Deep inelastic scattering from A = 3 nuclei: Nucleon structure and the EMC effect

A Dissertation Presented

by

Tyler Thomas Kutz

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Physics

Stony Brook University
December 2019

Stony Brook University
The Graduate School

Tyler Thomas Kutz

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Krishna Kumar
Professor, Department of Physics & Astronomy
Stony Brook University

Abhay Deshpande
Professor, Department of Physics & Astronomy
Stony Brook University

Dominik Schneble
Associate Professor, Department of Physics & Astronomy
Stony Brook University

Sergey Syritsyn
Assistant Professor, Department of Physics & Astronomy
Stony Brook University

Jin Huang
Physicist, Physics Department
Brookhaven National Laboratory

This dissertation is accepted by the Graduate School

Eric Wertheimer
Dean of the Graduate School

Abstract of the Dissertation

Deep inelastic scattering from A = 3 nuclei: Nucleon structure and the EMC effect

by

Tyler Thomas Kutz

Doctor of Philosophy

in

Physics

Stony Brook University

2019

The MARATHON experiment ran at Jefferson Laboratory in 2018, measuring deep inelastic scattering of 10.6 GeV electrons from the proton (p), deuterium (D), tritium (^{3}H), and helium-3 (^{3}He). MARATHON's experimental observables are cross section ratios, most notably ^{3}H/^{3}He. This ratio of mirror nuclei is used for a novel extraction of the nucleon structure function ratio F_2^n/F_2^p with minimal nuclear corrections. Determining leading-twist F_2^n/F_2^p at high x provides important constrains on the ratio of nucleon valence quark distributions d/u. The ratios ^{3}H/D and ^{3}He/D allow a determination of the EMC effect in the lightest nuclei sensitive to possible isospin dependence of nuclear effects. Using MARATHON's F_2^n/F_2^p to correct for unequal number of protons and neutrons, the isoscalar EMC ratios for tritium and helium-3 are obtained. Lastly, the D/p ratio allows a comparison of the MARATHON experiment to SLAC measurements, and provides data to which the A = 3 ratios can be renormalized. The MARATHON results emphasize the importance of higher-twist effects in extracting d/u from DIS measurements, and can help constrain said effects.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Most ordinary matter in the universe is made up of hadrons, composite particles of quarks bound by gluons. Notable among hadrons are the nucleons - protons and neutrons - which are the constituents of atomic nuclei. The discovery of the proton and neutron in the early 20th century was largely driven by a desire to explain the structure of matter (in this case, nuclei) in terms of fundamental constituents. It was not long before this same strategy to describe the structure of hadrons led to the prediction and discovery of quarks.

1.1 Quarks, gluons, and QCD

Quarks were proposed independently by Murray Gell-Mann [1] and George Zweig [2] in 1964. The motivation for the quark model was to provide an underlying structure to the proliferation of discovered hadrons, based on the flavor SU(3) symmetry group of the three lightest quark flavors: up, down, and strange. Flavor SU(3) is not a fundamental symmetry and was ultimately found to be broken. In 1968, pioneering experiments at the Stanford Linear Accelerator (SLAC) found evidence for small, point-like constituents in the proton. Due to initial hesitation that these point-like objects were the quarks of Gell-Mann and Zweig, they were referred to as "partons", a general term for hadronic constituents still used today. In the subsequent decades, a total of six quark flavors were theorized and discovered.

Flavor	Mass	Charge (e)	Flavor quantum number
up (u)	$2.16^{+0.49}_{-0.26}$ MeV	$+\frac{2}{3}$	$I_z = +\frac{1}{2}$
down (d)	$4.67^{+0.48}_{-0.17}$ MeV	$-\frac{1}{3}$	$I_z = -\frac{1}{2}$
strange (s)	93^{+11}_{-5} MeV	$-\frac{1}{3}$	Strangeness = -1
charm (c)	1.27 ± 0.02 GeV	$+\frac{2}{3}$	Charm = +1
bottom (c)	$4.18^{+0.03}_{-0.02}$ GeV	$-\frac{1}{3}$	Bottom = -1
top (t)	172.9 ± 0.4 GeV	$+\frac{2}{3}$	Top = +1

Table 1.1: Properties of the six known quarks.
All quarks have spin $S = \frac{1}{2}$ and baryon number $B = \frac{1}{3}$. Produced from [3].

The properties of the known quarks are listed in Table 1.1.

The theory of quarks, gluons, and the strong interaction is quantum chromodynamics (QCD), so named for the color charge carried by quarks and gluons. QCD arises from the local invariance of the color SU(3) symmetry group. Each quark flavor belongs to the fundamental color triplet, where the three states are identified with three values of color charge: red (r), green (g), and blue (b). Antiquarks are expressed in the conjugate representation, with charges antired ($\overline{r}$), antigreen ($\overline{g}$), and antiblue ($\overline{b}$). The massless gauge bosons of QCD, the gluons, belong to the color octet, carrying both color and anticolor charge. The strong interaction between quarks and gluons, sometimes called the "color force", is mediated by the exchange of gluons.

Color-charged particles are only observed as constituents of colorless objects, or more precisely, in color singlet states. This is known as color confinement. This can be achieved by combining one of each color or anticolor (rgb or $\overline{rgb}$) to form hadrons with baryon number $B = 1$, the baryons. Alternatively, combining color and anticolor ($r\overline{r}$, $g\overline{g}$, or $b\overline{b}$) results in hadrons with $B = 0$, the mesons. Exotic colorless states involving more than three quarks are theoretically allowed.

QCD has many unique features. Its gauge bosons themselves carry color charge, meaning

that in addition to quark-gluon couplings, gluon-gluon couplings are also allowed. This is in contrast to the photons of quantum electrodynamics (QED), which are electrically neutral and cannot couple to themselves. At low energy scales typical of nuclei, the strong coupling α_s is of order unity. This precludes the traditional perturbative approach of quantum field theory, which expands interactions in powers of the coupling. For $\alpha \ll 1$, such as in QED, such expansions rapidly converge. For $\alpha \approx 1$, they do not. This has led to the development of lattice QCD, which carries out low-energy QCD calculations on a spacetime lattice, without the need for a perturbative expansion. However, QCD is an asymptotically free theory. As energy scales increase and length scales decrease, the running coupling becomes weaker, eventually reaching $\alpha_s \ll 1$. This splits QCD into two regimes: a high-energy region where perturbative techniques are valid, and a low-energy region where lattice calculations are required.

1.2 Quarks in the nucleon

The nucleons are baryons, color singlet states formed by three quarks collectively carrying one of each color charge. In the proton, these three quarks have flavor uud, and in the neutron, udd. However, this simple picture belies the true nature of the nucleon's parton structure. The nucleons contain gluons, arising from both the exchange of gluons mediating the strong interaction (responsible for binding the quarks), and from the perturbative emission of gluons by quarks. These gluons can radiate additional gluons or split into quark-antiquark pairs. Thus, the "valence" quarks composing the color singlet actually exist in a "sea" of radiated gluons, quarks, and antiquarks.

Developing a rigorous description of the nucleons in terms of these quarks, antiquarks, and gluons is currently an active area of research. How do properties of the nucleon, such as mass and spin, emerge from partons? Of particular interest, and perhaps the most accessible experimentally, are fractional momentum distributions of partons. That is, what

fraction of a nucleon's momentum is carried by different types of partons? How does this change when nucleons are bound inside a nucleus? The momentum distributions of free and bound nucleons can provide important insight into the nature of QCD interactions. Precise measurements of these momentum distributions are critical to unraveling the emergence of nucleons and nuclei from QCD.

Illuminating this momentum structure, particularly in regions where a single quark carries most of the nucleon momentum, is one of the goals of the MARATHON[1] experiment, which ran spring 2018 in Experimental Hall A at Jefferson Laboratory. MARATHON studied high-energy scattering of electrons from hydrogen (proton), deuterium, tritium, and helium-3. The experimental observables are cross section ratios between various pairs of these nuclei. These ratios are sensitive to the parton structure of free nucleons, and how the structure of nucleons is modified when bound in nuclei (known as the EMC effect). MARATHON aims to directly extract (up to some theoretical model dependence) the ratio of free nucleon structure functions F_2^n/F_2^p from the tritium to helium-3 ratio, and to measure the EMC effect in A = 3 nuclei with the ratios of tritium and helium-3 to deuterium. With input from theory, F_2^n/F_2^p can be used to determine d/u, the ratio of down to up quarks in the nucleon. In addition to allowing the direct extraction of these quantities, the MARATHON data will add to the global data set that is used by many to perform fits of theoretical or phenomenological models to experimental data. This is particularly important for helium-3 and tritium, the lightest possible nuclei with mirror isospin symmetry. Tritium's radioactive nature complicates measurements, and therefore very few data on this nucleus currently exist.

Chapter 2 will discuss theoretical and experimental aspects of deep inelastic scattering, nucleon structure, and the EMC effect. Chapter 3 provides motivation for the MARATHON experiment and describes the apparatus used to carry out the measurement. Chapter 4 details the analysis procedure, with the various cuts and corrections developed to extract

[1] An "acronym" loosely derived from: **M**e**A**surement of the F_2^n/F_2^p, d/u **RA**tios and A = 3 EMC effect in deep inelastic scattering off the **T**ritium and **H**elium mirr**O**r **N**uclei.

the cross section ratios. Finally, Chapter 5 will show the experimental results and subsequent interpretation using input from theoretical models.

Chapter 2

Nucleon structure

2.1 Deep inelastic scattering

Since the discovery of the atomic nucleus, scattering experiments continue to be the most powerful tool for probing subatomic structure. A commonly used process is the electromagnetic scattering of an electron from a nucleus, where a beam of electrons incident on target nuclei scatter via the exchange of a virtual photon. The electron-photon vertex is well understood in QED, allowing clean separation of the nucleus-photon vertex and information on nuclear structure contained therein. Given the relatively weak strength of QED interactions, the process can be approximated by the exchange of a single photon, with only a small correction required to account for higher-order processes.

Figure 2.1 depicts a typical scattering process of an electron from a nucleus A. The initial and final electron four-momenta are $k = \left(E,\ \overline{k}\right)$ and $k' = \left(E',\ \overline{k}'\right)$, respectively. The four-momentum transfer between the electron and nucleus is carried by the photon, $q = k - k'$. P is the four-momentum of the target nucleus; for a fixed-target experiment, $P = (M_A, 0)$, where M_A is the mass of the nucleus. The final hadronic state X is, in general, not the same as the incoming nucleus A.

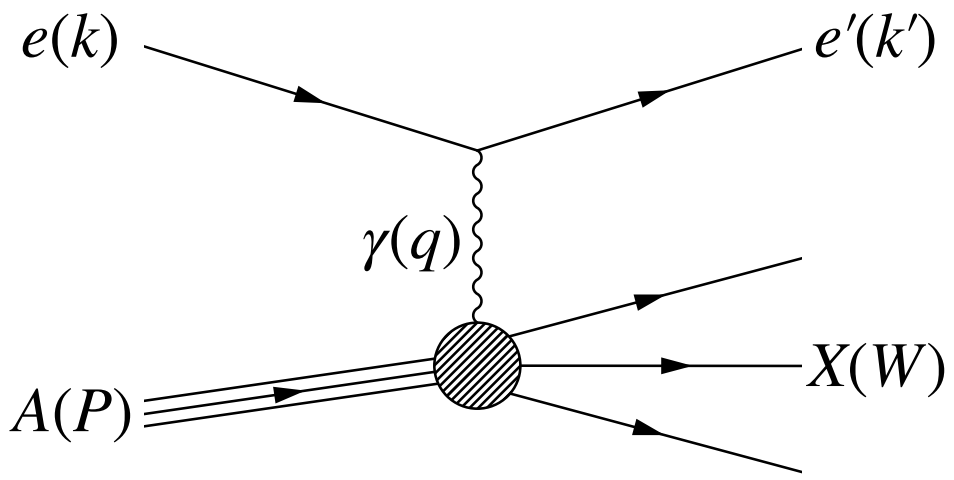


Figure 2.1: Feynman diagram for DIS process

It is useful to define several Lorentz invariant quantities to describe the scattering process:

- $\nu = (q \cdot P)/M = E - E'$ is the electron's energy loss.
- $Q^2 = -q^2 \approx 4EE'(1 - \cos\theta)$ is the momentum transfer squared. The electron mass is neglected, a valid approximation in the ultra-relativistic limit.
- $W^2 = (P + q)^2 = M_A^2 + 2M_A\nu - Q^2$ is the mass of the recoiling system.

The sensitivity of a scattering process scales with the wavelength of the photon probe, which is inversely proportional to the photons energy and momentum (ν and Q^2). Therefore, selecting processes with certain kinematics will change the effective degrees of freedom observed in the interaction. Low-energy elastic processes, where the electron scatters from the nucleus as a whole, are sensitive to the extended structure of the nucleus but not the individual nucleons within. Quasielastic processes occur in the low-energy limit of inelastic scattering, where the electron effectively scatters from quasi-free protons and neutrons. Inelastic processes with sufficiently large energy and momentum transfer become sensitive to the quarks and gluons within the nucleon. Here, it is common to rescale kinematic variables in terms of the proton mass M, rather than the nuclear mass M_A. These processes are referred to as deep inelastic scattering (DIS), as the exchanged photon is probing deep within the nucleon ($Q^2 \gg M^2$) in highly inelastic processes ($W^2 \gg M^2$).

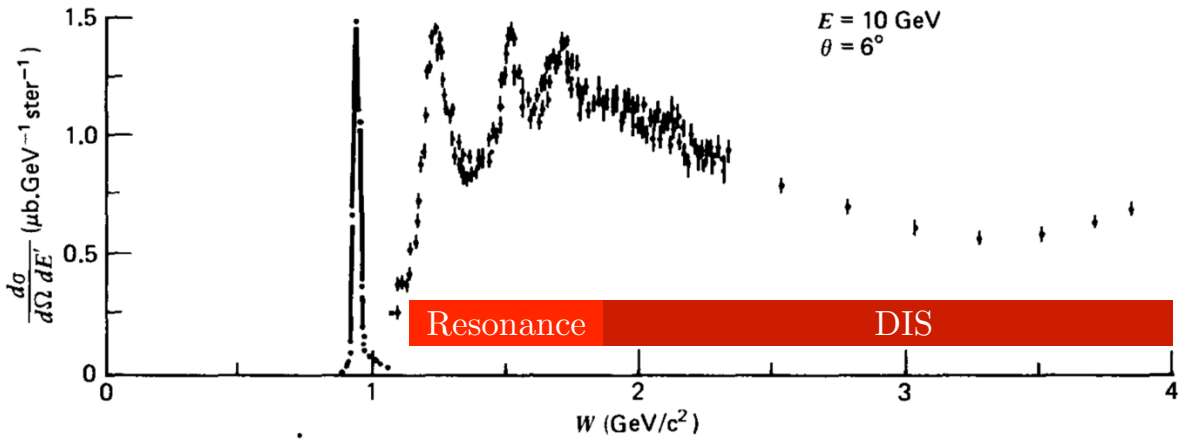


Figure 2.2: Final-state mass W spectrum of electron-nucleon scattering. Produced from [4].

The differential proton cross section as a function of W is shown in Figure 2.2. The elastic peak occurs at the proton mass, $W \approx M$. The resonance region contains peaks resulting from discrete excited proton states. Beyond the resonances, the complicated multi-body debris left by interactions with large W results in a smooth spectrum. This is the DIS region.

2.1.1 DIS differential cross section

The inclusive $eA \to eX$ cross section is given by the contraction of the leptonic and hadronic tensors:

$$\frac{d^2\sigma}{dE'\,d\Omega} = \frac{\alpha^2}{Q^4}\frac{E'}{E}L_{\mu\nu}W^{\mu\nu} \tag{2.1}$$

where α is the fine structure constant. Both tensors are process-dependent. The leptonic tensor is constructed from the independent momenta at the lepton vertex, k and k'. For single photon exchange,

$$L_{\mu\nu} = 2(k_\mu k'_\nu + k'_\mu k_\nu - k \cdot k' g_{\mu\nu}). \tag{2.2}$$

The hadronic tensor is constructed from the independent momenta at the hadron vertex, P and q. In contrast to the tensor for point-like leptons, the hadronic tensor for single photon

exchange contains independent structure functions W_1 and W_2,

$$W^{\mu\nu} = \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) W_1(\nu, Q^2) + \left(P^\mu - \frac{P\cdot q}{q^2} q^\mu\right)\left(P^\nu - \frac{P\cdot q}{q^2} q^\nu\right) \frac{W_2(\nu, Q^2)}{M^2}. \quad (2.3)$$

The structure functions parameterize our ignorance of the hadron's structure. In general, they are functions of any two independent Lorentz scaler variables, typically chosen to be ν and Q^2.

With the tensors defined, the DIS cross section can be written as

$$\frac{d^2\sigma}{dE'\, d\Omega} = \frac{4\alpha^2 E'^2}{Q^4} \cos^2\frac{\theta}{2}\left[W_2(\nu, Q^2) + 2W_1(\nu, Q^2)\tan^2\frac{\theta}{2}\right] \quad (2.4)$$

2.1.2 Scaling and the quark parton model

As DIS requires very large energy and momentum transfer, it is useful to consider the limit where $Q^2 \to \infty$, while Q^2/ν remains fixed. It was demonstrated by Bjorken [5], based on the sum rules of Callan and Gross [6], that in this asymptotic limit the structure functions exhibit scaling behavior,

$$MW_1(\nu, Q^2) \to F_1(x) \quad (2.5)$$

$$\nu W_2(\nu, Q^2) \to F_2(x) \quad (2.6)$$

where $x = Q^2/2M\nu$ is the Bjorken scaling variable. That is, the structure functions no longer depend independently on ν and Q^2, but on their ratio.

This leads to a useful, though ultimately incomplete, description of hadrons: the quark parton model (QPM). The QPM requires working in the infinite momentum frame (IMF), where the hadron is moving with infinite momentum. This satisfies several critical conditions. First, the IMF corresponds to the $Q^2 \to \infty$ limit required for the onset of scaling. Second, in the IMF the momenta of the parton constituents are parallel to the momentum of the

parent hadron, with a negligible transverse component. Lastly, in the IMF the interaction between the virtual photon and the quark occurs over a timescale much shorter than that of interactions between quarks. Thus, from the photon's perspective, the quarks are non-interacting.

Consider a hadron in the QPM with four-momentum $P = (p, 0, 0, p)$. A quark constituent carrying a fraction y of the hadron's momentum has four-momentum $yP = (yp, 0, 0, yp)$. Therefore, after a scattering interaction via the exchange of a virtual photon with four-momentum q, the quark's invariant momentum is:

$$\begin{aligned}(yP + q)^2 &= (yP)^2 - Q^2 + 2yq \cdot p \\ &= (yP)^2 - Q^2 + 2yM\nu \end{aligned} \tag{2.7}$$

The initial and final invariant momenta of the quark are small compared to Q^2 and $q \cdot p$ and can be neglected. This yields $y = Q^2/2M\nu$, the quantity previously identified as the Bjorken scaling variable x.

In the QPM, the exchanged virtual photon sees the hadron as a collection of non-interacting quarks, and can only couple to quarks carrying momentum fraction x. Thus, the hadronic cross section must be the incoherent sum of constituent quark cross sections and is dependent on the fractional momentum distributions of the quarks. The leads to structure functions that are charge-weighted sums of the quark fractional momentum distributions:

$$F_1(x) = \frac{1}{2} \sum_q e_q^2 [q(x) + \overline{q}(x)] \tag{2.8}$$

$$F_2(x) = x \sum_q e_q^2 [q(x) + \overline{q}(x)] \tag{2.9}$$

where the sum is over quark flavor q and e_q is Coulomb charge. The parton distribution functions (PDFs) $q(x, Q^2)$ and $\overline{q}(x, Q^2)$ describe the probability of finding a quark or antiquark of flavor q carrying fractional momentum x.

Physically, the onset of scaling can be thought of as resolving the smallest possible structure within the nucleon. If Q^2 is sufficiently large to resolve the point-like quarks, there is no additional structure to be resolved, and the cross section should become scale-invariant. At SLAC in 1968, *ep* DIS measurements revealed very weak Q^2 dependence of the cross section. This observation is considered to be the discovery of the nucelon's point-like constituents. In terms of the structure functions F_1 and F_2, which are functions of the scaling variable x, the DIS cross section is:

$$\frac{d^2\sigma}{dE'\,d\Omega} = \frac{4\alpha^2 E'^2}{Q^4}\cos^2\frac{\theta}{2}\left[\frac{F_2}{\nu} + \frac{2F_1}{M}\tan^2\frac{\theta}{2}\right] \tag{2.10}$$

2.1.3 Scaling violation

The simplicity of the QPM is useful, but neglects the nature of QCD. Within the nucleon, the quarks radiate gluons, which in turn radiate additional gluons or split into quark-antiquark pairs, making non-negligible contributions to the PDFs and structure functions. These perturbative processes depend on the strong coupling α_s, which is dependent on scale (in this case, Q^2). This results in PDFs and structure functions that are, in fact, sensitive to the scale at which they are probed. The scale dependence of PDFs is described by the DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equation,

$$\frac{d}{d\ln Q^2}q(x,Q^2) = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y}q(y,Q^2)P_{qq}\left(\frac{x}{y}\right) + \mathcal{O}(\alpha_s^2), \tag{2.11}$$

where the splitting function $P_{qq}(x/y)$ describes the probability that a quark carrying momentum fraction x resulted from a quark carrying momentum fraction $y > x$ that lost energy by radiating a gluon. The splitting functions for this and other processes (P_{gq}, P_{qg}, P_{gg}) are calculated in QCD and also depend on Q^2. It is important to note that QCD cannot predict the PDFs *a priori*. However, given a PDF measured at an input scale μ^2, the DGLAP equations can accurately evolve the PDFs to another scale Q^2.

While the quarks can be considered non-interacting during DIS processes, QCD reveals that they are surrounded by a perturbatively-radiated cloud of gluons, quarks, and anti-quarks. As the Q^2 scale increases, more and more of these previously virtual partons are resolved. The critical point: while the QPM is a function of x, QCD is a function of Q^2. The DIS cross section is still expressed by Equation (2.10), with the understanding that F_1 and F_2 (and the related PDFs) are functions of both x and Q^2.

2.1.4 Nuclear photoabsorption ratio $R = \sigma_L/\sigma_T$

The DIS cross section can equivalently be viewed as the probability that the exchanged virtual photon is absorbed by the nucleus. In this case, the cross section is expressed as

$$\frac{d^2\sigma}{dE'\,d\Omega} = \Gamma\left[\sigma_T(x,Q^2) + \epsilon\,\sigma_L(x,Q^2)\right], \tag{2.12}$$

where σ_T and σ_L are the nuclear absorption cross sections for transversely and longitudinally polarized virtual photons, respectively. The flux of transverse photons is

$$\Gamma = \frac{\alpha}{2\pi^2 Q^2}\frac{E'}{E}\frac{\nu(1-x)}{1-\epsilon} \tag{2.13}$$

and the ratio of longitudinal to transverse photons is

$$\epsilon = \left[1 + 2\left(1 + \frac{\nu^2}{Q^2}\right)\tan^2\frac{\theta}{2}\right]. \tag{2.14}$$

It is common to rewrite the cross section as

$$\frac{d^2\sigma}{dE'\,d\Omega} = \Gamma\,\sigma_T(1+\epsilon\,R), \tag{2.15}$$

where $R = \sigma_L/\sigma_T$ is ratio of the transverse to longitudinal photabsorption cross sections.

Equating this picture with the one presented in Section 2.1.1 yields the following rela-

tionship between the structure functions and photoabsorption cross sections:

$$F_1 = \frac{M\nu(1-x)}{4\pi^2\alpha}\sigma_T \tag{2.16}$$

$$F_2 = \frac{\nu^2(1-x)}{4\pi^2\alpha(1+\nu^2/Q^2)}\,\sigma_T(1+R) \tag{2.17}$$

The F_1 structure function is purely transverse, while the F_2 structure function is a combination of transverse and longitudinal. Thus, individually determining the structure functions requires a "separation" of the longitudinal and transverse contributions to the cross section. To obtain F_2 from a measured cross section, knowledge of R is required. To obtain the F_2 ratio of two nuclei from a measured cross section ratio, knowledge of the nuclear dependence of R is required. The L/T separation is achieved using a Rosenbluth separation style technique, where the reduced cross section $(d^2\sigma/dE'\,d\Omega)/\Gamma$ is measured at multiple values of ϵ and fixed x, Q^2. The intercept and slope of a linear fit yield σ_T and R. Cross section ratios of nucleus A and B are sensitive to the difference in R,

$$\frac{d\sigma_A}{d\sigma_B} \approx \frac{\sigma_A^T}{\sigma_B^T}\,[1+\epsilon'\Delta R] \tag{2.18}$$

where $\epsilon' = \epsilon/(1+\epsilon R_B)$ and $\Delta R = R_A - R_B$. This requires an absolute value of R for one of the nuclei (typically nucleus B is deuterium).

Early measurements of ΔR were consistent with zero [7, 8, 9]. This led to the assumption that R was independent of nucleus, and measuring the cross section ratio of two nuclei was equivalent to measuring the ratio of their F_2 structure functions. However, more recent measurements [10, 11] have have cast doubt on this assumption, with ΔR in some cases being $> 2\sigma$ from zero. This has prompted the proposal of a dedicated experiment in Hall C at Jefferson Lab to carry out high-precision L/T separations on several targets ranging from hydrogen to gold. A nuclear dependence of R would have serious implications for the interpretation of past and future cross section ratio measurements. However, for the time

being most cross section ratio analyses continue to assume that R is independent of nucleus.

2.2 Structure of the nucleon

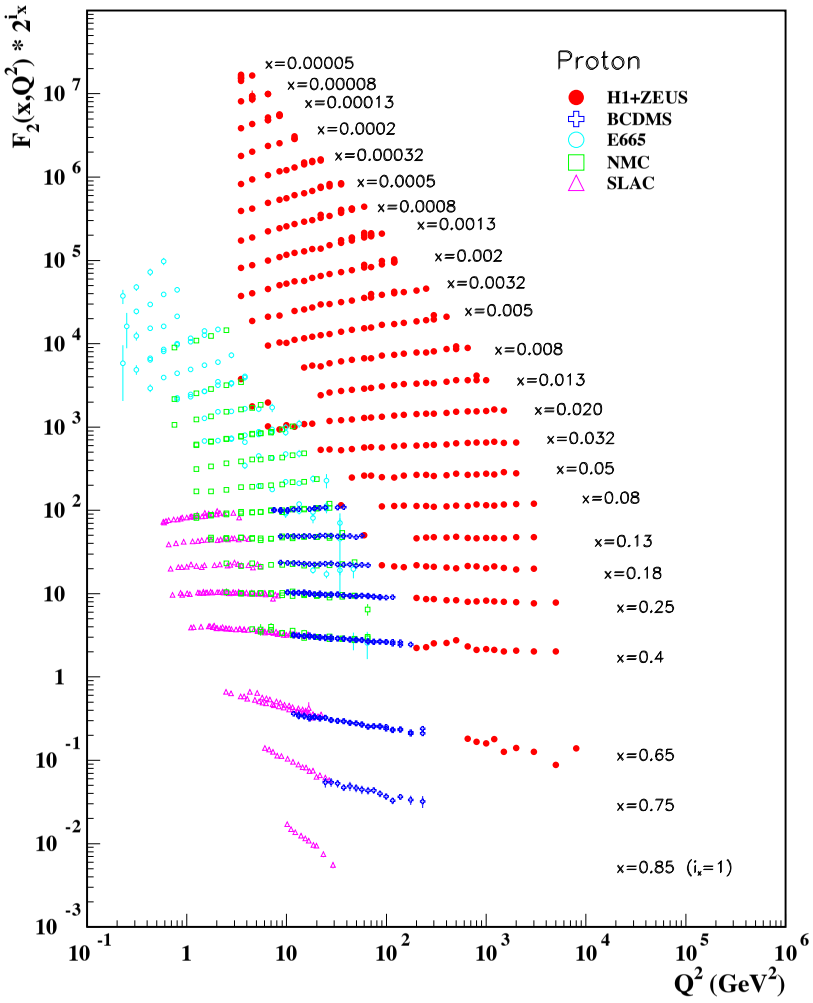
Electron scattering experiments are sensitive to the nucleon quark content through the structure functions. Although the nucleons contain only up and down valence quarks, their quark seas contain strange quarks as well. Heavier quark flavors are suppressed by the limited energy scales in the nucleon. Thus, the sums in Equations (2.8) and (2.9) run over $q = \{\text{u, d, s}\}$. Since u and d form an isospin doublet, it is typical to define $u^p = d^n \equiv u$, $d^p = u^n = d$, and $s^p = s^n \equiv s$ (with analogous relations for the antiquarks). The nucleon F_2 structure functions in the QPM can then be explicitly expressed as

$$F_2^n = x\left[\left(\frac{1}{9}\right)(u+\overline{u}) + \left(\frac{4}{9}\right)(d+\overline{d}) + \left(\frac{1}{9}\right)(s+\overline{s})\right] \tag{2.19}$$

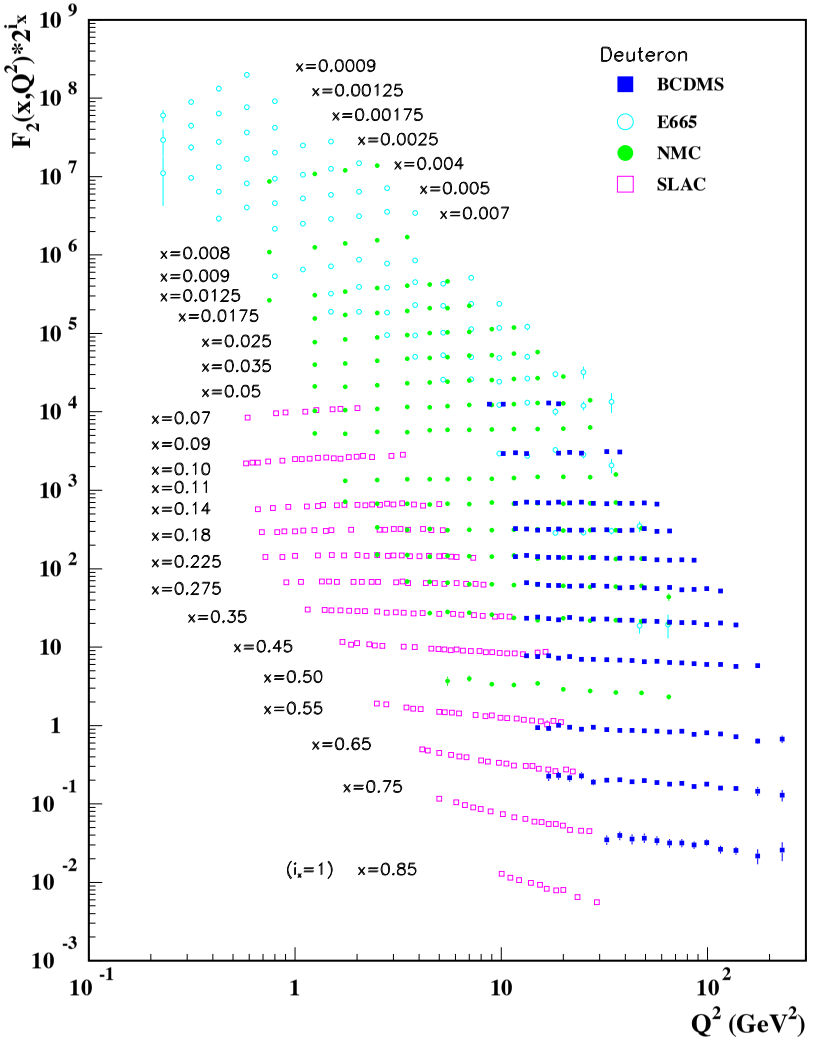
$$F_2^p = x\left[\left(\frac{4}{9}\right)(u+\overline{u}) + \left(\frac{1}{9}\right)(d+\overline{d}) + \left(\frac{1}{9}\right)(s+\overline{s})\right] \tag{2.20}$$

Using leptonic probes on free protons is readily achievable at both fixed-target and collider facilities, allowing precise determination of the proton structure functions. A representation of global F_2^p data is shown in Figure 2.3a, including measurements by H1/ZEUS at DESY [12], BCDMS and NMC at CERN [13, 14], E665 at Fermilab [15], and SLAC [16]. Several of these collaborations also measured the deuterium structure function, [17, 14, 15, 16], which is shown in Figure 2.3b. Note the Q^2 dependence of the structure functions arising from scaling violations. Due to the lack of free neutron targets, such measurements are not available for the neutron structure function F_2^n.

Recall that the structure functions are process-dependent, so measuring cross sections and structure functions of various interactions will probe different linear combinations of the PDFs. Combining measurements from multiple experiments, probes, and targets can allow a global fit of the nucleon PDFs. For example, the Coordinated Theoretical-Experimental



(a) Representation of global F_2^p data. F_2^p is scaled by 2^{i_x}, where i_x is the number of the x bin, ranging from $i_x = 1$ ($x = 0.85$) to $i_x = 24$ ($x = 0.00005$). Figure from [3]



(b) Representation of global F_2^d data. F_2^d is scaled by 2^{i_x}, where i_x is the number of the x bin, ranging from $i_x = 1$ ($x = 0.85$) to $i_x = 29$ ($x = 0.0009$). Figure from [3]

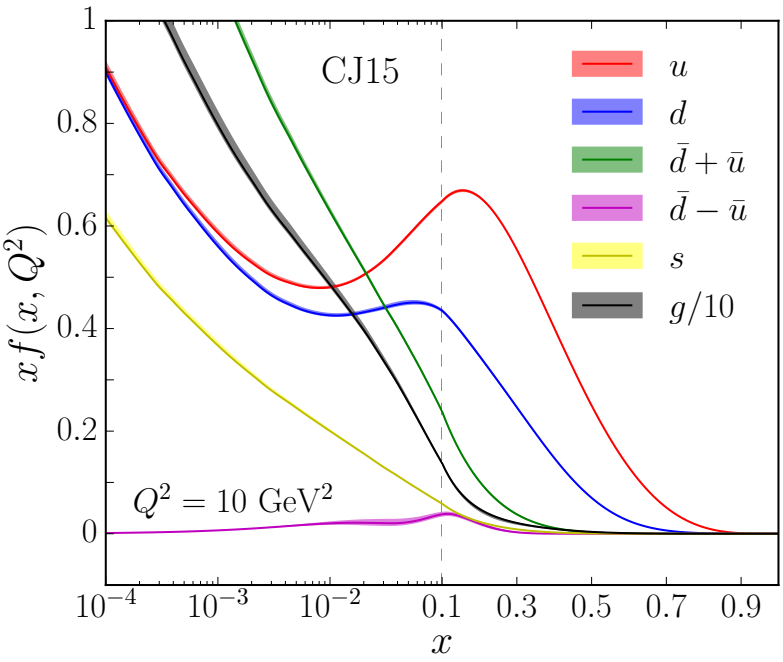


Figure 2.4: CJ15 PDFs for various flavors at $Q^2 = 10$ GeV^2. Plot from [19].

