Physical Implications of the Extrapolation and Statistical Bootstrap of Nucleon Structure Function Ratios $F_2^n/F_2^p$ for Mirror Nuclei $^3$He and $^3$H

Hannah Valenty,1 Jennifer Rittenhouse West,2 Fatiha Benmokhtar,1,*
Douglas W. Higinbotham,3 Asia Parker,1 and Erin Seroka4

1Duquesne University, Pittsburgh, PA 15282, USA
2Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
3Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA
4The George Washington University, Washington, DC 20052

(Dated: March 7, 2023)

A nuclear physics example of statistical bootstrap is used on the MARATHON nucleon structure function ratio data in the quark momentum fraction regions $x_B \to 0$ and $x_B \to 1$. The extrapolated $F_2$ ratio as quark momentum fraction $x_B \to 1$ is $F_2^n/F_2^p \to 0.4 \pm 0.05$ and this value is compared to theoretical predictions. The extrapolated ratio when $x_B \to 0$ favors the simple model of isospin symmetry with the complete dominance of seaquarks at low momentum fraction. At high-$x_B$, the proton quark distribution function ratio $d/u$ is derived from the $F_2$ ratio and found to be $d/u \to 1/6$. Our extrapolated values for both the $F_2^n/F_2^p$ ratio and the $d/u$ parton distribution function ratio are within uncertainties of perturbative QCD values from quark counting, helicity conservation arguments and a Dyson-Schwinger Equation with contact interaction model. In addition, it is possible to match the statistical bootstrap value to theoretical predictions by allowing two compatible models to act simultaneously in the nucleon wavefunction. One such example is nucleon wavefunctions composed of a linear combination of a quark-diquark state and a 3-valence quark correlated state with coefficients that combine to give the extrapolated $F_2$ ratio at $x_B = 1$.

I. INTRODUCTION

The ratio of the deep inelastic structure functions $F_2^n/F_2^p$ provides fundamental information about the quark distributions of nucleons [1, 2]. In particular, in the limit of the Bjorken scaling variable $x_B$ going to unity, this ratio provides a powerful tool for discriminating between different partonic models. Due to the lack of free neutron targets, it is very challenging to determine the ratio from scattering protons and neutrons; but using mirror nuclei, $^3$H and $^3$He, and exploiting isospin symmetry allows for an accurate determination of this ratio.

The MARATHON experiment at Jefferson Lab has recently published their measured ratios of nucleon structure functions $F_2^n/F_2^p$ with deep inelastic scattering of electrons off of $^3$H and $^3$He nuclei [4–11]. Their results cover the quark momentum fraction range $0.19 < x_B < 0.83$ and significantly improve on previous measurements [1, 12, 13]. The high and low Bjorken-x regions were not covered and yet are necessary, especially in the $x_B \to 1$ regime, in order to distinguish between theoretical predictions for quark and parton behavior at large momentum fraction. The theory predictions differ in their interpretation of the components of the nucleon wavefunction and how they act in the high-$x_B$ regime. One example is the scalar diquark model which predicts that at $x_B = 1$ only the scalar $[ud]$ diquark plus valence quark component contributes, yielding a structure function ratio of $1/4$. Another example is the SU(6) flavor model in which all vector diquarks (spin-1 and isospin-1 combinations) plus valence quark contribute to give a ratio of $2/3$. Table I presents an inexhaustive list of theoretical predictions, while theory reviews from 1996 [14] and 2010 [15] give comprehensive details and references.

Previous global analyses of parton distribution functions (PDFs) at next-to-next-to leading order (NNLO) by the CTEQ-TEA (Coordinated Theoretical/Experimental Project on QCD Phenomenology and Tests of the Standard Model) collaboration give a wide spread of $d/u$ parton distribution function (PDF) ratios [16, 17]. The extrapolation carried out in this work in the high-x region finds $d/u$ values differing from recent fits using the statistical Schlessinger point method, shown in Table II. Our value with uncertainties is within range of three of the theoretical models in Table I. However, it is possible that more than one theoretical model acts simultaneously in the high-$x_B$ region; for example, a quark-diquark configuration in the nucleon in linear combination with a 3-valence quark state can fit our extrapolated value. We discuss this possibility in the next section.

