

Isospin dependence of the nuclear EMC effect from a global QCD analysis

C. Cocuzza¹,[✉] T. J. Hague,² W. Melnitchouk,² N. Sato,² and A. W. Thomas³

(JAM Collaboration—PDF Analysis Group)

¹*Department of Physics, William & Mary, Williamsburg, Virginia 23187, USA*

²*Jefferson Lab, Newport News, Virginia 23606, USA*

³*CSSM and CDMPP, Department of Physics, Adelaide University, Adelaide SA 5005, Australia*



(Received 20 February 2026; accepted 30 July 2026; published 2 September 2026)

We perform a new global QCD analysis of unpolarized parton distribution functions (PDFs) in the nucleon from proton, deuteron, and $A = 3$ data, including recent measurements of $^3\text{He}/D$ and $^3\text{H}/D$ cross section ratios from the MARATHON experiment at Jefferson Lab. Simultaneously inferring the PDFs and nucleon off-shell corrections allows both to be determined consistently, without theoretical assumptions about the isospin dependence of nuclear effects. The analysis provides strong evidence for the need of nucleon off-shell corrections to describe the $A = 3$ data, with large isoscalar and a suggestion of nonzero isovector contributions in $A \leq 3$ nuclei. We find that the extracted EMC ratios of nuclear to nucleon structure functions for $A = 2$ and 3 differ from those naively extrapolated from heavy nuclei down to low A .

DOI: [10.1103/trd9-gtnr](https://doi.org/10.1103/trd9-gtnr)

I. INTRODUCTION

Understanding the structure of nuclei directly in terms of the fundamental quark and gluon degrees of freedom of QCD theory remains one of the ultimate objectives of nuclear physics. The first evidence for a modification of the parton structure of protons and neutrons inside nuclei was seen over four decades ago with the famous “nuclear EMC effect” [1,2], which revealed nontrivial deviations of the ratio of nuclear to deuteron structure functions from unity. While a universally accepted quantitative explanation has remained elusive, some approaches have related the observed structure function modifications to the strong scalar and vector mean fields in nuclei [3–8], while others have attributed the medium modification of the nucleon structure to short-range correlations [9–13] (see Refs. [14–16] for reviews of the different approaches).

Regardless of the dynamics responsible for the medium modifications, generally the nuclear structure functions can be represented as a sum of free-nucleon structure functions, convoluted with nucleon smearing functions that take into account nuclear binding and Fermi motion, and contributions that depend on modifications of the parton structure of the bound nucleons in the nucleus [17]. The latter can be

formally represented in terms of nucleon off-shell corrections [18–25], which can be either estimated in quark models or parametrized and extracted from experiment within a given theoretical framework.

While the general features of the nuclear EMC effect for heavy nuclei can be accommodated within more than one approach, it has been difficult to obtain unique experimental signatures that could potentially distinguish between different models. On the other hand, it has been noted that predictions for more specific aspects of the nuclear modifications, such as their spin and flavor or isospin dependence, can differ markedly depending on the dynamical mechanism assumed responsible for the modifications [16]. A particularly useful testing ground for studying the isospin dependence of the nuclear modifications are the light nuclei, such as the deuteron and the trinucleons, ^3He and ^3H , for which quantitative few-body calculations are more readily accessible than for nuclei with large mass number, A .

In fact, over 25 years ago Afnan *et al.* [26,27] (see also Refs. [28,29]) suggested that deep-inelastic scattering (DIS) from the ^3He and ^3H mirror nuclei could be a way to constrain the d/u quark distribution function ratio in the proton, assuming that the nuclear effects in the $A = 3$ nuclei mostly cancel. The availability of the high-precision data from the MARATHON experiment at Jefferson Lab, firstly on the ratio of $^3\text{He}/^3\text{H}$ cross sections [30] and more recently on the individual $^3\text{He}/D$ and $^3\text{H}/D$ cross section ratios [31], has allowed this assumption to be tested for the first time. In particular, by performing a global QCD

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

analysis of the world's high-energy scattering data, in which the nucleon parton distribution function (PDF) and off-shell parameters were fitted simultaneously, Cocuzza *et al.* [32] found the first indications for an isospin-dependent nuclear EMC effect in light nuclei, suggesting an enhanced nuclear effect on the d -quark PDF in the bound proton. With the measurement of the individual ${}^3\text{He}/D$ and ${}^3\text{H}/D$ ratios [31], an additional experimental constraint has become available, which could in principle provide an even more robust determination of the isospin dependence of the nuclear (off-shell) corrections.

In the analysis of the MARATHON data in Ref. [31], the ${}^3\text{He}/D$ and ${}^3\text{H}/D$ cross section ratios were normalized under the assumption of the same universal shape around $x = 0.3$ for the nuclear to nucleon cross section ratios across all nuclei, from Pb down to deuterium, as in the model of Kulagin and Petti (KP) [25,33]. While the binding energies per nucleon for nuclei from ${}^4\text{He}$ to ${}^{208}\text{Pb}$ are known to be in the range (7–8) MeV, those for the light nuclei ${}^2\text{H}$, ${}^3\text{He}$, and ${}^3\text{H}$ are considerably smaller, $\lesssim 3$ MeV, so that extrapolating the dynamics of heavy nuclear systems smoothly to $A \leq 3$ nuclei is a strong assumption.

As we demonstrate in the present work, this assumption introduces a significant bias into the analysis, leading to rather different conclusions about the nuclear EMC effect and the extracted neutron to proton structure function ratio, which was one of the goals of the MARATHON experiment. In contrast, we avoid assumptions about the shape and magnitude of the off-shell corrections in the $A \leq 3$ nuclei and allow these to be determined from the global QCD analysis. Consistent with the previous JAM analysis [32] of global data including the ${}^3\text{He}/{}^3\text{H}$ ratio, we find strong evidence for sizable off-shell contributions to the nuclear structure functions, and a flavor dependence of the EMC ratios in $A = 3$ nuclei.

In this paper we present the results of a global QCD analysis of unpolarized PDFs from the world's high-energy scattering data on protons, deuterons, and $A = 3$ nuclei using the JAM Bayesian Monte Carlo framework. In particular, we simultaneously fit the free-nucleon PDF and nucleon off-shell parameters, with the latter constrained by the MARATHON data on the D/p , ${}^3\text{He}/D$, and ${}^3\text{H}/D$ structure function ratios, along with other high-energy scattering data involving $A = 2$ and $A = 3$ nuclei. We begin in Sec. II with an overview of the theoretical framework employed in this analysis, focusing in particular on the formulation of nuclear structure functions in terms of convolutions of free-nucleon structure functions and nucleon off-shell corrections, along with some details of the methodology used in the analysis. In Sec. III we discuss the quality of the global fit to all the data, paying particular attention to the relative normalizations of the various datasets used in the analysis. The results for the PDFs and off-shell functions from the QCD analysis are

presented in Sec. IV, along with those for ratios of nuclear structure functions determined from the global fit. We also contrast our results with those obtained from the analysis in Ref. [31] that imposes the KP model [25,33] as the reference point for the MARATHON data normalization. Finally, in Sec. V we summarize our findings and outline future avenues for improving our understanding of the relation between PDFs in free nucleons and those bound inside nuclei.

II. THEORETICAL FRAMEWORK

In this section we outline the theoretical framework used in this analysis, focusing on the essential features related to the description of nuclear structure functions in lepton-nucleus DIS. We review the standard convolution formulas for the nuclear structure functions in the nuclear impulse approximation, and discuss in detail the nucleon off-shell contributions in terms of their isoscalar and isovector components. We further summarize the core aspects of the Bayesian Monte Carlo methodology employed in JAM global QCD analyses, along with details of the parametrizations employed for the various parton distributions.

A. Nuclear structure functions

In the nuclear impulse (or weak nuclear binding) approximation, DIS from a nucleus A (with 4-momentum P_A) takes place through the scattering of an exchanged virtual boson (4-momentum q) incoherently from individual off-shell nucleons (4-momentum p) in the nucleus [18–24]. In this case the nuclear F_2^A structure function can be written in convolution form as a sum of on-shell and off-shell nucleon contributions [17,21,22,25,32,33],

$$F_2^A(x_B, Q^2) = \sum_N [f_{N/A}^{(\text{on})} \otimes F_2^N + f_{N/A}^{(\text{off})} \otimes \delta F_2^{N/A}](x_B, Q^2), \quad (1)$$

where $x_B = (M_A/M)Q^2/(2P_A \cdot q)$ is the Bjorken scaling variable, scaled by the ratio of nuclear (M_A) to nucleon (M) masses to be in the range $x_B \in [0, M_A/M] \approx [0, A]$, and Q^2 is the 4-momentum squared of the exchanged boson.

The functions $f_{N/A}^{(\text{on})}$ and $f_{N/A}^{(\text{off})}$ are the on-shell and off-shell smearing functions of nucleons N in nucleus A [17], and F_2^N and $\delta F_2^{N/A}$ are the on-shell and off-shell nucleon structure functions, respectively. The symbol $\otimes$ in Eq. (1) represents the convolution

$$[f_{N/A} \otimes F_2](x_B, Q^2) \equiv \int_{x_B}^{M_A/M} dy f_{N/A}(y, \rho) F_2\left(\frac{x_B}{y}, Q^2\right), \quad (2)$$

where $y = (p_0 + \rho p_z)/M$ is the light-cone fraction of the nuclear momentum carried by the interacting nucleon, and

p_0 and p_z are the nucleon's energy and longitudinal momentum, respectively. The kinematical parameter $\rho^2 = 1 + 4M^2 x_B^2 / Q^2$ takes into account finite energy effects. Finally, we note that while Eqs. (1) and (2) are valid for all nuclei A , in practice we will consider only the specific cases of $A = 2$ (deuteron) and $A = 3$ (^3He and ^3H).