Project on QCD/Jefferson Lab (CJ) collaboration completes such fits. An illustrative plot of PDF fits from CJ15 are shown in Figure 2.4. The low-x region is dominated by the perturbative sea, while the high-x region is dominated by valence quarks. Of particular interest is the down to up quark ratio d/u in the limit $x \to 1$ (the significance of this quantity will be discussed in the following section). This ratio is shown in Figure 2.5 for various global PDF fits. As can be seen, the uncertainties grow quite large as $x \to 1$ due to the poorly constrained neutron structure function.

The PDFs, particularly at large x, are critical to hadronic physics. Predictions of cross sections at high-energy hadron colliders require PDFs as input to the calculation. The currently large uncertainties associated with PDFs at large x results in large uncertainties in the cross section predictions, limiting the impact of collider measurements. For example, a prediction of the inclusive Higgs cross section to NNLO varies as much as 13% depending on which PDF parameterization is used [18]. Providing experimental constraints on the PDFs at large x is essential to the future of hadronic physics.

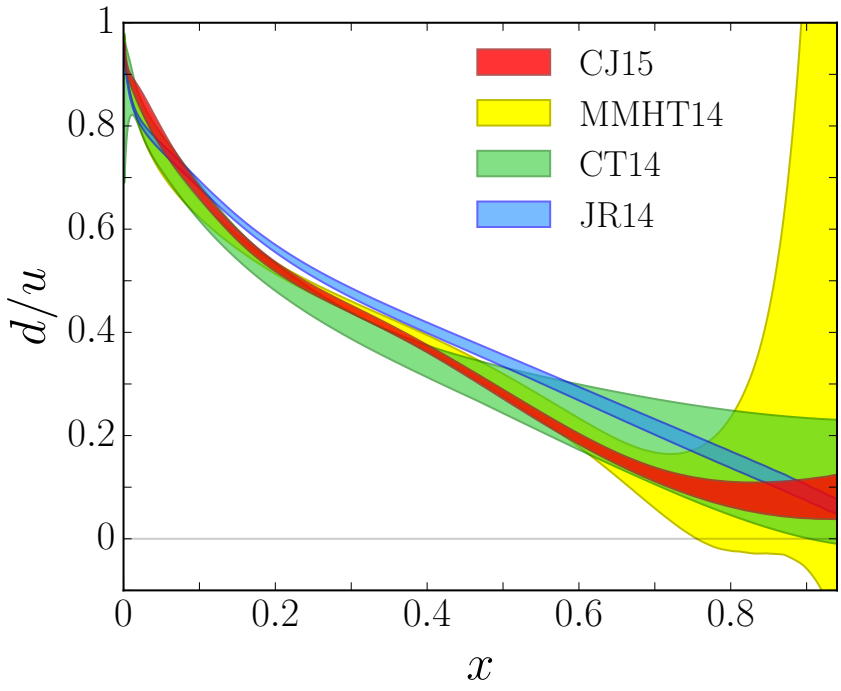


Figure 2.5: d/u ratio at $Q^2 = 10$ GeV^2 for various PDF fits. Plot from [19].

2.2.1 Flavor dependence of valence structure

Constraining d/u at large x does not necessarily require a global PDF fit. Rather, one can take advantage of the fact that in the $x \to 1$ limit, F_2 is primarily sensitive to the valence structure of the nucleon. This means that the strange contribution in Equation (2.19) and (2.20) can be neglected, and

$$\frac{F_2^n}{F_2^p} = \frac{u + 4d}{4u + d} = \frac{1 + 4(d/u)}{4 + (d/u)} \tag{2.21}$$

Therefore, a determination of F_2^n/F_2^p at high x can be used to extract information on the flavor structure of the nucleon's valence quarks. The value of d/u is driven by the QCD interactions of the valence quarks, which can lead to the suppression or enhancement of different quark states. Various theoretical and phenomenological predictions for valence quark dynamics predict different values for d/u and F_2^n/F_2^p in the $x \to 1$ limit. Some of these predictions are summarized below.

2.2.1.1 SU(6)

Under the naive assumption of spin-flavor SU(6) symmetry, QCD interactions do not discriminate between the spin and flavor states of the quarks. Therefore, the wavefunction for a spin-up (↑) proton can be expressed in terms a spin up or down (↑ or ↓) valence quark and its valence "diquark" partner with total spin S:

$$p_\uparrow = \frac{1}{\sqrt{2}} u_\uparrow (ud)_{S=0} + \frac{1}{\sqrt{18}} u_\uparrow (ud)_{S=1} - \frac{1}{3} u_\downarrow (ud)_{S=1} - \frac{1}{3} d_\uparrow (uu)_{S=1} - \frac{\sqrt{2}}{3} d_\downarrow (uu)_{S=1} \quad (2.22)$$

Note that in this case, there is equal probability of finding the diquark with $S = 0$ or $S = 1$. If SU(6) symmetry was exactly preserved, the up content of the proton would be identical to the down content of the neutron (and vice versa) for all x, leading to

$$\frac{F_2^n}{F_2^p} = \frac{2}{3}, \qquad \frac{d}{u} = \frac{1}{2} \qquad \text{[SU(6) symmetry]} \quad (2.23)$$

However, SU(6) symmetry is known to be broken: the up, down, and strange quarks have different masses, the nucleons and Δ have different masses, and measurements of F_2^n/F_2^p differ significantly from 2/3. What is the mechanism by which the flavor symmetry is broken, and what consequence would this mechanism have in the $x \to 1$ limit?

2.2.1.2 Scalar diquark

A phenomenological approach by Close [20] defines the spin-up proton wavefunction similarly to Equation (2.22), but with the diquark core specified by its isospin state $|I, I_z\rangle$ rather than flavor composition. Further, breaking of SU(6) symmetry is allowed by introducing

amplitudes $f_0(x)$ and $f_1(x)$ on the $S = 0$ and $S = 1$ states, respectively.

$$\psi = \frac{1}{\sqrt{2}}(\psi_0 + \psi_1) \tag{2.24}$$

$$\psi_0 = [u_\uparrow \, |0\ 0\rangle_{S=0}] \, f_0(x) \tag{2.25}$$

$$\psi_1 = \left[\frac{1}{3}\, p_\uparrow \, |1\ 0\rangle_{S=1} - \sqrt{\frac{2}{9}}\, p_\downarrow \, |1\ 0\rangle_{S=1} + \frac{2}{3}\, d_\downarrow \, |1\ 1\rangle_{S=1} - \sqrt{\frac{2}{9}}\, d_\uparrow \, |1\ 1\rangle_{S=1}\right] f_1(x) \tag{2.26}$$

In the case of $f_0 = f_1$, the SU(6) wavefunction is recovered.

Using the expectation value of the magnetic moment operator, the ratio of the $\Delta^+(1236)$ transition magnetic form factor G_M^Δ to the proton magnetic form factor G_M^p is found to be

$$\frac{G_M^\Delta}{G_M^p} \sim \frac{\psi_1}{\psi_0} = 1 - x \tag{2.27}$$

where the x dependence is parameterized by a fit to elastic data. Note that this behavior indicates a suppression of diquark states with $S = 1$ as $x \to 1$, leading to a dominance of scalar diquark states. The expectation value of the charge squared operator is used to find the structure functions:

$$F_2^n \sim \psi_0 + 3\psi_1 \tag{2.28}$$

$$F_2^n \sim 4\psi_0 + 2\psi_1. \tag{2.29}$$

Combining Equations (2.27) and (2.28)-(2.29) yields

$$\frac{F_2^n}{F_2^p} = \frac{1 + 3\psi_1/\psi_0}{4 + 2\psi_1/\psi_0} = \frac{4 - 3x}{6 - 2x}. \tag{2.30}$$

giving a prediction of

$$\frac{F_2^n}{F_2^p} \to \frac{1}{4}, \qquad \frac{d}{u} \to 0 \qquad \text{[Scalar diquark]} \tag{2.31}$$

as $x \to 1$.

2.2.1.3 Perturbative QCD (pQCD) and quark counting

Alternative explanations come from pQCD effects in the nucleon. These arguments are based on the fact that in order for a single valence quark to be carrying nearly all the nucleon momentum (i.e., $x \to 1$), the quark must have a large invariant mass. Since the normal bound-state wavefunction is dominated by quark states with low invariant mass, this highly off-shell configuration can only be achieved by the exchange of hard gluons.

Farrar and Jackson [21] note that gluon exchange is constrained by the spins of the diquark spectator. When the spectator quarks have antialigned spins, they can exchange transverse gluons, flipping their spins. When their spins are aligned, conservation of angular momentum requires that they only exchange longitudinal gluons. In the exchange of high invariant mass gluons between low invariant mass quarks, longitudinal gluons are suppressed by a factor of $(1 - x)$ relative to transverse gluons. This favors states where the diquark spectators have antialigned spins, and thus the struck quark's spin is aligned with the parent nucleon's.

The same result is obtained by Brodsky *et al.* [22]. Based on quark-counting rules for the exchange of two hard gluons between three valence quarks, it is found that quark distributions as $x \to 1$ have the limiting behavior $(1 - x)^{3+2\Delta S_z}$, where ΔS_z is the difference in helicity between the struck quark and parent nucleon. Thus, states with antialigned quark and nucleon helicities decrease by an extra power of $(1 - x)$ relative to states with aligned helicities.

By Equations (2.21) and (2.22), the suppression of states where the struck quark has

spin opposite the parent nucleon results in

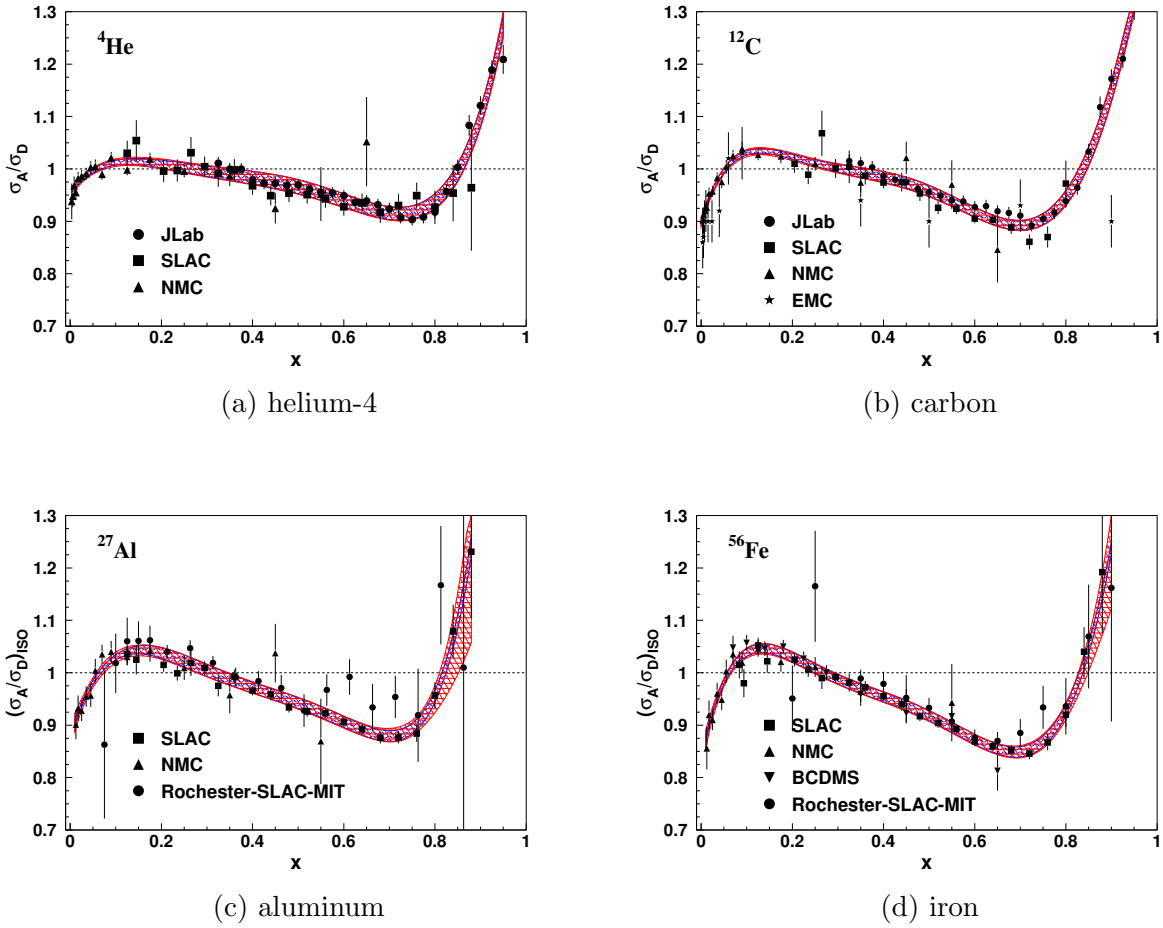
$$\frac{F_2^n}{F_2^p} \to \frac{3}{7}, \qquad \frac{d}{u} \to \frac{1}{5} \qquad \text{[pQCD]} \tag{2.32}$$

as $x \to 1$.

2.3 The EMC effect

It was originally assumed that the structure of free nucleons was identical to the structure of nucleons bound in nuclei. This assumption had both qualitative and quantitative motivation. First, the nuclear shell model - a model which describes atomic nuclei in terms of independent nucleons filling energy levels - was remarkably successful at reproducing, and even predicting, nuclear properties. Second, it was thought that the typical QCD energy scales of quarks and gluons (GeV) would be completely unaffected by the typical binding energies of atomic nuclei (MeV). In 1983, the European Muon Collaboration (EMC) found that the ratio of the per-nucleon F_2 structure function of iron to deuterium differed significantly from unity [23]. This surprising phenomenon, named the EMC effect after its discoverers, indicates that the quark and gluon structure of bound nucleons is altered in the nuclear environment.

The EMC effect has since been studied in a wide range of nuclei. Figure 2.6 shows the EMC effect in a sample of nuclei, using measurements by Jefferson Lab [24], SLAC [25], NMC [26, 27], EMC [28], Rochester-MIT-SLAC [29, 30], and BCDMS [31]. EMC ratios tend to share the same qualitative features which are generally described as: the shadowing region ($x < 0.1$) where the ratio is suppressed, the antishadowing region ($0.1 < x < 0.3$) where the ratio is enhanced, and the EMC region ($0.3 < x < 0.7$) where the ratio linearly decreases. For $x > 0.7$, the ratio increases dramatically. The suppression in the shadowing region tends to increase with A, while the enhancement in the anti-shadowing region is largely independent of A. In the EMC region, the strength of the effect is oftentimes quantified by the linear slope of the decreasing ratio, which becomes steeper with A. An advantage to this approach



(a) helium-4

(b) carbon

(c) aluminum

(d) iron

Figure 2.6: The EMC effect in various nuclei. Plots and fits (red crosshatch bands) from [32].

is that the slope is independent of normalization uncertainties, which can be quite large and vary between targets and experiments.

The difference of EMC ratios from unity isn't entirely due to the modification of individual nucleons. For instance, Fermi motion of nucleons in the nucleus smears the nuclear structure function. This so-called Fermi smearing is present at all x, but in particular is responsible for the steep rise in EMC ratios for $x > 0.7$. However, this does not completely account for modification of the nuclear structure functions, indicating that the structure of the nucleons themselves are also modified. There is currently no universally accepted theory explaining this nucleon modification.

2.3.1 The F_2^n dilemma

The important theoretical implications of determining F_2^n/F_2^p and d/u have motivated much effort to extract these quantities directly from experimental data. However, as previously mentioned, the lack of free neutron targets means that F_2^n cannot be directly measured, but must be obtained from measurements of bound neutrons. Historically, the standard approach has been to extract F_2^n/F_2^p from deuterium and proton data. But even weakly-bound deuterium - often treated as a proxy for a proton plus neutron target - is subject to nuclear effects that must be removed to accurately determine the neutron contribution. This challenge has been well-known since the beginning of the DIS era [33], although the full extent of nuclear effects was not immediately known. Prior to the discovery of the EMC effect, it was thought that the neutron structure functions could be obtained from proton and deuterium data by simply accounting for Fermi smearing in deuterium. In the conventional approach [34], the structure function for nucleus A is given as the convolution of the nucleon structure functions with the nucleon momentum distributions in the nucleus,

$$F_2^A(x) = \sum_N \int dy\ f_N^A(y)\ F_2^N\left(\frac{x}{y}\right), \tag{2.33}$$

where the sum is over the nucleons N. The momentum distributions $f_N^A(x)$ are obtained from a theoretical solution for the nuclear wavefunction. An extraction of F_2^n/F_2^p was carried out in 1979 by Bodek *et al.* [35] using SLAC DIS measurements of the proton and deuterium. The convolution formalism was used to "unsmear" the measured deuterium structure function, and the neutron structure function was obtained by subtraction of the proton structure function.

The discovery of the EMC effect made it clear that nuclear effects could not be accounted for with Fermi smearing alone. A variety of models were developed to explain the growing set of nuclear EMC effect measurements, each of them having unique implications for obtaining F_2^n/F_2^p from deuterium. A re-analysis of the SLAC data was performed in 1991 by Whitlow *et*

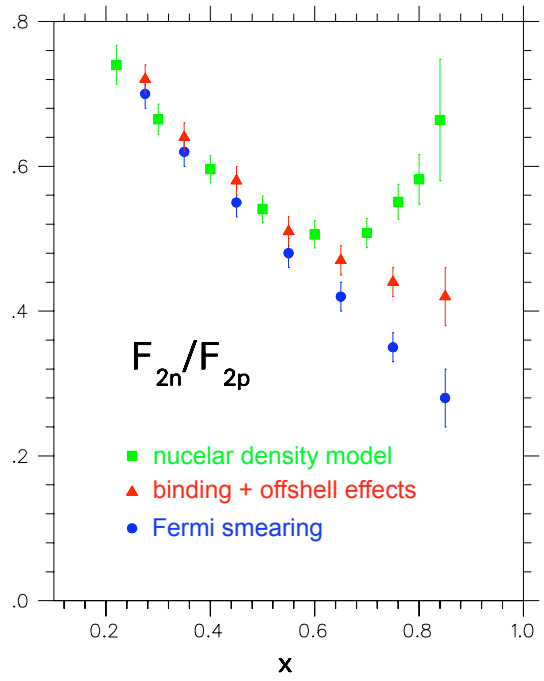


Figure 2.7: F_2^n/F_2^p extracted from SLAC deuterium and proton data using various models for nuclear effects. Produced from MARATHON proposal [38].

al. [16], using improved radiative corrections and more precise measurements of the nuclear photoabsorption ratio $R = \sigma_L/\sigma_T$. The analysis also extracted F_2^n/F_2^p using a "nuclear density model" [36], which assumed that the EMC effect scaled with nuclear density, and could be extrapolated to deuterium from measurements in heavier nuclei. Many modern theoretical models take a more fundamental approach, incorporating effects from nuclear binding and off-mass-shell nucleons [37]. Binding effects are well understood, and can be applied in the convolution formalism with the nuclear wavefunction. Nucleon off-shell effects, however, cannot be applied by convolution, and must be treated as an additive correction to the smeared structure functions. The offshell functions are empirically constrained, though some features (such as parameterization and normalization) have theoretical motivation. Figure 2.7 show the extraction of F_2^n/F_2^p from SLAC data using each of these models. As can be seen, differences in the treatment of nuclear effects in deuterium can cause large discrepancies in the determination of F_2^n/F_2^p. Therefore, current knowledge of F_2^n is highly model-dependent.

2.3.2 Isoscalar EMC ratios

The nuclear EMC effect is typically expressed as the ratio of per-nucleon structure functions of nucleus A to deuterium:

$$R(A) = \frac{F_2^A/A}{F_2^D/2} \tag{2.34}$$

Since deuterium is also subject to nuclear effects, this is not the same as the ratio to free nucleons. However, F_2^D can be directly measured, so it provides a convenient and consistent way to compare modification in heavier nuclei. Deuterium's EMC ratio can be expressed as

$$R(D) = \frac{F_2^D}{F_2^n + F_2^p}, \tag{2.35}$$

although this is model-dependent due to F_2^n.

Deuterium is an isoscalar nucleus, having one proton and one neutron. Many of the heavy nuclei used in EMC measurements do not have equal numbers of protons and neutrons. It is common to apply a multiplicative isoscalar correction to these EMC ratios, adjusting for the excess of protons or neutrons

$$f_{iso}^A = \frac{\frac{1}{2}(1 + F_2^n/F_2^p)}{\frac{1}{A}(Z + (A-Z)F_2^n/F_2^p)} \tag{2.36}$$

This correction is also model-dependent, as it contains the ratio F_2^n/F_2^p. Note that for symmetric nuclei, $f_iso^A = 1$.

Chapter 3

The MARATHON experiment

In 2006, MARATHON was formally proposed to Experimental Hall A at Thomas Jefferson National Accelerator Facility (Jefferson Lab). The experiment was designed to take advantage of the planned 12 GeV energy upgrade to the Continuous Electron Beam Accelerator Facility (CEBAF), Jefferson Lab's flagship electron accelerator. The proposal was conditionally approved pending the design and review of a safe tritium target, and an updated proposal was submitted in 2010 [38]. Several other experiments taking advantage of the tritium target were proposed and accepted. The tritium experiments program commenced December 2017 and concluded November 2018, with MARATHON taking production data January-April 2018.

3.1 Motivation

3.1.1 F_2^n/F_2^p from mirror A = 3 nuclei

The MARATHON experiment proposes to extract F_2^n/F_2^p from DIS measurements on tritium (^{3}H) and helium-3 (^{3}He). The novel extraction method, first proposed in 2000 [39], exploits the approximate mirror symmetry of tritium and helium-3. In terms of the nuclear and

nucleon structure functions, EMC-type ratios can be defined for the two nuclei:

$$R(^3\mathrm{H}) = \frac{F_2^{^3\mathrm{H}}}{2F_2^n + F_2^p} \tag{3.1}$$

$$R(^3\mathrm{He}) = \frac{F_2^{^3\mathrm{He}}}{F_2^n + 2F_2^p} \tag{3.2}$$

These equations can be used to define the so-called "super-ratio", a ratio of the helium-3 to tritium EMC ratios:

$$\mathcal{R} = \frac{R(^3\mathrm{He})}{R(^3\mathrm{H})} \tag{3.3}$$

Inverting these expressions yields the ratio of neutron to proton structure functions,

$$\frac{F_2^n}{F_2^p} = \frac{2\mathcal{R} - F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}}{2F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}} - \mathcal{R}} \tag{3.4}$$

Under the commonly made assumption that the nuclear photoabsorption ratio $R = \sigma_L/\sigma_T$ is independent of nucleus[1], the nuclear structure function ratio can be obtained simply by measuring the DIS cross section ratio,

$$\frac{F_2^{^3\mathrm{He}}}{F_2^{^3\mathrm{H}}} = \frac{\sigma^{^3\mathrm{He}}}{\sigma^{^3\mathrm{H}}} \tag{3.5}$$

Thus, combining an experimental measurement of $\sigma^{^3\mathrm{He}}/\sigma^{^3\mathrm{H}}$ with a theoretical calculation of $\mathcal{R}$ allows the extraction of the free nucleon structure function ratio, F_2^n/F_2^p. Unlike the extraction from proton and deuterium data, which depends on accurately calculating the absolute magnitude of nuclear effects in deuterium, this approach depends only on the relative magnitude of nuclear effects in helium-3 and tritium. This could significantly reduce the theoretical model dependence of the extraction.

[1]See Chapter 2.1.4 for a discussion on the validity of this assumption

3.1.2 EMC effect in A = 3 nuclei

The nuclear EMC effect remains poorly understood. Any universal explanation of the EMC effect must be able to reproduce the effect in the lightest nuclei, and tritium and helium-3 are of particular interest as they are the lightest nuclei sensitive to possible isospin dependence of the EMC effect. While the EMC effect has been extensively studied in many nuclei, very little data exists for A = 3 nuclei. Limited helium-3 measurements have been made at low x [40] and medium to high x [24]. Due to the practical complications of using radioactive tritium as a target, there exists very little tritium scattering data at all, and no measurements of the EMC effect. Precise measurements of these EMC ratios will test competing model calculations and provide critical constraints on theoretical explanations of the EMC effect.

3.2 Experimental apparatus

3.2.1 CEBAF

CEBAF is responsible for delivering beam electrons to Jefferson Lab's four experimental halls. The accelerator has a racetrack design, with two linear accelerators (linacs) joined on either end by circular arcs. Electrons originate in the injector, make multiple passes through the accelerator until reaching the desired energy, and then go to the beam switchyard where they are directed to individual halls. See Figure 3.1 for the layout of the accelerator.

The electron source in the injector is a gallium arsenide (GaAs) photocathode. Alternating pulses from four seed lasers (one for each hall) eject electrons from the crystal, creating a train of electron bunches at 1497 MHz. A -100 kV gun and set of cryomodules (discussed in the next paragraph) accelerate the electrons to 100 MeV, at which point they are injected into the north linac.

The north and south linacs each contain 200 radio-frequency (RF) cavities housed in 25 cryomodules, where they are cooled to superconducting temperatures with 2 K liquid

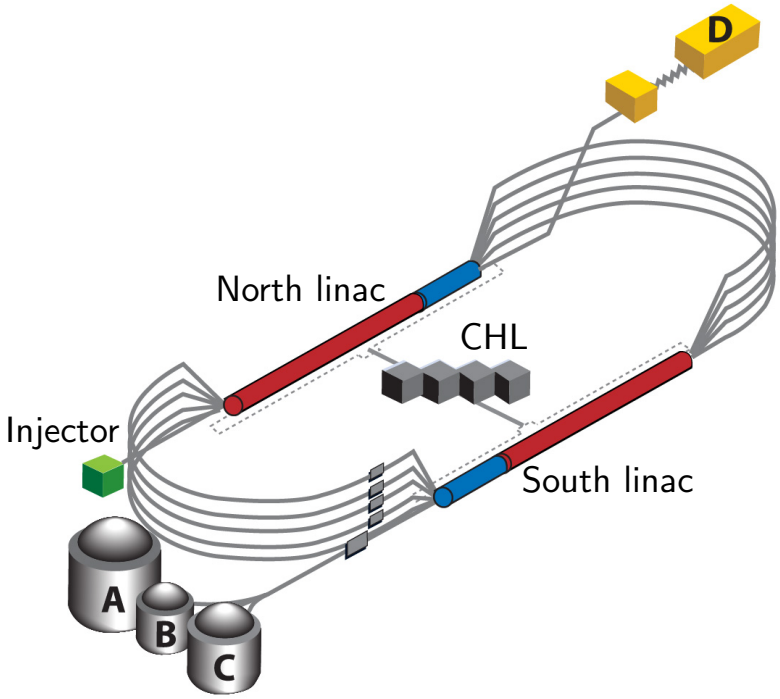


Figure 3.1: CEBAF layout. Red linac region represents 20 original cryomodules in 6 GeV accelerator. Blue linac region represents 5 new cryomodules for 12 GeV upgrade. Locations of Experimental Halls A, B, C, and D are shown.

helium-4. The RF cavities have a resonant frequency of 1497 MHz (which is why the electron bunch train is created with the same frequency). By precisely coordinating the arrival of the electron bunches with the oscillating electric fields in the RF cavities, the electrons are continuously accelerated as they pass through the cryomodules. During the MARATHON experiment, electrons gained approximately 1.05 GeV of energy per linac (2.1 GeV for one full pass).

The linacs are joined by circular arcs with magnetic fields to bend the electron trajectories. Since the electrons gain energy with every pass through a linac, there must be separate arcs and magnetic fields for bending electrons of different energies. There are ten total arcs, five on either side of the linacs, supporting 5-pass beam to Halls A, B, and C (at end of the south linac) and 5.5-pass beam to Hall D (at the end of the north linac). The arcs are vertically stacked. Before and after each arc are spreader and combiner magnets for sorting electrons into and out of their appropriate arcs.

Once electrons reach the desired energy, an RF separator directs them to the beam switch yard. Here, a second RF separator directs electron bunches to their respective Halls.

3.2.2 Hall A beamline

3.2.2.1 Beam current monitors

The beam current must be constantly monitored to determine the total number of electrons incident on the target. The Hall A beam current monitors (BCMs) consist of two RF cavities and an Unser monitor. The RF cavities are tuned to resonate at 1497 MHz and contain a large loop antenna coaxial to the cavity. When the electron beam passes through the cavity, it excites electromagnetic modes that are picked up by the antenna. The antenna output voltage is proportional to the beam current in the cavity. The Unser monitor is a parametric current transformer consisting of five toroids. When electrons pass through the monitor, the transformer's output voltage is proportional to the beam current.

The Unser monitor can be precisely calibrated for absolute current measurements using a wire carrying a well-defined DC current. However, the Unser's output signal drifts significantly over periods of several minutes, making it unsuitable to continuously monitor the beam current. Rather, the Unser monitor is used to provide the absolute calibration of the cavity monitor, which is stable over large periods of time. To calibrate the cavities, beam is requested at several current values between $3 < I < 22.5$ μA for approximately 1-2 minute intervals (see Figure 3.2). The absolute current for each interval (determined by the calibrated Unser monitor) is then plotted against the frequency response f of the cavity monitors, and the BCM gain g and offset b are obtained from the slope and intercept of a linear fit (see Figure 3.3). The beam current is determined by:

$$I = gf + b \tag{3.6}$$

The gains found from the three BCM calibrations performed during the MARATHON

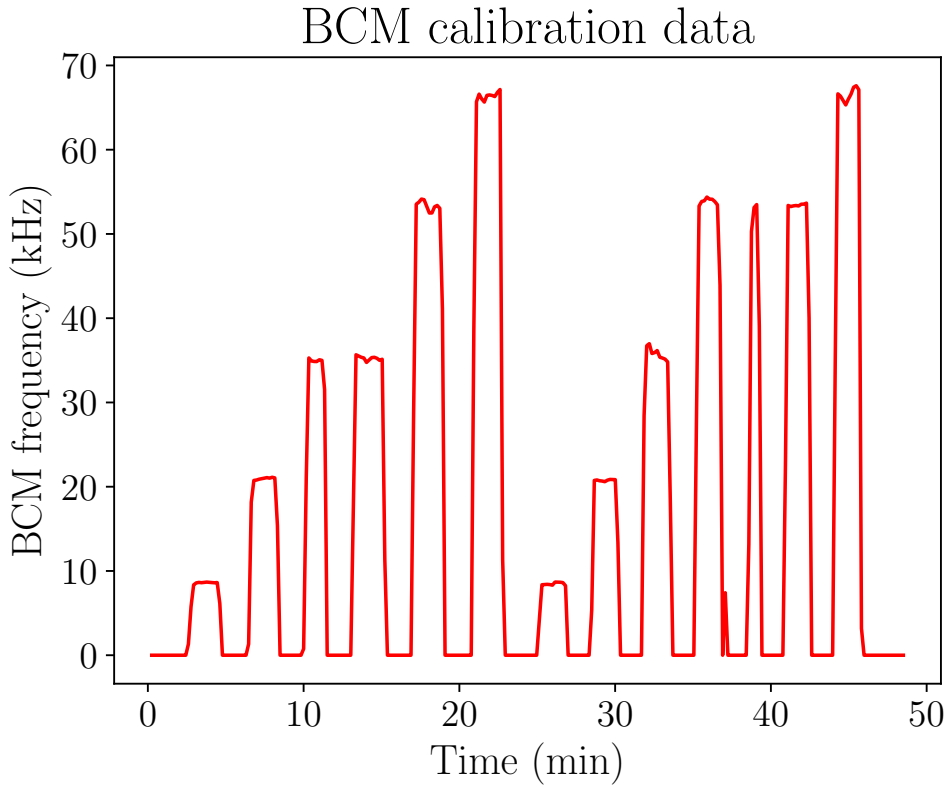


Figure 3.2: BCM calibration data.

run are plotted in Figure 3.4, showing good stability of the cavity monitor.

3.2.2.2 Beam position monitors

Two beam position monitors (BPMs), located 7.52 m and 1.29 m upstream of the target, are used to measure the position and direction of the beam. In the hall coordinate system, the beam travels along $\hat{z}$ and the transverse plane is defined by $\hat{x}$ (horizontal, beam left) and $\hat{y}$ (vertical). Each BPM contains four stripline antennae located at $\pm u$ and $\pm v$, where u and v are along axes rotated 45° about $\hat{z}$. The induced current in each antenna is proportional to the distance to the beam. The difference over sum of currents in opposing antennae determine the relative position of the beam to 100 μm for currents > 1 μA.

Calibrating to the absolute hall coordinates is completed with a Harp scan, an invasive method of determining the absolute position of the beam. The Harp consists of three wires (one vertical and two diagonal) that are scanned through a low-current beam at a known

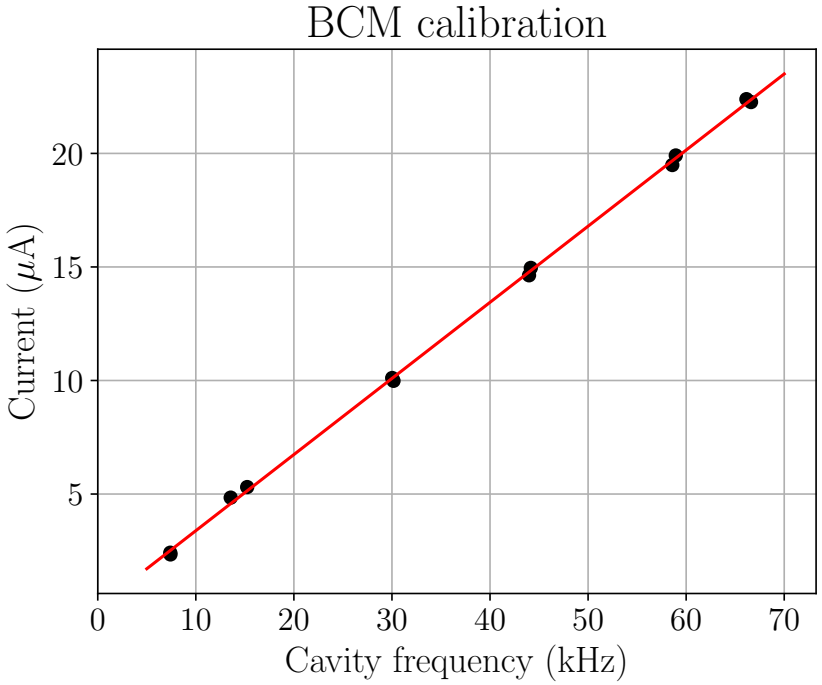


Figure 3.3: BCM calibration fit. The BCM gain g and offset b are the slope and intercept of the linear fit, respectively.

