveats in mind, the isovector nuclear momentum fractions were calculated precisely and found to



FIG. 4. The ratio $R^{(^{3}\text{He})}(x) = 3q_{3}^{(^{3}\text{He})}(x)/q_{3}^{(p)}(x)$ of the nNNPDF2.0 isovector PDF in ³He to that in the proton [31], as well as the same distribution with the LQCD moment constraint imposed into the global analysis as described in the text. 68% confidence intervals are shown.

be similar to that of the proton. In particular, the ratios $\langle x \rangle_{u-d}^{(pp)} / \langle x \rangle_{u-d}^{(p)} = 1.010(14)$ and $3 \langle x \rangle_{u-d}^{(^3\text{He})} / \langle x \rangle_{u-d}^{(p)} = 1.029(15)$ were determined and nuclear EFT arguments were used to connect the ³He result to global analyses of nuclear PDFs, providing important constraints on the flavor decomposition of nuclear PDFs that are complementary to those obtained from experiment.

While in its early stages, this work emphasizes the utility of LQCD in constraining less well-measured aspects of partonic structure in an analogous way to how LQCD inputs have been used to constrain the proton transversity PDFs [82]. Future calculations at the physical quark masses will consider higher moments of nuclear PDFs (or even directly study their x dependence) for a wider range of nuclei and provide a complete flavor decomposition. Calculations will also quantitatively address the full set of systematic uncertainties.

Acknowledgements — We thank Juan Rojo for comments and substantial help with the comparison to the nNNPDF2.0 global fits, and Silas Beane, Jiunn-Wei Chen, Zohreh Davoudi, Assumpta Parreño, Martin Savage and Brian Tiburzi for insightful discussions. This research used 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 number DE-AC05-00OR22725, as well as facilities of the USQCD Collaboration, which are funded by the Office of Science of the U.S. Department of Energy. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of EnergyOffice of Science User Facility operated under Contract No. DE-AC02-05CH11231. The Chroma [83], Qlua [84], QUDA [85, 86], QDP-JIT [87] and QPhiX [88] software libraries were used in data production and analvsis. WD, DJM and PES acknowledge support from the

U.S. DOE grant DE-SC0011090. WD is also supported within the framework of the TMD Topical Collaboration of the U.S. DOE Office of Nuclear Physics, and by the SciDAC4 award DE-SC0018121. PES is additionally supported by the National Science Foundation under CA-REER Award 1841699 and under EAGER grant 2035015. by the U.S. DOE Early Career Award DE-SC0021006, by a NEC research award, by the Carl G and Shirley Sontheimer Research Fund. MI is supported by the Universitat de Barcelona through the scholarship APIF, by the Spanish Ministerio de Economía y Competitividad (MINECO) under the project No. MDM-2014-0369 of ICCUB (Unidad de Excelencia María de Maeztu) and with additional European FEDER funds under the contract FIS2017-87534-P. KO was supported in part by U.S. DOE grant DE-FG02-04ER41302 and in part by the Jefferson Science Associates, LLC under U.S. DOE Contract DE-AC05-06OR23177. This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The authors thank Robert Edwards, Bálint Joó, and members of the NPLQCD collaboration for generating and allowing access to the ensembles used in this study, as well as for helpful discussions.

- J. J. Aubert *et al.* (European Muon), Phys. Lett. **B123**, 275 (1983).
- [2] M. Arneodo, Phys. Rept. 240, 301 (1994).
- [3] D. F. Geesaman, K. Saito, and A. W. Thomas, Ann. Rev. Nucl. Part. Sci. 45, 337 (1995).
- [4] G. Piller and W. Weise, Phys. Rept. **330**, 1 (2000).
- [5] P. R. Norton, Rept. Prog. Phys. 66, 1253 (2003).
- [6] O. Hen, D. W. Higinbotham, G. A. Miller, E. Piasetzky, and L. B. Weinstein, Int. J. Mod. Phys. E 22, 1330017 (2013).
- [7] N. Fomin, J. Arrington, R. Asaturyan, F. Benmokhtar, W. Boeglin, P. Bosted, A. Bruell, M. H. S. Bukhari, M. E. Christy, E. Chudakov, *et al.*, Phys. Rev. Lett. **108**, 092502 (2012).
- [8] O. Hen, E. Piasetzky, and L. B. Weinstein, Phys. Rev. C85, 047301 (2012), arXiv:1202.3452 [nucl-ex].
- [9] L. L. Frankfurt, M. I. Strikman, D. B. Day, and M. Sargsian, Phys. Rev. C48, 2451 (1993).
- [10] K. S. Egiyan *et al.* (CLAS), Phys. Rev. C68, 014313 (2003), arXiv:nucl-ex/0301008 [nucl-ex].
- [11] K. S. Egiyan *et al.* (CLAS Collaboration), Phys. Rev. Lett. **96**, 082501 (2006).
- [12] L. B. Weinstein, E. Piasetzky, D. W. Higinbotham, J. Gomez, O. Hen, and R. Shneor, Phys. Rev. Lett. 106, 052301 (2011).
- [13] O. Hen, E. Piasetzky, and L. B. Weinstein, Phys. Rev. C 85, 047301 (2012).
- [14] G. Martinelli and C. T. Sachrajda, Nucl. Phys. B306, 865 (1988).
- [15] G. Martinelli and C. T. Sachrajda, Nucl. Phys. B316, 355 (1989).