In QCD collinear factorization the nucleon structure functions can be written as convolutions of PDFs with the hard scattering coefficients, C_q ,

$$F_2^N(x_B, Q^2) = x_B \sum_q e_q^2 [C_q \otimes q_N^+](x_B, Q^2) + \dots, \quad (3)$$

where $q_N^+ = q_N + \bar{q}_N$ is the C -even PDF in the nucleon, and the ellipsis includes higher-order contributions from gluons, as well as power corrections associated with target mass effects and higher twist (HT) multiparton contributions. Note that here and in the following the PDFs q_N^+ are functions of the momentum fraction, x , of the nucleon carried by the parton, which at leading order coincides with the Bjorken scaling variable, but differs at higher orders. The convolution in this case is given by

$$[C_q \otimes q_N^+](x_B, Q^2) \equiv \int_{x_B}^1 \frac{dx}{x} C_q(x) q_N^+\left(\frac{x_B}{x}, Q^2\right). \quad (4)$$

When referring to the free nucleon in Eq. (4) and elsewhere, it is understood that x_B refers to the corresponding free-nucleon ($A = 1$) scaling variable.

Analogously, we can write the off-shell nucleon structure functions in the nucleus in factorized form in terms of the off-shell PDFs $\delta q_{N/A}$ as

$$\delta F_2^{N/A}(x_B, Q^2) = x_B \sum_q e_q^2 [C_q \otimes \delta q_{N/A}](x_B, Q^2) + \dots \quad (5)$$

Similar expressions exist for the longitudinal F_L , transverse F_1 , and parity-odd F_3 structure functions [17,25]. The nuclear convolution expression in Eq. (1) is obtained from the general virtual photon-off-shell nucleon scattering amplitude by Taylor expanding the bound nucleon structure function around the on-shell limit, in terms of the nucleon virtuality $v(p^2) \equiv (p^2 - M^2)/M^2$, where p^2 is the bound nucleon's 4-momentum squared. Our analysis will focus on the large- x_B region, where the structure functions are dominated by contributions from u and d quarks, and sea quark and gluon contributions to the off-shell functions can be effectively neglected.

In the leading twist approximation the nuclear structure function in Eq. (1) can also be written in terms of nuclear PDFs,

$$F_2^A(x_B, Q^2) = x_B \sum_q e_q^2 [C_q \otimes q_A^+](x_B, Q^2). \quad (6)$$

In the limit of exact charge symmetry, the u_A and d_A PDFs in the $A = 2$ and 3 nuclei are related to each other by

$$u_D = d_D, \quad (7)$$

and

$$u_{^3\text{He}} = d_{^3\text{H}}, \quad (8a)$$

$$d_{^3\text{He}} = u_{^3\text{H}}, \quad (8b)$$

respectively. In this case, the PDFs of a given flavor in different nuclei (for example, u_D , $u_{^3\text{He}}$, or $u_{^3\text{H}}$) are not related and are parametrized independently. Since in this work we wish to relate the free-nucleon PDFs to those in bound nucleons, we do not work with nuclear PDFs directly in Eq. (6), but use Eqs. (1)–(5) as the basis for the global analysis.

At finite Q^2 values the structure functions also receive contributions from target mass corrections [34,35], for which we use the collinear factorization framework [36,37] in order to consistently describe all of the high-energy scattering data and not just inclusive DIS. In addition, we allow for HT power corrections to the structure functions, for which we use an additive parametrization,

$$F_2^N(x_B, Q^2) = F_2^{N(\text{TMC})}(x_B, Q^2) + \frac{H^N(x_B)}{Q^2} + \mathcal{O}\left(\frac{1}{Q^4}\right), \quad (9)$$

where $F_2^{N(\text{TMC})}$ is the leading twist structure function with target mass corrections (TMCs) included, and $H^N(x_B)$ is the x_B -dependent HT function which will be parametrized in the numerical analysis.

In the literature, a multiplicative ansatz is sometimes used, in which the HT correction is multiplied by the leading twist contribution $F_2^{N(\text{TMC})}$. The advantage of an additive parametrization is that it allows a cleaner separation of the $1/Q^2$ power corrections in the HTs from the $\ln Q^2$ corrections in the leading twist structure functions, whereas a multiplicate form inevitably mixes the two behaviors. To compare with previous analyses, however, we also consider the multiplicative parametrization by modifying the HT function $H^N(x_B) \rightarrow H^N(x_B, Q^2) = H^N(x_B) \ln Q^2$. In this multiplicative parametrization, the evolution of H^N is determined by the leading twist evolution of $F_2^{N(\text{TMC})}$, for which the dependence on Q^2 scales with $\ln Q^2$, and which, therefore, allows one to mimic a multiplicative form.

We note also that while the HT contribution enters at the free-nucleon level, when applied to nuclei the HT part of the structure function is smeared by the nuclear light-cone momentum distribution, as in Eq. (2), in the same manner as the leading twist component. We do not include off-shell corrections to the HT term, however, since this doubly suppressed effect cannot be constrained by existing data.

B. Nucleon off-shell corrections

While the on-shell and off-shell smearing functions can be computed directly from known nuclear wave functions, the off-shell nucleon PDFs $\delta q_{N/A}$ are *a priori* unknown and need to be determined from experiment. In the absence of charge symmetry violation and Coulomb effects, the off-shell contribution from u quarks in a neutron inside a nucleus A is identical to the off-shell contribution from d quarks in a proton inside the corresponding mirror nucleus A^* (number of protons and neutrons interchanged), and similarly for d quarks in a neutron,

$$\delta u_{n/A} = \delta d_{p/A^*}, \quad \delta d_{n/A} = \delta u_{p/A^*}. \quad (10)$$

For the specific case of the deuteron D ,

$$\delta u_{n/D} = \delta d_{p/D}, \quad \delta d_{n/D} = \delta u_{p/D}, \quad (11)$$

while for the $A = 3$ nuclei ${}^3\text{He}$ and ${}^3\text{H}$, one has

$$\delta u_{n/{}^3\text{He}} = \delta d_{p/{}^3\text{H}}, \quad \delta d_{n/{}^3\text{He}} = \delta u_{p/{}^3\text{H}}, \quad (12a)$$

$$\delta u_{n/{}^3\text{H}} = \delta d_{p/{}^3\text{He}}, \quad \delta d_{n/{}^3\text{H}} = \delta u_{p/{}^3\text{He}}. \quad (12b)$$

To further classify the off-shell distributions, we can decompose the functions into contributions reflecting the nature of the interactions between the remnants of the struck nucleon and the rest of the nucleus during the scattering process. To illustrate this, it is instructive to consider the formal expression for the twist-two PDF of a nucleon at rest [38],

$$q(x) = \sqrt{2} \sum_k \delta(xM - (p^+ - p_k^+)) |\langle S_k(p_k) | \psi_+(0) | N(p) \rangle|^2, \quad (13)$$

where the state $|N(p)\rangle$ is a nucleon in the Lorentz scalar and vector fields generated by the rest of the nucleus, and $|S_k(p_k)\rangle$ are the intermediate (spectator) states with 4-momentum p_k and mass m_k left when a quark is removed by the operator ψ_+ [38]. Note that since DIS at leading twist involves the imaginary part of the virtual Compton scattering amplitude, the spectator states are necessarily on-shell. The key features of the valence quark PDFs of the free nucleons can be readily understood in terms of the contribution from the lowest-mass intermediate state, consisting of two valence quarks ($k = 2$), with mass m_2 and three-momentum $\mathbf{p}_2$ [39]. For ground state baryons, the most likely momentum is zero, $\mathbf{p}_2 = 0$, so the δ function in Eq. (13) will produce a peak in the $q(x)$ distribution from this lowest-mass intermediate state at $x = 1 - m_2/M$. The gluon hyperfine interaction, which lifts the mass of a pair of quarks with spin-1 by roughly 200 MeV above that for a spin-0 pair, then naturally explains, for example, why the d/u PDF ratio decreases rapidly as $x \rightarrow 1$.

In a nucleus, the off-shell corrections are expected to be dominated by the shifts in p^+ and p_2^+ resulting from their interaction with the scalar and vector fields generated by the rest of the nucleus. In a simple effective hadronic model these interactions may be considered as generated by the exchange of isoscalar σ and ω mesons and the isovector ρ meson. Labeling the potentials felt by each quark due to the exchange of these mesons as V_σ , V_ω , and V_ρ , the δ function in Eq. (13) now shifts the peak in the PDF by an amount proportional to $(V_\sigma + V_\omega + V_\rho)/M$. The isoscalar interactions depend only on the number of spectator nucleons, while the isovector term will depend on the third component of their isospin, for example, repulsive for a d -quark in a struck proton in ${}^3\text{H}$.

With this picture in mind, the off-shell PDF can in general be written as a sum of isoscalar “(0)” and isovector “(1)” contributions,

$$\delta q_{N/A} = \delta q_{N/A}^{(0)} + \delta q_{N/A}^{(1)}, \quad (14)$$

characterizing the possible isospin projections resulting from the interactions between the on-shell spectator quark system after a quark has been removed from the nucleon and the residual nuclear system that does not take part in the scattering. For the case of a proton, scattering from a u quark gives a spectator quark system that is dominated by a ud quark pair that has third component of isospin 0, whereas scattering from a d quark leaves behind a uu pair that has isospin projection +1, and similarly for the neutron.