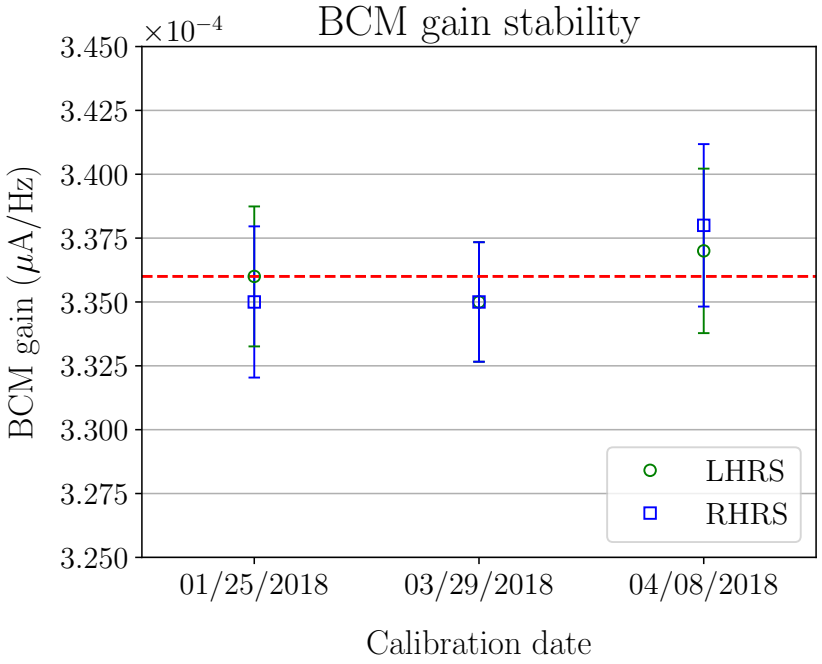


Figure 3.4: BCM gains from MARATHON calibrations.

rate. The Harp wire positions are precisely surveyed in hall coordinates, so when the wires intersect the beam, the absolute beam position is known. This information can be used to determine the rotation matrix $\boldsymbol{\theta}$ and offset $\boldsymbol{b}$ to transform from the BPM coordinates (u, v) to the hall coordinates (x, y):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \theta_{00} & \theta_{01} \\ \theta_{10} & \theta_{11} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} \tag{3.7}$$

Equation (3.7) expands into two equations, each with three unknowns. Therefore, a Harp scan must be completed at three unique values of absolute beam position to completely calibrate the BPMs.

3.2.2.3 Beam rastering

The small spot size of the beam (<400 μm) can cause significant heating in small portions of the target and the beam dump. To mitigate this effect, a raster is applied to the beam to deposit power over a larger area. This is done with a set of four rastering coils, two for each transverse direction, located 23 m upstream of the target. MARATHON only used a single pair of coils, one for each transverse direction. The coils are driven by a triangular wave current at 25 kHz, creating oscillating transverse magnetic fields that perturb the beam trajectory. This scans the beam over a rectangular footprint, typically 4 mm x 4 mm at the target. The driving currents were also recorded by an analog to digital converter (ADC).

Determining the position of the rastered beam at the target is critical to accurately reconstructing the z-vertex and momentum of each event. This requires calibrating the raster current to the absolute position at the target, allowing the determination of the absolute beam position at target, (x_{tg}, y_{tg}), in terms of the raster currents, (R_x, R_y), and offsets, (b_x, b_y). Since there is no coupling between the horizontal and vertical components,

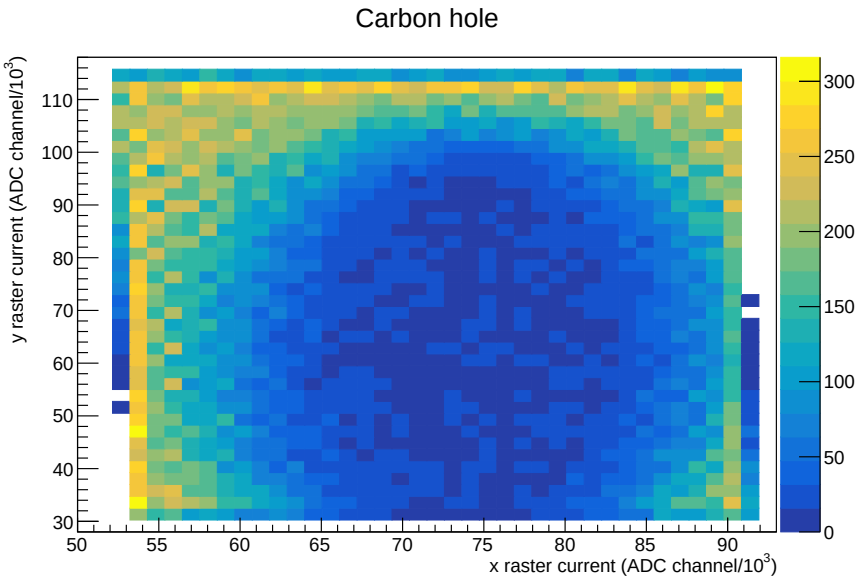


Figure 3.5: Distribution of all event triggers versus horizontal and vertical raster current with carbon hole target.

there is a single calibration factor S for each direction:

$$x_{tg} = R_x S_x + b_x \tag{3.8}$$

$$y_{tg} = R_y S_y + b_y \tag{3.9}$$

Prior to the implementation of the 25 kHz triangle wave raster, the mapping from raster current to absolute position at the target was found by directly comparing the mean and RMS of the raster current to the mean and RMS of the BPM measurements. However, some of the BPM electronics effectively act as a low-bandpass filter with a frequency cutoff of 30 kHz. This results in an attenuation of the BPM amplitude relative to the raster current amplitude. An alternative must be found for the absolute position calibration of the raster current. For MARATHON, the calibration was performed with two inputs: a 2D sigmoid fit to the distribution obtained from the carbon hole target, and the correlation between the horizontal raster current and the reconstructed z-position of the scattering event.

Among the solid targets used for MARATHON was a carbon hole target, a thin carbon

foil with a 2 mm diameter hole. Figure 3.5 shows a 2D histogram of carbon hole data binned by vertical and horizontal raster current (in ADC units). Since the absolute radius of the hole is known, fitting the distribution with an appropriate function can provide information on the absolute position calibration for the raster current. The function must account for the fact that misconstructed events smear the hole edge and populate the hole center, and that the relationship between raster current and absolute position might be different for the horizontal and vertical direction. These requirements are satisfied by a 2D sigmoid function with 7 free parameters p_i:

$$f(x,y) = \frac{p_0}{1 + e^{-p_1[p_2^2(p_3-x)^2 + p_4^2(p_5-y)^2 - r^2]}} + p_6 \tag{3.10}$$

p_0 and p_6 are the function values outside and inside the hole, respectively. p_1 describes the hardness of the hole edge in the reconstructed distribution. p_2 and p_4 are the horizontal and vertical calibration factors from raster current to absolute position, corresponding to S_x and S_y in Equations (3.8) and (3.9). p_3 and p_5 are the horizontal and vertical position of the hole center. Finally, $r = 1$ mm is the absolute radius of the hole. Although the 2D sigmoid fit to the carbon hole distribution was found to be quite stable, the best-fit p_2 and p_4 did not result in a good raster calibration. This was ostensibly due to ambiguity in the location of the true hole edge in the smoothly-varying sigmoid function. More information was required to fix the calibration.

A properly calibrated raster should result in no correlation between the horizontal raster current and the reconstructed z-position. This condition was not satisfied by using the calibration obtained from the sigmoid fit. In an effort to fix the horizontal raster calibration, the correlation between horizontal current and reconstructed z-position was compared for two incorrect raster calibration values (see Figure 3.6). The horizontal calibration value S_x was fixed by linearly interpolating between the two incorrect values to minimize the correlation between the horizontal current and reconstructed z-position (see Figure 3.7).

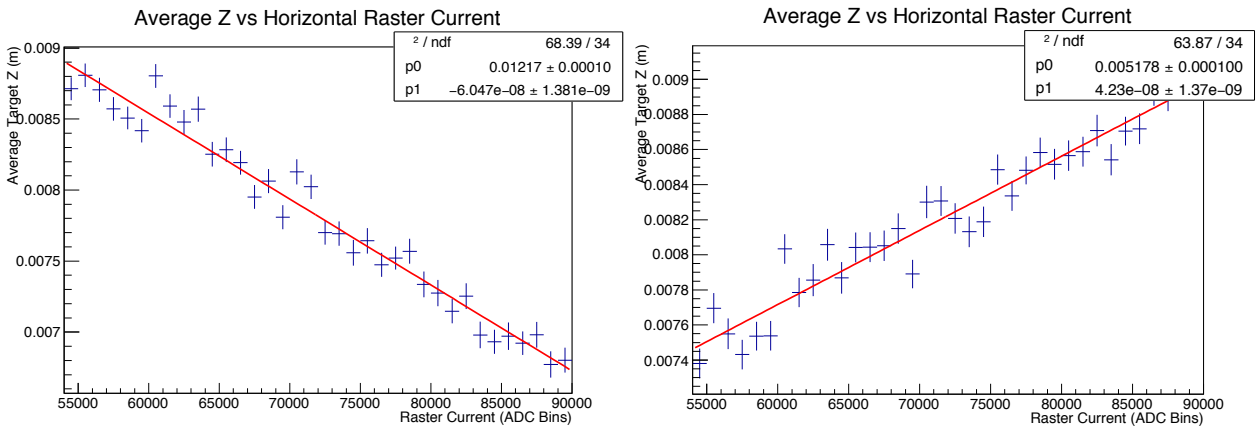


Figure 3.6: Horizontal raster current vs. reconstructed z for two incorrect raster calibrations.

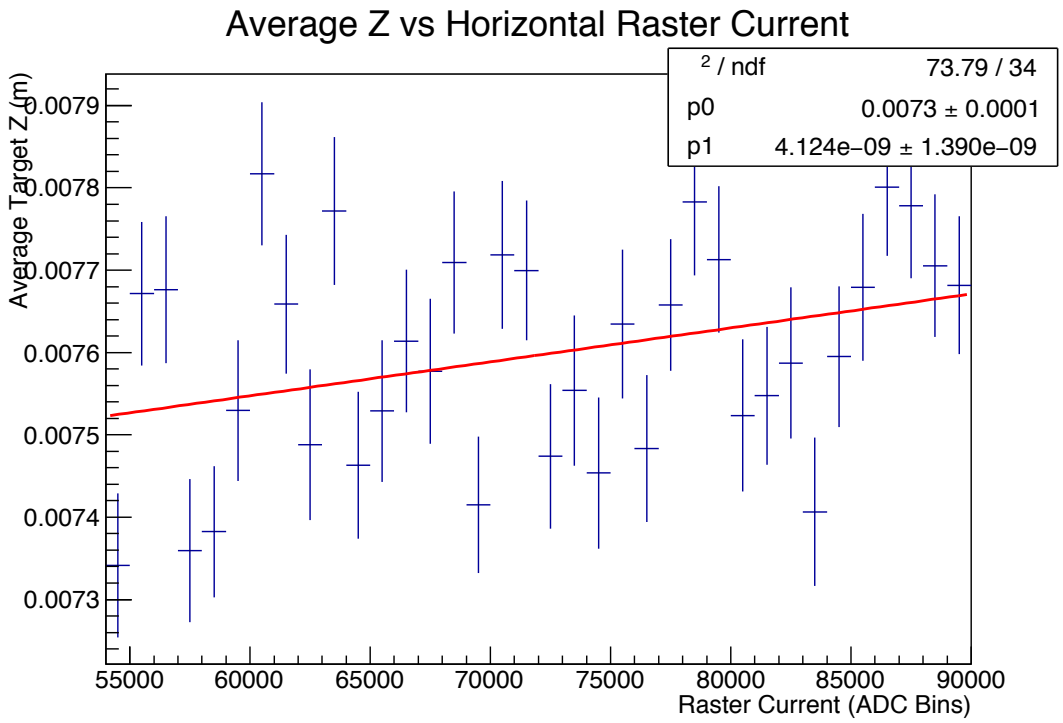


Figure 3.7: Horizontal raster current vs. reconstructed z for interpolated raster calibration.

A similar method could be used to calibrate the vertical raster, by minimizing the correlation of the vertical raster current with the reconstructed W^2 of an identifiable momentum feature (such as an elastic peak). However, MARATHON's DIS data had no such feature. With S_x fixed, the vertical raster calibration was obtained by returning to the 2D sigmoid fit. The function value at the hole edge f_{edge} was found from Equation (3.10), with $p_2 = S_x$ and $y = p_4$. The remaining parameters used their best-fit values. Since the physical hole is symmetric, the function value should be constant along the hole edge. This was exploited to find the vertical raster calibration. Setting Equation (3.10) equal to f_{edge} with $x = p_2$ allowed the expression to be solved for p_4. This was the value used for S_y.

The raster offsets were determined using the BPMs. As the rasters scan symmetrically, the time-averaged beam position should be equal to the beam position with zero raster current. By projecting a line to the target through the average beam positions measured by the two BPMs, Equations (3.8) and (3.9) (with $R_x = R_y = 0$) fix the offsets b_x and b_y.

3.2.2.4 Beam energy measurement

The beam energy is determined using the so-called ARC method, which takes advantage of the fact that the deflection angle of an electron through a magnetic field is proportional to its momentum:

$$p = (0.299792 \text{ GeV rad T}^{-1} \text{ m}^{-1}) \times \frac{\int \overline{B} \cdot d\overline{\ell}}{\theta}. \tag{3.11}$$

Eight identical dipoles are used to bend the beam through the Hall A arc from the switchyard into Hall A. BPMs before and after the dipoles determine the total deflection angle θ. A ninth identical dipole, powered in series with the ones in the arc, is used to determine the field integral. A platform containing two pick-up coils and a nuclear magnetic resonance (NMR) probe is moved through the center of the dipole. The flux through the coils is integrated, while the NMR probe fixes the absolute calibration of the field. It is assumed that the eight dipoles in the arc have the same field integrals as the one measured.

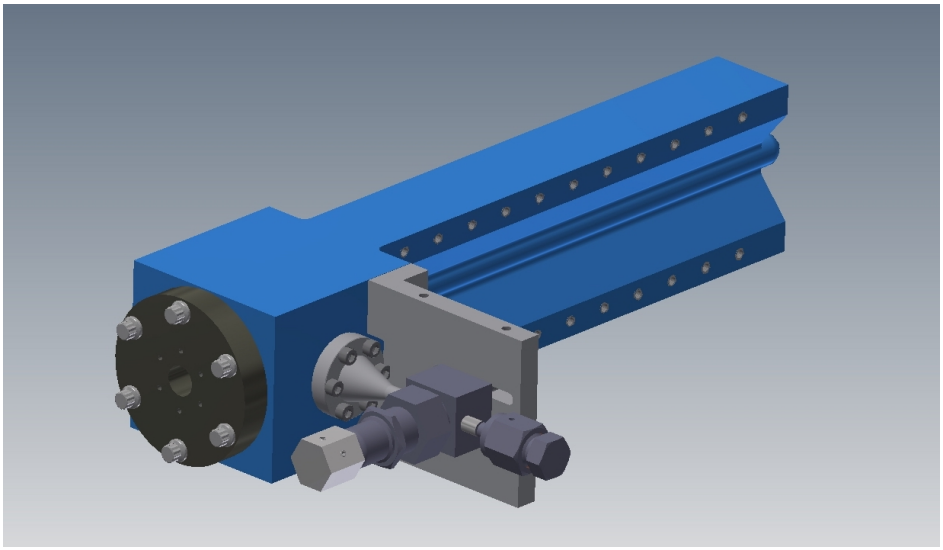


Figure 3.8: CAD model of gas target cell. Image from [41].

3.2.3 Target system

The target system used by MARATHON contained over 10 gas and solid targets used for physics, background, and calibration measurements. The design of the gas targets was driven primarily by safety considerations for the tritium target. Atypical for Hall A, the target gases were not circulated but rather completely contained within sealed, modular cells. See Figure 3.8. This allowed the tritium cell to be filled and sealed off-site, eliminating the need for Jefferson Lab to handle tritium directly. The cells were constructed from aluminum, with a cylindrical gas volume 25 cm in length and 12.7 mm in diameter. A flange attached to the upstream end of the cell sealed the gas volume and defined the entrance window. The incoming and outgoing electron beam had to pass through the cell entrance and exit windows (or endcaps), and electrons scattered into the spectrometers had to pass through the cell side walls. Therefore, these boundaries were designed to be as thin as possible (≈0.25-0.50 mm) without compromising the safe containment of the tritium gas. Due to machining tolerances, the cell thicknesses vary between cells, and the side thickness varies along the length of a cell.

MARATHON used five identical (up to machining tolerances) gas cells. Four contained

Target	Fill pressure (kPa)	Fill temperature (K)	Thickness (mg/cm^2)
^{1}H	3549	297.4 ± 0.1	70.8 ± 0.4
^{2}H	3549	296.1 ± 0.1	142.2 ± 0.8
^{3}H	1400	296.3 ± 0.1	85.1 ± 0.8
^{3}He	1742	294.5 ± 0.1	53.4 ± 0.6

Table 3.1: Fill parameters and thicknesses of MARATHON gas targets.

gases for the primary physics measurements: hydrogen (^{1}H), deuterium (^{2}H), tritium (^{3}H), and helium-3 (^{3}He). The cell volumes were measured at Jefferson Lab and found to be $V = 33.28 \pm 0.3$ cm^3. The tritium cell used for MARATHON was filled and sealed by Savannah River Tritium Enterprise at the Savannah River Site on October 23, 2017. The cell was shipped to Jefferson Lab and stored in triple containment until it was installed shortly before the experiment. The hydrogen, deuterium, and helium-3 cells were filled at Jefferson Lab. The fifth cell was left empty to measure the background from the aluminum endcaps. Table 3.1 shows the fill parameters and thicknesses of each gas target.

In addition to the gas cells, the target ladder included a variety of solid targets. A thin carbon foil with a 2 mm hole was used for beam centering checks and the raster calibration. A dummy target consisting of two aluminum foils spaced 25 cm apart (replicating the length of the gas cells), but much thicker than the gas cell endcaps, was used for additional background measurements. An optics target consisting of 11 equally spaced carbon foils was used for optics reconstruction checks. Finally, there was a thick aluminum foil and thin foils of carbon, titanium, and beryllium oxide.

All the targets were arranged on a vertical, motor-controlled ladder, allowing remote adjustment of the target in the beamline. See Figure 3.9. The target ladder was cooled to approximately 40 K using a 15 K liquid helium-4 supply and controlled heater. Various sensors were implemented for monitoring the temperature of the targets, the cryogen flow rate and temperature, and the heater power. In the event of a tritium leak or other disaster,

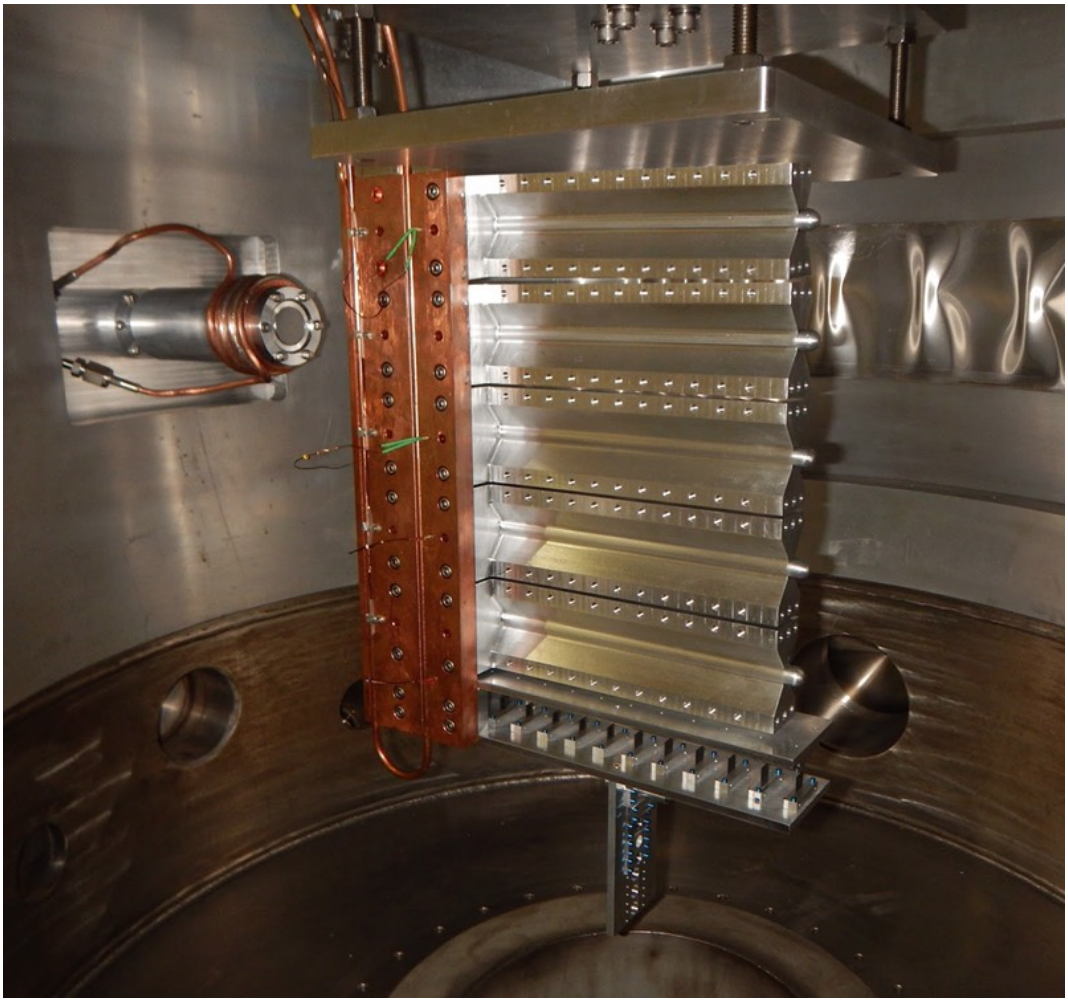


Figure 3.9: Photo of MARATHON target ladder.
Photo credit: D. Meekins and R. Cruz-Torres

the scattering chamber contained an emergency pump that vented to a high altitude above Hall A.

3.2.4 High resolution spectrometer

MARATHON required the detection of electrons that participated in large energy transfer DIS processes in the target. Selecting these electrons for detection was completed using Hall A's pair of high-resolution spectrometers (HRS). The left (L) and right (R) HRS rotate independently about the hall center, allowing coverage of central scattering angles ranging from 12.5° to 150° (LHRS) and 130° (RHRS). The spectrometers were designed to have a

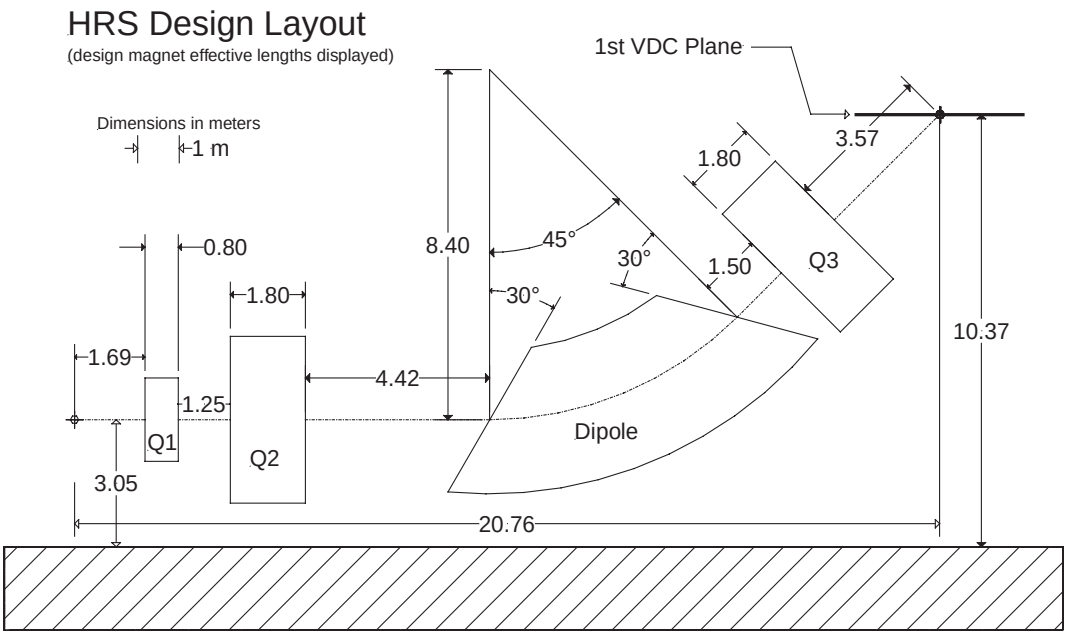


Figure 3.10: Schematic layout of HRS. Dashed line indicates central trajectory. From [42].

momentum resolution of 10^{-4} over the momentum range 0.8-4.0 GeV. Each HRS has three quadrupole magnets and one dipole magnet, ordered QQDQ, which are shown schematically in Figure 3.10. The dipole has a 45° vertical bend, defining the dispersive direction of the spectrometer. The first quadrupole (Q1) focusses in the dispersive direction, while the subsequent quadrupoles (Q2 and Q3) focus in the non-dispersive direction.

The trajectory of particles through the HRS is described in the transport coordinate system (TCS). The coordinate system is defined relative to the central trajectory, which is the idealized trajectory of a particle whose angle and momentum match exactly the central settings of the spectrometer. The z-axis is parallel to the central trajectory. The x-axis is perpendicular to the z-axis, pointing in the dispersive direction of the spectrometer dipole. The y-axis is perpendicular to both, $\hat{y} = \hat{x} \times \hat{z}$, and thus points in the non-dispersive direction. The direction of a trajectory is specified by the tangent of angles relative to the central trajectory, $\theta = dx/dz$ and $\phi = dy/dz$. Momentum is expressed as the fractional deviation from the momentum of central trajectory, $\delta p/p$. A particle is fully defined in the

TCS by the vector

$$\overline{x} = \begin{bmatrix} x \\ \theta \\ y \\ \phi \\ \delta p/p \end{bmatrix} \tag{3.12}$$

These variables are measured at the focal plane downstream of the spectrometer. However, determining the kinematics of an event requires knowledge of these variables at the original scattering vertex. To first order, the propagation of a particle through the HRS can be described by a single transport matrix $\mathbf{T}$. That is, a particle's variables at the focal plane, $\overline{x}_{fp}$, are related to the variables at the target, $\overline{x}_{tg}$, by:

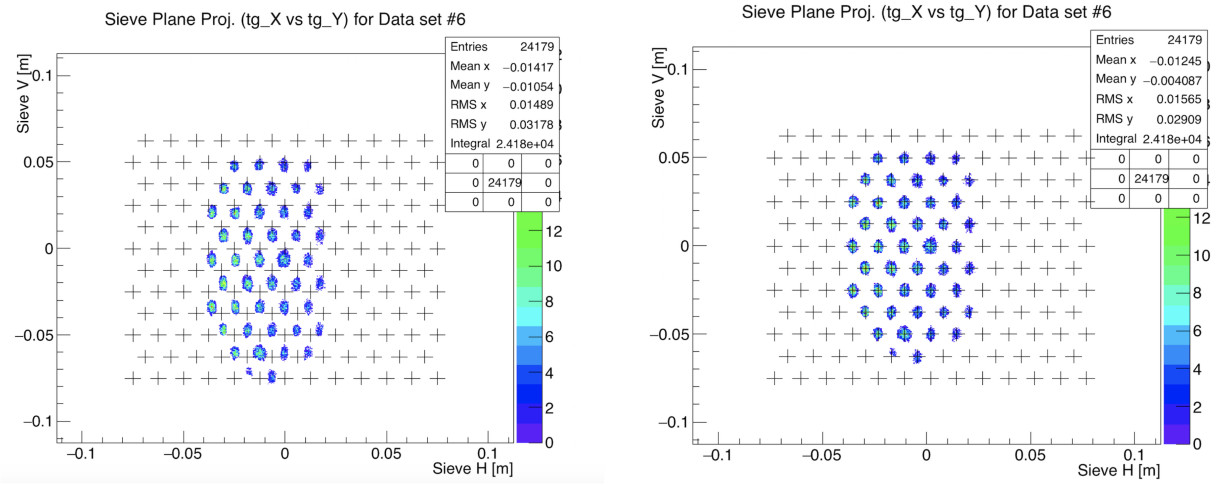
$$\overline{x}_{fp} = \mathbf{T}\overline{x}_{tg} \tag{3.13}$$

The inverse of the transport matrix, $\mathbf{T}^{-1}$, allows the reconstruction of the target variables based on the measured focal plane variables. In practice, the expansion in focal plane variables is performed up to fifth order. The single reconstruction matrix is replaced with a set of tensors that are polynomials of x_{fp}. Thus, the target variables x^{α} ($= x$, θ, y, ϕ, $\delta p/p$) are reconstructed by

$$x^{\alpha} = \sum_{j,k,l} \sum_{i} (C_i^{\alpha_{jkl}} x_{fp}^i) \theta_{fp}^j y_{fp}^k \phi_{fp}^l. \tag{3.14}$$

The superscripts indicate the power of each focal plane variable, and $C_i^{\alpha_{jkl}}$ are the corresponding optics matrix elements for variable α

The optics tensor elements $C_i^{\alpha_{jkl}}$ are tuned using a dataset that covers the entire space of θ_{tg}, y_{tg}, ϕ_{tg}, and $\delta p/p$ [43]. To reduce the number of unknowns at the target, x_{tg} is effectively fixed at zero by constraining the beam position at the target. The spectrometer setting is scanned over $-4.5\% < \delta p/p < 4.5\%$, moving the elastic peak across the focal plane. Further, data is taken with a special collimator called the sieve, placed downstream



(a) before optics calibration

(b) after optics calibration

Figure 3.11: Reconstructed sieve data from E12-07-108. The crosshairs indicate the absolute position of the sieve holes.

of the target prior to any magnetic fields. The sieve contains a grid of small holes that, when carefully positioned and surveyed, correspond to precise positions and angles in the TCS. Such a dataset allows the optimization of the tensor elements to reproduce the known variable values at the target.

The optics tensor elements are unique for a given spectrometer tune, defined by a particular combination of field settings for each magnet. Many experiments in Hall A, including MARATHON, use the standard HRS tune described in [42]. Completing a full optics calibration is time-consuming, requiring time to both take the data and to install and survey the sieve. For this reason, MARATHON did not perform an optics calibration, but rather relied on the calibration obtained from experiment E12-07-108. This experiment also used the standard HRS tune, and no configuration changes were made before MARATHON that would affect the resulting optics tensor. See Figure 3.11 for the reconstructed sieve data before and after the E12-07-108 optics calibration.

3.2.5 Detector package

Each HRS is equipped with a detector package for particle tracking, counting, and identification. Historically, the RHRS was used to detect final-state hadrons when required, while the LHRS detected electrons. This led to some small geometrical differences in the left and right detectors, although the overall package is functionally the same. A schematic of the detector setup is shown in Figure 3.12.

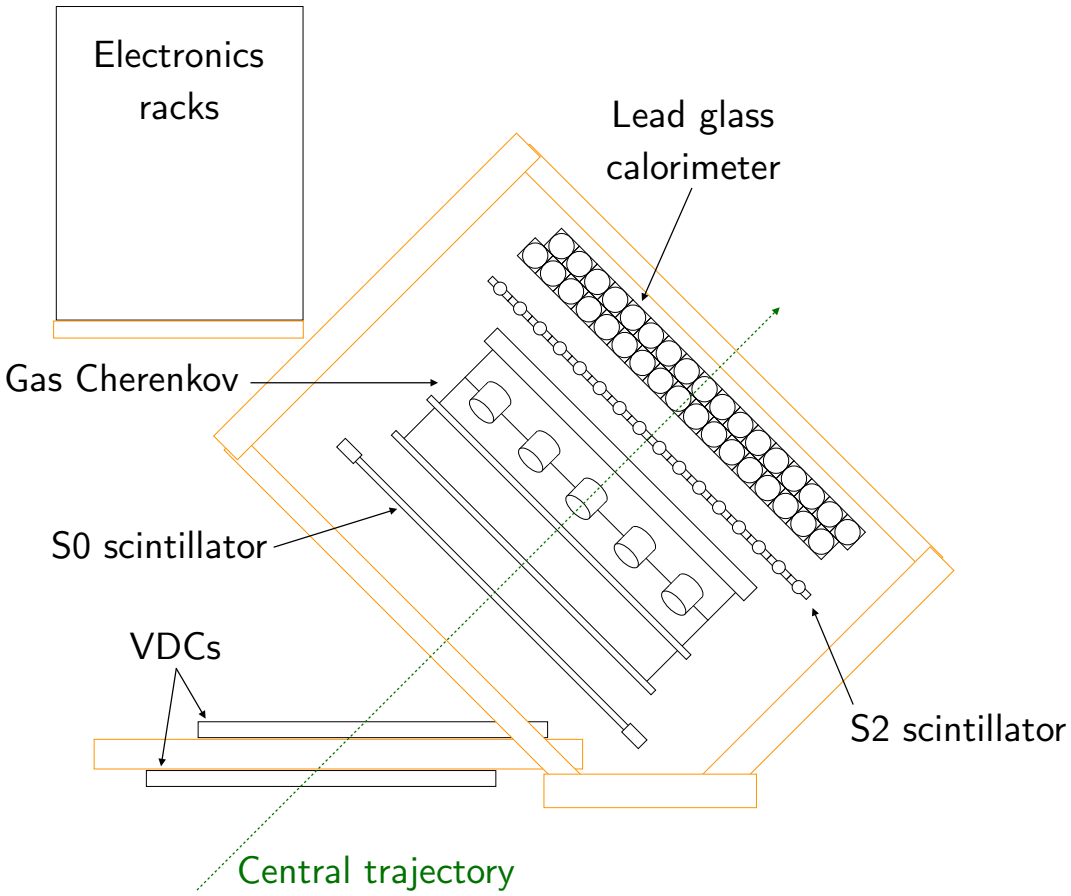


Figure 3.12: Schematic layout of detector package. The calorimeter is depicted in the LHRS configuration. Orange elements represent structural elements of the detector package.

3.2.5.1 Vertical drift chambers

As previously discussed, precisely determining the position and angle of particles at the focal plane is critical to reconstructing event kinematics. This was done using a pair of vertical drift chambers (VDCs) [44], located approximately 3.5 m downstream of the Q3 exit. Each VDC contains two grounded sense wire planes suspended between conducting foils held at -4 kV, creating a very strong electric field directed away from the wires. The chamber is circulated with an equal mix (by volume) of argon and ethane gas, which is ionized when charged particles pass through it. The freed electrons are accelerated by the field, and the resulting ionization avalanche results in charge that collects on the sense wires and is read out as the signal.

The wire directions of the two planes are orthogonal to each other and form a 45° angle with both the dispersive and non-dispersive axis of the TCS, such that the wire planes are horizontal in lab coordinates. Within each plane, the sense wires are separated by a perpendicular distance of 4.24 mm. The wire spacing and orientation of the wire planes ensures that the central trajectory of the HRS traverses 5 wire regions, and extreme tracks traverse 3 wire regions. A consequence of this orientation is that the ionization charge drifts vertically, hence the name vertical drift chamber. The two sense wire planes are separated by a vertical distance of 26 mm, and the two VDCs are positioned such that corresponding sense wire planes in each chamber are separated by a vertical distance of 0.335 m. See Figure 3.13.