- [16] H.-W. Lin *et al.*, Prog. Part. Nucl. Phys. **100**, 107 (2018), arXiv:1711.07916 [hep-ph].
- [17] R. Horsley, R. Millo, Y. Nakamura, H. Perlt, D. Pleiter, P. E. L. Rakow, G. Schierholz, A. Schiller, F. Winter, and J. M. Zanotti (UKQCD, QCDSF), Phys. Lett. B714, 312 (2012), arXiv:1205.6410 [hep-lat].
- [18] C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, H. Panagopoulos, and C. Wiese, (2016), arXiv:1611.06901 [hep-lat].
- [19] Y.-B. Yang, M. Gong, J. Liang, H.-W. Lin, K.-F. Liu, D. Pefkou, and P. Shanahan, Phys. Rev. **D98**, 074506 (2018), arXiv:1805.00531 [hep-lat].
- [20] P. E. Shanahan and W. Detmold, Phys. Rev. D99, 014511 (2019), arXiv:1810.04626 [hep-lat].
- [21] P. E. Shanahan and W. Detmold, Phys. Rev. Lett. 122, 072003 (2019), arXiv:1810.07589 [nucl-th].
- [22] X. Ji, Phys. Rev. Lett. **110**, 262002 (2013), arXiv:1305.1539 [hep-ph].
- [23] K. Cichy and M. Constantinou, Adv. High Energy Phys. 2019, 3036904 (2019), arXiv:1811.07248 [hep-lat].
- [24] X. Ji, Y.-S. Liu, Y. Liu, J.-H. Zhang, and Y. Zhao, (2020), arXiv:2004.03543 [hep-ph].
- [25] P. Hagler, Phys. Rept. 490, 49 (2010), arXiv:0912.5483 [hep-lat].
- [26] B. Yoon, M. Engelhardt, R. Gupta, T. Bhattacharya, J. R. Green, B. U. Musch, J. W. Negele, A. V. Pochinsky, A. Schfer, and S. N. Syritsyn, Phys. Rev. D 96, 094508 (2017), arXiv:1706.03406 [hep-lat].
- [27] P. Shanahan, M. Wagman, and Y. Zhao, Phys. Rev. D 102, 014511 (2020), arXiv:2003.06063 [hep-lat].
- [28] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, A. Scapellato, and F. Steffens, (2020), arXiv:2008.10573 [hep-lat].
- [29] H.-W. Lin, (2020), arXiv:2008.12474 [hep-ph].
- [30] Q.-A. Zhang *et al.* (Lattice Parton), (2020), arXiv:2005.14572 [hep-lat].
- [31] R. Abdul Khalek, J. J. Ethier, J. Rojo, and G. van Weelden, (2020), arXiv:2006.14629 [hep-ph].
- [32] M. Lüscher and P. Weisz, Commun. Math. Phys. 97, 59 (1985), [Erratum: Commun. Math. Phys.98,433(1985)].
- [33] B. Sheikholeslami and R. Wohlert, Nucl. Phys. B259, 572 (1985).
- [34] S. R. Beane, E. Chang, S. D. Cohen, W. Detmold, H. W. Lin, T. C. Luu, K. Orginos, A. Parreño, M. J. Savage, and A. Walker-Loud (NPLQCD), Phys. Rev. D87, 034506 (2013), arXiv:1206.5219 [hep-lat].
- [35] M. L. Wagman, F. Winter, E. Chang, Z. Davoudi, W. Detmold, K. Orginos, M. J. Savage, and P. E. Shanahan, Phys. Rev. D 96, 114510 (2017), arXiv:1706.06550 [hep-lat].
- [36] S. R. Beane et al., (2017), arXiv:1705.09239 [hep-lat].
- [37] S. R. Beane *et al.* (NPLQCD), Phys. Rev. C88, 024003 (2013), arXiv:1301.5790 [hep-lat].
- [38] S. R. Beane, E. Chang, S. Cohen, W. Detmold, H. W. Lin, K. Orginos, A. Parreño, M. J. Savage, and B. C. Tiburzi, Phys. Rev. Lett. **113**, 252001 (2014), arXiv:1409.3556 [hep-lat].
- [39] S. R. Beane, E. Chang, W. Detmold, K. Orginos, A. Parreo, M. J. Savage, and B. C. Tiburzi (NPLQCD), Phys. Rev. Lett. **115**, 132001 (2015), arXiv:1505.02422 [hep-lat].
- [40] E. Chang, W. Detmold, K. Orginos, A. Parreño, M. J. Savage, B. C. Tiburzi, and S. R. Beane (NPLQCD), Phys. Rev. **D92**, 114502 (2015), arXiv:1506.05518 [hep-

lat].