We begin with a theory overview and analysis in the following section and then describe the extrapolation method in detail (Sec. III) before concluding.

* benmokhtar@duq.edu
Reference 0

Page dimensions: 612.0x792.0

| Model                        | $F_2$ ratio prediction | Ref.  |
|------------------------------|------------------------|-------|
| Scalar diquark q[ud]         | 1/4                    | [18, 19] |
| Quark model/lsqur            | 1/4                    | [20] |
| Holographic LFQCD            | 0.282                  | [21, 22] |
| DSE with qq correlations     | 0.33                   | [23]* |
| DSE-II contact interaction   | 0.41                   | [24, 25] |
| pQCD helicity conservation   | 3/7                    | [26] |
| Quark counting rules         | 3/7                    | [27] |
| 3-quark correlations with CJ12MID | $\lesssim 0.51$    | [28]* |
| DSE-I dressed quark masses   | 0.49                   | [24] |
| SU(6) Flavor                 | 2/3                    | [29] |
| 3-quark correlations with MMHT2014 | $\geq 0.68$    | [28]* |

TABLE I. $F_2^n/F_2^p$ Model Predictions as $x \to 1$. An asterisk indicates that the reference gave a value for the valence quark ratio $d/u$, from which the $F_2$ is derived using the isospin symmetry assumptions $u^B = d^n = u$, $d^p = u^o = d$ and assuming the seaquark contribution is negligible.

| Method                        | $F_2$ ratio | Reference |
|------------------------------|-------------|-----------|
| Schlessinger point method     | $\approx 0.45$ | [30]*    |
| CTEQ                         | unconstrained | [16, 17] |
| JAM                          | $\approx 0.4$ but $x \gtrsim 0.8$ | [31]    |

TABLE II. $F_2^p/F_2^p$ fits as $x \to 1$. An asterisk indicates that the reference gave a value for the valence quark ratio $d/u$, from which the $F_2$ is derived using the isospin symmetry assumptions $u^B = d^n = u$, $d^p = u^o = d$ and negligible seaquark contributions.

II. STRUCTURE FUNCTION THEORY

We start with a review of relevant structure function definitions. The internal structure of nucleons is encoded in the form of structure functions $F_2(x_B, Q^2)$ describing quark and antiquark behavior in terms of parton momentum distribution functions weighted by the square of their electric charge.

Structure functions are components of the deep-inelastic scattering differential cross-section for charged lepton scattering, e.g. $eN \to eX$,

$$\frac{d^2\sigma}{dt \, du} = \frac{4\pi\alpha^2}{t^2} \frac{1}{2} \frac{1}{s^2(s + u)} \left[2x_F(s + u)^2 - 2usF_2\right],$$

(1)

given in terms of the Mandelstam variables for the 4-momenta of the incoming lepton $p_1$, target $p_2$, outgoing lepton $p_3$, and hadronized debris $p_4$, $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$, $u = (p_1 - p_4)^2$. The nucleon structure function $F_2$ is defined (up to leading order in the strong force coupling $\alpha_S$) as

$$F_2(x_B) = 2x_BF_1(x_B) = \sum_i Q_i^2x(q_i(x_B) + \bar{q}_i(x_B))$$

(2)

where the sum is over quark flavors, $Q_i^2$ is the charge on the $i$th flavor of quark, $q_i(x_B)$ are the quark momentum distribution functions, and Bjorken-x is the fraction of nucleon momentum carried by the struck quark. Bjorken-x is given in terms of experimental variables

$$x_B = \frac{Q^2}{2M_T\nu},$$

(3)

where $\nu$ is the energy lost by the lepton, $E - E'$, $M_T$ is typically the mass of the struck nucleon in the target, $Q^2$ is minus the square of the 4-momentum transfer (the virtual photon 4-momentum squared) $Q^2 = -q^2 = 2EE'(1 - \cos \theta)$ where $\theta$ is the lepton scattering angle. The subscript $B$ will be dropped for notational simplicity from here on out.