The VDC uses drift time measurements from the sense wires to determine the position of the traversing particle. The measured drift time from the particle trajectory to the sense wire corresponds to the perpendicular distance between them. Each VDC wire is connected to a dedicated time to digital converter (TDC) channel, with the wire signal acting as the start and the event trigger acting as the common stop for all channels. Figure 3.14 shows a typical TDC spectrum for one wire. The peak at short drift time is populated by events that pass very close to the sense wire, where the electric field has transitioned from contstant

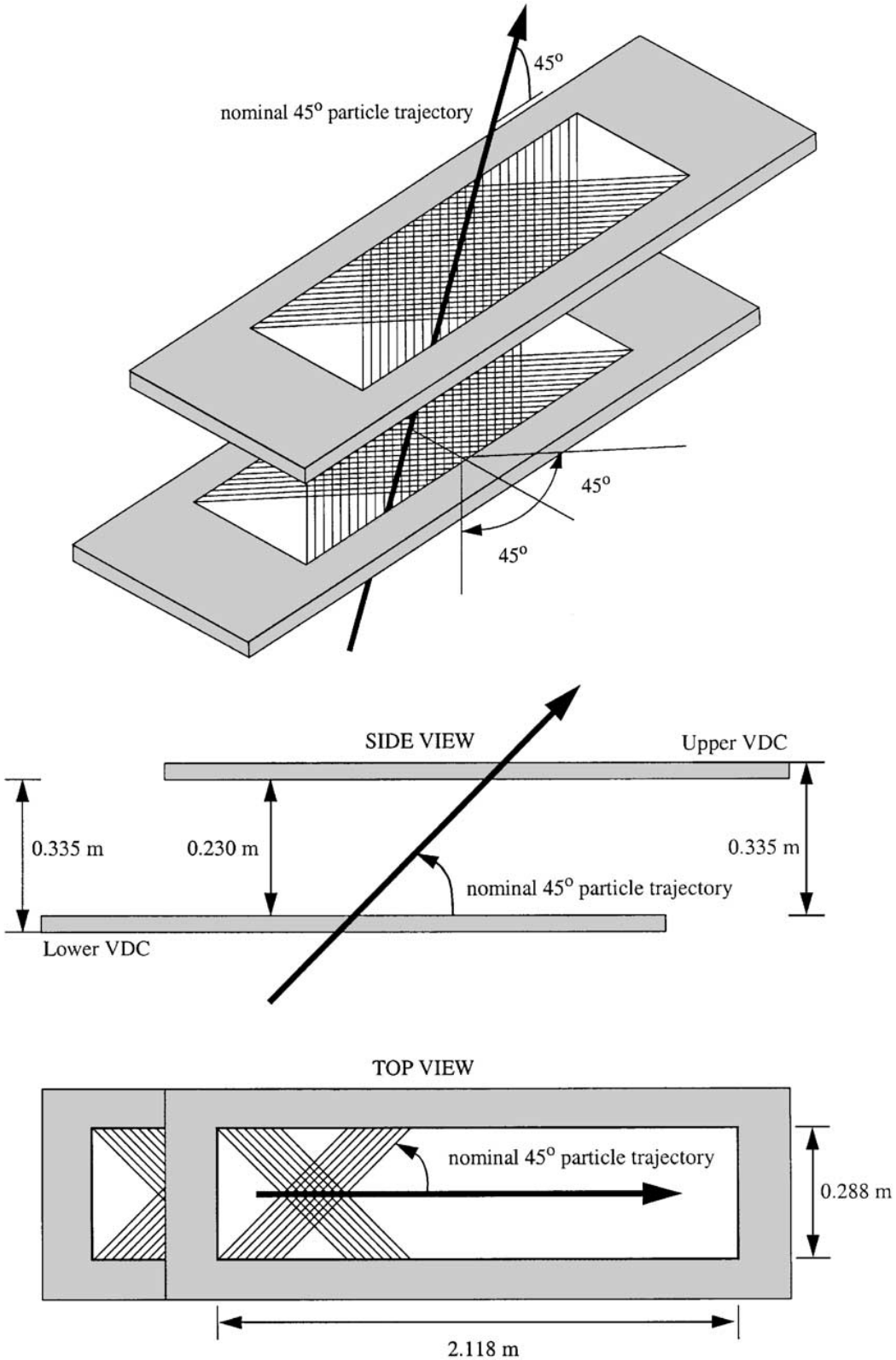


Figure 3.13: Schematic layout of VDCs. From [44].

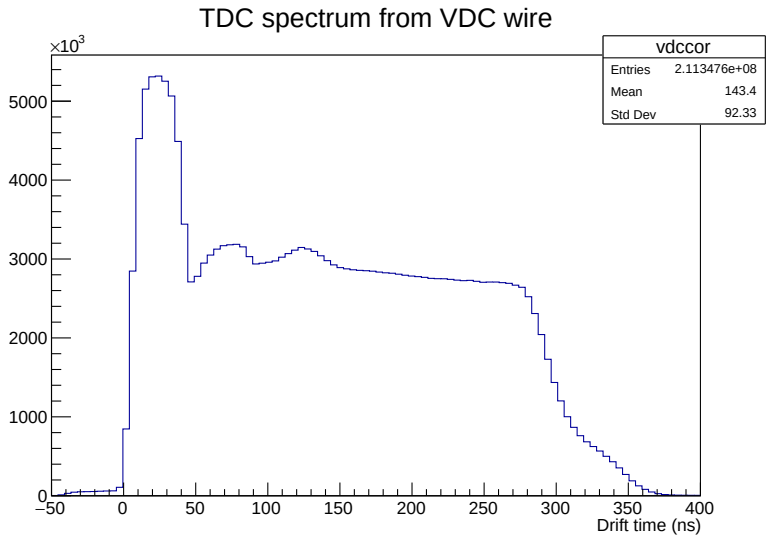


Figure 3.14: TDC spectrum of a VDC sense wire.

to quasiradial. The central region is populated by events that pass through the the VDC region with a uniform electric field, resulting in an approximately flat distribution. Events that pass very far from the sense wire have the largest drift time and decreasing probability of detection, resulting in the spectrum going to zero.

A group of adjacent sense wires with a TDC signal form a wire cluster. Timing measurements from a cluster can identify the 1D point (in that plane's coordinates) where the trajectory crossed the wire plane. For a nominal trajectory traversing five wire regions, the per-plane time resolution is 4.5 ns, corresponding to a position resolution of 225 μm. Since the sense wires in each plane are orthogonal, a cluster in both planes can determine the 2D intersection point. Correlated clusters in all four planes of both VDCs form a track, which is fully defined by its angle as well as position.

To compare drift times from different sense wires, timing offsets for each TDC channel need to be accounted for. This is completed by selecting a reference time t_0 to which all TDC spectra are aligned. Typically, t_0 is chosen to be the steep edge of low-drift time events at large TDC channel, identified by finding the drift time with the largest derivative. The TDC signals are offset (in software) to align t_0 for all wires. Figure 3.15 shows a two-dimensional histograms of calibrated TDC spectrum versus wire number. The edge of the short drift

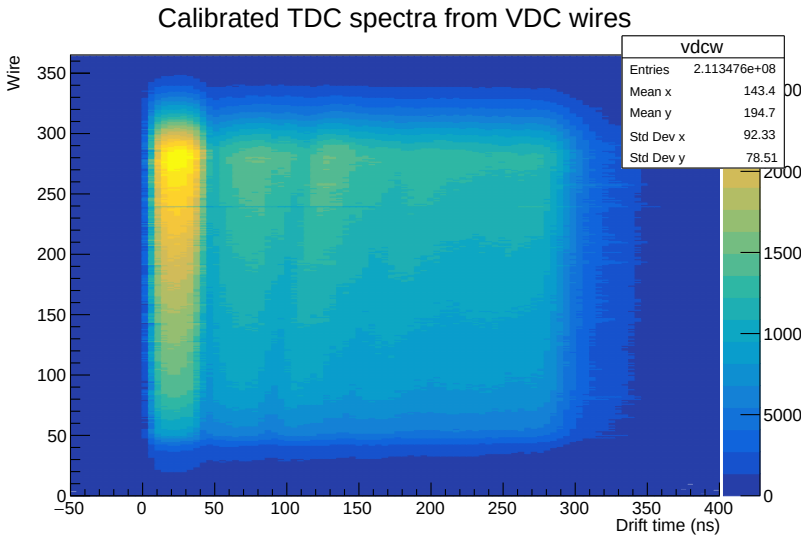


Figure 3.15: Calibrated TDC spectra of all VDC wires.

time peak is aligned for all wires.

3.2.5.2 Scintillators

Two plastic scintillating detectors were used for event triggering. The scintillation photons emitted by high energy particles in the plastic are detected by photomultiplier tubes (PMTs), indicating the incidence of a particle. The S0 scintillator consists of a single plastic paddle equipped with a PMT on the top (A) and bottom (B) of the paddle. The S2 scintillator consists of 16 adjacent paddles, each equipped with a PMT on the left (L) and right (R) side of the paddle. A schematic of the scintillators is shown in Figure 3.16. A scintillator trigger requires a coincidence between both PMTs viewing a paddle (A and B for S0, or L and R for any of the S2 paddles). See Section 3.2.5.5 for more details on the trigger.

3.2.5.3 Gas Cherenkov detector

A threshold gas Cherenkov detector [45] was used for both event triggering and particle identification. When high-energy charged particles pass through a material at speeds faster than the local phase velocity of light, Cherenkov radiation is emitted. The detector has a large gas volume that acts as the Cherenkov radiator. The depth of the detector is such that

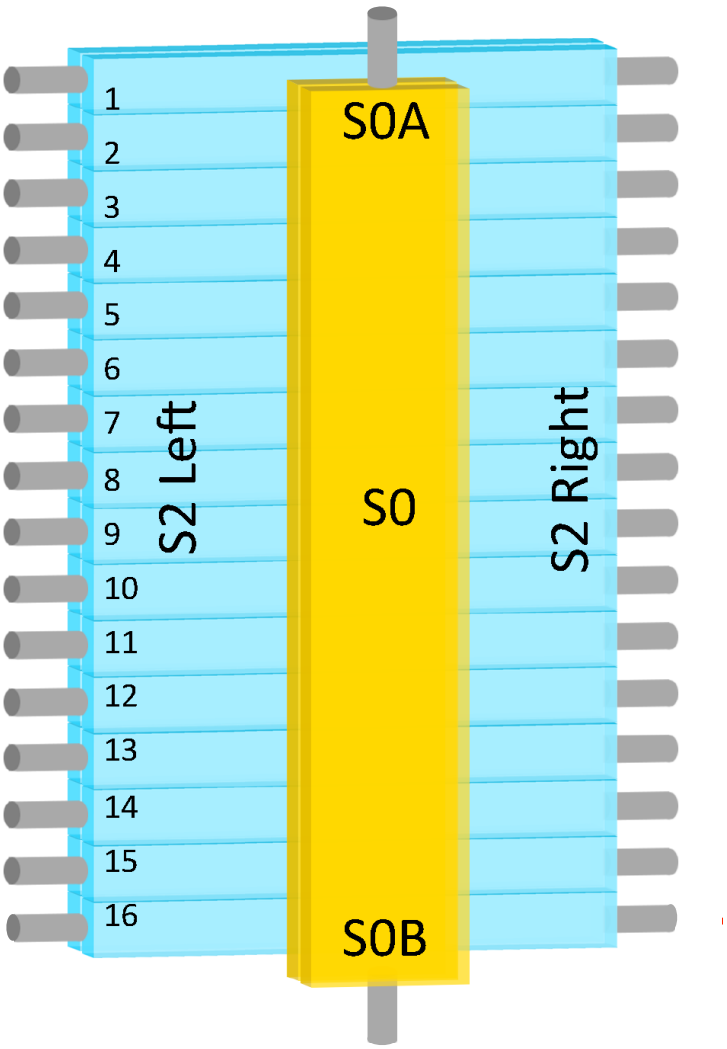


Figure 3.16: Schematic layout of scintillators.

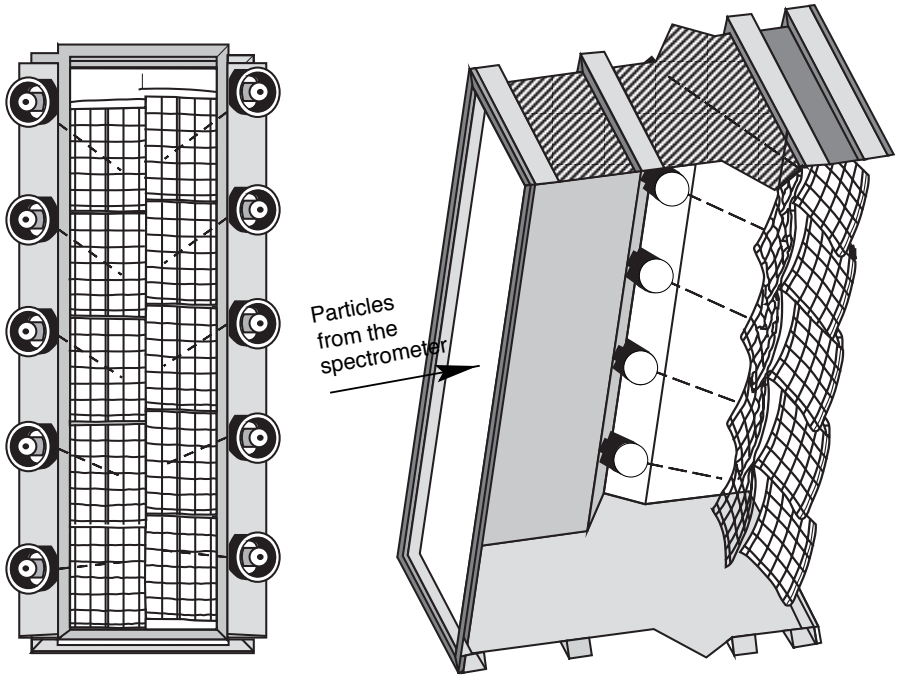


Figure 3.17: Front (left) and side cross section (right) schematic of gas Cherenkov detector. From [45].

the central trajectory of the LHRS (RHRS) has a pathlength of 1 m (1.5 m) through the gas. At the rear of the detector are ten spherical mirrors arranged in two columns, designed and positioned to reflect the Cherenkov radiation into PMTs on the sides of the detector. As the mirrors must be positioned in the path of the detected particle, they were designed to be very thin (5.5×10^{-3} radiation lengths) to minimize the effect on the particle's trajectory. See Figure 3.17 for a schematic of the detector. The detector is circulated with CO_2 gas at atmospheric pressure, which has a refractive index of $n = 1.00041$. Since the threshold for Cherenkov radiation is $\beta > 1/n$, this means that the minimum detection energy for electrons is 0.017 GeV. On the other hand, pions and protons require 4.8 and 32 GeV, respectively, which are outside of the HRS momentum range. This provides a very effective method of discriminating electronic from hadronic events.

Each PMT is connected to a dedicated ADC channel, with a typical spectrum shown in Figure 3.18. The large, sharp peak, is the pedestal, The small subsequent peak is the minimum possible response produced by the PMT, corresponding to the emission of one

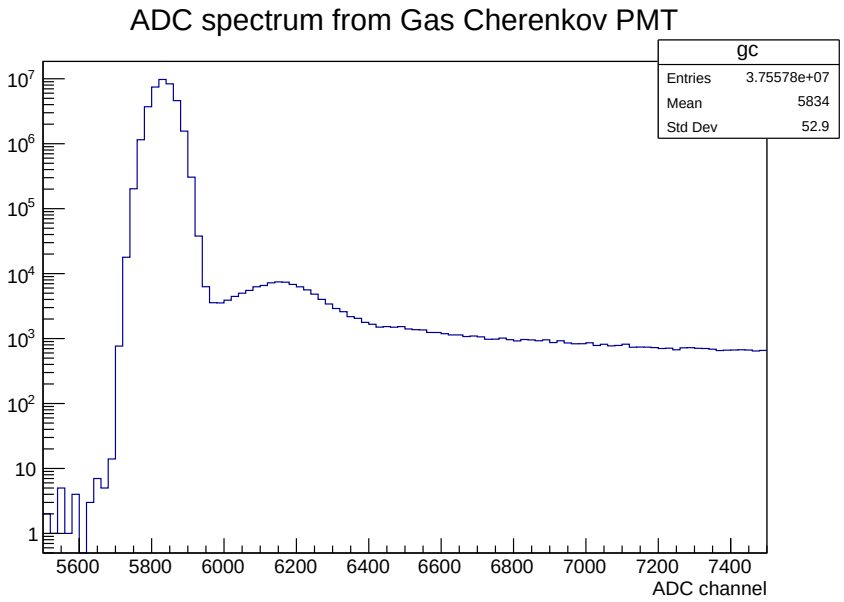


Figure 3.18: ADC spectrum of gas Cherenkov PMT.

photoelectron (PE) from the cathode. The continuous signal at higher channels is from the emission of multiple photoelectrons. The scaling between PEs and observed ADC signal depends on the PMT gain and the ADC sensitivity, which must be calibrated with some absolute scale. This is accomplished by extracting the gain coefficient required to scale the 1 PE peak for all PMTs to 300 channels above pedestal. For the ADC signal from each PMT, the pedestal and 1 PE peaks are fit with Gaussians, determining the mean positions $\overline{x}_{ped}$ and $\overline{x}_{\mathrm{1PE}}$. The gain coefficient is then determined by:

$$G = \frac{300}{\overline{x}_{\mathrm{1PE}} - \overline{x}_{ped}} \tag{3.15}$$

The total detector response is determined by summing the calibrated ADC signals from all PMTs.

3.2.5.4 Calorimeter

The final detector in each package is a lead glass calorimeter for particle identification. The calorimeters consist of two layers of lead glass blocks, with each block connected to a PMT.

High energy particles shower as they travel through the lead glass, and the resulting light is detected by the PMTs. Summing the shower signal from all blocks yields the energy deposited by the particle. Electrons deposit all their energy in the calorimeter, but hadrons do not.

In the LHRS, the first lead glass layer contains 34 blocks of dimension 15 cm × 15 cm × 30 or 35 cm, oriented perpendicular to the particle tracks. The blocks are arranged in two columns, with one column containing the shorter blocks and the other column containing the longer blocks. The second layer is nearly identical, with the short/long blocks swapped and the whole layer shifted by 50 mm. This ensures that no block gaps exist continuously through both layers. In the RHRS, the first layer contains 48 blocks of dimension 10 cm × 10 cm × 35 cm, oriented perpendicular to the particle tracks and arranged in two columns. The second layer contains 75 blocks of dimension 15 cm × 15 cm × 35 cm, oriented parallel to the particle tracks and arranged in a 5×15 grid. Schematics of both calorimeters are shown in Figure 3.19.

Each lead glass block is attached to a PMT connected to a dedicated ADC channel. The scaling between the absolute energy deposited in the block and the observed ADC signal depends on the PMT gain and ADC sensitivity, which must be calibrated using the absolute energy of each event. The energy deposited in a single block is simply the product of the block's coefficient C and ADC signal A. Thus, the total energy of a single event i can be expressed by a sum over all calorimeter blocks j:

$$E_i = \sum_{j}^{blocks} C_j A_{ij}. \tag{3.16}$$

The coefficients C_j can be calibrated by comparing the total calorimeter energy, E_i, to the event momentum obtained by the optics reconstruction of the VDC track, p_i. Unlike the gas Cherenkov calibration, where the gain coefficients could be individually determined, the calorimeter calibration requires the sum of all blocks. Thus, the block coefficients must be

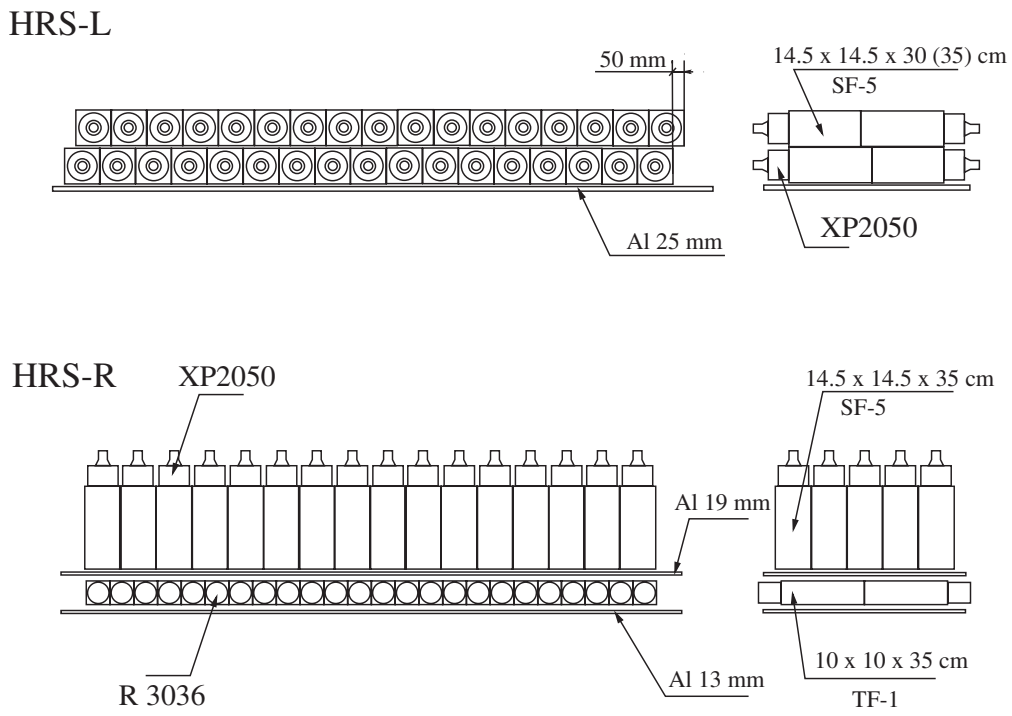


Figure 3.19: Schematic layout of shower calorimeters. Particles enter from the bottom. TF-1 and SF-5 refer to the type of lead glass. XP2050 and R3060 refer to the PMT model. From [42].

determined simultaneously.

The goodness of a calibration can be described by defining:

$$\chi^2 = \sum_{i}^{events} \left(\sum_{j}^{blocks} C_j A_{ij} - p_i \right)^2 \tag{3.17}$$

By minimizing χ^2 with respect to the block coefficients, one can find the set of coefficients required to convert observed ADC signal to absolute energy.

As a cross check, two methods were used for minimizing χ^2. The first method used the TMinuit class, a C++/ROOT implementation of the Minuit minimization program, with the block coefficients as free parameters. The calibration obtained with this method was found to be stable agains variation in initial parameter values and step sizes. The second method reduced the minimization to a set of linear equations by taking the derivative of Equation 3.17 with respect to the coefficient of block k and setting equal to zero to find the

minimum,

$$\frac{\partial \chi^2}{\partial C_k} = 0 \tag{3.18}$$

leading to the equation

$$\sum_{i}^{events} \left(\sum_{j}^{blocks} C_k A_{ij} A_{ik} \right) = \sum_{i}^{events} p_i A_{ik}. \tag{3.19}$$

This can be summarized for all blocks, $k = \{0, 1, 2...N\}$, with the matrix equation:

$$\sum_{i}^{events} \begin{pmatrix} A_{i0}A_{i0} & A_{i0}A_{i1} & A_{i0}A_{i2} & \dots & A_{i0}A_{iN} \\ A_{i1}A_{i0} & A_{i1}A_{i1} & A_{i1}A_{i2} & \dots & A_{i1}A_{iN} \\ A_{i2}A_{i0} & A_{i2}A_{i1} & A_{i2}A_{i2} & \dots & A_{i2}A_{iN} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{iN}A_{i0} & A_{iN}A_{i1} & A_{iN}A_{i2} & \dots & A_{iN}A_{iN} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ \vdots \\ C_N \end{pmatrix} = \sum_{i}^{events} \begin{pmatrix} p_i A_{i0} \\ p_i A_{i1} \\ p_i A_{i2} \\ \vdots \\ p_i A_{iN} \end{pmatrix} \tag{3.20}$$

Thus, by filling the leftmost and rightmost matrices with ADC and tracking information from each event, this expression can to solved for the column of block coefficients.

The calibrations obtained with these two methods were effectively identical. The calibrated energy spectrum from the calorimeter is shown in Figure 3.20. A minimum gas Cherenkov signal cut has been applied in order to emphasize the electron peak at 3.1 GeV, corresponding to the momentum setting of the spectrometer.

3.2.5.5 Trigger setup

The event triggers for MARATHON were formed using various combinations of the S0 and S2 scintillators and the gas Cherenkov detector. The LHRS and RHRS had separate triggers, but were formed from identical combinations of detectors. Three event triggers were defined, numbered T1-T3 for the LHRS and T4-T6 for the RHRS:

- T2/T4: Coincidence of both scintillators, S0 & S2

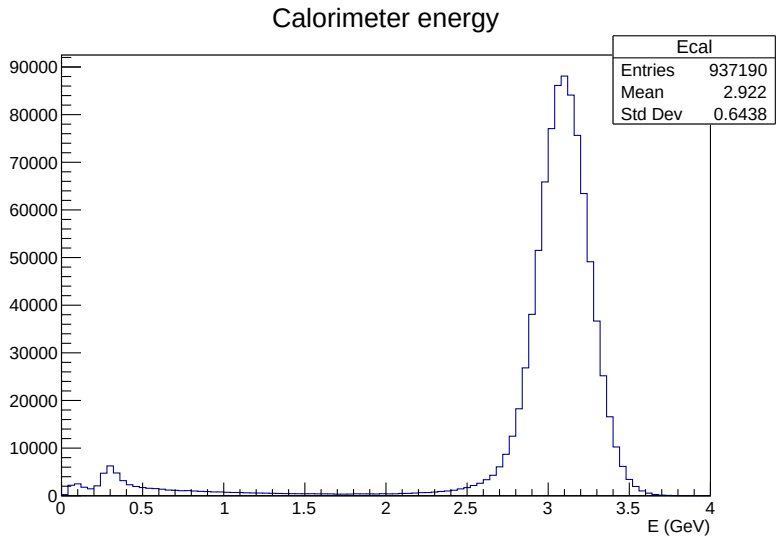


Figure 3.20: Energy spectrum from calorimeter.

- T3/T5: Coincidence of both scintillators and gas Cherenkov, (S0 & S2) & GC
- T4/T6: Coincidence of either scintillator and gas Cherenkov, (S0 — S2) & GC

Both HRS also had a clock trigger operating at 103.7 kHz.

3.3 Runplan

MARATHON production data was taken with a beam energy of 10.6 GeV, the maximum energy deliverable to Hall A during the run period. This beam energy, made possible by Jefferson Lab's 12 GeV energy upgrade, was critical to accessing the high momentum and energy transfers required to make high-x measurements in the DIS region. Different kinematic settings can be accessed by adjusting the HRS momentum and/or angle; during MARATHON, the HRS momentum was fixed, and higher x points were accessed by increasing the HRS angle. The LHRS momentum was set at 3.1 GeV. The cryogenic cooling on the RHRS dipole was found to be unstable when the magnet was set above 3 GeV, so the RHRS momentum was set at 2.9 GeV. These settings correspond to energy transfers of $\nu \approx 7.5$ GeV.

Due to safety considerations for the tritium target, limits were placed on beam operations

KIN	E (GeV)	E' (GeV)	θ (°)	x	Q^2 (GeV^2)	W^2 (GeV^2)
1	10.6	3.1	17.58	0.22	3.07	11.89
2	10.6	3.1	19.12	0.26	3.62	11.33
3	10.6	3.1	20.58	0.30	4.19	10.76
4	10.6	3.1	21.93	0.34	4.76	10.20
5	10.6	3.1	23.21	0.38	5.32	9.63
7*	10.6	3.1	25.59	0.46	6.45	8.51
9*	10.6	3.1	27.78	0.54	7.57	7.38
11*	10.6	3.1	29.92	0.62	8.76	6.20
13*	10.6	3.1	31.73	0.70	9.82	5.13
15*	10.6	3.1	33.56	0.78	10.96	4.00
16*	10.6	2.9	36.12	0.82	11.82	3.51

Table 3.2: MARATHON runplan.
KIN1-KIN15 were taken with the LHRS. KIN16 was taken with the RHRS.
*took data both before and after the two week beam shutdown

to Hall A. These included a maximum allowed beam current of 22.5 μA, and a mandatory 4 mm $\times$ 4 mm raster. These restrictions applied to all targets, not just the tritium target.

Table 3.2 shows the kinematic settings of production data. The KIN numbering is based on the original runplan, in which kinematic settings were equally spaced with $\Delta x = 0.04$. Due to time constrains, even-numbered settings above KIN5 were eliminated, with the exception of the highest x setting, KIN16. For nearly the whole production run, the RHRS was set to KIN16, where the event rate was lowest. The LHRS was used to scan the lower kinematics. Data for KIN7-KIN15 was taken in two passes. The first pass covered the kinematic range more quickly at the cost of not collecting the full desired statistics. From KIN1 to KIN5, data was taken on hydrogen, deuterium, tritium, helium-3, and the empty cell. For the remaining kinematics, hydrogen was omitted.

During each kinematic setting, the targets were alternated for periods of approximately one hour of beam on target. When the desired statistics for a target was reached, it would be removed from the rotation. Following beam downtime of 2 hours or more, the carbon hole target was used to verify the beam was centered on the target ladder.

MARATHON's kinematic coverage (combining data from all targets) is shown in Fig-

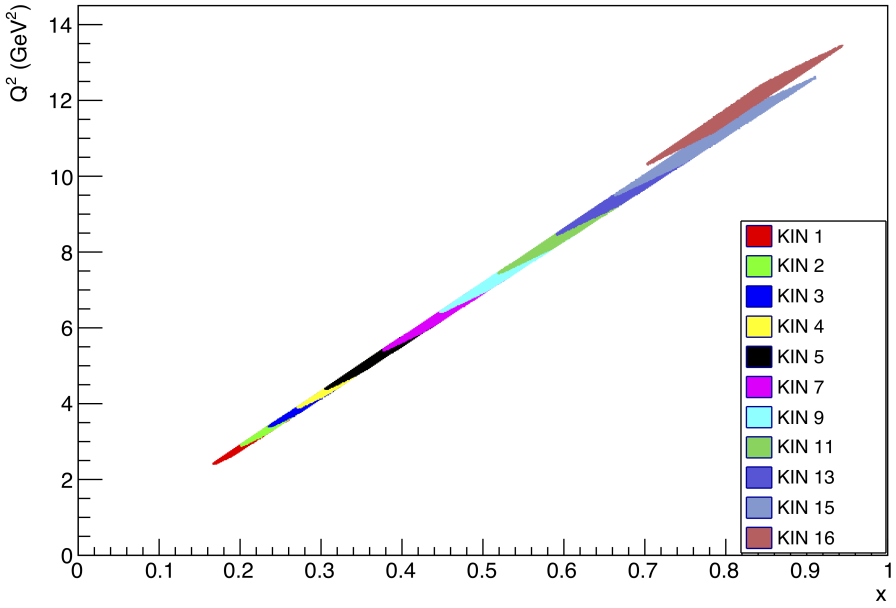
ure 3.21. The different colors represent different spectrometer settings. The Q^2 of the data depends strongly on x, scaling as:

$$Q^2 = (14\ \text{GeV}^2) \cdot x \tag{3.21}$$

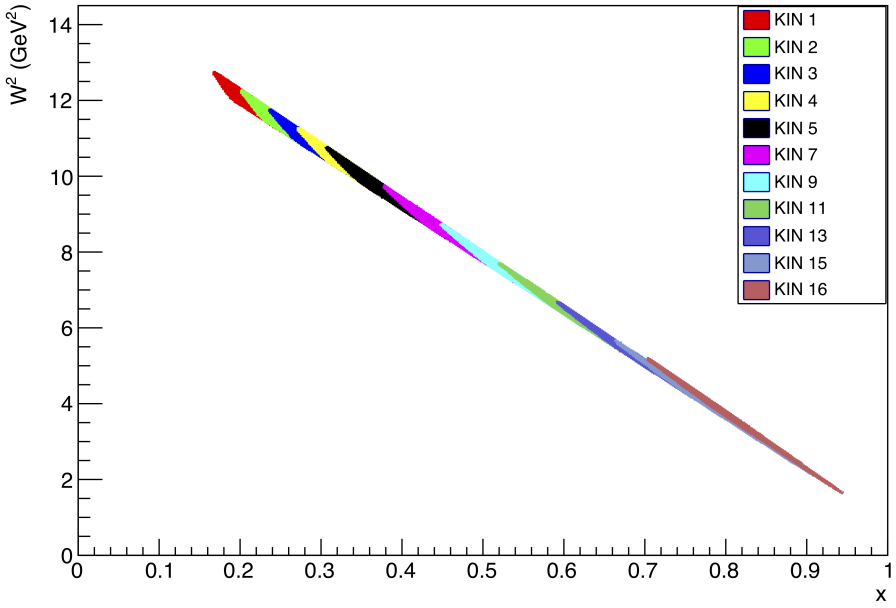
Requiring $W^2 > 3\ \text{GeV}^2$ eliminated a small portion of the highest-x data. As all the data had $Q^2 > 2\ \text{GeV}^2$, no Q^2 was necessary.

It should be noted that in the middle of the experimental run, multiple transformers were blown in the central helium liquefier (responsible for supplying 2 K helium-4 to the cryomodules). The accelerator was shut down for approximately two weeks while the transformers were replaced. This happened in between the first and second pass on the LHRS. No differences in the data quality before and after the shutdown were observed.

In addition to production data, a variety of specialized runs were taken for calibrations and systematic corrections. These will be addressed in the following chapter.



(a) x vs. Q^2 (GeV^2)



(b) x vs. W^2 (GeV^2)

Figure 3.21: MARATHON kinematic coverage.

Chapter 4

Data analysis

4.1 Cross section ratios

An absolute cross section can be extracted from experimental data by

$$\sigma \equiv \frac{d^2\sigma}{dE'\, d\Omega} = \left(\frac{1}{\Delta E'\, \Delta\Omega\, \varepsilon}\right)\left(\frac{N}{Q\, \eta}\right) \tag{4.1}$$

where $(\Delta E'\, \Delta\Omega)$ is the energy and solid angle acceptance of the spectrometer, ε is the detection efficiency of the detector system, N is the number of scattered electrons counted by the detector system, Q is the beam charge deposited on the target, and η is the thickness (nuclei $\cdot$ cm^{-2}) of the target. Note that for MARATHON, all the gas target cells had the same geometry, so η can be replaced by the number density n (nuclei $\cdot$ cm^{-3}), which differs by a factor of the cell length.

The quantities in the first parentheses depend only on intrinsic properties of the spectrometer and detector system. There are multiple methods that can be employed to find the spectrometer acceptance, which typically require a Monte Carlo simulation of the spectrometer and input from a cross section model. This process is complicated and can contribute significantly to systematic uncertainties. Additionally, as the acceptance can depend on the spectrometer settings, the acceptance must be determined for each kinematic setting.