- [41] W. Detmold, K. Orginos, A. Parreño, M. J. Savage, B. C. Tiburzi, S. R. Beane, and E. Chang, Phys. Rev. Lett. 116, 112301 (2016), arXiv:1508.05884 [hep-lat].
- [42] M. J. Savage, P. E. Shanahan, B. C. Tiburzi, M. L. Wagman, F. Winter, S. R. Beane, E. Chang, Z. Davoudi, W. Detmold, and K. Orginos, Phys. Rev. Lett. 119, 062002 (2017), arXiv:1610.04545 [hep-lat].
- [43] B. C. Tiburzi, M. L. Wagman, F. Winter, E. Chang, Z. Davoudi, W. Detmold, K. Orginos, M. J. Savage, and P. E. Shanahan, Phys. Rev. **D96**, 054505 (2017), arXiv:1702.02929 [hep-lat].
- [44] P. E. Shanahan, B. C. Tiburzi, M. L. Wagman, F. Winter, E. Chang, Z. Davoudi, W. Detmold, K. Orginos, and M. J. Savage, Phys. Rev. Lett. **119**, 062003 (2017), arXiv:1701.03456 [hep-lat].
- [45] F. Winter, W. Detmold, A. S. Gambhir, K. Orginos, M. J. Savage, P. E. Shanahan, and M. L. Wagman, Phys. Rev. D96, 094512 (2017), arXiv:1709.00395 [hep-lat].
- [46] E. Chang, Z. Davoudi, W. Detmold, A. S. Gambhir, K. Orginos, M. J. Savage, P. E. Shanahan, M. L. Wagman, and F. Winter (NPLQCD), Phys. Rev. Lett. 120, 152002 (2018), arXiv:1712.03221 [hep-lat].
- [47] M. Lüscher, Commun. Math. Phys. 104, 177 (1986).
- [48] M. Lüscher, Commun. Math. Phys. **105**, 153 (1986).
- [49] M. Lüscher, Nucl. Phys. B **354**, 531 (1991).
- [50] S. Beane, P. Bedaque, A. Parreño, and M. Savage, Phys. Lett. B 585, 106 (2004), arXiv:hep-lat/0312004.
- [51] Z. Davoudi and M. J. Savage, Phys. Rev. D 84, 114502 (2011), arXiv:1108.5371 [hep-lat].
- [52] S. Knig and D. Lee, Phys. Lett. B 779, 9 (2018), arXiv:1701.00279 [hep-lat].
- [53] R. A. Briceo, M. T. Hansen, and A. W. Jackura, Phys. Rev. D 100, 114505 (2019), arXiv:1909.10357 [hep-lat].
- [54] M. Gockeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. E. L. Rakow, G. Schierholz, and A. Schiller, Phys. Rev. D54, 5705 (1996), arXiv:hep-lat/9602029 [hep-lat].
- [55] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, Nucl. Phys. B 445, 81 (1995), arXiv:heplat/9411010.
- [56] J. Gracey, Nucl. Phys. B 667, 242 (2003), arXiv:hepph/0306163.
- [57] J. Gracey, Nucl. Phys. B 662, 247 (2003), arXiv:hepph/0304113.
- [58] C. Bouchard, C. C. Chang, T. Kurth, K. Orginos, and A. Walker-Loud, Phys. Rev. **D96**, 014504 (2017), arXiv:1612.06963 [hep-lat].
- [59] M. Albanese et al. (APE), Phys. Lett. B 192, 163 (1987).
- [60] W. Detmold and K. Orginos, Phys. Rev. D87, 114512 (2013), arXiv:1207.1452 [hep-lat].
- [61] S. Beane et al., (2020), arXiv:2003.12130 [hep-lat].
- [62] D. Dolgov *et al.*, Nucl. Phys. B Proc. Suppl. **94**, 303 (2001), arXiv:hep-lat/0011010.
- [63] J.-W. Chen and W. Detmold, Phys. Lett. B 625, 165 (2005).
- [64] J.-W. Chen, W. Detmold, J. E. Lynn, and A. Schwenk, Phys. Rev. Lett. **119**, 262502 (2017), arXiv:1607.03065 [hep-ph].
- [65] J. Lynn, D. Lonardoni, J. Carlson, J. Chen, W. Detmold, S. Gandolfi, and A. Schwenk, J. Phys. G 47, 045109 (2020), arXiv:1903.12587 [nucl-th].
- [66] W. Detmold, W. Melnitchouk, J. W. Negele, D. B. Renner, and A. W. Thomas, Phys. Rev. Lett. 87, 172001 (2001), arXiv:hep-lat/0103006.

- [67] D. Arndt and M. J. Savage, Nucl. Phys. A 697, 429 (2002), arXiv:nucl-th/0105045.
- [68] J.-W. Chen and X.-d. Ji, Phys. Lett. B 523, 107 (2001), arXiv:hep-ph/0105197.
- [69] NPLQCD collaboration (2020) in progress.
- [70] T. Uchiyama and K. Saito, Phys. Rev. C38, 2245 (1988).
- [71] I. R. Afnan, F. R. P. Bissey, J. Gomez, A. T. Katramatou, W. Melnitchouk, G. G. Petratos, and A. W. Thomas, Phys. Lett. **B493**, 36 (2000), arXiv:nucl-th/0006003 [nucl-th].
- [72] K. Saito, C. Boros, K. Tsushima, F. R. P. Bissey, I. R. Afnan, and A. W. Thomas, Phys. Lett. B493, 288 (2000), arXiv:nucl-th/0008017 [nucl-th].
- [73] W. Bentz, I. C. Cloet, J. T. Londergan, and A. W. Thomas, Phys. Lett. B693, 462 (2010), arXiv:0908.3198 [nucl-th].
- [74] I. C. Cloet, W. Bentz, and A. W. Thomas, Phys. Rev. Lett. 109, 182301 (2012), arXiv:1202.6401 [nucl-th].
- [75] A. J. Tropiano, J. J. Ethier, W. Melnitchouk, and N. Sato, Phys. Rev. C99, 035201 (2019), arXiv:1811.07668 [nucl-th].
- [76] B. Schmookler *et al.* (CLAS), Nature **566**, 354 (2019).
- [77] E. P. Segarra, A. Schmidt, D. W. Higinbotham, E. Piasetzky, M. Strikman, L. B. Weinstein, and O. Hen, (2019), arXiv:1908.02223 [nucl-th].
- [78] I. C. Cloet, W. Bentz, and A. W. Thomas, Phys. Rev. Lett. **102**, 252301 (2009), arXiv:0901.3559 [nucl-th].
- [79] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, Nucl. Phys. B 855, 608 (2012), arXiv:1108.1758 [hep-ph].
- [80] J. Mousseau *et al.* (MINERvA), Phys. Rev. **D93**, 071101 (2016), arXiv:1601.06313 [hep-ex].
- [81] G. G. Petratos *et al.*, Jefferson Lab PAC37 Proposal experiment E12-10-103, 2010.
- [82] H.-W. Lin, W. Melnitchouk, A. Prokudin, N. Sato, and H. Shows, Phys. Rev. Lett. **120**, 152502 (2018), arXiv:1710.09858 [hep-ph].
- [83] R. G. Edwards and B. Joó (SciDAC, LHPC, UKQCD), Lattice field theory. Proceedings, 22nd International Symposium, Lattice 2004, Batavia, USA, June 21-26, 2004, Nucl. Phys. Proc. Suppl. 140, 832 (2005), [,832(2004)], arXiv:hep-lat/0409003 [hep-lat].
- [84] A. Pochinsky, "Qlua, https://usqcd.lns.mit.edu/qlua.".
- [85] M. Clark, R. Babich, K. Barros, R. Brower, and C. Rebbi, Comput. Phys. Commun. 181, 1517 (2010), arXiv:0911.3191 [hep-lat].
- [86] R. Babich, M. A. Clark, and B. Joo, in SC 10 (Supercomputing 2010) (2010) arXiv:1011.0024 [hep-lat].
- [87] F. Winter, PoS LATTICE2013, 042 (2014).
- [88] B. Joó, D. D. Kalamkar, T. Kurth, K. Vaidyanathan, and A. Walden, in *High Performance Computing*, edited by M. Taufer, B. Mohr, and J. M. Kunkel (Springer International Publishing, Cham, 2016) pp. 415–427.
- [89] O. Ledoit and M. Wolf, Journal of Multivariate Analysis 88, 365 (2004).
- [90] E. Rinaldi, S. Syritsyn, M. L. Wagman, M. I. Buchoff, C. Schroeder, and J. Wasem, Phys. Rev. **D99**, 074510 (2019), arXiv:1901.07519 [hep-lat].
- [91] G. Golub and V. Pereyra, Inverse Problems 19, R1 (2003).
- [92] B. W. R. Dianne P. O'Leary, Computational Optimization and Applications 54, 579 (2013).
- [93] A. C. Davison and D. V. Hinkley, *Bootstrap Methods and*