For a deuteron target, the recoil nuclear system consists of an isospin-1/2 nucleon, regardless of whether the proton or neutron is probed in the scattering. In this case the spectator systems can interact via the exchange of either isoscalar or isovector quantum numbers. For ${}^3\text{H}$, scattering from a proton leaves behind two neutrons, which have isospin projection -1 , while scattering from a neutron involves a pn state with isospin projection 0. For ${}^3\text{He}$, on the other hand, scattering from a proton gives an isospin-0 pn state, while scattering from a neutron will leave two protons with isospin projection +1 in the intermediate state. Finally, to preserve the number of valence quarks in the bound nucleons in nuclei each of the isoscalar and isovector off-shell contributions to the off-shell functions, $\delta q_{N/A}$, must satisfy

$$\int_0^1 dx \delta q_{N/A}^{(0)}(x) = 0 = \int_0^1 dx \delta q_{N/A}^{(1)}(x). \quad (15)$$

In the previous JAM analysis [32] of the earlier MARATHON data [30] on the ${}^3\text{He}/{}^3\text{H}$ ratio, the off-shell functions $\delta q_{N/A}$ were assumed to arise purely from the isovector contributions in the nucleus. Here we generalize this approach by allowing both isoscalar and isovector contributions to the off-shell functions. From the above discussion, for the isoscalar contributions to the u and d off-shell functions, we have

$$\delta u_{p/{}^3\text{He}}^{(0)} = \delta u_{p/{}^3\text{H}}^{(0)} = 2\delta u_{p/D}^{(0)} \equiv \delta u_0, \quad (16a)$$

$$\delta d_{p/{}^3\text{He}}^{(0)} = \delta d_{p/{}^3\text{H}}^{(0)} = 2\delta d_{p/D}^{(0)} \equiv \delta d_0. \quad (16b)$$

In ${}^3\text{He}$, the isovector contributions will be zero since the spectator system for scattering from the proton in a ${}^3\text{He}$ nucleus is an isospin-0 pn pair. For ${}^3\text{H}$, on the other hand, isovector exchange is allowed,

$$\delta u_{p/{}^3\text{He}}^{(1)} = 0, \quad (17a)$$

$$\delta d_{p/{}^3\text{He}}^{(1)} = 0, \quad (17b)$$

$$\delta u_{p/{}^3\text{H}}^{(1)} = 2\delta u_{p/D}^{(1)} \equiv \delta u_1, \quad (17c)$$

$$\delta d_{p/{}^3\text{H}}^{(1)} = 2\delta d_{p/D}^{(1)} \equiv \delta d_1. \quad (17d)$$

Using Eqs. (16) and (17), we can therefore write the set of off-shell u and d PDFs in the proton in the $A = 2$ and $A = 3$ nuclei as

$$\delta u_{p/D} = \frac{1}{2}(\delta u_0 + \delta u_1), \quad \delta u_{p/{}^3\text{He}} = \delta u_0, \quad \delta u_{p/{}^3\text{H}} = \delta u_0 + \delta u_1, \quad (18a)$$

$$\delta d_{p/D} = \frac{1}{2}(\delta d_0 + \delta d_1), \quad \delta d_{p/{}^3\text{He}} = \delta d_0, \quad \delta d_{p/{}^3\text{H}} = \delta d_0 + \delta d_1, \quad (18b)$$

so that all of the off-shell functions can be expressed in terms of the four basis off-shell functions δu_0 , δd_0 , δu_1 , and δd_1 . However, since there are only three unique experimental constraints from the inclusive nuclear DIS data (on D , ${}^3\text{He}$, and ${}^3\text{H}$ nuclei), the number of functions that can be determined from data needs to be reduced to at most three. Naturally, the reduction of the number of functions to be determined should be made in the most physically motivated way possible. Setting δu_1 equal to δd_1 , for instance, would contradict expectations from mean field calculations of nuclear forces [3–6,8], so it is more reasonable to set the isoscalar functions equal, $\delta u_0 = \delta d_0 \equiv \delta q_0$. This allows us to obtain the most faithful representation of the uncertainties possible on quantities that are sensitive to off-shell corrections, such as nuclear EMC ratios, while taking into account the limits of the experimental data and keeping the model physically motivated.

We have also tested two other scenarios with isovector off-shell corrections switched off, $\delta u_1 = \delta d_1 = 0$, namely,

one with $\delta u_0 \neq \delta d_0$ and one with $\delta u_0 = \delta d_0$. For the former, we find that the experimental data cannot distinguish between the individual δu_0 and δd_0 flavors, and these end up being highly correlated. This further motivates the choice of setting the isoscalar off-shell functions equal and fitting the three independent functions δq_0 , δu_1 , and δd_1 . In practice, we find that the extracted isoscalar function δq_0 is very similar to the isoscalar off-shell contribution extracted when isovector effects are set to zero, which demonstrates the stability of the extracted isoscalar component.

C. Methodology

The methodology employed in this analysis is based on Bayesian inference using Monte Carlo techniques developed in previous JAM QCD global analyses [32,40–46]. In contrast to attempts to extract information on PDFs or nuclear effects from a single experiment [30,31], which inevitably requires model-dependent inputs and assumptions, our global analysis allows both the nucleon PDFs and nuclear effects to be determined simultaneously with reduced theoretical bias.

The Bayesian analysis consists of sampling the posterior distribution $\mathcal{P}$ for a set of parameters $\mathbf{a}$ given by

$$\mathcal{P}(\mathbf{a}|\text{data}) \propto \mathcal{L}(\mathbf{a}, \text{data})\pi(\mathbf{a}), \quad (19)$$

with a likelihood function of Gaussian form,

$$\mathcal{L}(\mathbf{a}, \text{data}) = \exp\left(-\frac{1}{2}\chi^2(\mathbf{a}, \text{data})\right), \quad (20)$$

and a prior function $\pi(\mathbf{a})$. The χ^2 function in Eq. (20) is defined for each replica as

$$\chi^2(\mathbf{a}) = \sum_{e,i} \left(\frac{d_{e,i} - \sum_k r_{e,k} \beta_{e,i}^k - T_{e,i}(\mathbf{a})/N_e}{\alpha_{e,i}} \right)^2, \quad (21)$$

where $d_{e,i}$ is the data point i from experimental dataset e , and $T_{e,i}(\mathbf{a})$ is the corresponding theoretical value. All uncorrelated uncertainties are added in quadrature and labeled by $\alpha_{e,i}$, while $\beta_{e,i}^k$ represents the k th source of point-to-point correlated systematic uncertainties for the i th data point weighted by $r_{e,k}$. The latter are optimized per values of the parameters $\mathbf{a}$ via $\partial\chi^2/\partial r_{e,k} = -2r_{e,k}$, which can be solved in closed form. We include normalization parameters N_e for each dataset e as part of the posterior distribution per dataset.

The prior $\pi(\mathbf{a})$ for each replica is given by

$$\pi(\mathbf{a}) = \left[\prod_{\ell} \Theta((a_{\ell} - a_{\ell}^{\min})(a_{\ell}^{\max} - a_{\ell})) \right] \left[\prod_e \prod_k \exp\left(-\frac{1}{2}r_{e,k}^2\right) \right] \left[\prod_e \exp\left(-\frac{1}{2}\left(\frac{1 - N_e}{\delta N_e}\right)^2\right) \right], \quad (22)$$

where $\prod_{\ell}$ is a product over the parameters. The step function Θ forces the parameters to be within the chosen range between $a_{\ell}^{\min}$ and $a_{\ell}^{\max}$. A Gaussian penalty controlled by the experimentally quoted normalization uncertainties δN_e is applied when fitting the normalizations N_e . Similarly, a penalty proportional to $r_{e,k}^2$ is applied when the experiment e has correlated sources of uncertainty.

As in most previous unpolarized PDF analyses by JAM and other groups, we parametrize the PDFs at the input scale $\mu_0^2 = m_c^2$ using a generic template function of the form

$$f(x, \mu_0^2) = \frac{N}{\mathcal{M}} x^{\alpha} (1-x)^{\beta} (1 + \gamma\sqrt{x} + \eta x), \quad (23)$$

where $\mathbf{a} = \{N, \alpha, \beta, \gamma, \eta\}$ is the set of parameters to be inferred. The normalization constant $\mathcal{M} = B[\alpha+2, \beta+1] + \gamma B[\alpha+\frac{5}{2}, \beta+1] + \eta B[\alpha+3, \beta+1]$, where B is the Euler beta function, normalizes the function to the second moment, and allows us to avoid large numerical values for the α and β parameters.

Our parametrization of the on-shell unpolarized PDFs is similar to that of the recent JAM analysis of Ref. [46]. To characterize the nucleon valence region and discriminate it from the sea components, we parametrize the light-quark and strange PDFs according to

$$\begin{aligned} u &= u_v + \bar{u}, & d &= d_v + \bar{d}, \\ \bar{u} &= S_1 + \bar{u}_0, & \bar{d} &= S_1 + \bar{d}_0, \\ s &= S_2 + s_0, & \bar{s} &= S_2 + \bar{s}_0, \end{aligned} \quad (24)$$

where the dependence on x and the scale μ_0^2 has been suppressed for convenience. The input quark distributions $u_v, d_v, \bar{u}_0, \bar{d}_0, s_0$, and $\bar{s}_0$, as well as the gluon distribution g , are parametrized individually, as in Eq. (23). For the sea quark PDFs, the additional functions S_1 and S_2 are also parametrized via Eq. (23), and are designed to allow a more singular small- x behavior compared to the valence distributions by restricting the corresponding α parameter to more negative values. For u_v, d_v, s , and g the normalization parameters N are fixed by the valence and momentum sum rules. For $s, \bar{s}, S_1$, and S_2 we also set $\gamma = \eta = 0$. This leads to a total of 33 free parameters for the on-shell PDFs.

For the off-shell functions, we fit the isoscalar δq_0 and isovector δu_1 and δd_1 functions independently using Eq. (23), with $\gamma = 0$ and η fixed by the sum rule (15). This leads to three independent parameters for each function, and a total of nine for the entire set of off-shell functions.

For the HT corrections to the DIS structure functions, we use the additive parametrization in Eq. (9), in contrast to the previous JAM analysis of the MARATHON data [30] where a multiplicative parametrization was used. In practice, since our HT parametrization is sufficiently flexible, we find that the choice of additive or multiplicative ansatz

does not substantially affect the final results. However, as discussed in Sec. II A, we choose the additive form in order to maintain a cleaner separation between the $1/Q^2$ power and logarithmic in Q^2 corrections in the HTs. The HT corrections for protons and neutrons in the present work are parametrized independently, using Eq. (23) but without the factor $\mathcal{M}$. This leads to a total of 10 independent parameters for the HT corrections. In addition to these parameters, there are a further 22 fitted normalizations in the analysis, leading to a total of 74 free parameters in the global fit.

III. QUALITY OF FIT AND DATA NORMALIZATION

The baseline data used in this analysis are very similar to those used in the previous JAM global analysis [32] that included the earlier MARATHON data [30], but with a few updates. Specifically, the W^2 cut for unpolarized DIS data was increased from 3.0 GeV² to 3.5 GeV², which significantly improved the χ^2_{red} for some datasets. Inclusive W +charm data were also included, as in the recent JAM study in Ref. [46]. In addition to the previous baseline, we also include the recent D/p inclusive DIS data from the E12-10-002 experiment in Hall C at Jefferson Lab [47].