Detector efficiencies should not depend on the spectrometer position, as long as the detector settings remain unchanged. Conversely, the second parentheses contains quantities that depend on the target. These depend on the underlying DIS processes of interest: given a certain number of scattering opportunities (also called luminosity, $L = Qn$), how many electrons are scattered and detected at particular kinematics?

MARATHON is primarily interested in determining cross section ratios between various nuclear targets, based on simultaneous measurements carried out with identical spectrometer and detector settings. This has an important impact on the required analysis. When forming cross section ratios, quantities related to the spectrometer and detector will cancel, leaving only the quantities that depend on the target. Therefore, the cross section ratio between nucleus A and B will reduce to

$$\frac{\sigma_A}{\sigma_B} = \left(\frac{N_A}{Q_A\, n_A}\right) \Big/ \left(\frac{N_B}{Q_B\, n_B}\right) \equiv \frac{Y_A}{Y_B}, \tag{4.2}$$

where $Y = N/Qn$ is referred to as the normalized yield. Thus, the analysis only needs to determine the normalized yields for each target, and the ratio of these are equivalent to the ratio of the corresponding cross sections.

Practically, the normalized yields will be extracted from a series of many data acquisition (DAQ) runs. These data must be combined to give a single value of the measured yield. This can be done by summing the total number of detected electrons and normalizing by the total luminosity over the course of the whole experimental run. That is,

$$Y = \frac{\sum N_i}{\sum Q_i n_i}, \tag{4.3}$$

where the sum is over run number i.

First, cuts must be developed to ensure that only appropriate electron scattering events are counted (Section 4.2). Next, various corrections must be developed for both physical effects and sources of contamination (Section 4.3). Once this is complete, the normalized

yields and cross section ratios can be determined.

4.2 Cuts

4.2.1 Kinematics

To ensure that only DIS events were used in the MARATHON analysis, a cut was placed on the invariant mass W^2 of the final hadronic state:

$$W^2 > 3 \text{ GeV}^2 \tag{4.4}$$

This cut eliminated a relatively small number of KIN16 events (the highest x data). All of the MARATHON data satisfied the desired minimum momentum transfer $Q^2 > 2$ GeV^2, so no explicit cut on Q^2 was required.

4.2.2 Acceptance

A set of cuts on the target observables θ_{tg}, ϕ_{tg}, and $\delta p/p$ were applied to avoid the edges of the spectrometer acceptance. This is necessary as the optics reconstruction (see Section 3.2.4) is over-constrained, leading to erratic behavior on the acceptance boundary. For the HRS, the standard cuts typically used are:

$$\begin{aligned} -0.06 \leq \theta_{tg} \leq 0.06 \\ -0.03 \leq \phi_{tg} \leq 0.03 \\ -0.04 \leq \delta p/p \leq 0.04 \end{aligned} \tag{4.5}$$

where $\theta_{tg} = \phi_{tg} = \delta p/p = 0$ corresponds to the central trajectory of the spectrometer. Recall that the HRS have a fractional momentum resolution of approximately 10^{-4}. In addition to avoiding the acceptance edges, the acceptance cuts for MARATHON must satisfy an addi-

tional condition. As described in Section 4.1, extracting cross section ratios from normalized yields assumes that the acceptance is the same for both sets of scattering data. Therefore, MARATHON must ensure that within the applied acceptance cuts, the acceptance function is the same for each target and will cancel in cross section ratios.

The yield in a particular bin will be proportional to the product of the acceptance function and cross section:

$$N \propto A(\nu,\ \theta)\sigma(\nu,\ \theta) \tag{4.6}$$

Therefore, the shape of the acceptance function can be obtained by taking the ν vs. θ event distribution and normalizing by the cross section of each bin (obtained from a model). The ratio of acceptance functions for different targets should be flat. The acceptance ratios for each cross section ratio measured by MARATHON are shown in Figure 4.1. As expected, the ratios are mostly flat. The large deviations on the edges are statistical fluctuations in less populated bins. Note that if the full normalization, including luminosity, was taken into account, each of these ratios should be unity.

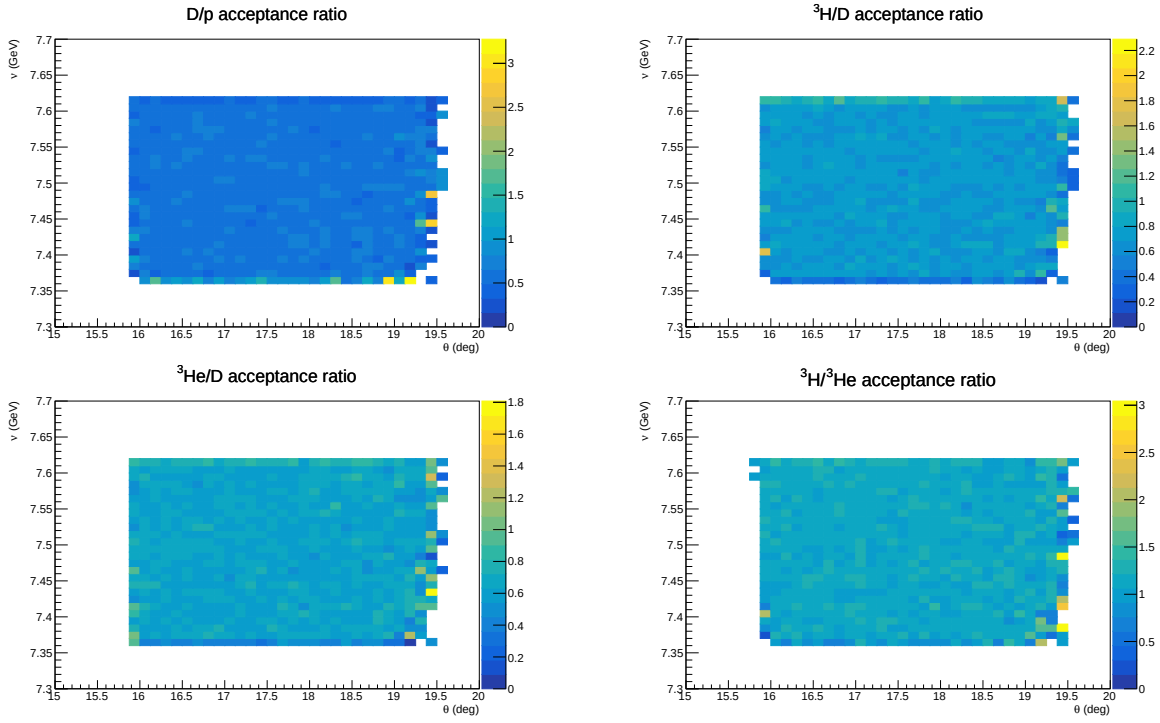


Figure 4.1: Acceptance function ratio for each cross section ratio.

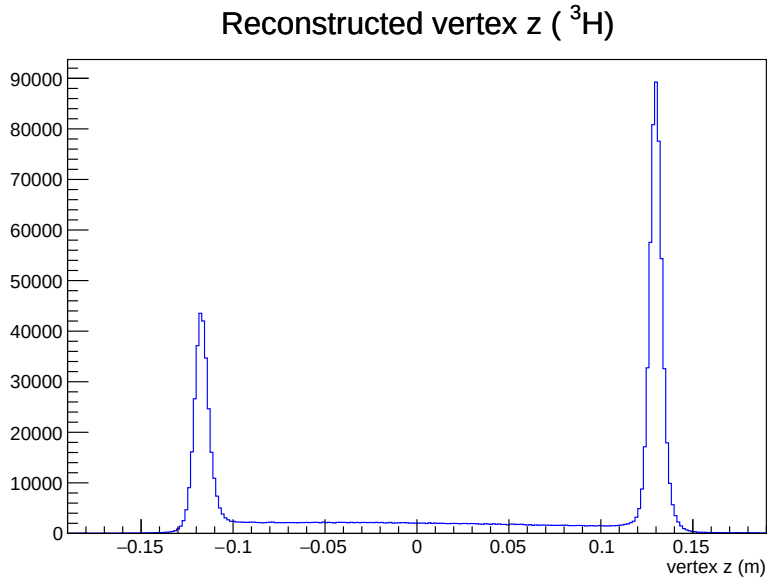


Figure 4.2: Reconstructed vertex z for tritium target.

4.2.3 z vertex

The reconstructed scattering vertex z is shown in Figure 4.2. The most dominant feature in the distribution are two peaks arising from electrons scattering off the aluminum endcaps of the target cell. In order to isolate events scattered from the gas target nuceli, a cut is placed on vertex z to exclude the endcap peaks. Despite this cut, there is a residual contamination from endcap scattering events that are mis-reconstructed in the center of the target (see Section 4.3.3). As the contamination depends on the HRS angle, a different vertex z cut was allowed for each kinematic setting. When the HRS are set to a more forward angle, the contamination from mis-reconstructed events is larger. However, the event rate is high enough that a stricter vertex z cut can be applied while retaining adequate statistics. For larger HRS angles, the contamination is smaller, while the event rate is lower. This allows a more relaxed vertex z cut to increase the included statistics. The vertex z cuts for each kinematic setting are shown in Table 4.1.

KIN	z_{min} (cm)	z_{max} (cm)
1	-8.0	10.0
2	-8.0	10.0
3	-8.0	10.0
4	-8.0	10.0
5	-9.0	10.0
7	-9.0	10.0
9	-9.5	10.0
11	-9.5	10.0
13	-10.0	10.5
15	-10.0	10.5
16	-10.5	11.0

Table 4.1: Vertex z cut used for each kinematic setting.

4.2.4 Tracking

Two cuts are applied using the tracking information from the VDCs. A cut on $\beta = v/c$ is applied to separate up-going particles exiting the spectrometer from down-going cosmic rays (the sign of β depends on which VDC detects the particle first). Also, a cut is applied selecting events with strictly one track.

$$\begin{aligned} \beta &> 0 \\ n_{track} &= 1 \end{aligned} \tag{4.7}$$

Obviously, a good electron event must have $n_{track} \geq 1$. For events with multiple tracks, a series of more complicated cuts can be used to determine which track should be used. However, MARATHON observed that only approximately 1% of all events meeting good electron cuts contained more than one track. It was decided that this negligible gain in statistics was not worth making the more complicated cuts, so only single-track events were included.

4.2.5 Particle identification

The gas Cherenkov detector and lead glass calorimeter are used to make particle identification (PID) cuts to separate electron from hadron events in the detector. In particular, these PID cuts must exclude the pions that are commonly part of the final hadronic state of DIS scattering. Due to the β threshold for Cherenkov radiation, requiring a minimum signal from the gas Cherenkov detector effectively excludes hadrons, which are much more massive than electrons. In the calorimeter, the electromagnetic shower caused by an electron is expected to deposit all of the electron's energy in the detector, while hadrons only leave a portion of their energy. Requiring a minimum E/p, where E is the total energy deposited in the calorimeter and p is the reconstructed track momentum, is also effective at separating electrons from hadrons. MARATHON used the following PID cuts on the gas Cherenkov signal GC and calorimeter E/p to isolate electrons:

$$\begin{aligned} GC &> 1500 \text{ (LHRS)} \\ GC &> 2000 \text{ (RHRS)} \\ E/p &> 0.7 \end{aligned} \tag{4.8}$$

These cuts are visualized in Figure 4.3. The efficiency of these cuts must be determined to estimate the residual contamination from non-electron events. Traditionally, this is done by applying a tight cut on one detector to select a clean electron (or pion) sample, and determining the electron (or pion) efficiency of the cut on the other detector by

$$\epsilon = \frac{N_{cut}}{N_{sample}} \tag{4.9}$$

where N_{cut} is the number of electrons (or pions) passing the cut, and N_{sample} is the number of particles in the sample. The total PID efficiency is then found by the product of the

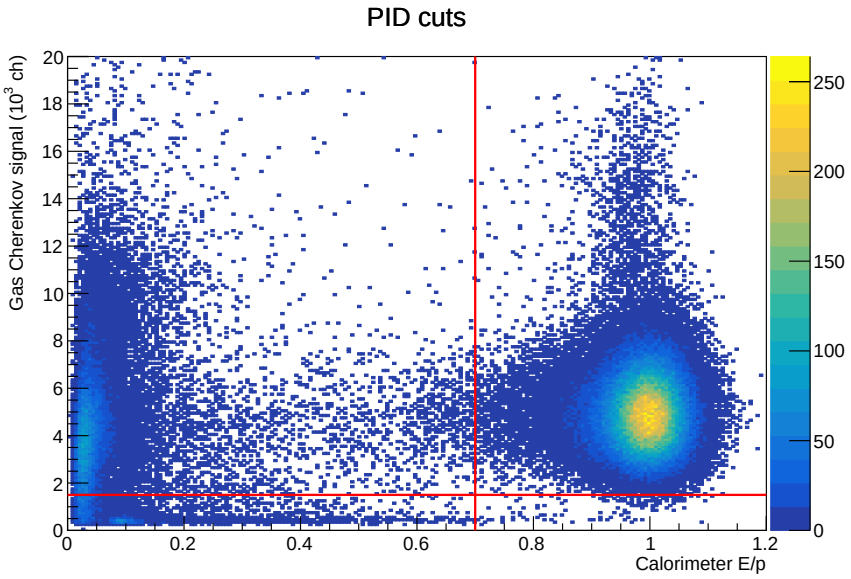


Figure 4.3: PID cuts used for MARATHON. LHRS gas Cherenkov cut is shown.

individual gas Cherenkov and calorimeter efficiencies:

$$\epsilon_{PID} = \epsilon_{GC}\epsilon_{cal} \tag{4.10}$$

In additional to observing the expected pion peak, leaving a low energy signal in both the gas Cherenkov and calorimeter, MARATHON observed an unexpected background signal. This signal, while leaving just trace amounts of energy in the calorimeter, was able to produce large signals in the gas Cherenkov detector (see again Figure 4.3). Some hypotheses regarding the origin of this signal are that high-energy neutral particles (produced by the upgraded CEBAF beam energy) are scattered from the HRS dipole into the detectors, or that charged pions produced in the target decay into muons. While this signal can still be completely excluded with a cut on E/p, it complicates the calorimeter efficiency calculation as it is impossible to get a clean electron sample with a gas Cherenkov cut alone.

An alternative approach was taken to determine the efficiencies, using four populations of events. N_0 is the total number of events passing the acceptance and tracking cuts, with no PID cut applied. N_{GC} is the subset of N_0 that passes the gas Cherenkov cut. N_{cal} is

the subset of N_0 that passes the calorimeter E/p cut. Finally, N_{PID} is the subset of N_0 that passes both the gas Cherenkov and E/p cut. The total population N_0 is the sum of the number of electrons N_e and non-electrons N_x. The probabilities that an electron passes the individual detector cuts are P_e^{GC} and P_e^{cal}, with analogous probabilities P_x^{GC} and P_x^{cal} for non-electrons. These probabilities are equivalent to the detector cut efficiencies. The relationship between these quantities is summarized in the following equations.

$$N_0 = N_e + N_x \tag{4.11}$$

$$N_{GC} = P_e^{GC} N_e + P_x^{GC} N_x \tag{4.12}$$

$$N_{cal} = P_e^{cal} N_e + P_x^{cal} N_x \tag{4.13}$$

$$N_{PID} = P_e^{GC} P_e^{cal} N_e + P_x^{GC} P_x^{cal} N_x \tag{4.14}$$

Note that since a clean electron sample can be obtained with a cut on E/p, P_e^{GC} and P_x^{GC} can be calculated using the traditional approach. This leaves four unknowns to be found using Equations (4.11)-(4.14). The resulting relative contamination from non-electron events that pass both PID cuts can be calculated by

$$\frac{x}{e} = \frac{P_x^{GC} P_x^{cal} N_x}{N_{PID}}, \tag{4.15}$$

and was found to be negligibly small for all targets.

4.3 Systematic corrections

4.3.1 DAQ deadtime

When the DAQ gate is active following an event, additional event triggers attempting to activate the gate are ignored. This is known as DAQ deadtime. Every event trigger is

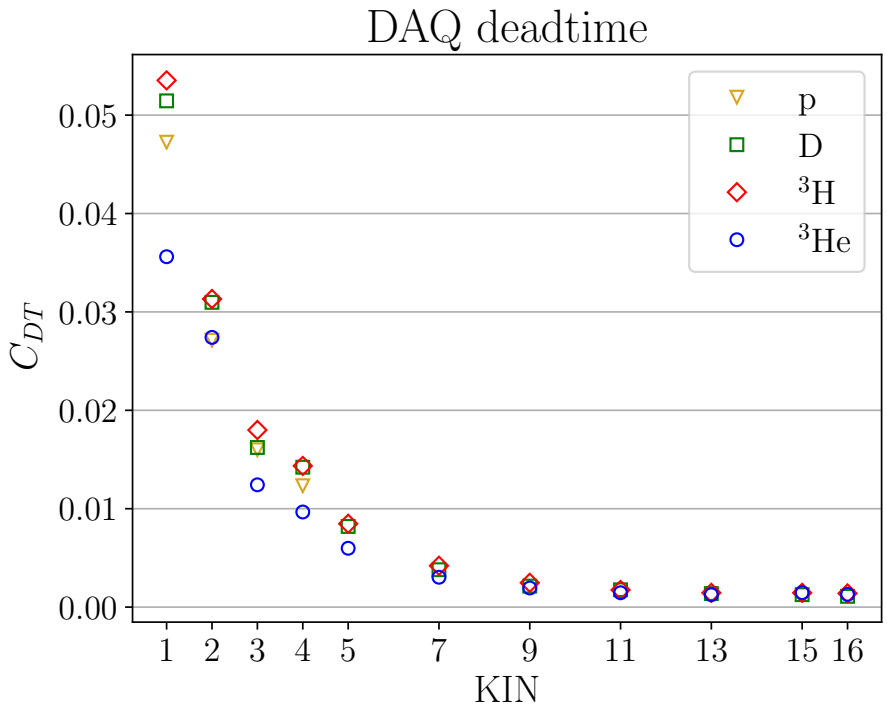


Figure 4.4: Average DAQ deadtime for each kinematic setting.

counted by a scaler, regardless of whether or not the event trigger activated the DAQ gate to record the event information. This allows the percentage of events excluded due to deadtime to be determined by:

$$C_{DT} = \frac{N_{scaler} - N_{DAQ}}{N_{scaler}} \tag{4.16}$$

where N_{scaler} is the total number of event triggers counted by the scalar, and N_{DAQ} is the total number of events recorded by the DAQ. The number of good electron events meeting all cuts can then be corrected to account for events that were excluded due to deadtime. The average deadtime for each of MARATHON's kinematic settings is shown in Figure 4.4.

4.3.2 Target density fluctuation

As the beam passes through the target, it deposits energy in the cell walls and gas itself, raising the temperature of the target. The gas targets are sealed, and therefore the overall gas density must remain constant. However, the gas volume immediately surrounding the path of the beam is heated more than the rest of the gas, causing a localized decrease in gas

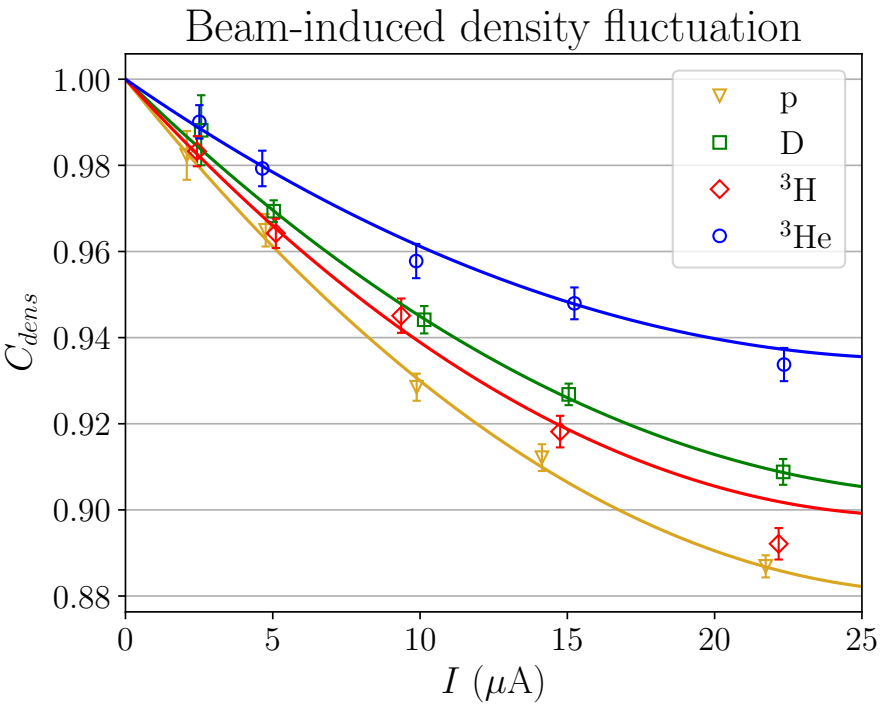


Figure 4.5: Beam-induced density fluctuation with fit for each MARATHON target.

density that must be accounted for when calculating luminosity. This is oftentimes referred to as target "boiling", although the target is not actually boiling. Since the heating effect depends on the power being deposited in the target, it will be a function of the beam current.

To quantify this correction, several runs were taken on each gas target with currents ranging from 2.5 μA to the maximum allowed 22.5 μA. The normalized yield (corrected for DAQ deadtime and detector efficiencies) was calculated for each run. Any dependence of the yield on current must be due to beam-induced changes in the local target density. As a proof of principle, this method was tested on a solid target, which does not experience these density fluctuations. The solid target showed a normalized yield that was independent of beam current. On the other hand, the gas target data shows a clear decrease (see Figure 4.5) that is fit with a second-order polynomial in beam current I:

$$C_{dens}(I) = aI^2 + bI + c \tag{4.17}$$

Since the correction must reflect the relative decrease in target density, the data for each

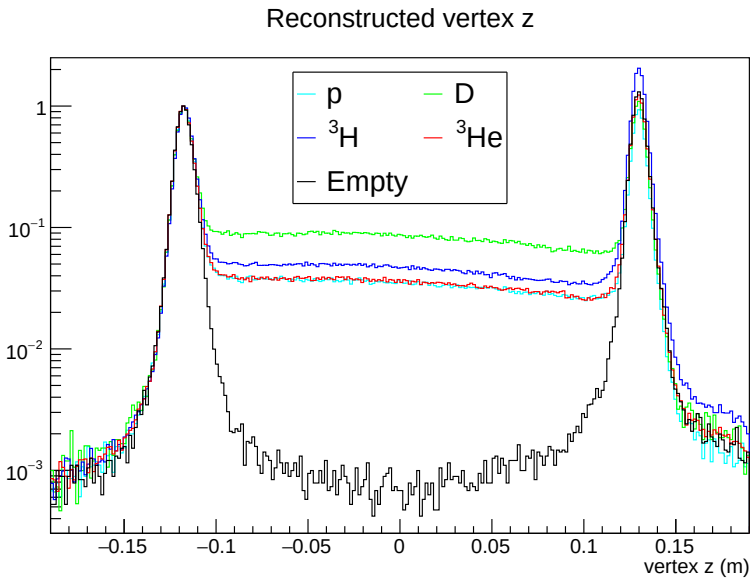


Figure 4.6: Yield-normalized reconstructed vertex z for all target cells.

target is normalized such that $c = 1$. When the luminosity for each run is calculated, the measured target density must be corrected by a factor of $C_{dens}(I_{avg})$, where I_{avg} is the average beam current for that run.

4.3.3 Endcap contamination

Although cuts are placed on vertex z to eliminate the endcap peaks, due to the resolution of reconstruction, it is possible that events scattered from the endcaps are reconstructed in the gas volume of the cell. Figure 4.6 shows the vertex z distribution for all gas cells, including the empty target cell. While there is no gas in the cell volume for electrons to scatter from, there is a continuous distribution of events reconstructed between the endcaps.

To quantify the size of this contamination, a common normalization must be found to compare the empty cell data to the gas cell data. This was done by finding the normalized yield of the upstream endcap for each cell, and scaling them to match (as shown in Figure 4.6). The same cuts used for the primary yield analysis (neglecting the vertex z cut) were used to determine the endcap yields. The relative endcap contamination C_{EC} for each target was found by integrating the empty cell distribution over the accepted vertex z range, and

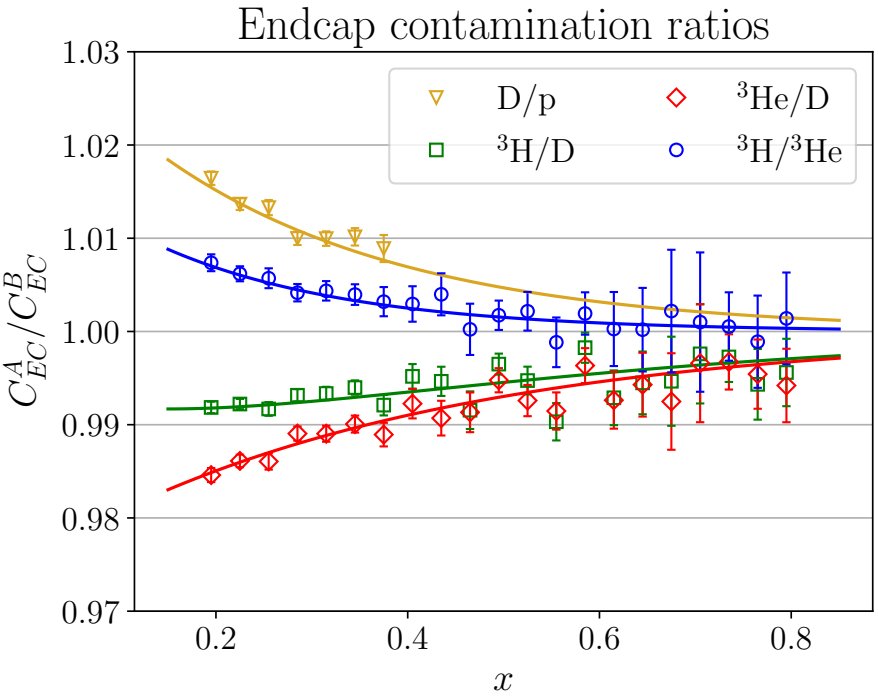


Figure 4.7: Endcap contamination data and fit for each target ratio.

dividing by the integral of the gas cell distribution over the same range. The contamination is dependent on spectrometer angle, and must be determined for each of MARATHON's kinematic settings. The ratio of the endcap contaminations for target A and B is fit with an exponential function,

$$\frac{C^A_{EC}}{C^B_{EC}} = 1 \pm e^{ax+b} \tag{4.18}$$

where the $+$ or $-$ depends on which target in the ratio had the larger contamination. The contamination ratio and fits for each MARATHON cross section ratio are shown in Figure 4.7. These fits were used to calculate and correct the endcap contamination for each bin, using the corresponding x-position of that bin.

4.3.4 Radiative effects

4.3.4.1 External radiation

When electrons pass through material, both before and after scattering, they can lose energy through processes such as ionization and bremsstrahlung. For electrons that lose energy,

simply calculating kinematic variables using the beam energy (for E) and detected energy (for E') does not reflect the kinematics of the actual scattering event that took place. This results in events being binned incorrectly. Note that since electrons can only *lose* energy through such processes, these events will always appear to have a larger energy transfer than reality.

The four gas target cells used for MARATHON were designed to be identical, however, due to machining tolerances ended up having variations in the thickness of the cell walls. Further, each cell contained a different density of a different gas, each with a unique radiation length. It was therefore possible that the energy loss was not the same between targets, resulting in net bin migrations that did not cancel in the cross section ratios. In order to evaluate this effect, a GEANT4 simulation of the target system was developed. The simulation geometry included all material that the electrons pass through between the accelerator and spectrometer:

- 0.2 mm beryllium (window separating accelerator beamline from Hall A beamline)
- Gas target
- 0.4 mm aluminum (scattering chamber exit window)
- 81.6 cm air (gap between scattering chamber and spectrometer)
- 0.305 mm Kapton (spectrometer entrance window)

Items with specified thicknesses were common to all targets. Most importantly, the simulation could switch between different targets, defined by the measured thicknesses of each cell and fill parameters of each gas. See Figure 4.8 for a visualization of events and geometry in the simulation.

Electrons were generated upstream of the target at the MARATHON beam energy of 10.6 GeV, and propagated to a random scattering vertex uniformly sampled within the target. The kinematics of the scattering event were chosen by sampling uniformly in $\cos\theta$ and Q^2.

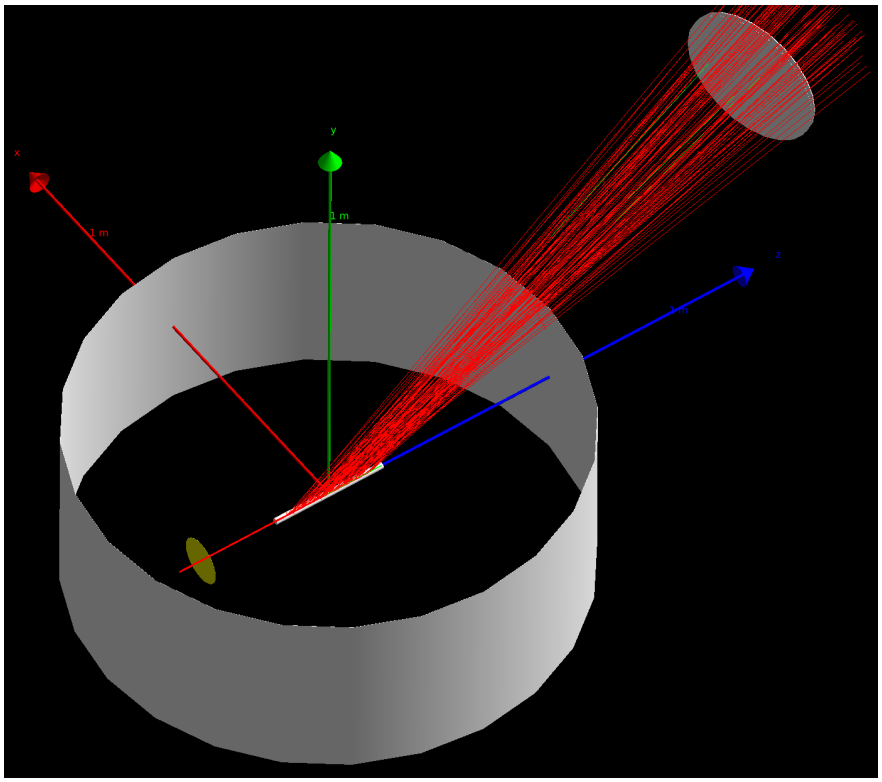


Figure 4.8: Geometry and events in GEANT4 energy loss simulation. Grey and yellow volumes indicate material the electrons must pass through. The red lines are simulated electron tracks.

This is, of course, unphysical. However, since the relevant quantity to be extracted is the relative probability that an event migrate from one bin to another, physical sampling (i.e., from a cross section model) was not necessary. After scattering, the electron was propagated through the remaining material before entering the spectrometer. Relevant energy loss and multiple scattering processes were included in the GEANT4 physics list.

The simulation output included two sets of kinematic variables. The *observed* values were calculated using the initial beam energy (excluding pre-scattering energy loss) and the final detected energy (including post-scattering energy loss). The *true* values were calculated using the energy entering the scattering process (including pre-scattering energy loss) and the energy immediately after scattering (excluding post-scattering energy loss). This allowed a comparison of the true event kinematics, representing the energies of the actual scattering process, to the observed event kinematics that have been smeared by energy loss effects. Finally, the output included the electron position and angle entering the spectrometer, allowing acceptance cuts (replicating cuts on real data) to be applied.

The simulation was run for the first kinematic setting, where effects from differences in cell geometry would be maximized due to the forward scattering angle. An illustrative histogram of the results is shown in Figure 4.9. The vertical axis is x calculated with the true variables; the horizontal axis is x calculated with the observed variables. Both axes have bins of width $\Delta x = 0.02$, comparable to the bin size used for the data analysis. The bins outlined in red indicate events that are observed in the same bin as their true kinematics. All other bins are populated by events that migrated from their true bin to another due to energy loss.