The ratio of neutron to proton structure functions $F_2^n/F_2^p$ as measured by the MARATHON experiment and in the regions extrapolated by the statistical bootstrap method on MARATHON data are shown in Fig.2 and Fig.3, respectively. Physical implications and simplifying assumptions of the low and high Bjorken-x behavior are discussed next.
A. Isospin symmetry analysis of extrapolated behavior of $F^n_2/F^p_2$: $x \to 0$ region

In the low-$x$ region, the process of gluon bremsstrahlung following by gluon splitting to quark-antiquark pairs dominates the nucleon structure function due to a $\frac{dk}{k}$ probability for gluon splitting, Fig.1 [32]. The structure function in this region

$$F_2(x) = \sum_i Q_i^2 x(q_i(x) + \bar{q}_i(x))$$

contains contributions from the light quark $q\bar{q}$ pairs, $u\bar{u}, d\bar{d}, s\bar{s}$. Gluon splitting to the heavy mass quarks $c, b, t$ is neglected as these are highly virtual processes and contribute negligibly to the scattering amplitude.

![Fig. 1. Example of low-x gluon splitting in a 5-quark Fock state of a nucleon. Adapted from [33].](image)

The structure functions are therefore written as

$$\frac{1}{x} F^n_2(x) = \frac{4}{9} u^n(x) + \frac{1}{9} d^n(x) + \frac{4}{9} [u^n_s(x) + \bar{u}^n_s(x)]$$

$$+ \frac{1}{9} [d^n_s(x) + \bar{d}^n_s(x)] + \frac{1}{9} [s^n_s(x) + \bar{s}^n_s(x)]$$

and

$$\frac{1}{x} F^p_2(x) = \frac{4}{9} u^p(x) + \frac{1}{9} d^p(x) + \frac{4}{9} [u^p_s(x) + \bar{u}^p_s(x)]$$

$$+ \frac{1}{9} [d^p_s(x) + \bar{d}^p_s(x)] + \frac{1}{9} [s^p_s(x) + \bar{s}^p_s(x)]$$

with subscript $s$ denoting sea quark distributions.

The following assumptions are now made for the quark momentum distribution functions. Strong isospin symmetry assumes $u$ and $d$ form an SU(2)$_{iso}$ doublet (as do $p$ and $n$) which implies the relations [33]

$$u^p = d^n \equiv u$$

$$d^p = u^n \equiv d$$

$$s^p = s^n \equiv s.$$ (7)

The rest of the analysis is therefore given in terms of the proton’s parton distribution functions, the same approximations as used in MARATHON [4].

The structure functions can be rewritten as

$$\frac{1}{x} F^n_2 = \frac{4}{9} u + \frac{1}{9} d + \text{sea terms}$$

$$\frac{1}{x} F^p_2 = \frac{4}{9} d + \frac{1}{9} u + \text{sea terms}.$$ (8)

For simplicity, assume that the sea quark distributions are the same, as in [33],

$$u_s(x) = \bar{u}_s(x) = d_s = \bar{d}_s = s_s = \bar{s}_s \equiv K.$$ (9)

This assumption does not affect the results.

Then we have

$$\frac{1}{x} F^n_2 = \frac{1}{9} (u + 4d) + \frac{12}{9} K$$

$$\frac{1}{x} F^p_2 = \frac{1}{9} (d + 4u) + \frac{12}{9} K.$$ (10)
When sea quark momentum distribution functions dominate the structure functions, $K \gg u,d$, we find
\[ \frac{F_n^u}{F_p^u}(x) \to 1 \text{ for } x \to 0. \tag{11} \]
This conclusion agrees with the extrapolated value shown in Fig.3. Note that it does not depend on the assumption of equal sea quark distribution functions. It requires only that the sum of sea quark distribution functions to be the same for the proton and the neutron, which is essentially a sea quark version of the isospin symmetry assumptions of Eq. 7.