While the MARATHON data on the D/p and ${}^3\text{He}/{}^3\text{H}$ ratios [30] were included in the earlier JAM global analysis [32], the current study includes the new data on the individual ${}^3\text{He}/D$ and ${}^3\text{H}/D$ cross section ratios [31] in place of the ${}^3\text{He}/{}^3\text{H}$ measurement. The restriction on the W^2 range removes the highest- x_B point in both of the ${}^3\text{He}/D$ and ${}^3\text{H}/D$ measurements, so that 21 of the 22 original data points for both ratios are included. Crucially, whereas the ${}^3\text{He}/D$ and ${}^3\text{H}/D$ ratios in the MARATHON analysis [30] were normalized by factors of 1.021 and 0.996, respectively, to match with the KP model [33], in our analysis we do not impose any model-dependent normalization on the MARATHON data, but fit the overall normalizations with χ^2 penalties dependent on the experimental normalization uncertainties, as detailed in Sec. II C above. This allows us to perform an analysis that is not biased by the assumed veracity of the KP model that was tuned to large- A nuclear data [33].

Note that the MARATHON experiment measured ratios of inclusive cross sections, without separating the transverse and longitudinal contributions. At the 11 GeV Jefferson Lab kinematics at which the data were taken, we expect the contributions from longitudinal cross sections to be much smaller than the transverse cross sections, and the nuclear dependence of the longitudinal to transverse ratio to be weak. Consistent with other studies of nuclear EMC ratios, in this analysis we take the MARATHON cross section ratios to be the same as the structure function ratios, and refer to these interchangeably.

TABLE I. Reduced χ^2_{red} values for the data used in this analysis, for the full fit including off-shell corrections and for the on-shell only fit. The MARATHON and JLab E03-103 data are listed separately from the rest of the fixed target DIS data.

Process	N_{dat}	χ^2_{red}	
		On-shell fit	Full fit
DIS			
MARATHON D/p [30]	7	0.68	0.70
MARATHON $^3\text{He}/D$ [31]	21	5.10	1.04
MARATHON $^3\text{H}/D$ [31]	21	2.20	0.83
JLab E03-103 $^3\text{He}/D$ [48]	13	0.33	0.22
Other fixed target	2793	0.99	0.99
HERA	1185	1.17	1.16
Drell-Yan	205	1.18	1.18
W -lepton asymmetry	70	0.82	0.82
W charge asymmetry	27	1.03	1.03
Z rapidity	56	1.10	1.11
Jet	200	1.12	1.10
Total	4598	1.07	1.04

In Table I we summarize the χ^2_{red} values for the new analysis, as well as for a purely on-shell analysis as a reference point. Between these two scenarios there is almost no change for non-DIS processes, nor for HERA or the fixed target data outside of MARATHON. For the MARATHON data, little change is seen for the D/p ratio, while the description of the $^3\text{H}/D$ data improves. The largest change is seen in the description of the $^3\text{He}/D$ data, which improves dramatically with inclusion of off-shell effects. This improvement provides the strongest evidence to date for off-shell effects when describing DIS from $A \leq 3$ nuclei.

We stress that the three $A = 3$ observables from MARATHON ($^3\text{He}/^3\text{H}$, $^3\text{He}/D$, and $^3\text{H}/D$) are not independent. To avoid double-counting of the data, only two of these can be used in the fit. Since the $^3\text{He}/D$ and $^3\text{H}/D$ ratio data are the most recent, in practice we choose to include these in the main analysis. We have checked that selecting other combinations of the data (namely, $^3\text{He}/^3\text{H}$ and $^3\text{He}/D$, or $^3\text{He}/^3\text{H}$ and $^3\text{H}/D$) in the fit does not affect any of the conclusions.

A comparison of the results of the global analysis with the MARATHON data is shown in Fig. 1, where the results of the full fit are shown together with an on-shell fit in which all off-shell corrections are set to zero. As is standard practice in global QCD analyses, the fitted normalizations are used to shift the theory curves, while the experimental data are left unmodified. Consistent with the observation in Ref. [32], for the D/p ratio there is little change between the two fits. For the new $^3\text{H}/D$ data, the on-shell fit slightly undershoots the data at lower x_B and slightly overshoots it at larger x_B . With off-shell corrections included, the data are described very well across all x_B . As expected from the χ^2_{red}

values, the largest impact occurs for the $^3\text{He}/D$ ratio, where the on-shell only results severely overshoot the data at large x_B . The inclusion of off-shell corrections resolves this discrepancy.

The normalizations for the MARATHON structure function ratios for the full off-shell fit are shown in Table II. The experimental normalization uncertainties used here were taken from Ref. [49], where they were determined by the uncorrelated (from target to target) target thickness uncertainties combined in quadrature with the uncorrelated beam heating effects. In contrast, the normalization uncertainties reported in the MARATHON publications [30,31] were taken from the uncertainty on the size of the normalization shifts required for the derivative F_2^n/F_2^p ratio to agree at $x_B = 0.31$ when extracted from deuterium and $A = 3$ ratios using the KP model. This normalization procedure assumes that the MARATHON F_2^D/F_2^p data require no normalization, and that the nuclear effects for all nuclei vanish at $x_B = 0.31$, as in the KP model [25,33]. We stress that while the latter behavior has been observed for ratios of structure functions of heavy nuclei to deuterium [2], there is no direct evidence to suggest that it also holds for $A \leq 3$ nuclei, and, as we show below, enforcing this constraint introduces a significant bias into the analysis.

For the $^3\text{He}/D$ and $^3\text{H}/D$ ratios, the fitted normalizations are within 1σ using the experimental normalization uncertainties [49]. For the D/p ratio, the fitted normalization is 2.2σ away from 1, indicating some tension of the MARATHON deuteron data with other datasets in the global analysis. We also compare our normalizations to those derived using the KP model applied to the MARATHON data in Refs. [30,31]. For $^3\text{He}/D$ we find differences of 3.2σ , indicating significant tension between our extracted normalization and that obtained using the KP model. For $^3\text{H}/D$, on the other hand, the two normalizations are in better agreement within the uncertainties.

In Fig. 2 we compare the $^3\text{He}/D$ data from MARATHON and the earlier data from the Jefferson Lab E03-103 experiment [48], and with the fitted results of our global analysis. No normalizations have been applied to any of the data here, but the theory curves have been divided by the normalizations listed in Table II. We find excellent agreement between the experimental results and the fitted values, as well as consistency between the two experiments. In contrast, the results in Fig. 1 of Ref. [31] indicate the MARATHON $^3\text{He}/D$ ratio lying significantly higher, with essentially no overlap with the Jefferson Lab E03-103 data. This is a direct result of the 2.1% upward normalization that was applied to the MARATHON data in Ref. [31] in order to match with the KP model result [25,33]. From our global analysis, we find that this normalization is not necessary, and in fact introduces a significant bias into the subsequent extraction of the neutron structure function and determination of the nucleon off-shell contributions.

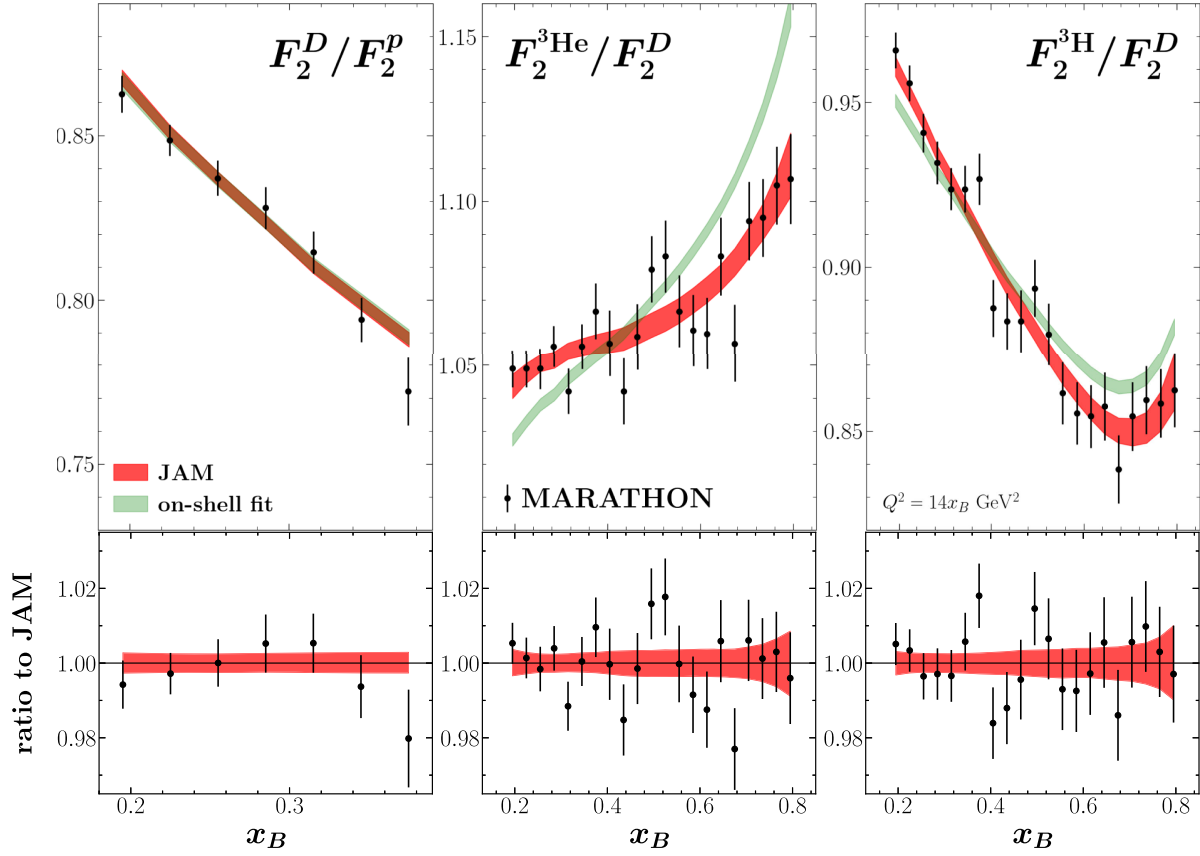


FIG. 1. Ratios F_2^D/F_2^p (left), $F_2^{3\text{He}}/F_2^D$ (middle), and $F_2^{3\text{H}}/F_2^D$ (right) from MARATHON [30,31] (black circles) at the experimental kinematics $Q^2 = 14x_B \text{ GeV}^2$ compared to the full JAM fit (red 68% CI bands) and to the on-shell fit (green 68% CI bands). The bottom panels show the ratios of the results relative to the mean JAM results. The theory curves are normalized by the fitted values 1.017(4), 0.996(6), and 0.991(5) for the D/p , $^3\text{He}/D$, and $^3\text{H}/D$ ratios, respectively.