their Application, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, 1997).

- [94] W. I. Jay and E. T. Neil, (2020), arXiv:2008.01069 [stat.ME].
- [95] B. Blossier, P. Boucaud, M. Brinet, F. De Soto, Z. Liu, V. Morenas, O. Pene, K. Petrov, and J. Rodriguez-Quintero, Phys. Rev. D 83, 074506 (2011), arXiv:1011.2414 [hep-ph].

Supplementary Material

MATRIX ELEMENT CALCULATION

The Euclidean finite difference form of the twist-two operator that determines the quark momentum fraction is

$$\mathcal{T}_{\mu\nu}^{(\Box)}(x) = \frac{1}{4a} \left(\overline{q}(x) \tau_3 \gamma_{\{\nu} [U_{\mu\}}(x) q(x+\hat{\mu}) - U_{\mu\}}^{\dagger}(x-\hat{\mu}) q(x-\hat{\mu}) \right] - [\overline{q}(x+\hat{\mu}) U_{\{\mu}^{\dagger}(x) - \overline{q}(x-\hat{\mu}) U_{\{\mu}(x-\hat{\mu})] \tau_3 \gamma_{\nu\}} q(x) \right).$$
(S1)

The combination $\mathcal{T}^{(\Box)} = \frac{1}{\sqrt{2}}(\mathcal{T}^{(\Box)}_{33} - \mathcal{T}^{(\Box)}_{44})$, which belongs to the $\tau_1^{(3)}$ irreducible representation of the hypercubic group [54], is used in this work. To compute matrix elements of this operator, the compound propagator technique is generalized from that previously used in Refs. [42, 46]. The operator insertion point is used as a sequential source, and three-point correlation functions are formed by first calculating a (smeared) point-to-all quark propagator extending from the hadronic/nuclear source to the operator insertion point and subsequently calculating additional quark propagators from the operator insertion point to the sink. Since the finite difference form of the operator contains shifts, three different sequential inversions are utilized. Taking the appropriate linear combinations of displaced sources to implement $\mathcal{T}^{(\Box)}$ results in fixed-order background-field compound propagators that include the operator insertions throughout spacetime. These compound propagators are used to construct the two-point correlation functions in Eq. (4). The background-field two-point correlation functions $G_h(t; \lambda)$ have a spectral representation as a sum of exponentials. This, in turn, determines the full *t*-dependent form of the ratio of two-point correlation functions in zero and non-zero background fields, $\mathcal{R}_h(t)$ (defined in Eq. (5)) in terms of the eigenenergies, interpolating operator overlap factors, and ground- and excited-state matrix elements of the lattice operator in Eq. (S1). The ground-state matrix elements of interest for each nuclear system can thus be extracted by fitting LQCD results for $\mathcal{R}_h(t)$ to the form arising from the spectral representations.

Zero background field two-point correlation function have the spectral representation

$$G_h^{ss'}(t;\lambda=0) = \sum_n Z_n^s (Z_n^{s'})^* e^{-E_n t},$$
(S2)

where $\{s, s'\} \in \{S, P\}$ specifies the source and sink smearing, E_n is the energy of the *n*-th energy eigenstate, and Z_n^s is an overlap factor defined by $Z_n^s = \sqrt{V} \langle n | \chi_h^s(0) | 0 \rangle$, where $V = a^4 \prod_{\mu} L_{\mu}$ is the (dimensionful) lattice volume and L_{μ} is the extent of the lattice geometry in the μ direction. The corresponding background-field two-point functions at $\mathcal{O}(\lambda)$ have the spectral representation

$$G_{h}^{ss'}(t;\lambda)\Big|_{\mathcal{O}(\lambda)} = a \sum_{\tau=0}^{b} \sum_{n,m} Z_{n}^{s} (Z_{m}^{s'})^{*} e^{-E_{n}\tau} \left\langle n \middle| \tilde{T}^{(\Box)} \middle| m \right\rangle$$