A comparison between targets can be made by quantifying the bin migration. Let N_{true} be the number of events (meeting acceptance cuts) whose true kinematics place it in a particular x bin. Let N_{obs} be the number of events (meeting acceptance cuts) whose observed kinematics place it in the same bin. The net effect of bin migration due to energy loss can

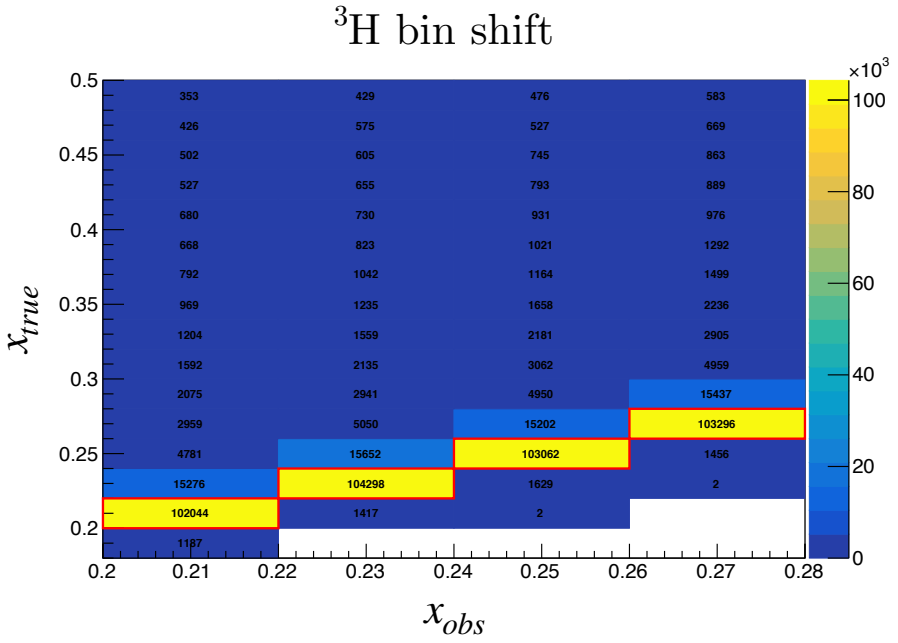


Figure 4.9: Distribution of simulated events binned by true and observed x values

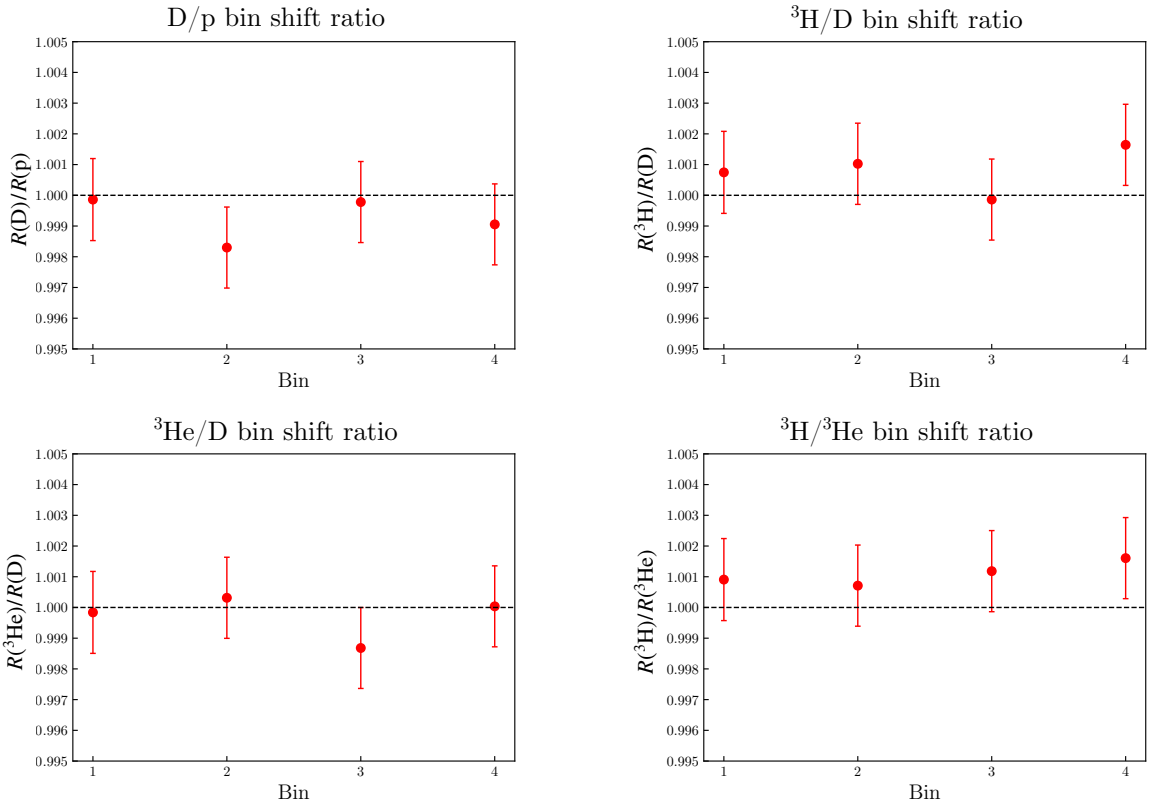


Figure 4.10: Ratio of simulated bin migration magnitudes for each cross section ratio.

then be quantified by:

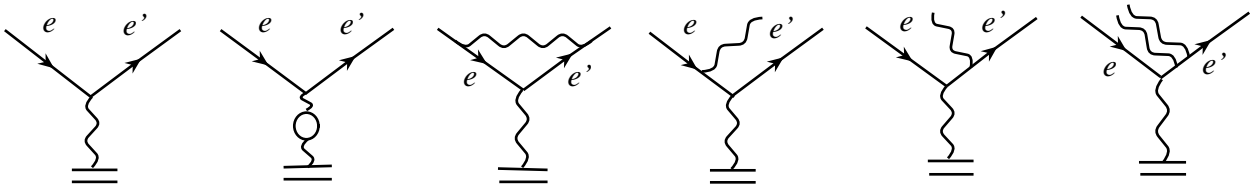
$$R = \frac{N_{true}}{N_{obs}} \tag{4.19}$$

The simulation was used to determine R for each target. Figure 4.10 shows the ratio of R for each of the four cross section ratios measured by MARATHON. The results are consistent with one, indicating that bin migration from energy loss and multiple scattering cancels in cross section ratios.

4.3.4.2 Internal radiation

Figure 2.1 depicts the lowest-order QED interaction between an electron and a nucleus. The cross section for this $\mathcal{O}(\alpha^2)$ process, sometimes referred to as the Born cross section, is the most convenient to theoretically analyze. Interpreting the DIS cross section in terms of structure functions and PDFs assumes this leading-order interaction. However, the probability of this exact process occurring is zero [46]. Various higher-order processes contribute to the measured cross section (see Figure 4.11), and experimentally there is no way to distinguish the true nature of each event. These must be removed using input from QED.

A standard prescription for removing these radiative effects is given by Mo & Tsai [46, 48]. While the photon-hadron coupling is unknown (indeed, this is what the experiment seeks to measure), the electron-photon couplings are well understood in QED. This allows the portion of the radiative interactions depicted in Figure 4.11 to be calculated theoretically,



(a) Born (b) Vacuum polarization (c) Vertex correction (d) Bremsstrahlung (e) Multi-photon emission

Figure 4.11: Born and higher order contributions to total QED cross section. Reproduced from [47].

regardless of the details of the omitted hadronic vertex. Keeping only $\mathcal{O}(\alpha^3)$ terms allows the radiated cross section to be expressed as an integral over the Born cross section. Thus, with a reasonable input for the Born cross section and the radiative matrix elements calculated in QED, the full radiated cross section can be calculated. Of course, this introduces model dependence based on the choice of Born cross section model.

Radiative corrections to the DIS cross section can arise in two ways. First, events from the elastic peak (including the $\mathcal{O}(\alpha^3)$ contributions) can radiate energy, forming a large tail that extends to the DIS region. This contribution is calculated by treating the elastic peak as a discrete state and accounting for all possible ways that an elastically scattered electron can lose energy. Second, the $\mathcal{O}(\alpha^3)$ interactions contribute directly to the DIS cross section. This contribution is calculated by treating the DIS region as a continuum of discrete states, and integrating the radiative effects over all states above the elastic peak.

In this approach, the model radiated cross section σ_{mod}^{rad} can be calculated, given a model Born cross section σ_{mod}^{Born} as input. Thus, the experimental cross section σ_{exp}, which includes all radiative effects, can be corrected by:

$$\sigma_{exp}^{Born} = \sigma_{exp}\left(\frac{\sigma_{mod}^{Born}}{\sigma_{mod}^{rad}}\right) \tag{4.20}$$

The fractional contribution of the Born cross section to the full cross section will be referred to as the radiative correction factor:

$$C_{rad} = \frac{\sigma_{mod}^{Born}}{\sigma_{mod}^{rad}} \tag{4.21}$$

MARATHON used an existing FORTRAN implementation of the Mo & Tsai prescription to provide radiative corrections for all gas targets. This required absolute cross sections for the proton, deuterium, tritium, and helium-3 as input. In order to assess the model dependence of the radiative corrections, multiple models were used to obtain cross sections for each nucleus. The proton and deuterium cross sections were obtained from measurements

of F_2 structure functions with corresponding measurements of $R = \sigma_L/\sigma_T$. For tritium and helium-3, the cross section was obtained from models of the EMC ratio (σ_A/σ_D) multiplied by the absolute deuterium cross section. If the EMC model only provided the isoscalar ratio for an A = 3 nucleus, a model for F_2^n/F_2^p to remove the isoscalar correction. The various references used for these cross section inputs are given below.

- Proton and deuterium
  1. F_2 from SLAC with $R = \sigma_L/\sigma_T = 0.18$ [35],
  2. F_2 from NMC [49], $R = \sigma_L/\sigma_T$ from R1998 [50]
- EMC ratios $F_2^{A=3}/F_2^D$
  1. Nuclear model by S. Kulagin and R. Petti (KP) [51, 52, 53]
  2. Isoscalar EMC ratio for A = 3 nucleus from SLAC data [25]
- F_2^n/F_2^p
  1. Linear parameterization, $F_2^n/F_2^p = 1 - 0.8x$ [25]
  2. Structure functions from QCD extraction of PDFs by CJ15 [19]
  3. Extraction from NMC measurements of deuterium and hydrogen [54] (no nuclear corrections applied!)

While the absolute radiative correction for each target could be as large as 15%, the net correction to the MARATHON cross section ratios was significantly smaller (see Figure 4.12). For D/p, the correction is $\leq 4\%$; for all other ratios, it is $\leq 1\%$. As the primary MARATHON measurement is cross section ratios, the model dependence should be evaluated on the ratio of radiative corrections between targets. It was found that the variation between models was $\leq 0.35\%$, with the exception of combinations using F_2^n/F_2^p from NMC, which reached 1% at high x. However, the NMC model does not account for any nuclear effects when extracting F_2^n/F_2^p from deuterium and proton data. While this could be a reasonable approximation

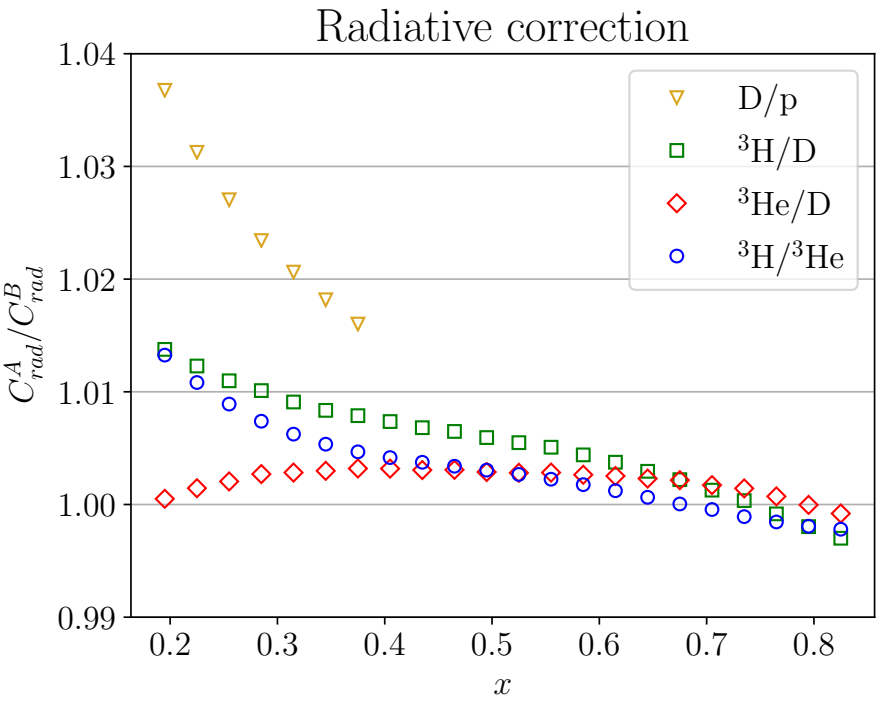


Figure 4.12: Radiative correction to cross section ratios used for MARATHON.

at low x, it leads to large inaccuracy at high x. The high x behavior of the NMC model will not be considered for the MARATHON systematics. MARATHON has assigned a 0.5% uncertainty to each cross section ratio due to model variations. The official radiative correction for MARATHON used F_2^D and F_2^p from SLAC, EMC ratios from KP, and the linear parameterization of F_2^n/F_2^p.

4.3.5 Charge-symmetric background

High energy photons are created in the target, either directly radiated by the beam electrons or resulting from the decay of neutral pions produced by DIS processes. These photons can decay into e^-e^+ pairs, some of which enter the acceptance of the spectrometer. When taking standard production data, the spectrometer is set to negative polarity to select the DIS electrons. This rejects the positrons, but allows their electron pairs to be detected. The measured electron yield is includes the sum of electrons from DIS and the background from

pair-produced electrons:

$$Y^{e^-}_{meas} = Y^{e^-}_{DIS} + Y^{e^-}_{BG} \tag{4.22}$$

These background electrons are indistinguishable from DIS electrons. In order to quantify and correct this background, a series of runs were taken with the spectrometer switched to positive polarity. With this setting, both the background electrons and DIS electrons are rejected by the spectrometer, while the positrons are detected. Since the pair production is charge symmetric, the positron yield observed with positive spectrometer polarity is equal to the electron background yield with negative spectrometer polarity:

$$Y^{e^+}_{BG} = Y^{e^-}_{BG} \tag{4.23}$$

The fractional contribution of background electrons to the measured electron yield can then be expressed in terms of the positron yield,

$$\frac{Y^{e^-}_{BG}}{Y^{e^-}_{meas}} = \frac{Y^{e^+}_{BG}}{Y^{e^-}_{meas}}, \tag{4.24}$$

allowing the background to be removed from the measured yield:

$$Y^{e^-}_{DIS} = Y^{e^-}_{tot}\left(1 - \frac{Y^{e^+}_{BG}}{Y^{e^-}_{meas}}\right) \tag{4.25}$$

As the rate of pair production could depend on the target, positron data was taken on all four gas targets (plus the empty cell to measure the endcap contamination). The distribution of pair-produced electrons and positrons is not uniform, so the size of the background will depend on the angle of the spectrometer. Positron data was taken at multiple kinematic settings to allow the kinematic dependence to be determined. However, due to time constraints, no positron data was taken beyond KIN3.

Extraction of the positron yield was nearly identical to the extraction of the electron yields. Identical acceptance and PID cuts were applied. The data was corrected for DAQ

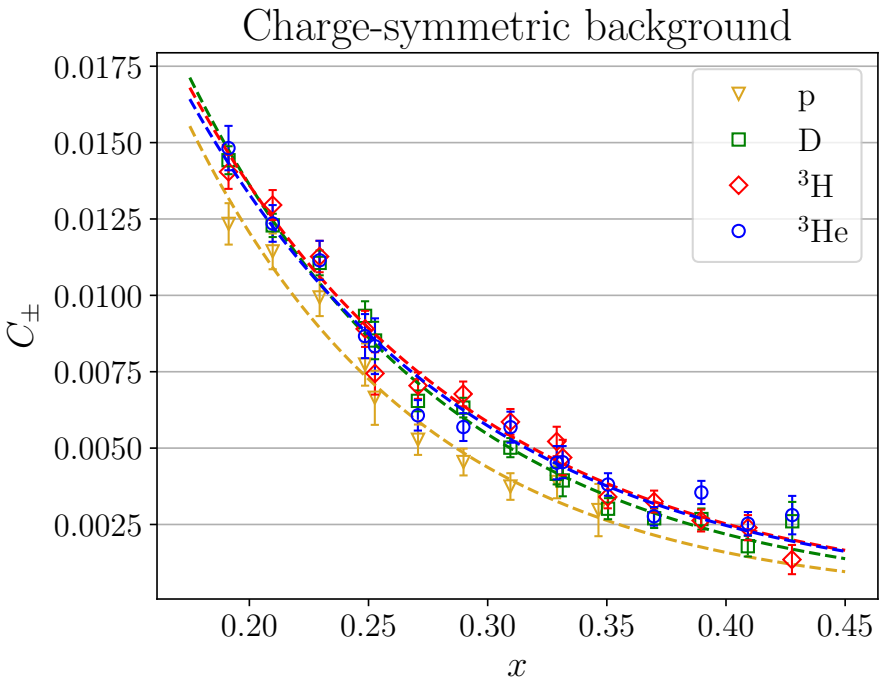


Figure 4.13: Charge-symmetric background contamination for each target.

deadtime and endcap contamination, in the same manner described by Sections 4.3.1 and 4.3.3. An additional correction was applied for the non-positron contamination left by the PID cut. This was done by fitting the non-positron tail in the calorimeter spectrum and subtracting the integrated contribution above the E/p cut. This correction was necessary due to the limited number of positron events, resulting in a higher percent contamination allowed by the PID cuts. The relative contaminations of positron to total electron yield $C_\pm$ for each target were fit with an exponential function:

$$C_\pm = e^{ax+b} \tag{4.26}$$

The data and fits are shown in Figure 4.13. The fits (including extrapolation to high x) were used to calculate and correct the electron background contamination for each bin, using the corresponding x-position of that bin.

4.3.6 Tritium β-decay

Tritium is radioactive, β-decaying to helium-3 ($^3\mathrm{H} \rightarrow^3 \mathrm{He} + e^- + \overline{\nu}_e$) with a half-life of $\tau_{1/2} = 4500 \pm 8$ days [55]. As previously mentioned, this required special precautions for the safe handling and operation of the tritium target. In addition, it meant that the composition of the tritium target was continuously changing. When calculating the luminosity, the exponentially decreasing tritium density must be accounted for. When counting scattered electron events, events scattered from the exponentially increasing number of helium-3 nuclei must be removed. A correction accounting for both of these effects had to be developed. For cleaner notation, in this section the subscripts T and H will refer to tritium (^{3}H) and helium-3 (^{3}He), respectively.

The time evolution of the tritium and helium-3 number densities can be described simply with exponential functions,

$$n_T(t) = n_T^0 e^{-t/\tau} \tag{4.27}$$

$$n_H(t) = n_T^0 \left(1 - e^{-t/\tau}\right) \tag{4.28}$$

where n_T^0 is the initial tritium number density and the initial helium-3 number density is zero. Since the decay process conserves the total number of nuclei, the total nuclear number density n_{tot} should remain constant in time:

$$n_{tot} = n_T(t) + n_H(t) \tag{4.29}$$

$$= n_T^0 \tag{4.30}$$

The helium-3 contamination in the tritium target can be quantified as the helium-3 fraction of total nuclei,

$$f_H(t) = \frac{n_H(t)}{n_{tot}} = \left(1 - e^{-t/\tau}\right), \tag{4.31}$$

which approaches 1 as $t \rightarrow \infty$. The helium-3 fraction, shown in Figure 4.14, grew to

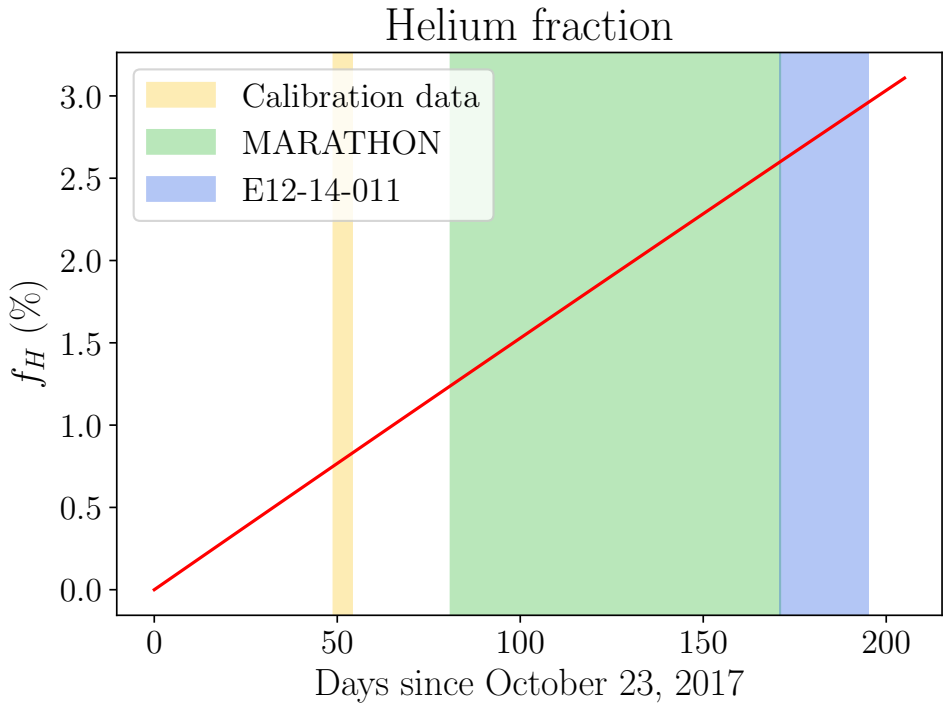


Figure 4.14: Fractional helium-3 contamination in tritium target. The shaded bands represent the different experimental periods during which tritium data was taken.

approximately 3% by the end of the spring run.

The raw normalized yield obtained from the tritium target is

$$Y_{raw} = \frac{\sum(N_{T,i} + N_{H,i})}{\sum Q_i n_{tot}}, \tag{4.32}$$

where N_T and N_H are the number of detected events where the electron scattered from a tritium or helium-3 nucleus, respectively. These events look identical to the spectrometer and detectors and therefore cannot be eliminated with any cut on the data. The tritium cross section ratios require the clean normalized tritium yield,

$$Y_T = \frac{\sum N_{T,i}}{\sum Q_i n_T(t_i)}, \tag{4.33}$$

and removing the contamination from helium-3 scattering will require the normalized helium-

3 yield,

$$Y_H = \frac{\sum N_{H,i}}{\sum Q_i n_H(t_i)}, \tag{4.34}$$

Some simple algebraic manipulation of Equation 4.32, outlined below, allows the extraction of Y_T in terms of Y_{raw} and Y_H.

$$\begin{aligned}
Y_{raw} &= \frac{\sum N_{T,i} + N_{H,i}}{Q_i n_{tot}} \\
&= \frac{\sum N_{T,i}}{n_{tot} \sum Q_i} + \frac{\sum N_{H,i}}{n_{tot} \sum Q_i} \\
&= \left(\frac{\sum N_{T,i}}{\sum Q_i n_T(t_i)}\right)\left(\frac{\sum Q_i n_T(t_i)}{n_{tot} \sum Q_i}\right) + \left(\frac{\sum N_{H,i}}{\sum Q_i n_H(t_i)}\right)\left(\frac{\sum Q_i n_H(t_i)}{n_{tot} \sum Q_i}\right) \\
&= Y_T \left(\frac{\sum Q_i (1 - f_H(t_i))}{\sum Q_i}\right) + Y_H \left(\frac{\sum Q_i f_H(t_i)}{\sum Q_i}\right) \\
&= Y_T(1 - \langle f_H \rangle) + Y_H \langle f_H \rangle
\end{aligned}$$

In the last line, the charge-averaged helium-3 fraction is defined as:

$$\langle f_H \rangle = \frac{\sum Q_i f_H(t_i)}{\sum Q_i} \tag{4.35}$$

This leads to the final expression for the normalized tritium yield:

$$Y_T = Y_{raw} \left(\frac{1}{1 - \langle f_H \rangle}\right) - Y_H \left(\frac{\langle f_H \rangle}{1 - \langle f_H \rangle}\right) \tag{4.36}$$

Until this point, only data from the tritium target cell has been considered; references to helium-3 referred exclusively to helium-3 within the tritium cell resulting from tritium β-decay. However, by Equation (4.36), the normalized helium-3 yield Y_H is required for the correction. Clearly, Y_H cannot be determined from the tritium target data. However, MARATHON also required the measurement of the normalized helium-3 yield with a dedicated helium-3 target. This can be used to correct the tritium data.

4.4 Normalized yields

4.4.1 Extracting from data

The normalized yields were extracted in bins covering the range $0.18 < x < 0.84$, with a bin width of $\Delta x = 0.03$. Combining all of the cuts and corrections previously discussed, the yield for a bin centered at x_{bin} and Q^2_{bin} is determined by the following expression and quantities

$$Y(x_{bin}) = \frac{\sum N_i/(1 - C_{DT,i})}{n \sum C_{dens}(I_i) Q_i} \times (1 - C_{EC}(x_{bin})) \times C_{RC}(x_{bin}, Q^2_{bin}) \times (1 - C_{\pm}(x_{bin})) \quad (4.37)$$

- N_i is the number of events from run i that are contained in the bin range and satisfy all kinematic, acceptance, tracking, and PID cuts.
- $C_{DT,i}$ is the DAQ deadtime correction for run i.
- n is the target density.
- $C_{dens}(I_i)$ is the beam-induced target density fluctuation correction, evaluated at the average run current $I = I_i$.
- Q_i is the total charge on target for run i (see discussion below).
- $C_{EC}(x_{bin})$ is the relative endcap contamination, evaluated at $x = x_{bin}$.
- $C_{RC}(x_{bin}, Q^2_{bin})$ is the radiative correction, evaluated at $x = x_{bin}$ and $Q = Q^2_{bin}$.
- $C_{\pm}$ is the relative charge-symmetric background, evaluated at $x = x_{bin}$.

The total charge Q is simply the beam current integrated over the length of the run. The BCM signal is recorded by a scalar, where the total number of counts N_{BCM} represents the integrated BCM signal. The clock is also connected to a scalar, where the number of counts N_{clock} is divided by the clock frequency f_{clock} to obtain the total time of the run. In terms

of these measurements and the calibrated BCM parameters (see Section 3.2.2.1), the total charge can be calculated by:

$$Q = gN_{BCM} + b\left(\frac{N_{clock}}{f_{clock}}\right) \tag{4.38}$$

Adjacent kinematic settings produced data in the same bin. Due to the condition that all spectrometer settings be identical in order for the yield ratio to be equivalent to the cross section ratio, data from different kinematic settings were never combined at the yield level. Rather, cross section ratios for the overlapping bin were determined separately for each kinematic setting, and then combined by a weighted average.

Additionally, the tritium yield must be corrected for β-decay using Equation (4.36). The helium-3 yield extracted from helium-3 target data using Equation (4.37) was used for the β-decay correction.

4.4.2 Uncertainties

4.4.2.1 Bin-by-bin uncertainties

For a given bin, the statistical uncertainty of the yield measurement is determined by:

$$\delta Y_{stat} = \sqrt{N} \tag{4.39}$$

where N is the total number of events observed in that bin.

Reasonable variations in the applied acceptance cuts resulted in cross section ratio variations on the order of $\leq 0.2\%$, which is assigned as the acceptance cut uncertainty. Reasonable variations in the PID cuts resulted in cross section ratio fluctuations that were negligibly small ($\leq 0.02\%$).

For all of the corrections that rely on fitting a function to data (beam-induced target density fluctuation, endcap contamination, and charge-symmetric background), the uncer-

tainties of the individual data points were used as weights in the fitting procedure. This allows the covariance matrix of each fit to be used to estimate the uncertainty on the correction. For a correction parameterized by a function f with N free parameters (p_0, $p_1 ... p_N$), the uncertainty in the correction can be found by:

$$\delta C^2 = \boldsymbol{d}^\top \boldsymbol{\sigma}_{cov} \boldsymbol{d} \tag{4.40}$$

where $\boldsymbol{d}$ is a column matrix with $d_i = \partial f / \partial p_i$, and $\boldsymbol{\sigma}$ is the covariance matrix from the fit.

Uncertainty in the calculated charge can arise from uncertainty in the calibrated values of the BCM gain and offset. Mathematically, the BCM calibration values do not cancel exactly in the cross section ratios. However, when the uncertainty from the BCM calibration values are propagated through the ratio of total charge calculated for two different targets, the results are negligibly small. For this reason, uncertainty in the calculated charge is neglected.

Although the radiative corrections are applied individually to the extracted yields, the uncertainty arising from model dependence is applied to the cross section ratio (see Section 4.3.4.2). An uncertainty of 0.5% will be added in quadrature to the total uncertainties of the cross section ratios.

The tritium β-decay correction depends on the helium-3 fraction and charge. Note that in Equation (4.35), the charge enters in both the numerator and denominator. Thus, by the same reasoning that the charge calculation uncertainty is neglected in the yield ratios, it can be neglected in the β-decay correction. The only other quantity that could cause uncertainty in the β-decay correction is the tritium half-life (8 days, corresponding to $< 0.2\%$). When this uncertainty is propagated to the helium-3 fraction, it is negligibly small for the duration of the experiment. Therefore, no uncertainty will be included for this correction.

4.4.3 Normalization uncertainty

One of the major contributions to uncertainty in the normalized yield measurement is the overall normalization of the data by the target density. Clearly, for experiments that form ratios from the yields of different nuclear targets (such as MARATHON), these uncertainties do not cancel. The nominal relative uncertainties in density for the MARATHON targets, determined by the precision of the fill parameters of the cell, are 0.56% for hydrogen, 0.56% for deuterium, 0.94% for tritium, and 1.12% for helium-3. Pairs of these relative uncertainties are added in quadrature to obtain the normalization uncertainty for a cross section ratio. For the ^{3}H/^{3}He ratio, which have the largest uncertainties, this corresponds to a normalization uncertainty of nearly 1.5%. This strongly limits the impact of the F_2^n/F_2^p extraction. For this reason, the MARATHON analysis used a model-driven normalization procedure to fix the normalization of the A = 3 targets. This is discussed in detail in the following chapter.

Chapter 5

Results

5.1 Cross section ratios

Normalized yields for each target were extracted following the procedure described in Chapter 4. Ratios of these yields were formed to obtain four cross section ratios:

- $\sigma(\mathrm{D})/\sigma(\mathrm{p})$, for comparison to SLAC data and normalization of MARATHON data
- $\sigma(^3\mathrm{H})/\sigma(^3\mathrm{He})$, for the A = 3 super-ratio extraction of F_2^n/F_2^p
- $\sigma(^3\mathrm{H})/\sigma(\mathrm{D})$, the tritium EMC ratio
- $\sigma(^3\mathrm{He})/\sigma(\mathrm{D})$, the helium-3 EMC ratio

As discussed in Chapter 2, the MARATHON analysis assumes the nuclear independence of $R = \sigma_L/\sigma_T$. Therefore, these cross section ratios will be treated interchangeably with the ratios of the corresponding F_2 structure functions. Subsequent references to the cross section/structure function ratios will simply indicate the relevant nuclei (for instance, tritium to deuterium will be indicated by ^{3}H/D). All ratios will be shown per-nucleon unless otherwise noted. Please note that these cross section ratios are considered preliminary by the MARATHON collaboration, and that the systematic uncertainties are currently being

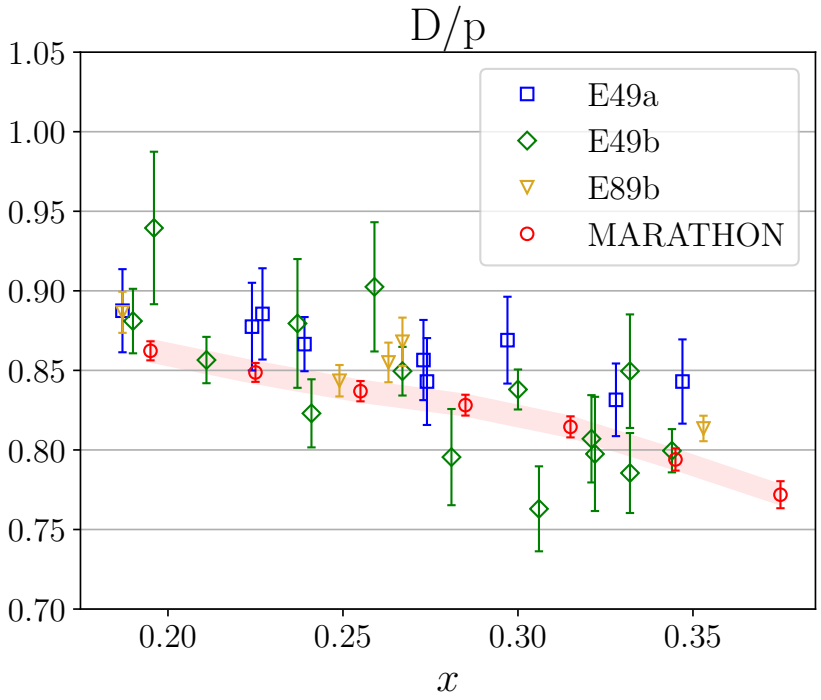


Figure 5.1: Preliminary D/p cross section ratio measured by MARATHON, compared to previous measurements.. Shaded band indicates normalization uncertainty.

finalized. The official results will be contained in a pending MARATHON publication. See Appendix A, Table A.1 for the numeric values displayed in the following figures.

The MARATHON result for D/p is shown in Figure 5.1. Also shown on the plot are selected data points from F_2^D/F_2^p measurements made at SLAC by experiments E49a, E49b, and E89a. A tight cut was placed on the SLAC data to only include points that satisfy the condition

$$\left|Q^2 - (14\ \text{GeV}^2) \times x\right| < 1\ \text{GeV}^2, \tag{5.1}$$

ensuring that the SLAC data being compared were at similar kinematics to the MARATHON data (see Equation (3.21)). The normalization uncertainty of the MARATHON data is 0.8% and is shown in the figure. The normalization uncertainty of the SLAC data is quoted at 1%. Within uncertainties, the MARATHON measurement is consistent with the SLAC measurements, albeit with much smaller point-to-point uncertainties.

The MARATHON result for ^{3}H/^{3}He is shown in Figure 5.2. This is the first measurement

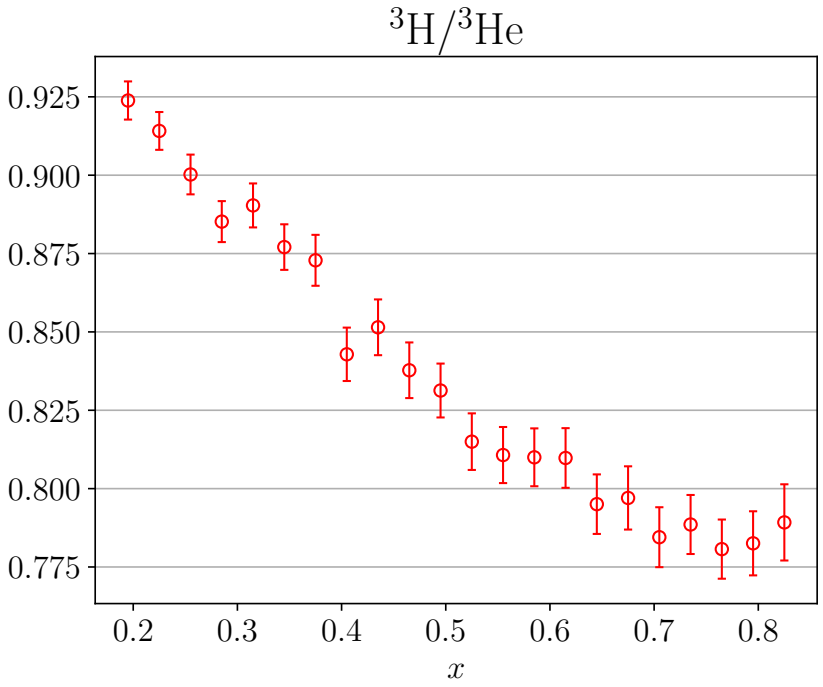


Figure 5.2: Preliminary ^{3}H/^{3}He cross section ratio measured by MARATHON. No normalization uncertainty is shown, as this will be determined by the renormalization procedure.

of this ratio, so no comparison can be made to previous data. A distinct feature of the ratio is the plateau at large x, the region where F_2^n/F_2^p constrains d/u.