In Fig. 3 we show the isoscalar EMC ratio $(F_2^{3\text{He}} + F_2^{3\text{H}})/2F_2^D$, as in Fig. 2 of Ref. [31]. We compare the JAM result, unmodified by any fitted normalizations, to the MARATHON data that are also unmodified by any normalizations. The results agree reasonably well, with JAM slightly undershooting it.

We emphasize, however, that the JAM ratio in Fig. 3 is a pure theory result unmodified by fitted normalizations, which cannot be cleanly translated from the three MARATHON observables (F_2^D/F_2^p , $F_2^{3\text{He}}/F_2^D$, and $F_2^{3\text{H}}/F_2^D$) to this new ratio. The slight underestimate can

therefore be explained by the non-negligible normalization on F_2^D/F_2^p of 1.017(4), which would decrease F_2^D and thus increase the ratio, which is not applied in Fig. 3. The result from Ref. [31], including their normalizations, is also shown and compared to the KP model. These also agree well, which is unsurprising given that the normalizations are matched by construction. It is interesting to observe that the disagreement between the JAM results and the KP model reflects very closely the difference between the MARATHON data without and with the applied normalizations. This difference shows clearly the bias that is introduced by applying normalizations to the experimental

TABLE II. Fitted normalizations from the current JAM analysis compared to the normalizations used in Refs. [30,31] from the KP model, along with the σ differences between these (with uncertainties summed in quadrature). The experimental normalization uncertainties from Ref. [49] are also shown.

Dataset	Experimental norm. uncertainty [49] (%)	Fitted normalization	MARATHON + KP model	Significance of difference
MARATHON D/p [30]	0.8	1.017(4)	...	...
MARATHON $^3\text{He}/D$ [31]	1.2	0.996(6)	1.021(5)	3.2σ
MARATHON $^3\text{H}/D$ [31]	0.8	0.991(5)	0.996(5)	0.7σ

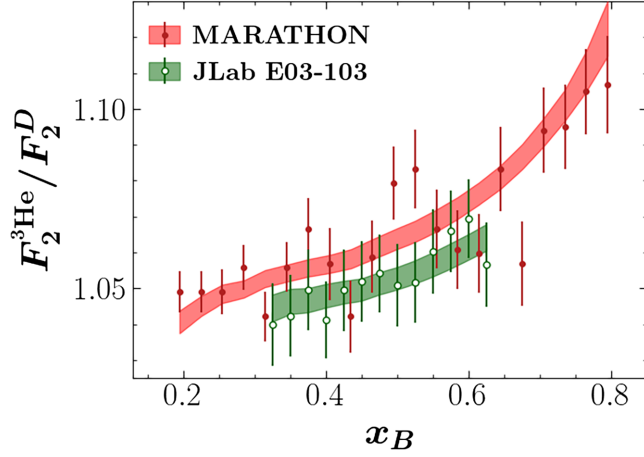


FIG. 2. Ratio $F_2^{3\text{He}}/F_2^D$ from the Jefferson Lab MARATHON [31] (red circles) and Hall C E03-103 [48] (green open circles) experiments compared to the JAM fit (colored 68% CI bands). No normalizations have been applied to any of the data here, while the theory curves have been divided by the normalizations listed in Table II.

data that are based on the assumed validity of the KP model for the $A = 3$ system.

IV. QCD ANALYSIS

Having demonstrated good agreement between the global fit and the MARATHON data, in this section we discuss the results of our global QCD analysis, which is based on an ensemble of over 700 replicas, for the on-shell

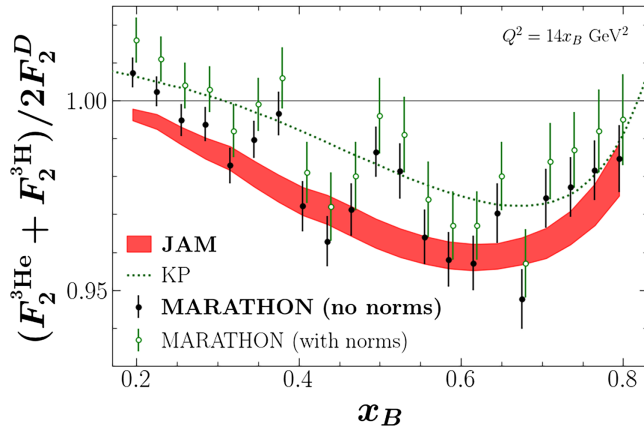


FIG. 3. Isoscalar nuclear EMC ratio $(F_2^{3\text{He}} + F_2^{3\text{H}})/2F_2^D$ at the MARATHON kinematics $Q^2 = 14x_B \text{ GeV}^2$. The JAM result, unmodified by any normalizations (red 68% CI band), is compared to the result directly from the MARATHON paper [31] (green open circles) and the result where the 1.021 and 0.996 normalizations are removed from ^3He and ^3H , respectively (black circles). The errors on the latter are computed using the quadrature sum of the uncorrelated errors on the $F_2^{3\text{He}}/F_2^D$ and $F_2^{3\text{H}}/F_2^D$ ratios. The result is also compared to that of the KP model with isoscalar off-shell effect [31,33] (green dotted line).

u and d quark PDFs, the isospin dependence of the off-shell corrections, and the systematics of the nuclear EMC ratios for $A = 2$ and $A = 3$ nuclei.

A. Impact on proton PDFs

The results for the u and d quark PDFs in the proton from our new analysis are shown in Fig. 4 as ratios to the JAM central values, at a scale of $Q^2 = 10 \text{ GeV}^2$, for analyses with and without the inclusion of the MARATHON data. We observe very little impact of the MARATHON data on the PDFs, with almost no change in the central values and similar uncertainties for the two analyses, suggesting that these are already well constrained by other data. The only visible impact is a small reduction in uncertainties in the d quark PDF at $x \gtrsim 0.6$. This is consistent with the findings of the previous JAM analysis [32] that included the earlier $^3\text{He}/^3\text{H}$ MARATHON data [30]. As expected, since the u quark PDF is more strongly constrained by proton DIS and other data, the PDF uncertainties are $\lesssim 5\%$ over the entire region of $x \lesssim 0.85$ constrained by data. For the d quark PDF, the uncertainties are larger, $\approx 20\%$ for $x \approx 0.75$, and increase significantly at higher x .

B. Isospin dependence of off-shell corrections

The off-shell functions δq_i from our analysis are shown in Fig. 5 for the isoscalar component δq_0 and the isovector contributions δu_1 and δd_1 , as well as for the difference $\delta u_1 - \delta d_1$, at the scale $Q^2 = 10 \text{ GeV}^2$. Notably, we observe that the isoscalar off-shell correction is nonzero and positive, with $x\delta q_0$ peaking at $x \approx 0.4$. The isovector contributions have larger uncertainties, with $x\delta d_1$ mostly consistent with zero, but $x\delta u_1$ displaying a negative trend at large x . This trend is enhanced in the difference $x(\delta u_1 - \delta d_1)$, which indicates a clear negative behavior for $x \gtrsim 0.5$, suggesting the presence of an isovector effect.

The trends in the δq_i off-shell functions can be understood by examining their contributions to the nuclear structure functions, as shown in Fig. 6. Here, the off-shell contribution to F_2^A is given by

$$F_2^{A(\text{off})}(x_B, Q^2) = \sum_N [f_{N/A}^{(\text{off})} \otimes \delta F_2^{N/A}](x_B, Q^2), \quad (25)$$

and is shown as a ratio to the total structure function, at a scale of $Q^2 = 10 \text{ GeV}^2$, along with the separate contributions from the isoscalar and isovector off-shell components. The isoscalar contribution is found to be negative for all nuclei considered. For deuterium it is suppressed by a factor of ≈ 4 relative to ^3He and ^3H because of the weaker binding [see Eq. (30) below] and the combinatorial factor 2 in Eq. (16), which effectively halves the isoscalar correction for deuterium relative to helium and tritium. The effect is a direct consequence of δq_0 being positive (as the off-shell smearing functions are negative) and the fact that the

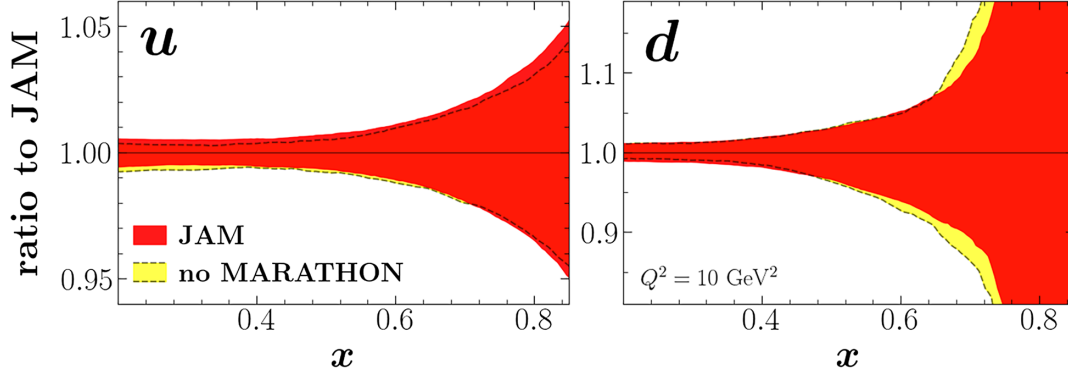


FIG. 4. Ratios of PDFs for u (left) and d (right) quarks relative to the mean JAM result, for the full JAM analysis (red 68% CI band) and the analysis with no MARATHON data included (yellow 68% CI band) at $Q^2 = 10 \text{ GeV}^2$.

isoscalar off-shell corrections affect all three nuclei in a similar manner.

The isovector components, on the other hand, are more varied. They are relatively large and positive for the deuteron and tritium, leading to strong cancellations with the isoscalar component. This cancellation is almost total for the deuteron, while it leaves a small, negative contribution to ^3H at large x_B . For ^3He , the isovector component is very small, leaving the effect here to be dominated by the large and negative isoscalar component.