$$= \sum_{n} Z_{n}^{s} (Z_{n}^{s'})^{*} t e^{-E_{n}t} \left\langle n \middle| \tilde{T}^{(\Box)} \middle| n \right\rangle$$

$$+ \sum_{n} \sum_{m \neq n} Z_{n}^{s} (Z_{m}^{s'})^{*} a \left(\frac{e^{-E_{n}t}}{1 - e^{(E_{n} - E_{m})a}} + \frac{e^{-E_{m}t}}{1 - e^{(E_{m} - E_{n})a}} \right) \left\langle n \middle| \tilde{T}^{(\Box)} \middle| m \right\rangle$$
 (S3)

where $\tilde{T}^{(\Box)} = \sum_{\mathbf{x}} T^{(\Box)}(\mathbf{x}, 0)$ and thermal effects from the finite temporal extent of the lattice have been neglected, see Refs. [42, 43, 58] for further discussions. The spectral representation for $\mathcal{R}_{h}^{ss'}(t)$ follows from inserting Eqs. (S2) and (S3) into Eq. (5) and can be expressed as

$$\mathcal{R}_{h}^{ss'}(t) = \sum_{n} \left\langle n \right| T^{(\Box)} \left| n \right\rangle Z_{n}^{s} (Z_{n}^{s'})^{*} \left(\frac{(t+a)e^{-E_{n}(t+a)}}{\sum_{k} Z_{k}^{s} (Z_{k}^{s'})^{*} e^{-E_{k}(t+a)}} - \frac{t e^{-E_{n}t}}{\sum_{k} Z_{k}^{s} (Z_{k}^{s'})^{*} e^{-E_{k}t}} \right) + \sum_{n} \mathcal{N}_{n}^{ss'} \left(\frac{e^{-E_{n}(t+a)}}{\sum_{k} Z_{k}^{s} (Z_{k}^{s'})^{*} e^{-E_{k}(t+a)}} - \frac{e^{-E_{n}t}}{\sum_{k} Z_{k}^{s} (Z_{k}^{s'})^{*} e^{-E_{k}t}} \right)$$
(S4)

where $\mathcal{N}_{n}^{ss'} = a \sum_{m \neq n} \langle n | T^{(\Box)} | m \rangle Z_{n}^{s}(Z_{m}^{s'})^{*}/(1 - e^{(E_{n} - E_{m})a}) + \langle m | T^{(\Box)} | n \rangle Z_{m}^{s}(Z_{n}^{s'})^{*}/(1 - e^{(E_{m} - E_{n})a})$ involves excited-state transition matrix elements as well as combinations of overlap factors not determined from fits to Eq. (S2). In order to extract the ground-state matrix elements of interest in this work, combined fits to $G_{h}^{ss'}(t;0)$ and $\mathcal{R}_{h}^{ss'}(t)$



FIG. S1. Left: Effective matrix elements (Eq. (6) of the main text), with best fit result (dark grey band) and final fit result (light grey band). Right: The results of all fits included in the weighted average to obtain the systematic and statistical uncertainties of the corresponding matrix element.

are used to fix the free parameters $\{\langle n | T^{(\Box)} | n \rangle, E_n, Z_n^s(Z_n^{s'})^*, \mathcal{N}_n^{ss'}\}$, for $ss' \in \{SS, SP\}$. It is noteworthy that if this spectral representation is truncated at $N_{\text{states}} = 1$, then the second sum in Eq. (S4) vanishes and $\mathcal{R}_h^{ss'}(t)$ is independent of $\mathcal{N}_0^{ss'}$. For $N_{\text{states}} > 1$, $\mathcal{R}_h^{ss'}(t)$ is similarly independent from a linear combination of the $\mathcal{N}_n^{ss'}$ that can be used to eliminate one redundant parameter, chosen here to be $\mathcal{N}_0^{ss'}$. Further, if the sums and finite-difference derivatives in Eq. (S3) and Eq. (S4) were replaced by continuum integrals and derivatives, then $\mathcal{R}_h^{ss'}(t)$ would be independent of $\mathcal{N}_n^{ss'}$, suggesting that fits might not be sensitive to $\mathcal{N}_n^{ss'}$ in practice if lattice artifacts are small. In this work, fits are performed both using Eq. (S4) and also neglecting these lattice artifacts by setting $\mathcal{N}_n^{ss'} = 0$ in Eq. (S4). The Aikiake Information Criterion (AIC) is used to select whether fits with or without these lattice artifacts are preferred for each choice of N_{states} and fit range that is considered. In all cases, fits without these lattice artifacts are preferred by the AIC and used for subsequent analysis.

Care must be taken in order to account for systematic uncertainties associated with the fit-range choice and



FIG. S2. The ratios of effective matrix elements, Eq. (S5), for pp and ³He to that in the proton. The gray bands show the reconstructed value of the matrix element ratio from the fits to the individual nuclear states.