The MARATHON results for tritium and helium-3 EMC ratios, ^{3}H/D and ^{3}He/D, are shown in Figures 5.3 and 5.4. Once again, there are no previous measurements with which to compare the tritium result. However, the EMC effect in helium-3 has been previously measured by HERMES at DESY [40] and by E03-103 at Jefferson Lab [24], and are shown alongside the MARATHON measurement. Minimum cuts of $Q^2 > 2$ GeV^2 and $W^2 > 3$ GeV^2 have been applied, which eliminated some low-x HERMES points and some high-x E03-103 points. Note that HERMES only published isoscalar-corrected data (see Section 2.3.2), using F_2^n/F_2^p from NMC measurements of the proton and deuterium. For the purposes of this comparison, the isoscalar correction has been removed using the NMC parameterization of F_2^n/F_2^p. In addition, the HERMES ratio was published with a +0.9% renormalization to the measured value in order to agree with NMC measurements of the EMC effect in

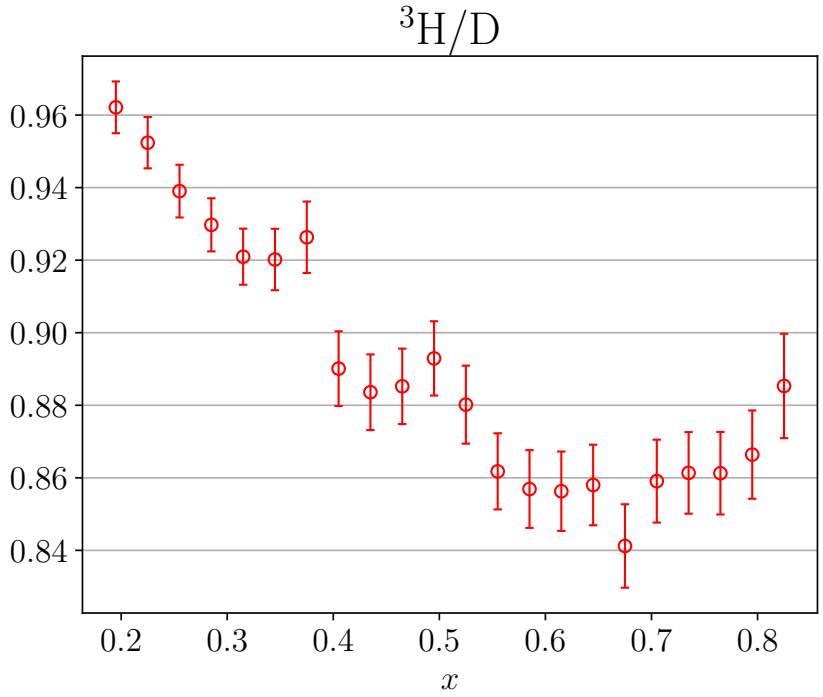


Figure 5.3: Preliminary ^{3}H/D cross section ratio measured by MARATHON. No normalization uncertainty is shown, as this will be determined by the renormalization procedure.

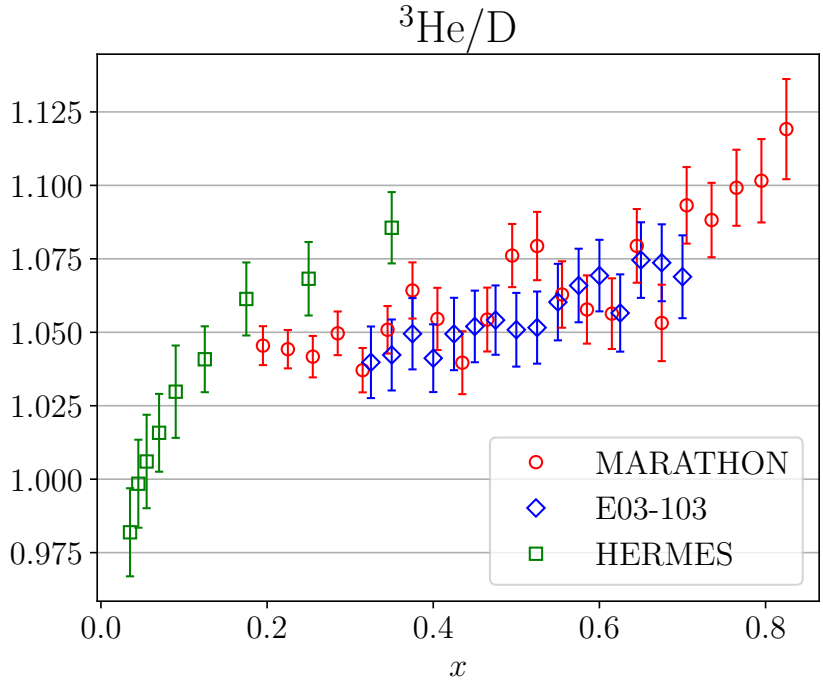


Figure 5.4: Preliminary ^{3}He/D cross section ratio measured by MARATHON, compared to previous measurements.. No normalization uncertainty is shown, as this will be determined by the renormalization procedure. See text for comments on the HERMES data.

^{4}He. Although the motivation for this renormalization is questionable, the as-published HERMES points are included in the figure. The MARATHON measurement agrees well with the E03-103 result, while both Jefferson Lab measurements are systematically lower than the HERMES measurements.

5.2 Nuclear models

5.2.1 Overview

This analysis used two nuclear models to provide the predictions necessary to extract F_2^n/F_2^p from the measured cross section ratios. The first model was developed by S. Kulagin and R. Petti (KP) [51, 52, 53]. The KP model is a phenomenological fit to global structure function ratio data. Well-established nuclear effects (i.e., Fermi motion and binding effects) are applied to the nuclear structure functions using the convolution formalism. Off-shell nuclear effects are described by a function with three free parameters which are empirically constrained. It is assumed that these off-shell effects are isospin-symmetric, altering the bound neutron and proton structure functions equally. Further, the off-shell function is considered universal for all nuclei. As the KP model is fit to structure function data of nuclei ranging from hydrogen to lead, predictions for light nuclei are strongly influenced by observations in heavy nuclei.

The second model is based on a global QCD fit of PDFs completed by the CJ15 collaboration [19]. The CJ15 analysis only uses structure function data from the proton and deuterium. In deuterium, Fermi motion and binding effects are accounted for in the convolution formalism, and the parameterization of off-shell effects is identical to that of the KP model. Additionally, CJ15 includes constraints from observables in measurements of pp and pD collisions. Of particular impact are measurements of the W-boson charge asymmetry made by the DØ collaboration at the Tevatron [56, 57], which provide constraints on the PDFs at large x. In a separate analysis [58], the CJ15 PDFs and off-shell correction were

used to calculate structure functions for A = 3 nuclei. For this reason the model will be referred to as CJ, although it should be noted that the A = 3 structure functions are not a part of the official CJ15 analysis.

5.2.2 Model inputs for F_2^n/F_2^p extraction

In order to extract F_2^n/F_2^p from a measurement of bound nucleons, a model prediction of the nuclear effects is required. Section 3.1.1 described how F_2^n/F_2^p can be obtained from $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$, given a model prediction of the super-ratio $\mathcal{R}$ (Equation (3.3)) as input. Extracting F_2^n/F_2^p from $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$, in particular at high x, is the primary goal of MARATHON. The A = 3 EMC effect and corresponding super-ratio predictions from KP and CJ are shown in Figures 5.5 and 5.6. Note that while the individual A = 3 EMC predictions can grow quite large and disagree between models, both models predict a super-ratio within 2.5% of unity over the MARATHON x range.

Extracting F_2^n/F_2^p from other cross section ratios will be required for MARATHON's renormalization procedure (described in the following section), each of which will require and addition model prediction. Define an EMC-type ratio for nucleus A as

$$R(A) = \frac{F_2^A}{NF_2^n + ZF_2^p} \tag{5.2}$$

where N is the number of neutrons and Z is the number of protons.

The EMC-type ratio for deuterium, $R(D)$, can be inverted to find F_2^n/F_2^p:

$$\frac{F_2^n}{F_2^p} = \frac{F_2^D/F_2^p}{R(D)} - 1 \tag{5.3}$$

requiring a model prediction of $R(D)$ as input. The $R(D)$ predictions from KP and CJ are shown in Figure 5.7.

The ratio of the EMC-type ratio for nucleus A to deuterium can be defined as the so-called

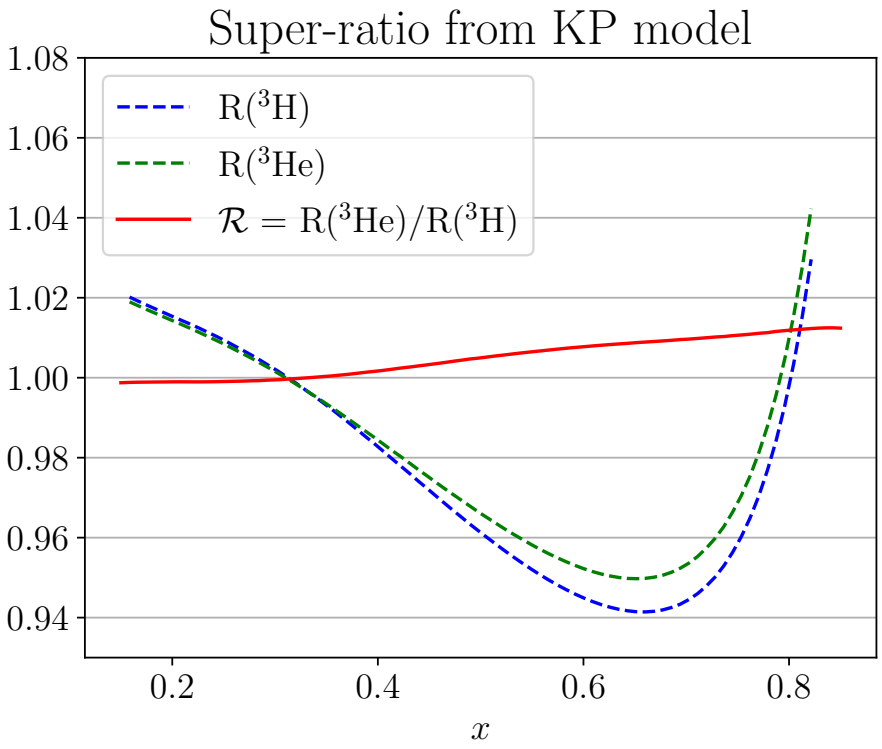


Figure 5.5: Super-ratio $\mathcal{R}$ prediction from KP nuclear model.

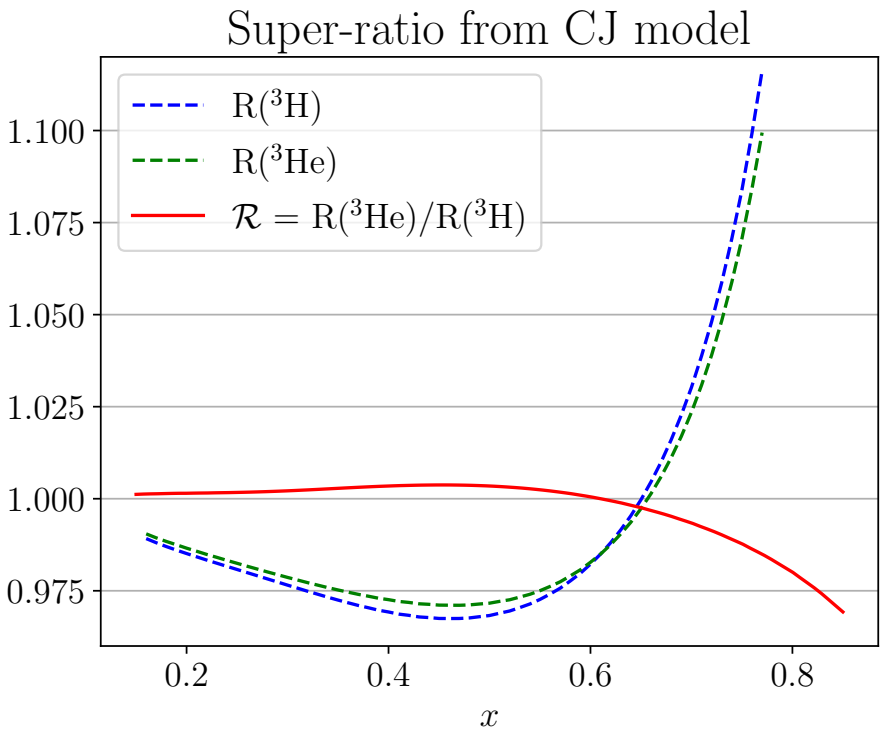


Figure 5.6: Super-ratio $\mathcal{R}$ prediction from CJ nuclear model.

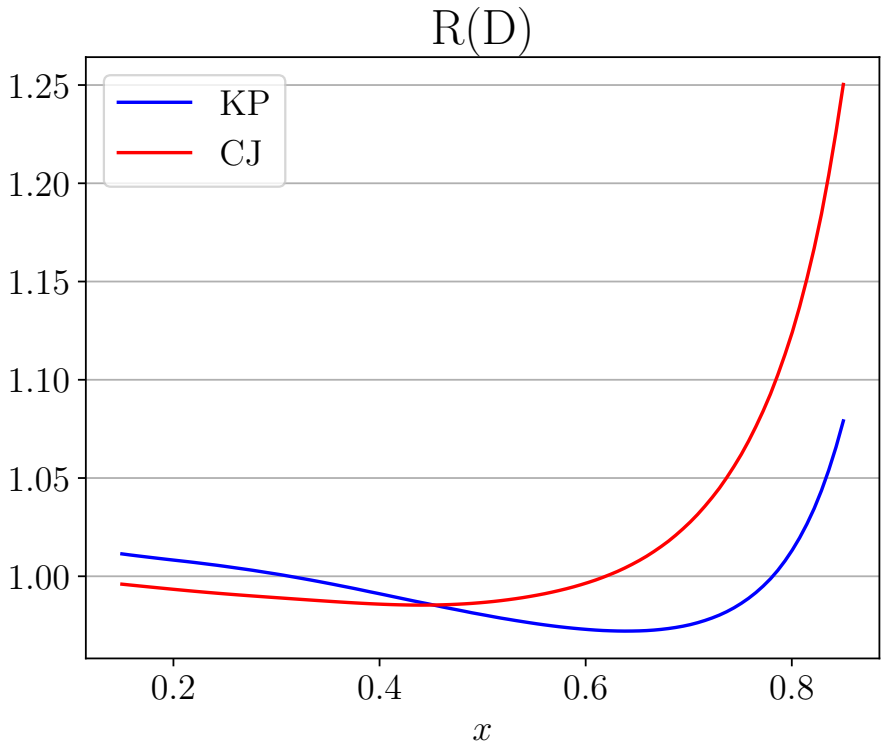


Figure 5.7: $R(D)$ predictions from KP and CJ nuclear models.

EMC super-ratio:

$$\mathcal{R}_{EMC} = \frac{R(A)}{R(D)} \tag{5.4}$$

This can be inverted to find F_2^n/F_2^p,

$$\frac{F_2^n}{F_2^p} = \frac{Z\mathcal{R}_{EMC} - F_2^A/F_2^D}{F_2^A/F_2^D - N\mathcal{R}_{EMC}} \tag{5.5}$$

which requires a model input for $\mathcal{R}_{EMC}$. The EMC super-ratio predictions for tritium and helium-3 from KP and CJ are shown in Figures 5.8 and 5.9.

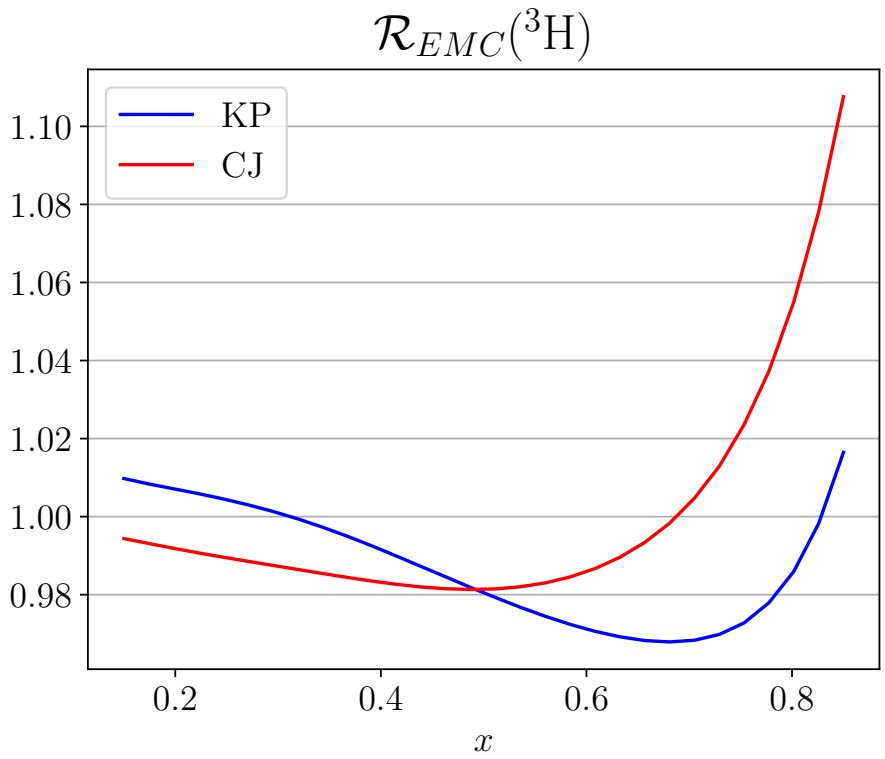


Figure 5.8: $\mathcal{R}_{EMC}(^3\mathrm{H})$ predictions from KP and CJ nuclear models.

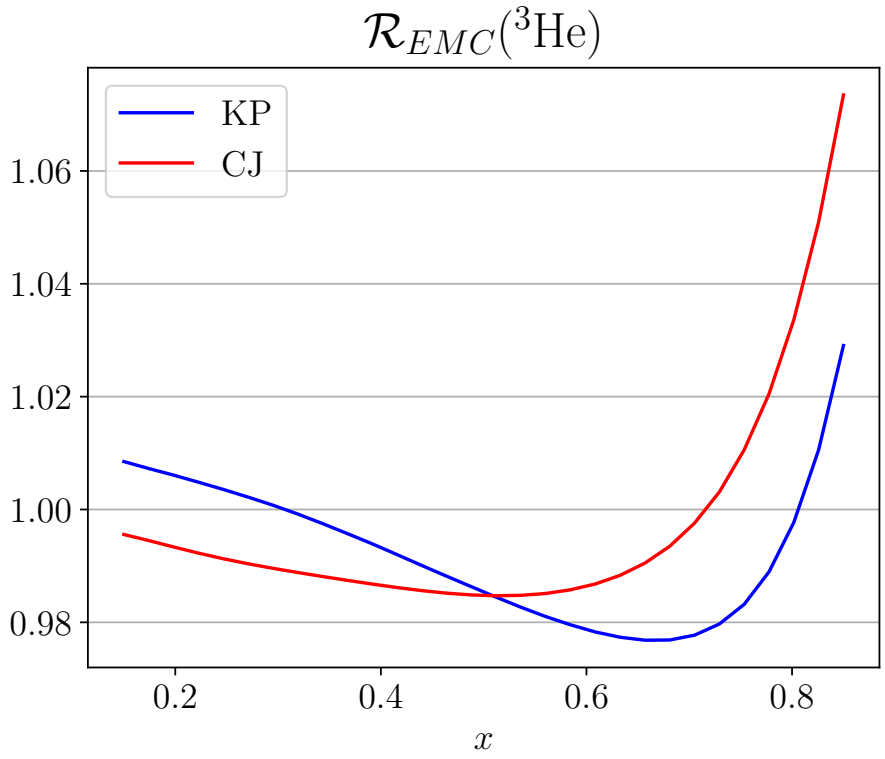


Figure 5.9: $\mathcal{R}_{EMC}(^3\mathrm{He})$ predictions from KP and CJ nuclear models.

5.3 Renormalization

5.3.1 Procedure

Calculating the luminosity used to normalize detected electron counts requires knowledge of the target density. Determination of target densities can suffer from large uncertainties, due to both uncertainty in measuring the physical parameters of the target (temperature, pressure, volume) and uncertainty in how these parameters are used to calculate the target density (i.e., the equation of state). As previously mentioned, the A = 3 targets used by MARATHON suffered from larger density uncertainties than the proton and deuterium targets. For this reason, the MARATHON proposal argued that cross section ratios including A = 3 targets should be renormalized by comparing to D/p data. This renormalization procedure is described below.

The nucleon structure function ratio F_2^n/F_2^p is a universal property of free neutrons and protons. If nuclear effects are accurately accounted for, extractions of F_2^n/F_2^p from measurements of bound nucleons must be the same, regardless of which nuclei were used in the measurement. For MARATHON, this means that F_2^n/F_2^p extracted from all four cross section ratios should agree. Of course, discrepancies in F_2^n/F_2^p from different cross section ratios could imply that the model is not properly accounting for nuclear effects in one or more nucleus. However, MARATHON will assume that discrepancies are due to incorrect normalization of the A = 3 targets, and use the universality of F_2^n/F_2^p to renormalize the data.

First, a cross section ratio must be chosen to which the A = 3 data is normalized. For MARATHON, the D/p ratio was chosen. Not only does this cross section ratio have the smallest normalization uncertainty of all the MARATHON ratios, but this ratio has also been well-studied at SLAC, allowing comparisons to another measurement. Second, a region in x must be chosen at which to perform the renormalization. In general, nuclear effects (and their uncertainties) increase with increasing x, making a low-x region ($x < 0.4$)

preferable. Measurements of the EMC effect in many nuclei tend to cross unity around $x \approx 0.3$, indicating a minimization of nuclear effects. For this reason, MARATHON has chosen to normalize at $x = 0.3$.

The normalization procedure is as follows. The relevant cross section ratios will be fit in the region $0.25 < x < 0.35$ with a linear function. Fitting the measured points not only allows a normalization exactly at $x = 0.3$ (regardless of bin location), but also allows a normalization to SLAC D/p (which is binned differently than MARATHON). From each fit, F_2^n/F_2^p is extracted using the equations described in Section. Note that extracting F_2^n/F_2^p from a fit to the cross section ratio points, is identical to performing a fit to directly to F_2^n/F_2^p extracted from the points themselves. The F_2^n/F_2^p fits is then evaluated exactly at $x = 0.3$. From the difference between each A = 3 extraction with the D/p extraction, the required renormalization of the A = 3 cross section ratios to obtain agreement can be determined. To estimate the uncertainty of the renormalization, the procedure will be performed four times, with each combination of D/p measurement (MARATHON and SLAC) and nuclear model (KP and CJ).

This is an over-constrained problem. The three ratios involving A = 3 nuclei (^{3}H/^{3}He, ^{3}H/D, ^{3}He/D) can be independently renormalized with this method. However, these three measurements used the same two A = 3 targets. Therefore, the individual renormalizations of the tritium and helium-3 targets obtained from the EMC ratios must combine to yield the same renormalization obtained from ^{3}H/^{3}He. This can be used to verify the consistency of the results.

Fits to the D/p cross section ratios are shown in Figure 5.10. The points from various SLAC experiments have been combined and fit simultaneously. Note that the F_2^n/F_2^p extraction requires the full structure function ratio, so these are no longer the per-nucleon structure functions shown previously. The point-to-point uncertainties were used as weights in the fit, allowing the covariance matrix to be estimated.

Comparisons of F_2^n/F_2^p extracted from each A = 3 ratio to D/p are shown in Figures 5.11,

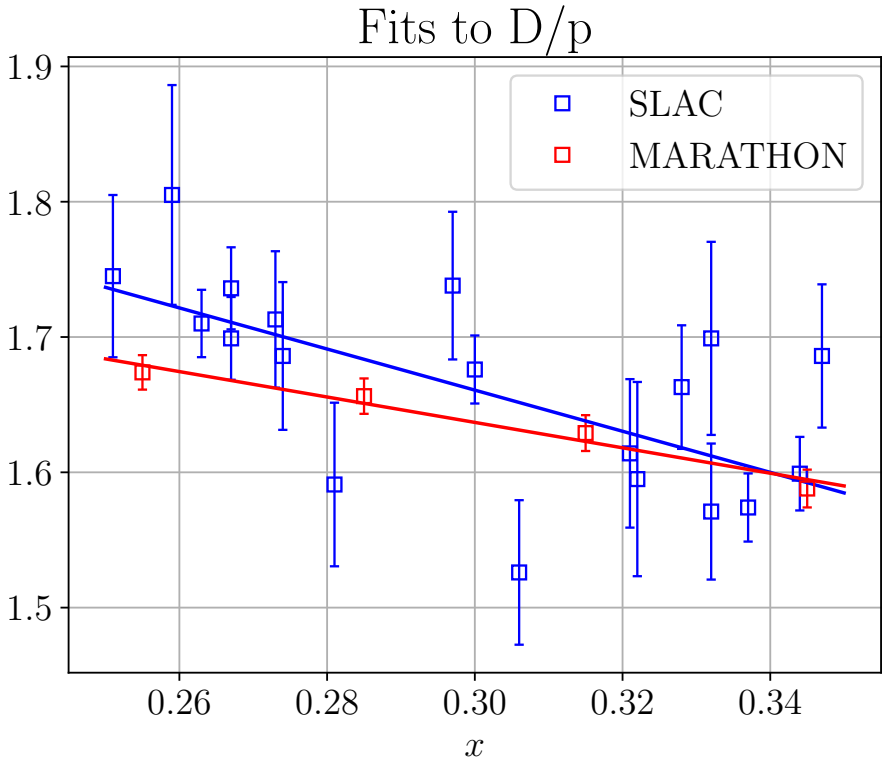


Figure 5.10: Fits to D/p from SLAC and MARATHON.

5.12 and 5.13. In each figure, the subfigures in the left and right columns show F_2^n/F_2^p extracted using the KP and CJ models, respectively. The subfigures in the top and bottom row compare to F_2^n/F_2^p extracted from MARATHON and SLAC D/p, respectively. To obtain the normalization for each ratio, the fits are evaluated at $x = 0.3$ and the difference in F_2^n/F_2^p is determined. Using the equations used to extract F_2^n/F_2^p, the required shift to the cross section ratios to obtain agreement in F_2^n/F_2^p can be calculated. The relative renormalization of the data can then be found.

5.3.2 Final normalization and uncertainty

The resulting renormalizations for the EMC ratios are shown in Figures 5.14 and 5.15. Each figure shows four renormalization values for each combination of nuclear model and D/p dataset. For tritium, the obtained renormalizations are very close to zero ($\leq 1\%$). The SLAC D/p gives a higher renormalization than SLAC D/p, and KP gives a higher renormalization

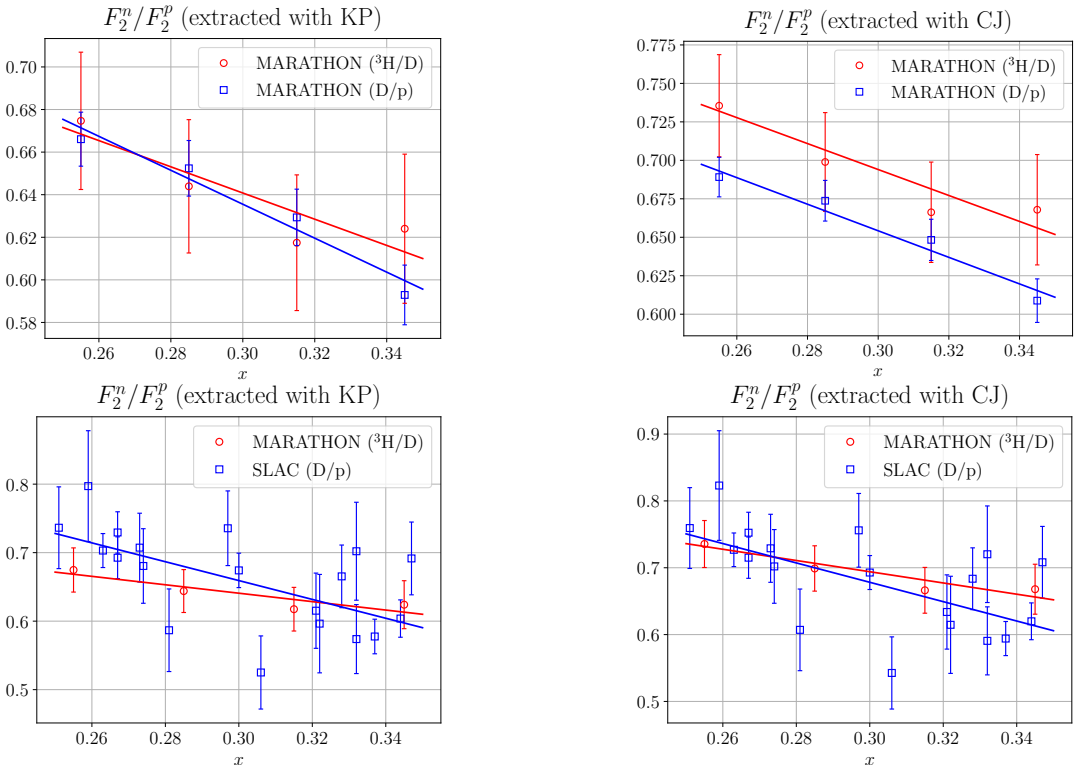


Figure 5.11: Comparison of F_2^n/F_2^p extractions from D/p and ^{3}H EMC ratio.

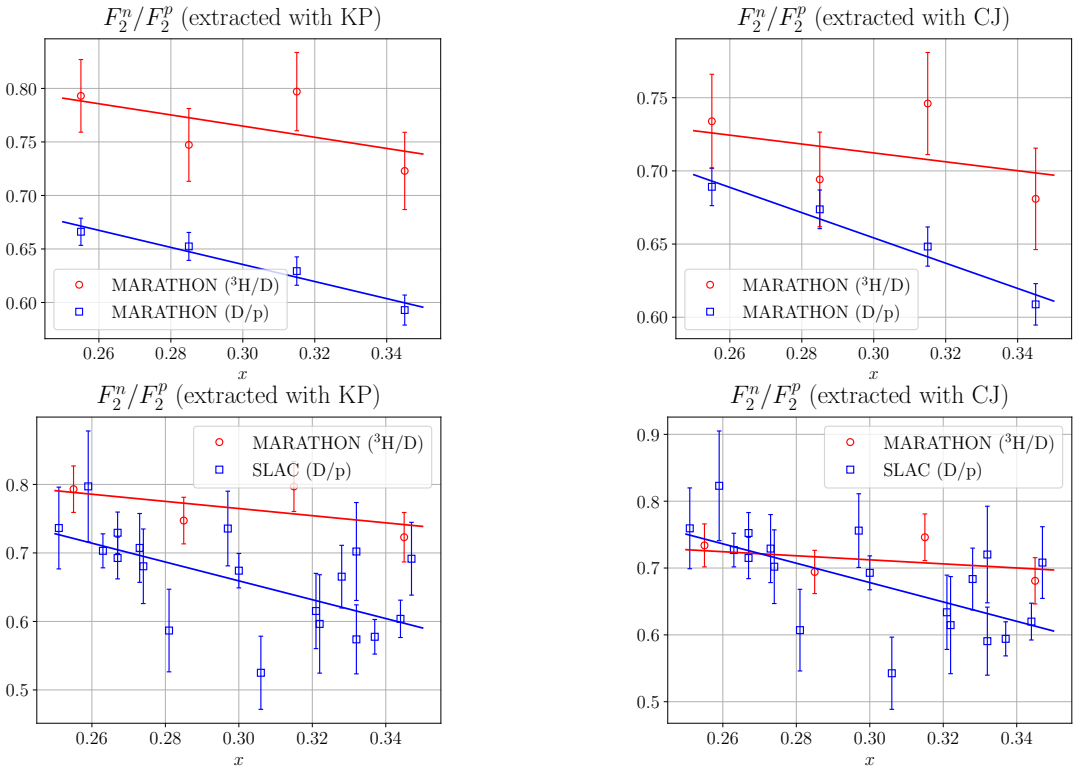


Figure 5.12: Comparison of F_2^n/F_2^p extractions from D/p and ^{3}He EMC ratio.

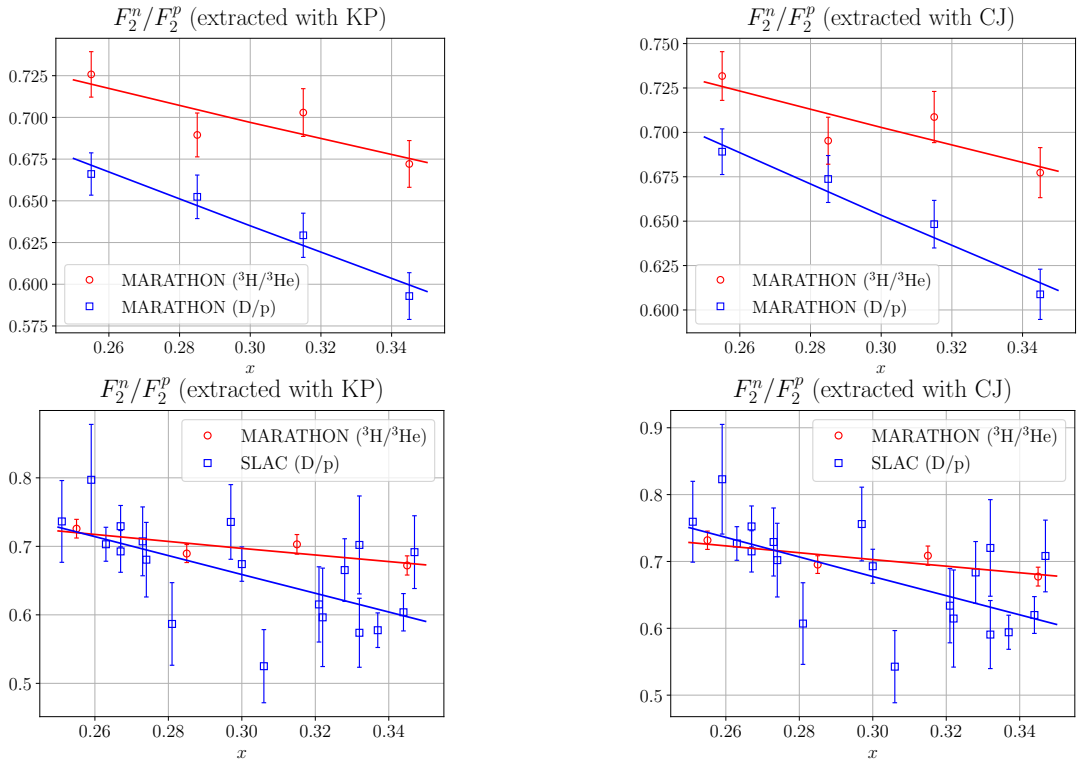


Figure 5.13: Comparison of F_2^n/F_2^p extractions from D/p and ^{3}H/^{3}He ratio.

than CJ. For helium-3, the obtained renormalizations are somewhat larger (between 1-3%). Here, the MARATHON D/p gives a higher renormalization than SLAC D/p, and KP gives a higher renormalization than CJ.