These trends can be directly related to those observed in the data versus theory comparisons for the structure function ratios in Fig. 1. The small effects for the deuteron

that largely cancel explain why there is no significant change in the MARATHON D/p data, or any of the other deuteron data in this analysis. For tritium, the isoscalar contribution alone makes the $F_2^{^3\text{H}}/F_2^D$ ratio too small at larger x_B . The cancellation with the isovector contribution raises the $F_2^{^3\text{H}}/F_2^D$ ratio slightly at larger x_B . For helium, the largest contribution by far is from the isoscalar, which

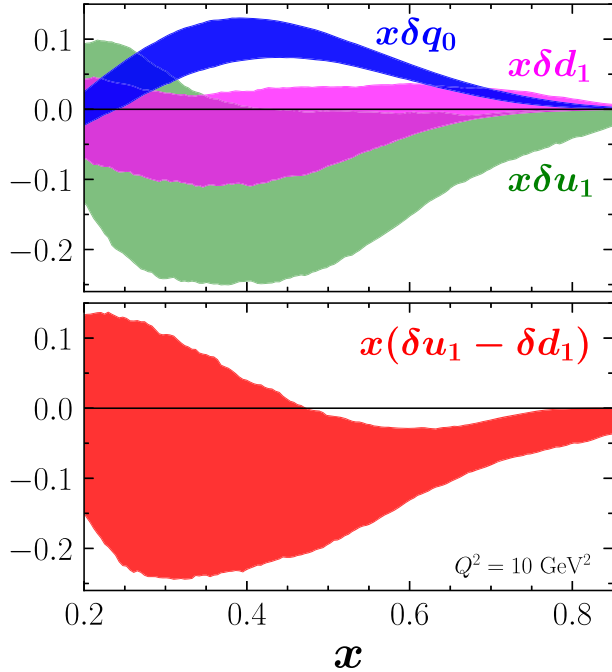


FIG. 5. Off-shell functions $x\delta q_i$ versus x at $Q^2 = 10 \text{ GeV}^2$. Top panel: isoscalar $x\delta q_0 \equiv x\delta u_0 = x\delta d_0$ (blue), isovector $x\delta u_1$ (green), and isovector $x\delta d_1$ (orange) distributions. Bottom panel: difference $x(\delta u_1 - \delta d_1)$ (red). The bands indicate 68% CIs.

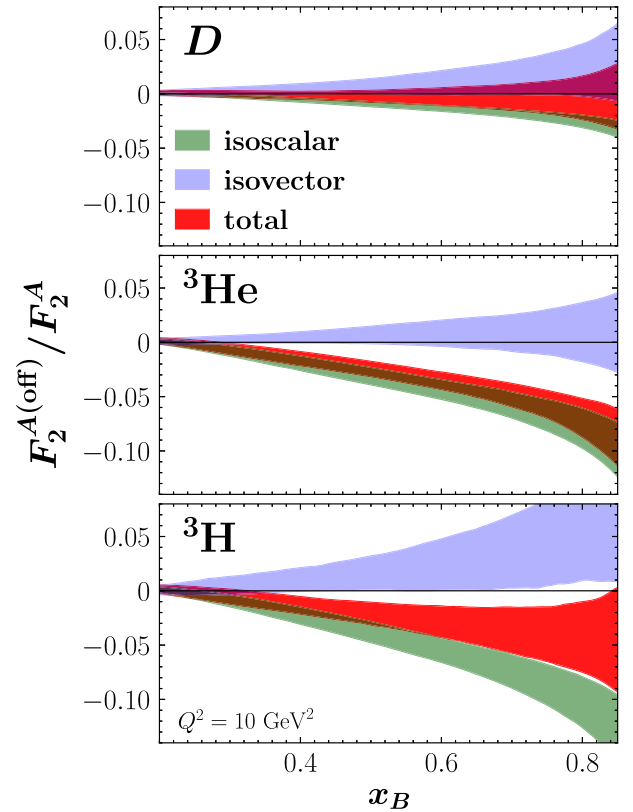


FIG. 6. Off-shell contributions $F_2^{A(\text{off})}$ to the nuclear structure functions relative to the total F_2^A for $A = D$ (top), ^3He (middle), and ^3H (bottom) at $Q^2 = 10 \text{ GeV}^2$ (68% CI red bands). The mean values of the isoscalar (68% CI blue bands) and isovector (68% CI green bands) contributions are shown separately.

significantly reduces $F_2^{3\text{He}}$, and thus the $F_2^{3\text{He}}/F_2^D$ ratio. When the isovector component is added, the $F_2^{3\text{He}}/F_2^D$ ratio decreases slightly as x_B becomes larger. While the isovector contribution to $F_2^{3\text{He}}$ is negligible, the slight positive isovector contribution to F_2^D causes the $F_2^{3\text{He}}/F_2^D$ ratio to decrease slightly.

C. Nuclear EMC ratios

Since the MARATHON experiment only measured ratios of cross sections (structure functions) and not their absolute values, we shall be primarily interested in EMC-type ratios of structure functions of nuclei A to free nucleons. In particular, for the $A = 2$ and $A = 3$ nuclei probed in the MARATHON experiment, we define the “isoscalar” structure function ratios

$$R_D = \frac{F_2^D}{F_2^p + F_2^n}, \quad (26)$$

for the deuteron, and

$$R_{3\text{He}} = \frac{F_2^{3\text{He}}}{2F_2^p + F_2^n}, \quad R_{3\text{H}} = \frac{F_2^{3\text{H}}}{F_2^p + 2F_2^n}, \quad (27)$$

for ^3He and ^3H , respectively. Further, taking ratios of the nuclear EMC ratios, we define the helium to tritium super-ratio as

$$\mathcal{R}_{3\text{He}/^3\text{H}} = \frac{R_{3\text{He}}}{R_{3\text{H}}}, \quad (28)$$

and the analogous super-ratios with respect to the deuteron as

$$\mathcal{R}_{3\text{He}/D} = \frac{R_{3\text{He}}}{R_D}, \quad \mathcal{R}_{3\text{H}/D} = \frac{R_{3\text{H}}}{R_D}. \quad (29)$$

In Fig. 7 we compare the x_B dependence of the deuteron to free-nucleon EMC ratio R_D at $Q^2 = 10 \text{ GeV}^2$ from the current JAM analysis with results from the CJ22ht [50] and AKP17 [25,33,51] global analyses, the latter of which utilizes constraints from the KP model. At small x_B the JAM ratio lies slightly below the CJ22ht and AKP17 results, while at larger $x_B \gtrsim 0.6$ the ratio is at the lower end of the CJ22ht band and just above the AKP17 band. Most notably, there is no indication for a cross of the ratio through unity at $x_B = 0.31$, as assumed in the KP model (see discussion below). Comparing to the JAM results without the MARATHON data, there is a very small reduction in uncertainties on R_D below $x_B \approx 0.4$, approximately matching the upper x_B limit of the MARATHON F_2^D/F_2^p data.

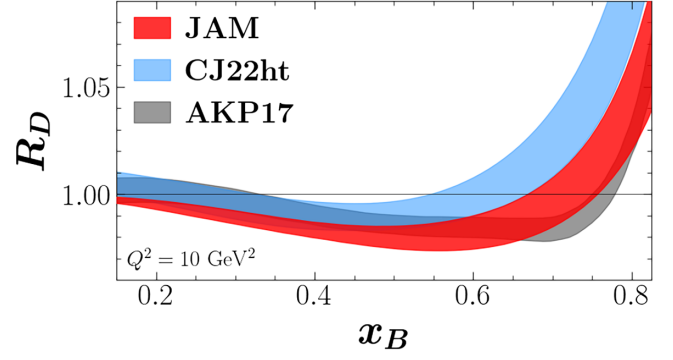


FIG. 7. Deuteron to free-nucleon EMC ratio R_D versus x_B at $Q^2 = 10 \text{ GeV}^2$ from the present JAM analysis (red 68% CI band), compared with that from the CJ22ht [50] (blue 90% confidence band) and AKP17 [51] (gray 1 σ band) PDF parametrizations.

For the $A = 3$ super-ratio $\mathcal{R}_{3\text{He}/^3\text{H}}$, shown in Fig. 8 at the experimental kinematics $Q^2 = 14x_B \text{ GeV}^2$, at intermediate x_B values our result largely overlaps with the KP model that was used to extract the F_2^n/F_2^p ratio in Ref. [30], but deviates at larger values of x_B . In particular, our result tends toward values lower than 1, whereas the KP model remains above unity up to $x_B \approx 0.8$, similar to what was observed in earlier calculations of the super-ratio that assumed flavor-independent off-shell corrections in $A = 3$ nuclei [26–29]. Our fitted result is consistent with the findings in the previous JAM analysis [32], except now the super-ratio has larger errors because of the increased flexibility of the off-shell model. Comparing to the result without the MARATHON data, one sees a significant reduction in uncertainties when MARATHON data are included. This is expected, given that prior to the inclusion of the MARATHON data there were no direct constraints on $F_2^{3\text{H}}$ at all, and only the Jefferson Lab E03-103 $^3\text{He}/D$ data [48] existed to constrain $F_2^{3\text{He}}$.

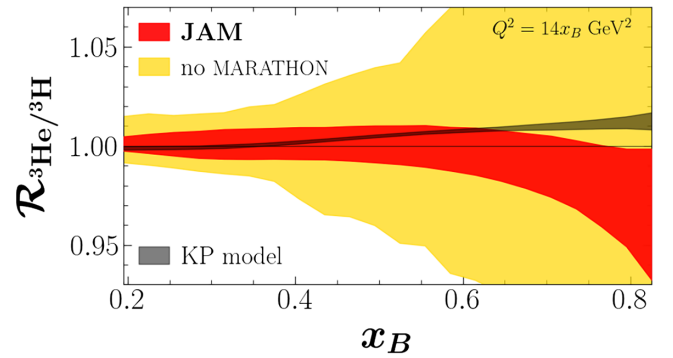


FIG. 8. Super-ratio $\mathcal{R}_{3\text{He}/^3\text{H}}$ of the ^3He to ^3H nuclear EMC ratios, as a function of x_B for $Q^2 = 14x_B \text{ GeV}^2$, from the present JAM analysis (red 68% CI band), compared with the ratio without the MARATHON data (yellow 68% CI band), and the result from the KP model (gray 1 σ band) [25,33] that was used to extract the F_2^n/F_2^p ratio from the MARATHON data in Ref. [31].