the model selection (i.e., the number of states to retain in truncated approximations to the infinite sums of states in Eqs. (S2)-(S4)). In this work, an automated fitting procedure analogous to the procedure described in Ref. [61] is used to assess these systematic fitting uncertainties. Results for SS and SP interpolating operators are fit simultaneously and maximum source-sink separations $\{t_{\max}^{SS}, t_{\max}^{SP}\}$ are chosen so that the signal-to-noise ratio of each correlation function is always greater than a tolerance of 0.5. A minimum separation $t_{\min}^{\min} = 4$ is required for the spectral representation to be well-defined for three-point correlation functions for the action and operator considered here. Within these constraints, $N_{\text{max}} = 200$ choices of minimum separations for both sources $(t_{\text{min}}^{\text{SS}}, t_{\text{min}}^{\text{SP}})$ are chosen randomly. For each fit range, a one-state fit is performed simultaneously for both sources. A two-state fit is subsequently performed, and if the two-state fit results in an AIC score of $\Delta AIC \leq -0.5$, then the two-state fit is accepted. This procedure is repeated until an N_{state} -state fit is rejected, at which point an $(N_{\text{state}}-1)$ -state fit is used. For the correlation functions analyzed in this work, this procedure results in accepted fits with $N_{\text{state}} \in \{1, 2, 3\}$. Correlated χ^2 -minimization is used for all fits with covariance matrices regularized using optimal shrinkage [89, 90]. The χ^2 depends linearly on the overlap factors, so these parameters are solved for using variable-projection methods [91, 92]. The excited-state energy gaps are then determined using nonlinear optimization (Nelder-Mead and gradient-based Newton solvers from the Julia package Optim are used and verified to reproduce the same minimum energy gaps within an absolute tolerance of 10^{-5}). The fit is then repeated for $N_{\text{boot}} = 200$ bootstrap samples, and the marginalized parameter uncertainty on the ground-state matrix element is determined using rank-based bootstrap confidence interval estimation [93] to add robustness to outlier bootstrap samples. Various checks on the fit result are then performed: an uncorrelated fit must reproduce the result within a tolerance of 5-sigma, the bootstrap median must reproduce the mean within a tolerance of 2-sigma, and the goodness-of-fit must be below a tolerance of $\chi^2/N_{dof} < 2$. All fit results passing these criteria are included in an ensemble of acceptable fits. A weighted average of these acceptable fits is used to determine the final central value and total systematic uncertainty, where the (in principle arbitrary - see Ref. [94] for a discussion of this in a Bayesian framework) weights are taken to be the ratio of the *p*-value to the variance of each fit as in Ref. [90]. Final uncertainties are obtained by adding in quadrature the statistical uncertainty of the highest weight fit with the systematic uncertainty obtained from the weighted average. The fitting procedure is fully specified by the parameters and tolerance values above, as well as by a random seed for bootstrap resampling that is fixed to allow correlated ratios of matrix elements for different hadrons to be formed.

Fig. S1 shows summaries of the fits of $\mathcal{R}_h^{\text{3pt}}$ for each state studied in this work using the approach described above. It is also convenient to define the further ratio for a nuclear state h:

$$\overline{\mathcal{R}}_h(t) = \mathcal{R}_h(t) / \mathcal{R}_p(t) \tag{S5}$$

which determines the ratio of the nuclear momentum fractions to that of the proton. Fig. S2 shows results for $\overline{\mathcal{R}}_h(t)$ including results obtained by fitting $\mathcal{R}_h(t)$ and $\mathcal{R}_p(t)$ independently as described above, using correlated bootstrap resampling to determine ratios of the ground-state matrix elements from all successful pairs of fit ranges, and taking a weighted average of the results.

NONPERTURBATIVE RENORMALIZATION

An ensemble of $N_{\text{meas}} = 101$ field configurations is used to compute the nonperturbative renormalization of the operator studied in this work. The parameters of the ensemble match those used for the nuclear matrix element calculations in the main text, except they have a smaller volume of $L^3 \times T = 24^3 \times 24$.

The renormalization coefficient of the local operator $\mathcal{T}_{\mu\nu}$ in Eq. (2) is computed in the RI'MOM scheme [55] by equating the Landau-gauge dressed vertex function

$$\Gamma_{\mu\nu}(p) \equiv \frac{1}{N_C} \operatorname{Tr}_C \left[S^{-1}(p) G_{\mu\nu}(p) S^{-1}(p) \right]$$
(S6)

with its tree level value at a fixed momentum p. Here, S is the quark propagator and $G_{\mu\nu}$ is the quark three-point correlation function for the zero-momentum projected operator $\tilde{\mathcal{T}}_{\mu\nu}^{(\Box)}$, and the trace is over color degrees of freedom. Defining $\tilde{\mathcal{T}}_{\mu\nu}^{(\Box)} = \sum_{z} \mathcal{T}_{\mu\nu}^{(\Box)}(z) = \sum_{z,z'} \bar{q}(z) J_{\mu\nu}^{(\Box)}(z,z') q(z')$, this zero-momentum-projected three-point correlation function is

$$G_{\mu\nu}(p) = \frac{1}{V} \sum_{x,y,z,z'} e^{ip(x-y)} S(x,z) J^{(\Box)}_{\mu\nu}(z,z') S(z',y),$$
(S7)

which is computed with the sequential source technique applied through the operator [15]. At tree-level, the vertex function is proportional to two tensor structures [56]:

where $\tilde{p}_{\mu} \equiv \frac{2}{a} \sin\left(\frac{a}{2}p_{\mu}\right)$ is the lattice momentum corresponding to p_{μ} . In the continuum, only $\Lambda^{1}_{\mu\nu}$ appears, however $\Lambda^{2}_{\mu\nu}$ enters as an $\mathcal{O}(a)$ correction [56]. Expressing $\Gamma_{\mu\nu}$ in the space spanned by $\{\Lambda^{(1)}, \Lambda^{(2)}\}$ amounts to imposing the renormalization condition:

$$\Gamma_{\mu\nu}(p)|_{p^2=\mu_0^2} = \mathcal{Z}_q(p) \left[\mathcal{Z}_1^{-1}(p) \Lambda_{\mu\nu}^1(p) + \mathcal{Z}_2^{-1}(p) \Lambda_{\mu\nu}^2(p) \right] \Big|_{p^2=\mu_0^2} , \qquad (S9)$$

where $\mathcal{Z}_a^{-1}(p)$ for $a \in \{1, 2\}$ are operator renormalization coefficients and the quark field renormalization \mathcal{Z}_q is defined as:

$$\mathcal{Z}_{q}(p)|_{p^{2}=\mu_{0}^{2}} = \frac{i}{12\tilde{p}^{2}} \operatorname{Tr}\left[S^{-1}(p)\vec{p}\right]\Big|_{p^{2}=\mu_{0}^{2}}.$$
(S10)

The renormalization coefficient of primary interest is $\mathcal{Z}^{\text{RI'MOM}}(\mu_0, a) \equiv \mathcal{Z}_1(p)|_{p^2 = \mu_0^2}$.