B. Isospin symmetry analysis of extrapolated behavior of $F_n^u/F_p^u$: $x \to 1$ region

In the high-x region, for $x \to 1$, the extrapolated ratio $F_n^u/F_p^u \to 0.4 \pm 0.05$. The sea quark distribution functions become negligible in this region of Bjorken-x and the structure functions are given by
\[ \frac{1}{x} F_n^u = \frac{1}{9} (u + 4d) \]
\[ \frac{1}{x} F_p^u = \frac{1}{9} (d + 4u) \tag{12} \]
where the symmetries of Eq.7 are again assumed.

For this analysis, we set the ratio equal to the extrapolated value of 0.4 and solve for the ratio of quark distribution functions which may then be compared to data. The result is
\[ \frac{F_n^u}{F_p^u} = \frac{\frac{1}{9} (u + 4d)}{\frac{1}{9} (d + 4u)} = 0.4 \Rightarrow \frac{d}{u} = \frac{1}{6} \tag{13} \]
This value is close to the lower limit of recent work [30],
\[ \frac{d}{u} \big|_{x \to 1} = 0.230(57) \tag{14} \]
and is due only to the isospin symmetry arguments of Eq.7 and the vanishing of sea quark distributions at high-x. Earlier theoretical work argued that a scalar $[ud]$ diquark forms within the nucleon in the $x_B \to 1$ limit as a way to lower the energy of the system [19] (Table I). In this version of the scalar diquark model, the highest probability state with a single quark containing all of the nucleon momentum must minimize the energy of the remaining constituents. Therefore, as $x \to 1$, the lowest energy proton state is that which has the $u$ quark carrying all the longitudinal momentum with the remaining valence quarks forming the spin-0 isospin-0 $[ud]$ diquark [34] with an estimated binding energy of nearly 150 MeV [35]. Similarly for the neutron with momentum distribution functions swapped $u \leftrightarrow d$, which gives $d/u = 0$ in the limit of $x \to 1$. The structure function ratio in this case is given by
\[ \lim_{x \to 1} \frac{F_n^u(x)}{F_p^u(x)} = 0.25. \tag{15} \]
Our analysis finds the limiting behavior
\[ \lim_{x \to 1} \frac{F_n^u(x)}{F_p^u(x)} = 0.4 \pm 0.05 \tag{16} \]
which disfavors this value. We most closely match the perturbative QCD (pQCD) values using helicity conservation, hard gluon exchange [26] and quark counting rules [27]. The latter two models give the same ratio limit,
\[ \lim_{x \to 1} \frac{F_n^u(x)}{F_p^u(x)} = \frac{3}{7} \approx 0.43, \tag{17} \]
and these are within the bootstrap uncertainties as shown by the blue band of Fig.3. There is another physical implication, however, because while some models are incompatible with each other, it is possible that two compatible models could act over the same physical region to give the extrapolated value. One such example is the scalar diquark model with $F_n^u/F_p^u \to 0.25$ [19] together with the 3-quark correlation plus MMHT2014NNLO with $F_n^u/F_p^u \geq 0.68$ (or CJ12MID with ratio $\lesssim 0.51$) from [28], with different weights given to scalar diquark vs. 3-quark correlation contributions in order to yield a final ratio of 0.4. The nuclear wave function for $A=3$ nuclei has been proposed as a linear combination of nucleons in quark-diquark and 3-valence quark configurations, with a range of predictions for short-range correlations [35] which matches the data [36]. We now move on to describing the extrapolation method.
FIG. 2. Deep inelastic scattering $F_2^n/F_2^p$ ratios vs. $x_B$ from MARATHON [4] compared to theoretical predictions. Error bars include overall systematic uncertainties. All curves correspond to MARATHON kinematics. Plot credit: T.Kutz.

FIG. 3. Fit of the Marathon data making use of a standard extrapolating function with the uncertainty band determined from statistical bootstrapping. The result at $x_B = 1$ indicates a ratio of $0.4 \pm 0.05$.