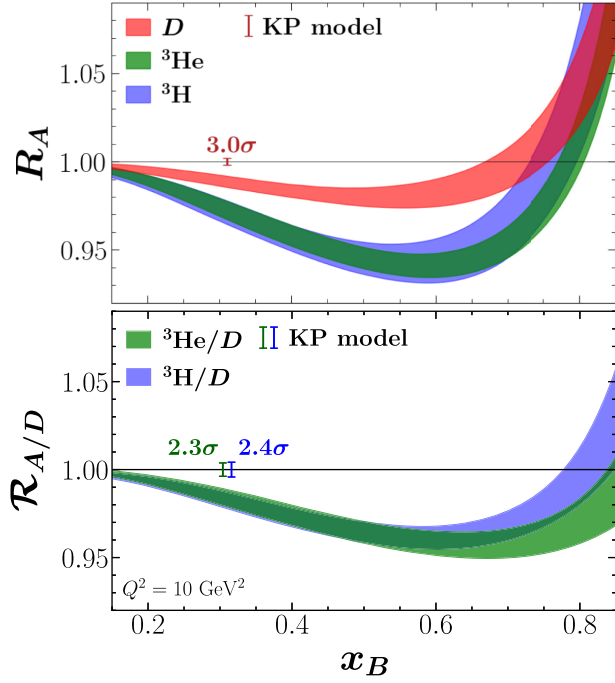


FIG. 9. Top panel: nuclear EMC ratios R_A for $A = D$ (red band), ${}^3\text{He}$ (green band), and ${}^3\text{H}$ (blue band) for the full JAM analysis. The σ difference between the JAM band for D at $x_B = 0.31$ and the KP result (denoted by the red error bar) is calculated using the quadrature sum of the errors. Bottom panel: super-ratios $\mathcal{R}_{A/D}$ for ${}^3\text{He}/D$ (green band) and ${}^3\text{H}/D$ (blue band), with σ differences with the KP points for ${}^3\text{He}/D$ (green error bar) and ${}^3\text{H}/D$ (blue error bar) at $x_B = 0.31$ indicated. All ratios are given at $Q^2 = 10 \text{ GeV}^2$, and all bands represent 68% CIs.

The EMC ratios R_A for $A = D$, ${}^3\text{He}$, and ${}^3\text{H}$, as well as the super-ratios $\mathcal{R}_{A/D}$ for ${}^3\text{He}/D$ and ${}^3\text{H}/D$, are shown in Fig. 9 for the full JAM analysis, along with the on-shell only results and the predictions from the KP model. The KP model gives unity for each of the ratios R_D , $\mathcal{R}_{{}^3\text{He}/D}$, and $\mathcal{R}_{{}^3\text{H}/D}$ at $x_B = 0.31$, with uncertainties of 0.20%, 0.38%, and 0.42%, respectively [31]. Comparing our uncertainty bands at $x_B = 0.31$ to these points, we find differences of 3.0σ , 2.3σ , and 2.4σ , respectively. Our result for R_D is $\approx 1\%$ below unity, while the results for $\mathcal{R}_{{}^3\text{He}}$ and $\mathcal{R}_{{}^3\text{H}}$ are a further 1% – 2%

below R_D , leading to super-ratios $\mathcal{R}_{{}^3\text{He}/D}$ and $\mathcal{R}_{{}^3\text{H}/D}$ that are $\approx 1\%$ – 2% below 1 at $x_B = 0.31$. Note that the effects for helium and tritium being similar and significantly larger than that for deuterium reflects the larger binding energies and smearing effects for the $A = 3$ system. For the Paris [52] deuteron wave function and the KPSV [53] ${}^3\text{He}$ spectral function, for example, we find

$$\langle f_{p/D}^{(\text{off})} \rangle \approx -4.3\%, \quad \langle f_{p/{}^3\text{He}}^{(\text{off})} \rangle \approx -6.8\%, \quad \langle f_{p/{}^3\text{H}}^{(\text{off})} \rangle \approx -9.5\%. \quad (30)$$

Comparing the full JAM results with the mean on-shell results in Fig. 10, the uncertainties for R_D increase significantly when off-shell corrections are included, while the central value remains largely unchanged. For $\mathcal{R}_{{}^3\text{He}}$ and $\mathcal{R}_{{}^3\text{H}}$, the uncertainties also increase while the curves move downward significantly. These observations are consistent with those seen in Fig. 1 when off-shell corrections were included.

In a recent analysis, Alekhin, Kulagin, and Petti [54] claimed that the results of the previous JAM analysis [32] of the ${}^3\text{He}/{}^3\text{H}$ MARATHON data may be biased by the choice of multiplicative HT parametrization and by correlations between off-shell PDF parameters. While the previous JAM analysis [32] used a multiplicative ansatz for the HT correction, an additive parametrization was also tested and produced no qualitative changes in any of the results. In the current analysis, we use additive HT corrections, finding results to be consistent with a multiplicative ansatz, and again in disagreement with the KP model.

The authors of Ref. [54] also claim that there is an increase in correlations between parameters introduced when including nuclear dependence in the off-shell functions for individual quark flavors, as in the current and previous analyses. We do find that the central values of the nuclear EMC ratios, as well as their uncertainties, are significantly influenced by the modeling of the off-shell functions, with added flexibility increasing the uncertainties on the ratios. This can already be seen in Fig. 10 when comparing the on-shell and off-shell results. To obtain the most faithful representation of the uncertainties on the

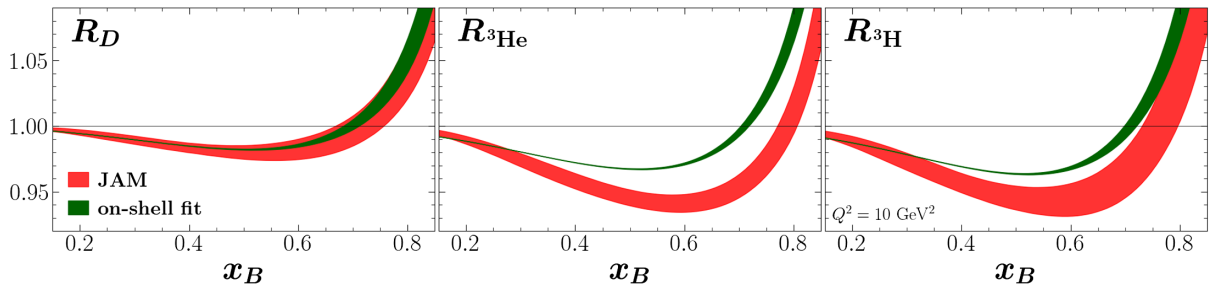


FIG. 10. Nuclear EMC ratios R_A for $A = D$ (left panel), ${}^3\text{He}$ (middle panel), and ${}^3\text{H}$ (right panel) for the full JAM analysis (red 68% CI bands) and for the on-shell only fit (green 68% CI bands) at $Q^2 = 10 \text{ GeV}^2$.

nuclear EMC ratios, it is important to use an off-shell model that is sufficiently flexible given the physical constraints and currently available experimental data, as done in this analysis. In fact, choosing an insufficiently flexible model, as in Ref. [54], would lead to a biased result with underestimated uncertainties on the nuclear EMC ratios.

V. CONCLUSIONS

We have performed a global QCD analysis of the world's unpolarized proton, deuteron, and $A = 3$ data within the JAM Bayesian Monte Carlo framework, including for the first time the latest results from the MARATHON experiment at Jefferson Lab on the $^3\text{He}/D$ and $^3\text{H}/D$ structure function ratios [31]. Most notably, we find that the $^3\text{He}/D$ ratio data cannot be described without the inclusion of nucleon off-shell corrections, without which we obtain a χ^2_{red} of 5.10 for these data, while inclusion of off-shell corrections greatly improves the χ^2_{red} to 1.04. These data clearly indicate the necessity of off-shell contributions to describe all of the world DIS data on $A \leq 3$ nuclei. A smaller improvement is also observed in the description of the $^3\text{H}/D$ data with the inclusion of off-shell corrections. Within our theoretical framework, which allows for both isoscalar and isovector off-shell corrections, we find that the combination of $^3\text{He}/D$ and $^3\text{H}/D$ data suggest a sizable isoscalar contribution, with a slight preference for a non-zero isovector off-shell effect, as indicated by a nonzero $\delta u_1 - \delta d_1$ distribution.

Our results differ significantly from those in the KP model, which is assumed in the analyses of the MARATHON data in Refs. [30,31]. Our fitted normalization for $^3\text{He}/D$ of 0.996(6) is consistent with 1, and differs by 3.2σ from the KP model result of 1.021(5). For the nuclear EMC ratios, we find a different trend for $\mathcal{R}_{ht}$ at large x_B compared to that in the KP model, consistent with the findings in the previous JAM analysis [45]. The KP model also enforces each of the ratios R_D , $\mathcal{R}_{hd}$, and $\mathcal{R}_{td}$ to be unity 1 at $x_B = 0.31$, while our global analysis disagrees with those results by 3.0σ , 2.3σ , and 2.4σ , respectively. The fact that our analysis describes the MARATHON data well (along with all other experimental data), and leads to different results for quantities such as the nuclear EMC ratios, suggests that the additional normalizations imposed on the MARATHON data to be consistent with the KP model needlessly bias the conclusions of the experimental analysis [30,31].

Of course, our conclusions must be viewed within the theoretical framework on which our analysis is based. This includes the basic assumption that DIS from nuclei proceeds through incoherent scattering from individual (off-shell) nucleons, and that multiple scattering from more than one nucleon or scattering from non-nucleonic constituents of the

nucleus are relatively rare events, which for light nuclei and at $x_B \gg 0$ should be a very good approximation [55–63]. A further underlying assumption in our analysis is that of charge symmetry for both the on-shell and off-shell PDFs: while the former could be tested with future data on $\pi^\pm$ leptonproduction or charge-changing neutrino interactions from isoscalar targets [64,65], constraining charge symmetry-violating effects in the off-shell functions is unlikely to be realized in the near future. We also do not consider off-shell corrections to the HT contributions, which would not be possible to determine from existing data. Finally, while we acknowledge potential bias from the choice of a particular functional form for our PDF parametrization, the use of Monte Carlo sampling in generating many hundreds of replicas for the PDFs in practice reduces any bias from adopting this particular functional form.