Eq. (S9) is solved by constructing a linear functional on the space of Dirac and Lorentz matrices. A bilinear form $\langle \cdot, \cdot \rangle$ is defined whose action is:

$$\langle \lambda^1, \lambda^2 \rangle \equiv \sum_{i \in \tau_1^{(3)}} \operatorname{Tr}_{\mathcal{D}} \left\{ \lambda_i^1 \lambda_i^2 \right\},\tag{S11}$$

where λ^1, λ^2 are objects with two Dirac and two Lorentz indices and the sum runs over elements of a basis for the irreducible representation of the hypercubic group that contains \mathcal{T} , as defined in Eq. (2). The functionals $\langle \Lambda^1, \cdot \rangle$ and $\langle \Lambda^2, \cdot \rangle$ are applied to Eq. (S9) to yield a system of equations:

$$\mathcal{Z}_q(p) \sum_{b \in \{1,2\}} \langle \Lambda^a(p), \Lambda^b(p) \rangle \mathcal{Z}_b^{-1}(p) = \langle \Lambda^a(p), \Gamma(p) \rangle, \tag{S12}$$

which can be solved for $\mathcal{Z}_{1,2}(p)$.

An $\overline{\text{MS}}$ renormalization factor can be constructed from the RI'MOM factor computed as described above as

$$\mathcal{Z}^{\overline{\mathrm{MS}}}(\mu) = \mathcal{R}^{\overline{\mathrm{MS}}/\mathrm{RI'MOM}}(\mu,\mu_0)\mathcal{Z}^{\mathrm{RI'MOM}}(\mu_0,a),$$
(S13)

where the matching factor $\mathcal{R}^{\overline{\text{MS}}/\text{RI'MOM}}(\mu,\mu_0)$ has been computed to 3-loop order in lattice perturbation theory [56, 57] (using the two-loop value produces a statistically indistinguishable result). While $\mathcal{Z}^{\overline{\text{MS}}}(\mu)$ defined in Eq. (S13) is



FIG. S3. Renormalization coefficients in the $\overline{\text{MS}}$ scheme, color-coded by the amount of hypercubic breaking at each point. $\mathcal{Z}^{\overline{\text{MS}}}$ is computed on each ensemble for each mode p_{μ} with $0.5 \leq (ap)^2 \leq 10$, and averaged over hypercubic orbits.

in principle independent of the matching scale $p^2 = \mu_0^2$, in practice there are discretization artifacts which arise as contamination in the form of dependence on the hypercubic invariants $p^{[2n]} \equiv \sum_{\mu} p_{\mu}^{2n}$, resulting in the classic jellyfishbone structure shown in Fig. S3. \tilde{p}_{μ} can be expanded in a Taylor series in $p^{[2n]}$, so either $p^{[2n]}$ or $\tilde{p}^{[2n]}$ may be used to perform the fit; $p^{[2n]}$ is used for this analysis, but consistent results are obtained by fitting to $\tilde{p}^{[2n]}$.

To obtain the renormalization factor, the hypercubic artifacts which depend on $p^{[2n]}$ for n > 1 are fit using the one-window-fit approach, detailed in Ref. [95]. Running terms and remaining artifacts which depend only on p^2 are then fit separately. To allow a quantification of systematic uncertainties, the fits are performed over a range of windows, with a range of functional forms. In particular, artifacts of the form

$$\left\{H_1, H_2, H_3\right\} = \left\{c_1 \frac{a^2 p^{[4]}}{p^2} + c_2 \frac{a^2 p^{[6]}}{(p^2)^2}, c_3 \frac{a^2 p^{[4]}}{p^2} \log(a^2 p^2), c_4 a^4 p^{[4]} + c_5 \frac{a^4 p^{[6]}}{p^2} + c_6 a^4 \left(\frac{p^{[4]}}{p^2}\right)^2\right\}$$
(S14)

$$\{R_1, R_2, R_3, R_4\} = \left\{ d_1 a^2 p^2, \frac{d_2}{a^2 p^2}, d_3 a^2 p^2 \log(a^2 p^2), d_4 \left(a^2 p^2\right)^2 \right\},\tag{S15}$$

are considered, where the H_i denote hypercubic terms and the R_i are contributions from running. Terms are grouped by their order in a^2 and by whether or not they contain logarithmic corrections. The full functional form is truncated at order a^4 . Adding further logarithmic terms into the functional form does not increase the fit quality. A fit form F, constructed from these components, is chosen for a specific data window to maximize the goodness of fit while preventing overfitting using the following procedure. Given a window of momenta, the first fit form is initialized to be $F^{(1)} = 0$. Given a fit form $F^{(n)}$, the subsequent form $F^{(n+1)}$ is determined by considering all possible forms

$$F_j^{(n+1)} = F^{(n)} + X_j, (S16)$$

which can be built from $F^{(n)}$ using only the terms X_j which are not currently present in $F^{(n)}$, and $X \in \{H, R\}$ is as appropriate for the fit to the hypercubic or running artifacts. A fit form $F_j^{(n+1)}$ is accepted if and only if $A(F_j^{(n+1)}) < A(F_j^n)$, where $A(F) \equiv 2N_{\text{param}}(F) + \chi^2(F)$ is the value of the AIC of the fit F. If no forms are accepted or there are no fit forms left, iteration stops and $F \equiv F^{(n)}$. The subsequent fit form $F^{(n+1)}$ is chosen out of the accepted fit forms $\{F_j^{(n+1)}\}_j$ by maximizing the *p*-value of the fit.