III. STATISTICAL BOOTSTRAP METHOD

Bootstrapping is a technique that allows an estimation of the sampling distribution of almost any statistical distribution using the method of sampling data with replacement. It is any test or metric that falls under the broader class of re-sampling methods. Bootstrapping assigns measures of accuracy such as variance, confidence interval and prediction errors to sample estimates [37].

In the case of the MARATHON experiment, a rational function $R_f$ of the form of

$$R_f = n_0 \frac{1 + x n_1}{1 + x m_1}$$

was used to fit the 22 data points corresponding to $x_B$ values between 0.195 and 0.825. $n_0$, $n_1$ and $m_1$ are the three fit parameters of this function. The fit was then extrapolated to lower and higher limits of $x_B$ to cover the full range from 0 to 1.

One fit line is created by randomly selecting a set of twenty-two points from the 22 existing data values allowing one, two or three substitutions of any of the data points with a randomly chosen value from the same data set. A total of one-thousand fits were performed using this method then plotted to determine the error band. The fits were then extrapolated to lower and higher values of $x_B$ to cover the full range from $x_B \to 0$ to $x_B \to 1$. Panels of selected fits using this statistical bootstrap method are presented in Fig.4. The blue error band shown in Fig.3 is the result of the accumulation of one-thousand such fits.
IV. CONCLUSION

We have used a statistical bootstrap of the MARATHON deep inelastic scattering data to extrapolate the high and low quark momentum fraction regimes of the $F_n^2/F_P^2$ ratio. As $x_B \to 1$, the ratio $F_n^2/F_P^2$ approaches $0.4 \pm 0.05$, which, under the usual isospin symmetry arguments for valence quark parton distribution functions of $u^p = d^n = u$ and $d^p = u^n = d$, corresponds to a ratio of $d/u \to 1/6$. Our extrapolation favors the perturbative QCD helicity (pQCD) conservation, quark counting rules and DSE with contact interactions models. The scalar diquark model...
value of the $x_B \to 1$ limit $F_2^a/F_2^P = 1/4$ is disfavored as are all other models from Table I. We note that while many models are incompatible with each other, it is possible that two compatible models could act over the same physical region to give the extrapolated value. One example is the scalar diquark model with $F_2^a/F_2^P \to 0.25$ together with the 3-quark correlation plus MMHT2014NNLO with $F_2^a/F_2^P \geq 0.68$ from [28], with different weight given to the scalar diquark vs. 3-quark correlation contributions in order to yield a final ratio of 0.4. Such a scenario would agree with recent results on the nucleon wavefunction as a linear combination of a 3-valence quark state and a quark-diquark state with unequal coefficients in $A=3$ nuclei [35, 36].

ACKNOWLEDGMENTS

We thank Wally Melnitchouk for helpful discussions and Tyler Kutz for his MARATHON plot. H.V and F.B acknowledge support from the National Science Foundation, Award No.Benmokhtar-2012413. J.R.W is supported by the LDRD programs of LBNL, the EIC Center at Jefferson Lab and by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under contract number DE-AC02-05CH11231. This research was funded in part by Department of Energy grant number DE-AC05-06OR23177 under which the Jefferson Science Associates operates the Thomas Jefferson National Accelerator Facility.
[33] F. E. Close, *An Introduction to Quarks and Partons* (Academic Press, 1979).

[34] R. L. Jaffe, Exotica, *Phys. Rept.* **409**, 1 (2005), arXiv:hep-ph/0409065.

[35] J. Rittenhouse West, Diquark induced short-range nucleon-nucleon correlations & the EMC effect, *Nucl. Phys. A* **1029**, 122563 (2023), arXiv:2009.06968 [hep-ph].

[36] S. Li et al., Revealing the short-range structure of the mirror nuclei $^3\text{H}$ and $^3\text{He}$, *Nature* **609**, 41 (2022).

[37] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall/CRC, 1993).