On the experimental front, in the near future constraints on neutron structure and the d/u PDF ratio at large x may come from the BONuS12 experiment at Jefferson Lab, which tags spectator protons in semi-inclusive DIS from the deuteron [66], as well as parity-violating DIS measurements on protons through the γZ interference process [67,68]. Future data on inclusive and semi-inclusive DIS from $A = 3$ nuclei may also provide further information about the isospin dependence of the nucleon structure modifications in nuclei, and shed light on the elusive dynamical origin of the EMC effect.

ACKNOWLEDGMENTS

We thank C. E. Keppel and A. Metz for helpful discussions. This work was supported by the U.S. Department of Energy, Contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab, and the National Science Foundation under Grant No. PHY-1516088. The work of A. W. T. was supported by Adelaide University and by the Australian Research Council through the Discovery Project DP230101791 and the ARC Centre of Excellence for Dark Matter Particle Physics CE200100008. The work of N. S. was supported by the DOE, Office of Science, Office of Nuclear Physics in the Early Career Program.

DATA AVAILABILITY

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

- [1] J. J. Aubert *et al.*, *Phys. Lett.* **123B**, 275 (1983).
- [2] J. Gomez *et al.*, *Phys. Rev. D* **49**, 4348 (1994).
- [3] A. W. Thomas, A. Michels, A. W. Schreiber, and P. A. M. Guichon, *Phys. Lett. B* **233**, 43 (1989).
- [4] K. Saito, A. Michels, and A. W. Thomas, *Phys. Rev. C* **46**, R2149 (1992).
- [5] H. Mineo, W. Bentz, N. Ishii, A. W. Thomas, and K. Yazaki, *Nucl. Phys. A* **735**, 482 (2004).
- [6] I. C. Cloet, W. Bentz, and A. W. Thomas, *Phys. Lett. B* **642**, 210 (2006).
- [7] J. R. Smith and G. A. Miller, *Phys. Rev. C* **65**, 055206 (2002).
- [8] P. A. M. Guichon, J. R. Stone, and A. W. Thomas, *Prog. Part. Nucl. Phys.* **100**, 262 (2018).
- [9] B. Schmookler *et al.*, *Nature (London)* **566**, 354 (2019).
- [10] X. G. Wang, A. W. Thomas, and W. Melnitchouk, *Phys. Rev. Lett.* **125**, 262002 (2020).
- [11] E. P. Segarra, J. R. Pybus, F. Hauenstein, D. W. Higinbotham, G. A. Miller, E. Piasetzky, A. Schmidt, M. Strikman, L. B. Weinstein, and O. Hen, *Phys. Rev. Res.* **3**, 023240 (2021).
- [12] W. Xing, X. G. Wang, and A. W. Thomas, *Phys. Lett. B* **846**, 138195 (2023).
- [13] D. N. Kim, O. Hen, G. A. Miller, E. Piasetzky, M. Strikman, and L. Weinstein, *Phys. Rev. C* **111**, 065201 (2025).
- [14] D. F. Geesaman, K. Saito, and A. W. Thomas, *Annu. Rev. Nucl. Part. Sci.* **45**, 337 (1995).
- [15] P. R. Norton, *Rep. Prog. Phys.* **66**, 1253 (2003).
- [16] A. W. Thomas, *Int. J. Mod. Phys. E* **27**, 1840001 (2019).
- [17] A. J. Tropiano, J. J. Ethier, W. Melnitchouk, and N. Sato, *Phys. Rev. C* **99**, 035201 (2019).
- [18] G. V. Dunne and A. W. Thomas, *Nucl. Phys. A* **455**, 701 (1986).
- [19] S. V. Akulinichev, S. A. Kulagin, and G. M. Vagradov, *Phys. Lett.* **158B**, 485 (1985).
- [20] R. P. Bickerstaff and A. W. Thomas, *J. Phys. G* **15**, 1523 (1989).
- [21] W. Melnitchouk, A. W. Schreiber, and A. W. Thomas, *Phys. Rev. D* **49**, 1183 (1994).
- [22] W. Melnitchouk, A. W. Schreiber, and A. W. Thomas, *Phys. Lett. B* **335**, 11 (1994).
- [23] S. A. Kulagin, W. Melnitchouk, G. Piller, and W. Weise, *Phys. Rev. C* **52**, 932 (1995).
- [24] S. A. Kulagin, G. Piller, and W. Weise, *Phys. Rev. C* **50**, 1154 (1994).
- [25] S. A. Kulagin and R. Petti, *Nucl. Phys. A* **765**, 126 (2006).
- [26] I. R. Afnan, F. R. P. Bissey, J. Gomez, A. T. Katramatou, W. Melnitchouk, G. G. Petratos, and A. W. Thomas, *Phys. Lett. B* **493**, 36 (2000).
- [27] I. R. Afnan, F. R. P. Bissey, J. Gomez, A. T. Katramatou, S. Liuti, W. Melnitchouk, G. G. Petratos, and A. W. Thomas, *Phys. Rev. C* **68**, 035201 (2003).
- [28] E. Pace, G. Salme, S. Scopetta, and A. Kievsky, *Phys. Rev. C* **64**, 055203 (2001).
- [29] M. M. Sargsian, S. Simula, and M. I. Strikman, *Phys. Rev. C* **66**, 024001 (2002).
- [30] D. Abrams *et al.*, *Phys. Rev. Lett.* **128**, 132003 (2022).
- [31] D. Abrams *et al.*, *Phys. Rev. Lett.* **135**, 062502 (2025).
- [32] C. Cocuzza, C. E. Keppel, H. Liu, W. Melnitchouk, A. Metz, N. Sato, and A. W. Thomas, *Phys. Rev. Lett.* **127**, 242001 (2021).
- [33] S. A. Kulagin and R. Petti, *Phys. Rev. C* **82**, 054614 (2010).
- [34] L. T. Brady, A. Accardi, T. J. Hobbs, and W. Melnitchouk, *Phys. Rev. D* **84**, 074008 (2011); **85**, 039902(E) (2012).
- [35] I. Schienbein *et al.*, *J. Phys. G* **35**, 053101 (2008).
- [36] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, *Phys. Rev. D* **50**, 3102 (1994).
- [37] E. Moffat, T. C. Rogers, W. Melnitchouk, N. Sato, and F. Steffens, *Phys. Rev. D* **99**, 096008 (2019).
- [38] R. L. Jaffe, *Nucl. Phys. B* **229**, 205 (1983).
- [39] F. E. Close and A. W. Thomas, *Phys. Lett. B* **212**, 227 (1988).
- [40] N. Sato, W. Melnitchouk, S. E. Kuhn, J. J. Ethier, and A. Accardi, *Phys. Rev. D* **93**, 074005 (2016).
- [41] N. Sato, J. J. Ethier, W. Melnitchouk, M. Hirai, S. Kumano, and A. Accardi, *Phys. Rev. D* **94**, 114004 (2016).
- [42] J. J. Ethier, N. Sato, and W. Melnitchouk, *Phys. Rev. Lett.* **119**, 132001 (2017).
- [43] N. Sato, C. Andres, J. J. Ethier, and W. Melnitchouk, *Phys. Rev. D* **101**, 074020 (2020).
- [44] E. Moffat, W. Melnitchouk, T. C. Rogers, and N. Sato, *Phys. Rev. D* **104**, 016015 (2021).
- [45] C. Cocuzza, W. Melnitchouk, A. Metz, and N. Sato, *Phys. Rev. D* **104**, 074031 (2021).
- [46] T. Anderson, W. Melnitchouk, and N. Sato, *Phys. Rev. D* **112**, 094011 (2025).
- [47] D. Biswas *et al.*, *Phys. Rev. Lett.* **135**, 151902 (2025).
- [48] J. Seely *et al.*, *Phys. Rev. Lett.* **103**, 202301 (2009).
- [49] S. Li *et al.*, *Nature (London)* **609**, 41 (2022).
- [50] M. Cerutti, A. Accardi, I. P. Fernando, S. Li, J. F. Owens, and S. Park, *Phys. Rev. D* **111**, 094013 (2025).
- [51] S. I. Alekhin, S. A. Kulagin, and R. Petti, *Phys. Rev. D* **96**, 054005 (2017).
- [52] M. Lacombe, B. Loiseau, R. Vinh Mau, J. Cote, P. Pires, and R. de Tournell, *Phys. Lett.* **101B**, 139 (1981).
- [53] A. Kievsky, E. Pace, G. Salme, and M. Viviani, *Phys. Rev. C* **56**, 64 (1997).
- [54] S. I. Alekhin, S. A. Kulagin, and R. Petti, *Proc. Sci. DIS2024* (2025) 067.
- [55] B. M. Badelek and J. Kwiecinski, *Nucl. Phys. B* **370**, 278 (1992).
- [56] W. Melnitchouk and A. W. Thomas, *Phys. Rev. D* **47**, 3783 (1993).
- [57] W. Melnitchouk and A. W. Thomas, *Phys. Lett. B* **317**, 437 (1993).
- [58] B. Badelek and J. Kwiecinski, *Phys. Rev. D* **50**, R4 (1994).
- [59] W. Melnitchouk and A. W. Thomas, *Phys. Rev. C* **52**, 3373 (1995).
- [60] G. Piller, W. Ratzka, and W. Weise, *Z. Phys. A* **352**, 427 (1995).
- [61] J. Edelman, G. Piller, and W. Weise, *Phys. Rev. C* **57**, 3392 (1998).

-
- [62] F. R. P. Bissey, V. A. Guzey, M. Strikman, and A. W. Thomas, [Phys. Rev. C **65**, 064317 \(2002\)](#).
- [63] G. Piller and W. Weise, [Phys. Rep. **330**, 1 \(2000\)](#).
- [64] J. T. Londergan, J. C. Peng, and A. W. Thomas, [Rev. Mod. Phys. **82**, 2009 \(2010\)](#).
- [65] T. J. Hobbs, J. T. Londergan, D. P. Murdock, and A. W. Thomas, [Phys. Lett. B **698**, 123 \(2011\)](#).
- [66] I. Albayrak *et al.*, [Nucl. Instrum. Methods Phys. Res., Sect. A **1062**, 169190 \(2024\)](#).
- [67] T. Hobbs and W. Melnitchouk, [Phys. Rev. D **77**, 114023 \(2008\)](#).
- [68] R. M. Whitehill, M. M. Dalton, T. Liu, W. Melnitchouk, and N. Sato, [arXiv:2606.26056](#).