This procedure is applied to a range of fitting windows. Windows are chosen with $(ap)^2 \in [(ap)^2_{\min}, (ap)^2_{\max}]$ and $p^{[4]}/(p^2)^2 \leq h_0$ by independently adjusting $(ap)^2_{\min}$ and $(ap)^2_{\max}$ between 0.5 and 10 in increments of 0.5, always keeping a window size $(ap)^2_{\max} - (ap)^2_{\min} \geq 5$, and taking $h_0 \in \{0.5, 0.6, ..., 1.0\}$. For each fit window f, the optimal fit form as defined above is accepted if and only if its p value $p_f \geq 0.01$, and each accepted fit is given a weight:

$$w_f \equiv p_f (\delta \mathcal{Z}_f^{\rm MS})^{-2}, \tag{S17}$$

where $\delta Z_f^{\overline{\text{MS}}}$ denotes the statistical uncertainty on $Z_f^{\overline{\text{MS}}}$, which is the fit result for $Z^{\overline{\text{MS}}}$ that is obtained in fit f. The weights are normalized so that the maximum weight is 1. Given the large number of accepted fits, the $N_w \equiv 100$



FIG. S4. $N_w = 100$ highest-weight accepted fits, Z_f , plotted as a function of their fit weight, w_f . Individual fits are color-coded by the maximum amount of hypercubic breaking in the window, and the gray band shows the final weighted average.

highest-weight fits are combined in a weighted average, with the result:

$$\mathcal{Z}^{MS}(\mu = 2 \text{ GeV}) = 0.885(42).$$
 (S18)

The final result and uncertainties are consistent under variation of N_w , with larger systematic uncertainties when additional fits with lower weights are included (e.g., with $N_w = 200$, the result is 0.895(55)). As a consistency check, the computation was also performed on an ensemble with identical action and parameters with lattice dimensions $16^3 \times 48$. A result of $\mathcal{Z}^{\overline{\text{MS}}}(\mu = 2 \text{ GeV}) = 0.910(57)$ is obtained, indicating finite volume effects are negligible within uncertainties.

CONNECTION TO PHENOMENOLOGY: MOMENTUM FRACTION EXTRAPOLATION

In order to connect to phenomenology, the lattice result for $\langle x \rangle_{u-d}^{(^{3}\text{He})}$ at quark masses corresponding to $m_{\pi} = 806$ MeV is extrapolated to the physical masses using the assumption of weak mass dependence of the short-distance two-body counterterms in nuclear EFT. The isovector twist-two operators in Eq. (1) match on to hadronic operators in nuclear EFT as [63]

$$\mathcal{O}_{\mu_1\dots\mu_n} \to \langle x^n \rangle_{u-d}^{(p)} v_{\mu_1}\dots v_{\mu_n} N^{\dagger} \tau_3 N (1 + \alpha_{3,n} N^{\dagger} N) + \dots$$
(S19)

where N is a nucleon field, v is the nuclear velocity, and the ellipsis denotes higher order nucleonic and pionic operators. Defining the nuclear factor $\mathcal{G}_3(N,Z) = \langle N, Z | N^{\dagger} \tau_3 N N^{\dagger} N | N, Z \rangle$, the two-nucleon counterterm $\alpha_{3,2}$ relates the nuclear and proton momentum fractions as [63]

$$\alpha_{3,2}\mathcal{G}_3(N,Z) \equiv \langle x \rangle_{u-d}^{(N,Z)} - \frac{Z-N}{A} \langle x \rangle_{u-d}^{(p)}.$$
(S20)

This LEC can be determined from the LQCD results for 3 He most precisely by re-expressing it in terms of the quantities in Table I as

$$\alpha_{3,2}\mathcal{G}_3({}^{3}\mathrm{He}) = \frac{1}{3} \left(3 \frac{\langle x \rangle_{u-d}^{({}^{3}\mathrm{He})}}{\langle x \rangle_{u-d}^{(p)}} - 1 \right) \langle x \rangle_{u-d}^{(p)}.$$
 (S21)

Since renormalization effects cancel in the ratio, it is more precisely determined than the individual momentum fractions themselves and, with a naive error propagation, computing $\alpha_{3,2}\mathcal{G}_3(^3\text{He})$ via Eq. (S21) rather than Eq. (S20) achieves smaller uncertainties.

The matching of the LEC, and extrapolation to physical quark masses proceeds in the following steps:

7

- 1. Determine the counterterm $\alpha_{3,2}\mathcal{G}_3({}^{3}\text{He})$ at $m_{\pi} = 806$ MeV via Eq. (S21), using the LQCD calculations of the ratio of $\langle x \rangle_{u-d}^{({}^{3}\text{He})} / \langle x \rangle_{u-d}^{(p)}$, and $\langle x \rangle_{u-d}^{(p)}$ itself.
- 2. While the momentum fractions themselves have nonanalytic quark mass dependence, the counterterms $\alpha_{3,2}$ and nuclear factors $\mathcal{G}_3(N, Z)$ are expected to have only mild quark mass dependence and so their their combination is extrapolated to the physical quark masses by assuming the same central value and increasing the uncertainty by 50%.
- 3. The extrapolated value of $\alpha_{3,2}\mathcal{G}_3({}^{3}\text{He})$ is combined with the value of $\langle x \rangle_{u-d}^{(p)}$ from phenomenology to produce a physical-point value of $\langle x \rangle_{u-d}^{({}^{3}\text{He})} / \langle x \rangle_{u-d}^{(p)}$.