The uncertainty of each point includes error calculated from the covariance matrices of the fits, which include the point-to-point uncertainties of each dataset. In addition, the uncertainty includes the normalization uncertainty of each D/p dataset. When normalizing to D/p from MARATHON, the same deuterium data is used for both EMC ratios (^{3}H/D, ^{3}He/D) and D/p. This means that the uncertainties in the deuterium normalization are correlated between each F_2^n/F_2^p extraction. Analytically propagating the deuterium normalization uncertainty to the renormalization uncertainty is non-trivial...rather, a numerical approach was used. By maximally varying the deuterium normalization within its uncertainty (adjusting both the EMC and D/p ratios), the corresponding shift to the EMC normalizations was calculated. This allowed the uncertainty from the deuterium normalization to be determined,

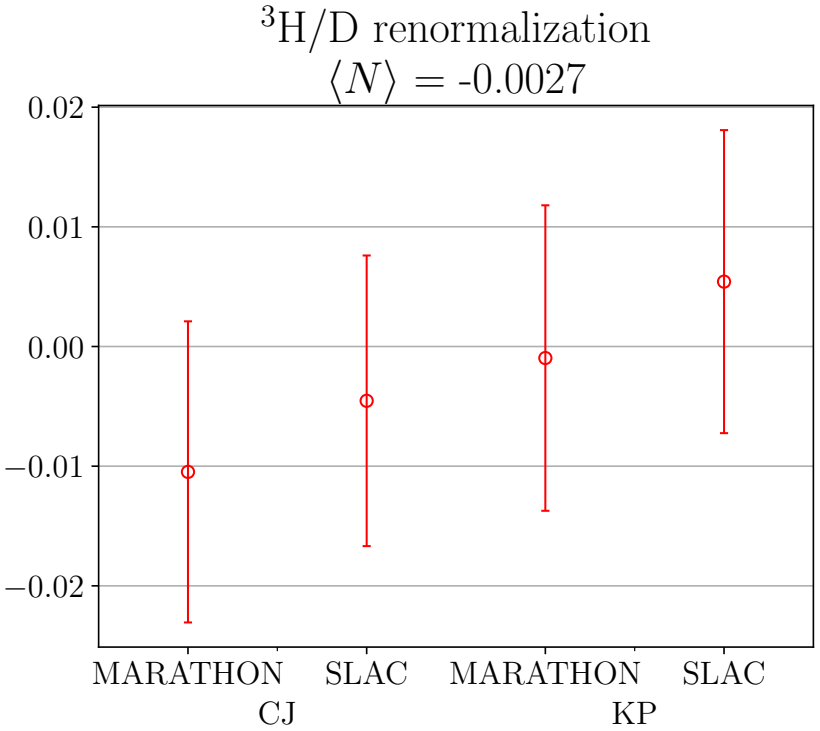


Figure 5.14: Renormalization to ^{3}H EMC ratio.

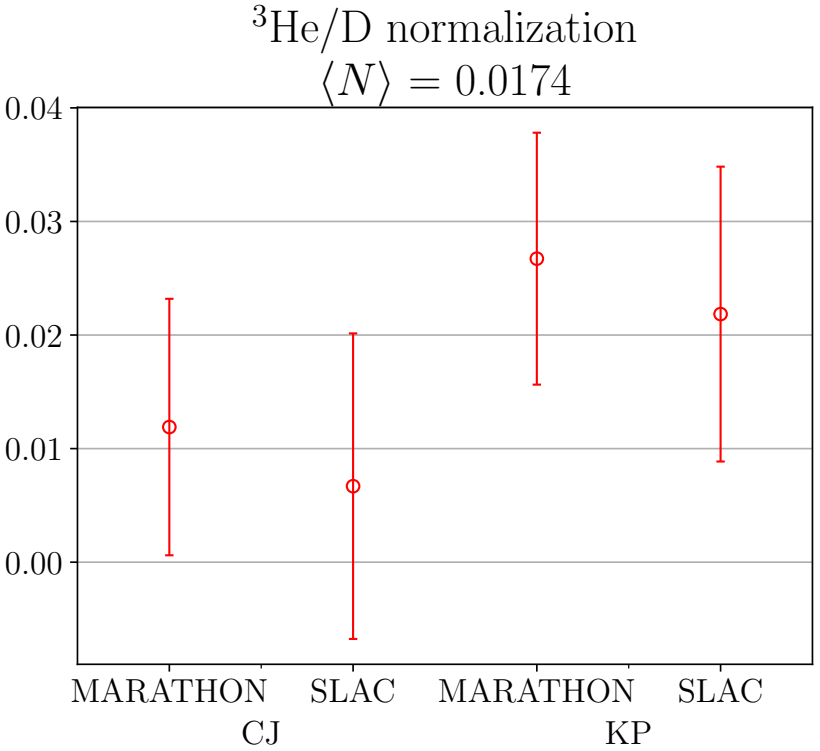


Figure 5.15: Renormalization to ^{3}He EMC ratio.

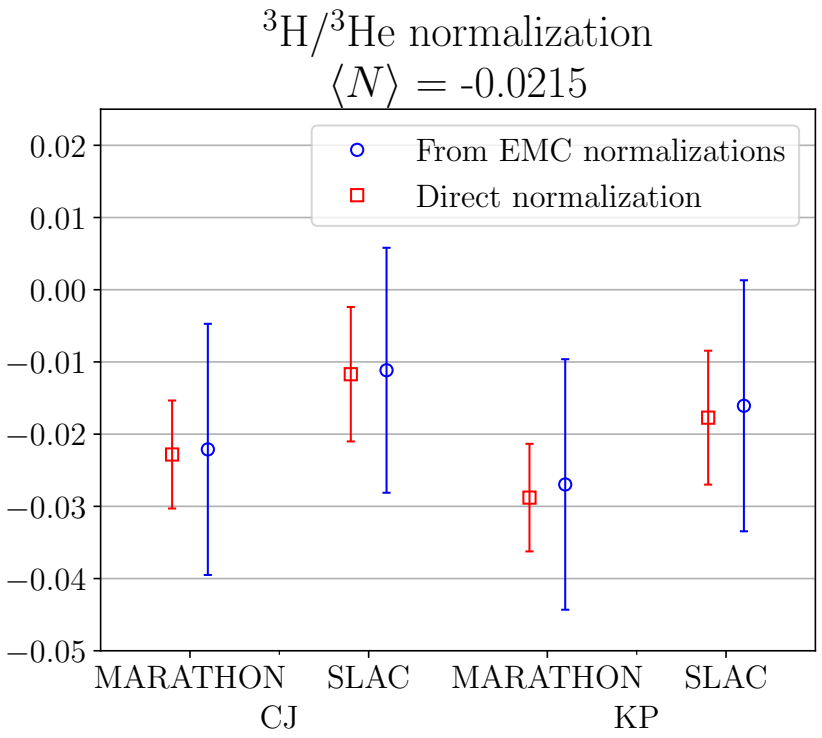


Figure 5.16: Renormalization to ^{3}H/^{3}He.

including the correlation between ratios.

Figure 5.16 shows the renormalizations to the ^{3}H/^{3}He ratio. For each combination of nuclear model and D/p dataset, two points are shown. The red squares indicate the the direct renormalization of the ^{3}H/^{3}He ratio; the blue circles indicate the ^{3}H/^{3}He renormalization resulting from the individual renormalizations of the EMC ratios. As can be seen, they are effectively the same, although the direct normalization has significantly smaller uncertainties.

To determine the renormalization of each ratio, a weighted average of each of the points shown in Figures 5.14-5.16 is performed. The uncertainty in the normalization arising from the data itself is found by the simple arithmetic mean of the uncertainties of each point. The resulting renormalizations are:

$$N(^3\text{H/D}) = (-0.27 \pm 1.25)\% \tag{5.6}$$

$$N(^3\text{He/D}) = (1.74 \pm 1.22)\% \tag{5.7}$$

$$N(^3\text{H}/^3\text{He}) = (-2.15 \pm 0.84)\% \tag{5.8}$$

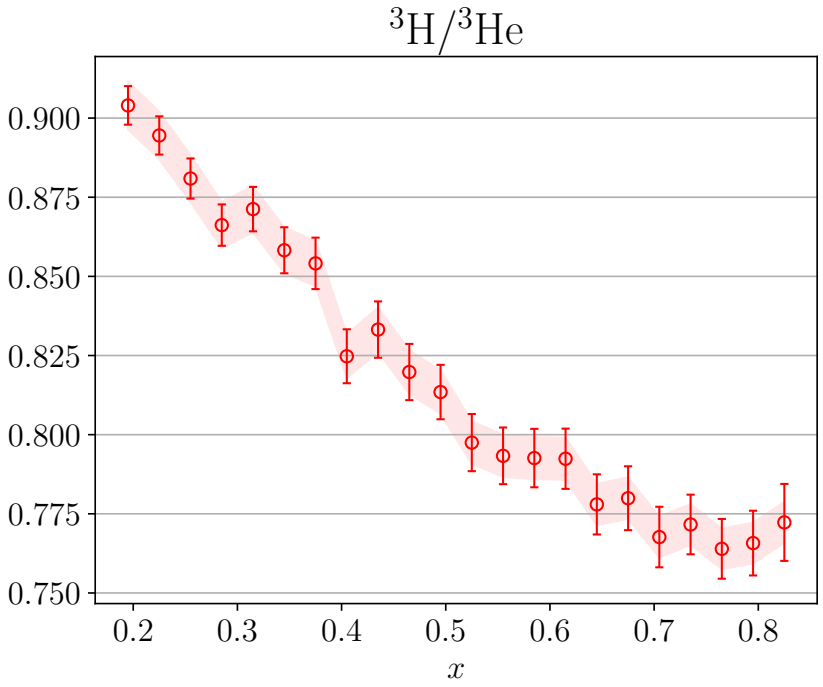


Figure 5.17: Renormalized preliminary ^{3}H/^{3}He cross section ratio. The shaded band indicates the normalization uncertainty.

5.4 Renormalized results

5.4.1 Cross section ratios

The renormalized ^{3}H/^{3}He and EMC ratios (^{3}H/D, ^{3}He/D) are shown in Figures 5.17, 5.18, and 5.19, respectively. Note that these are still based on the preliminary MARATHON results. See Appendix A, Table A.2 for the numeric values displayed in the these figures.

5.4.2 Isoscalar EMC effect in A = 3 nuclei

It is typical to apply an isoscalar correction to EMC ratios of asymmetric nuclei to account for unequal numbers of protons and neutrons (see Equation (2.36)). MARATHON's preliminary F_2^n/F_2^p (see Section 5.4.3) was fit with a second-degree polynomial to calculate the isoscalar correction at each EMC ratio point. The preliminary isoscalar-corrected ratios are shown in Figures 5.20 and 5.21. The figures also include predictions from the KP and CJ models for each ratio. The KP model tends to provide a better description of the MARATHON data.

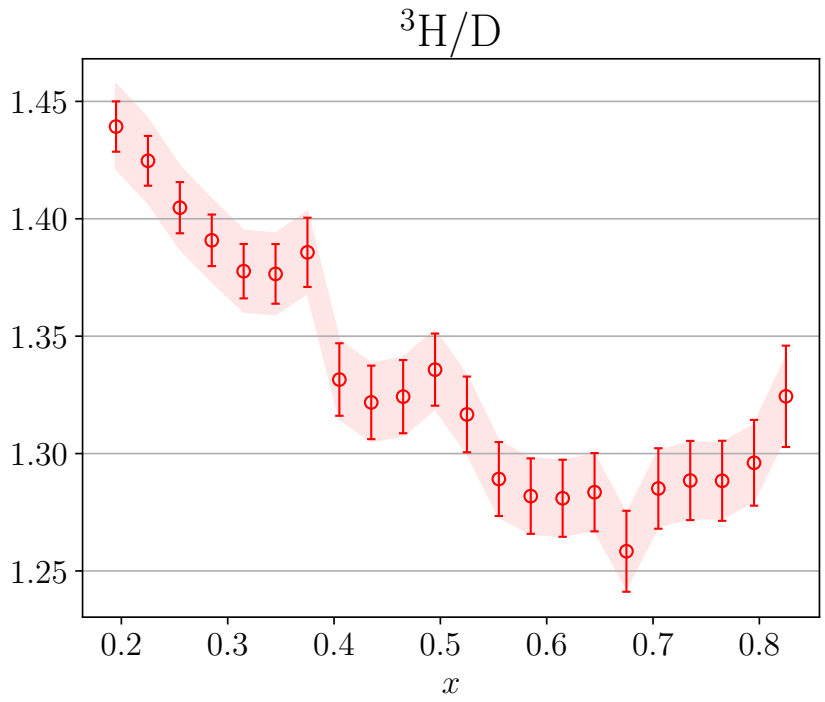


Figure 5.18: Renormalized preliminary ^{3}H/D cross section ratio. The shaded band indicates the normalization uncertainty.

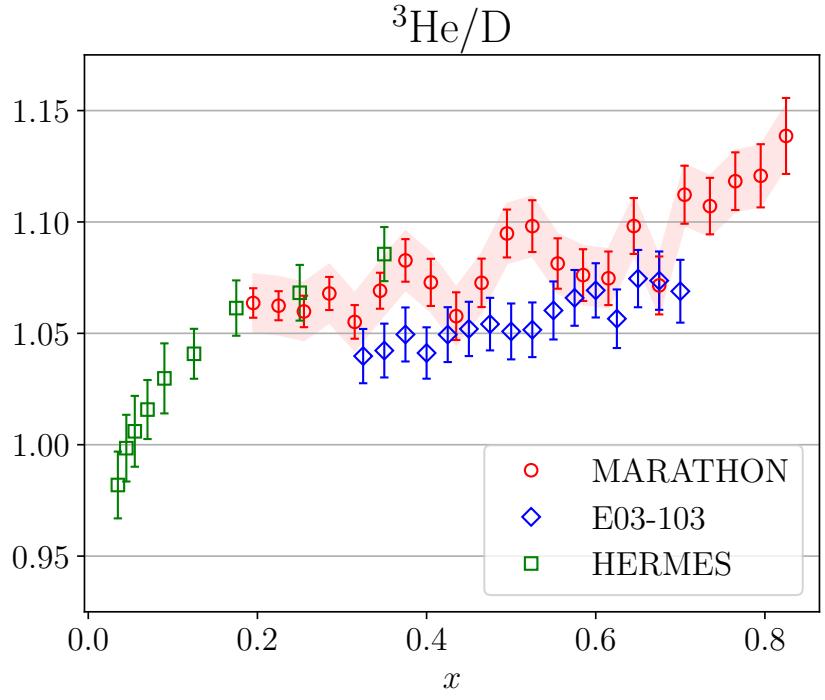


Figure 5.19: Renormalized preliminary ^{3}He/D cross section ratio. The shaded band indicates the normalization uncertainty.

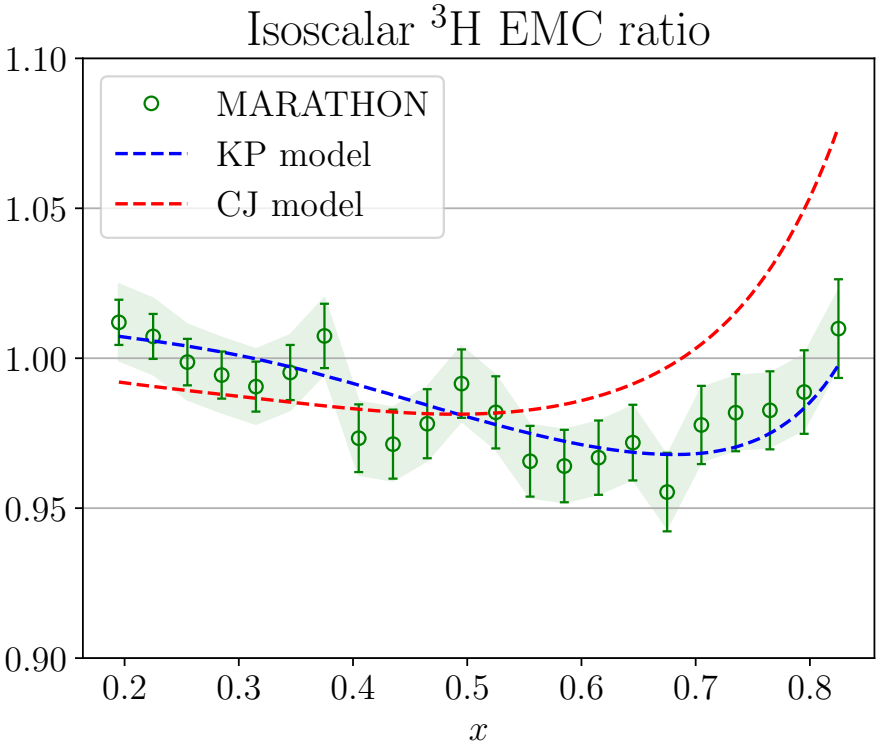


Figure 5.20: Preliminary isoscalar ^{3}H EMC ratio.

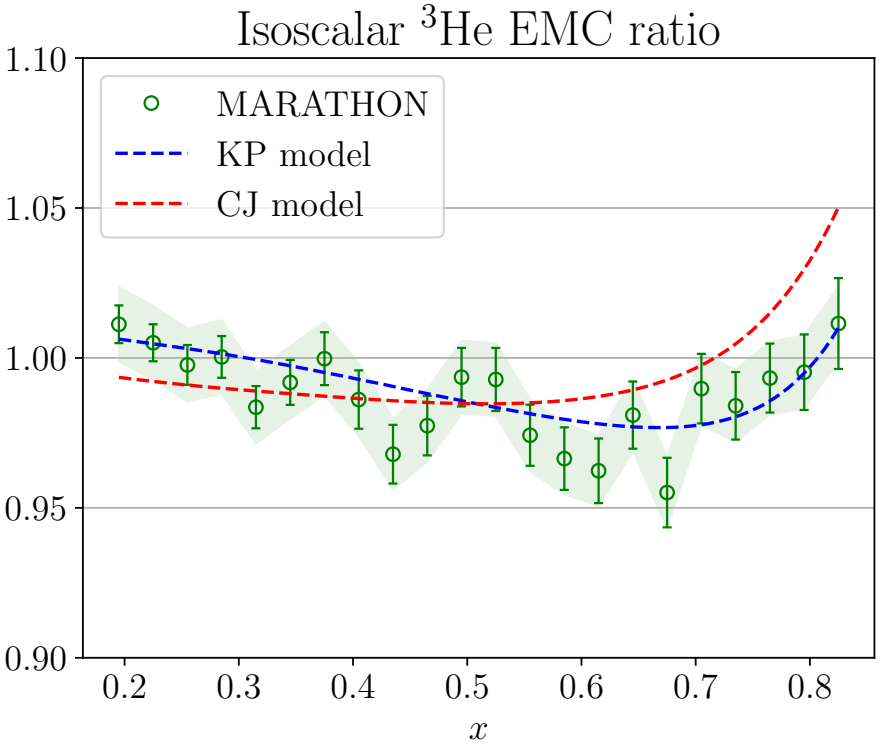


Figure 5.21: Preliminary Isoscalar ^{3}He EMC ratio.

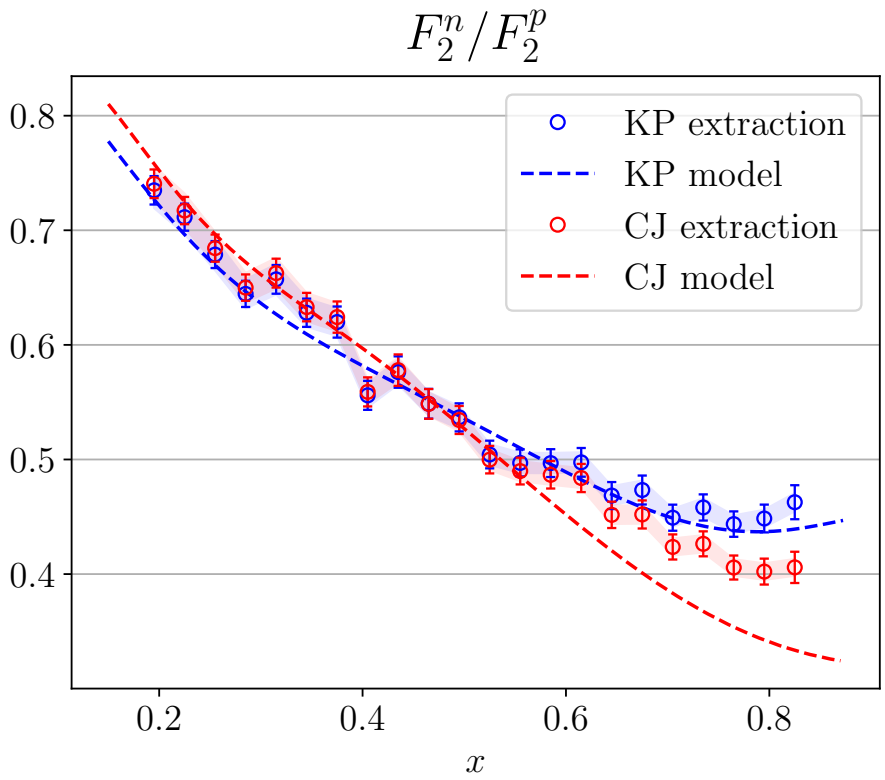


Figure 5.22: F_2^n/F_2^p extracted from preliminary renormalized ^{3}H/^{3}He using KP and CJ model. The shaded bands indicate the normalization uncertainty.

5.4.3 F_2^n/F_2^p

F_2^n/F_2^p was extracted from the renormalized ^{3}H/^{3}He ratio using both the KP and CJ nuclear model. These extractions are shown in Figure 5.22, along with the corresponding prediction of F_2^n/F_2^p from each model. The points extracted with the KP model are very close to the KP prediction, though appear to be systematically higher at large x. On the other hand, a larger discrepancy exists between the points extracted with the CJ model and the CJ prediction.

Ideally, the discrepancies between extraction with different models, and between the extractions and predictions, can be eliminated using an iterative extraction. In this procedure, the model parameterization of F_2^n/F_2^p is replaced with the F_2^n/F_2^p extracted with that model. The super-ratio is recalculated with the new F_2^n/F_2^p, and the extraction is repeated. This should eventually result in the extracted F_2^n/F_2^p matching the F_2^n/F_2^p parameterization of

that model. If the procedure is successful, it should also result in the convergence of F_2^n/F_2^p parameterization and extraction from different nuclear models. MARATHON could employ this procedure with Kulagin & Petti and the CJ15 collaboration in order to verify convergence of the F_2^n/F_2^p extraction.

Although completing the iterative extraction will require collaboration with the theorists responsible for each model, the effects of the procedure can be estimated. Comparing the F_2^n/F_2^p extraction from each nuclear model to the prediction from that model, KP clearly produces better agreement than CJ at high x. The iterative procedure would minimally shift the KP model, while the first iteration on CJ would shift the model by nearly 25% at high x. Given how small the initial discrepancy between the KP extraction and prediction is, the iterative procedure would shift the extraction at the sub-percent level, well within the uncertainties of the data. Iterating the CJ model should result in convergence very close to the KP extraction. For the following analysis, the KP extraction will be taken as the F_2^n/F_2^p result from MARATHON.

The importance of the F_2^n/F_2^p measurement at high x is to constrain the nucleon valence quark distributions, d/u. One final step remains to obtain d/u from F_2^n/F_2^p. The QPM is formulated in the limit of infinite momentum transfer Q^2. Obviously, experiments are performed at finite Q^2. There exist $1/Q^2$ suppressed effects that are relevant at the finite Q^2 values of MARATHON ($Q^2 < 14$ GeV^2). The structure functions in the QPM are so-called leading-twist, where the DIS process involves the interaction of a virtual photon with a single parton. However, dynamic higher-twist interactions involving multiple partons can contribute to the DIS cross section at MARATHON kinematics. Target mass effects (also known as kinematic higher-twist effects), arising from the fact that at finite Q^2 the struck parton's momentum fraction is not identically equal to Bjorken x, are also relevant. The relationship between F_2^n/F_2^p and d/u from Equation (2.21) is valid for the leading-twist nucleon structure functions, not the full structure functions that are measured experimentally. The highter-twist contributions to the extracted F_2^n/F_2^p must be removed, allowing

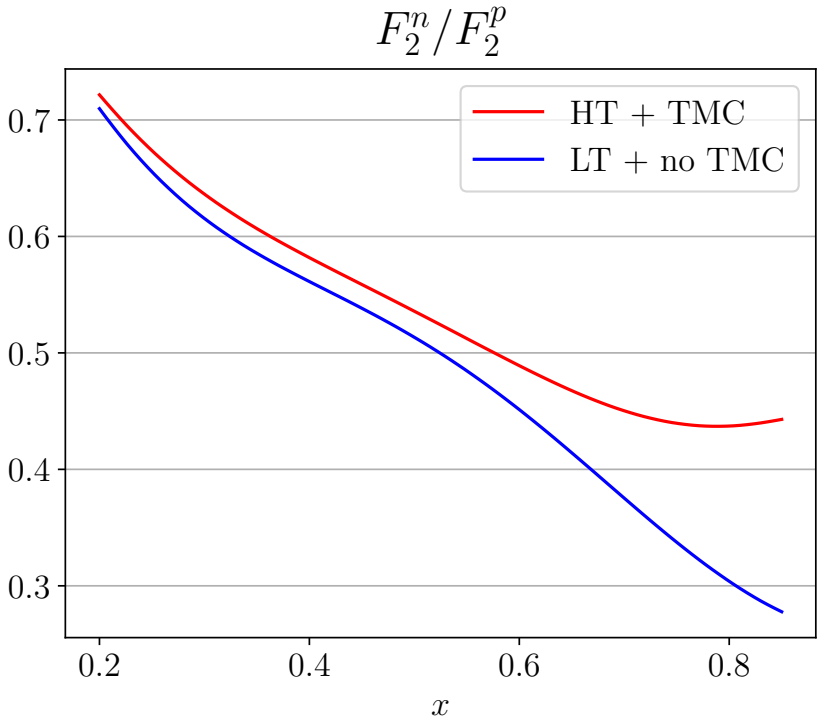


Figure 5.23: KP prediction of F_2^n/F_2^p with and without higher-twist corrections.

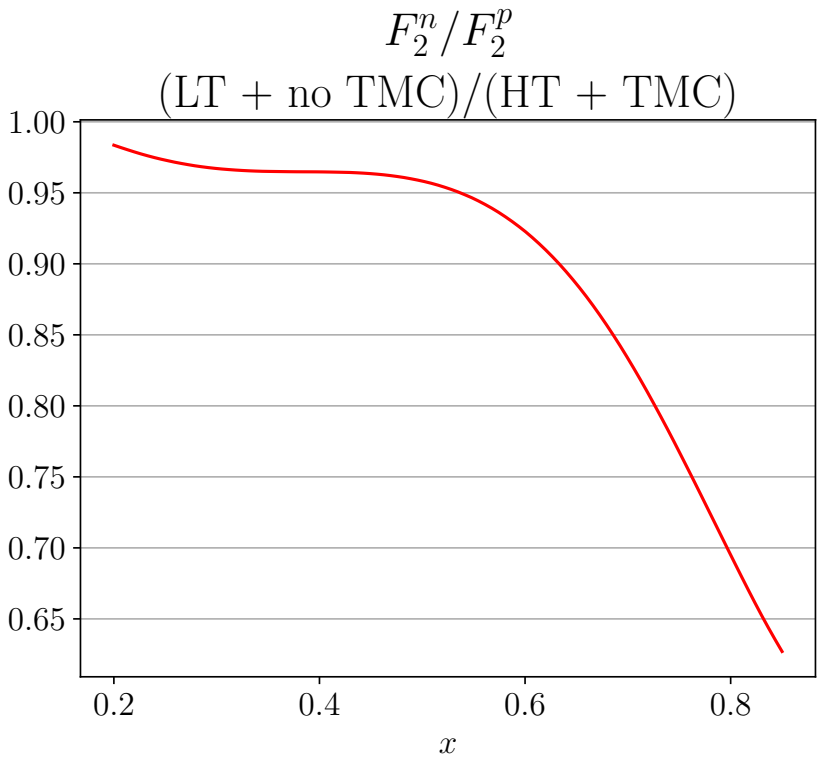


Figure 5.24: Ratio of KP prediction of F_2^n/F_2^p with and without higher-twist corrections.

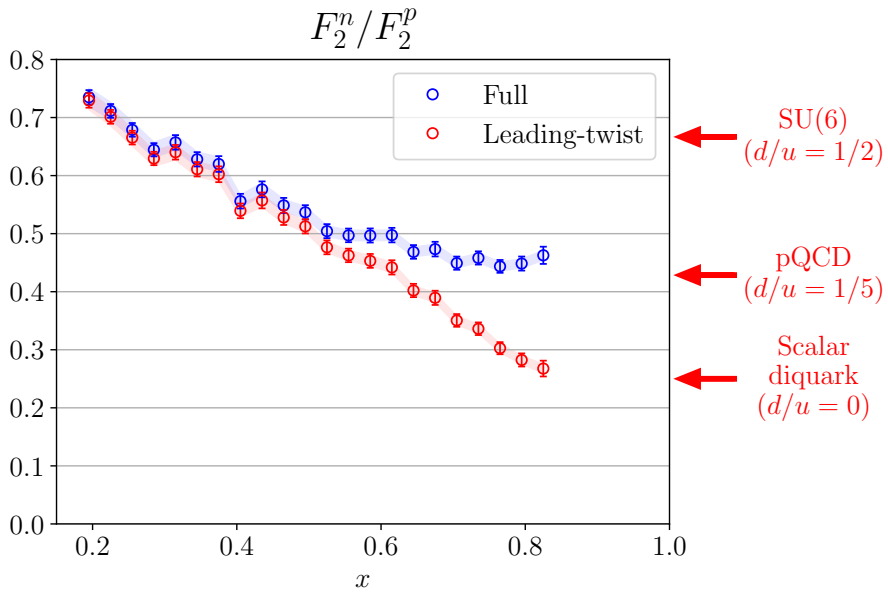


Figure 5.25: Estimation of leading-twist F_2^n/F_2^p from preliminary MARATHON data. Blue points are the full extraction of F_2^n/F_2^p, while the red points are corrected to remove higher-twist effects. Various predictions of F_2^n/F_2^p in the $x \to 1$ limit, and the corresponding d/u, are shown on the right axis. Note that the $x = 1$ predictions should be compared to the leading-twist points.

d/u to be determined. Some studies indicate that both the magnitude and uncertainties of these corrections are large [59]. This could potentially obscure the d/u interpretation of MARATHON's F_2^n/F_2^p measurement.

A full treatment of higher-twist effects will require the assistance of theorists. The MARATHON collaboration will work with Kulagin and Petti to obtain the leading-twist F_2^n/F_2^p from the data. In the meantime, a simple estimation of the leading-twist structure functions can be made using the KP nuclear model. The calculations by KP were completed both with (HT) and without (LT) dynamic higher-twist effects, and also with and without target mass corrections (TMC). The resulting predictions for F_2^n/F_2^p are shown in Figure 5.23. The ratio of the leading-twist to full F_2^n/F_2^p, shown in Figure 5.24, can be used to scale F_2^n/F_2^p from MARATHON for a preliminary estimate of the leading-twist contribution. This is a sizable correction, reaching approximately 30% at the highest x MARATHON point. The preliminary estimate of leading-twist F_2^n/F_2^p from MARATHON is shown as red points in Figure 5.25. For comparison, the uncorrected values are shown as blue points.

5.5 Discussion

The MARATHON experiment has carried out important new DIS measurements to aid the understanding of nucleon and nuclear structure. The low-x measurement of the D/p cross section ratio allowed a comparison to previous SLAC measurements and provided a normalization benchmark for the remaining cross section ratios. The ^{3}H/^{3}He cross section ratio allowed a novel extraction of F_2^n/F_2^p with minimal nuclear corrections, exploiting the mirror symmetry of A = 3 nuclei. Lastly, measurements of the ^{3}H/D and ^{3}He/D cross section ratios determined the EMC effect in A = 3 nuclei.

MARATHON has made the first measurement of the EMC effect in tritium (Figure 5.18). Additionally, MARATHON has made a new high Q^2 and W^2 measurement of the EMC effect in helium-3 (Figure 5.19). When the data is renormalized following the procedure described in Section 5.3.1, and MARATHON's F_2^n/F_2^p is used for the isoscalar correction, the isoscalar EMC ratios of tritium and helium-3 (Figures 5.20 and 5.21) are very similar.

The flagship result of MARATHON is the extraction of F_2^n/F_2^p from A = 3 nuclei. The full extraction (blue points in Figure 5.25) plateaus at large x, reaching a limit of $F_2^n/F_2^p \approx 0.45$ at $x \approx 0.85$. While this limit suggests a nonzero value of the valence quark ratio d/u, a naive estimate of the leading-twist F_2^n/F_2^p (red points in Figure 5.25), required to determine d/u, is quite different. Instead of plateauing, the ratio continues to decrease. If the estimated trend were to continue to $x = 1$, an extrapolation suggests that $F_2^n/F_2^p \approx 0.2$, indicating $d/u = 0$. A full leading-twist extraction must be carried out with input from theorists for the final determination d/u from MARATHON data.

The currently large uncertainties in higher-twist effects have the potential to obscure MARATHON's constraints on d/u, but this emphasizes the importance of the MARATHON data. Unraveling nucleon structure, including higher-twist effects, requires precise experimental data to which various theories and models can be compared. MARATHON's precise measurements of the ^{3}H/^{3}He cross section ratio and corresponding extraction of F_2^n/F_2^p are critical to this process.

Bibliography

[1] Gell-Mann, M. *Physics Letters* 8, 214 , 1964. ISSN 0031-9163
http://www.sciencedirect.com/science/article/pii/S0031916364920013

[2] Zweig, G. *Developments in the Quark Theory of Hadrons, Volume 1. Edited by D. Lichtenberg and S. Rosen. pp. 22-101* 22–101, 1964

[3] Tanabashi, M., et al. *Phys. Rev. D* 98, 030001, 2018
https://link.aps.org/doi/10.1103/PhysRevD.98.030001

[4] Halzen, F. and Martin, A.D. *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984. ISBN 0471887412, 9780471887416

[5] Bjorken, J.D. *Phys. Rev.* 179, 1547, 1969
https://link.aps.org/doi/10.1103/PhysRev.179.1547

[6] Callan, C.G. and Gross, D.J. *Phys. Rev. Lett.* 21, 311, 1968
https://link.aps.org/doi/10.1103/PhysRevLett.21.311

[7] Amaudruz, P., et al. *Physics Letters B* 294, 120 , 1992. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/037026939291650X

[8] Dasu, S., et al. *Phys. Rev. D* 49, 5641, 1994
https://link.aps.org/doi/10.1103/PhysRevD.49.5641

[9] Arneodo, M., et al. *Nuclear Physics B* 481, 23 , 1996. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/S0550321396901194

[10] Tvaskis, V., et al. *Phys. Rev. Lett.* 98, 142301, 2007
https://link.aps.org/doi/10.1103/PhysRevLett.98.142301

[11] Tvaskis, V., et al. *Phys. Rev. C* 97, 045204, 2018
https://link.aps.org/doi/10.1103/PhysRevC.97.045204

[12] Abramowicz, H., et al. *The European Physical Journal C* 75, 580, 2015. ISSN 1434-6052
https://doi.org/10.1140/epjc/s10052-015-3710-4

[13] Benvenuti, A., et al. *Physics Letters B* 223, 485 , 1989. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269389916377

[14] Arneodo, M., et al. *Nuclear Physics B* 483, 3 , 1997. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/S055032139600538X

[15] Adams, M.R., et al. *Phys. Rev. D* 54, 3006, 1996
https://link.aps.org/doi/10.1103/PhysRevD.54.3006

[16] Whitlow, L., et al. *Physics Letters B* 282, 475 , 1992. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/037026939290672Q

[17] Benvenuti, A., et al. *Physics Letters B* 237, 592 , 1990. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/037026939091231Y

[18] Accardi, A., et al. *The European Physical Journal C* 76, 471, 2016. ISSN 1434-6052
https://doi.org/10.1140/epjc/s10052-016-4285-4

[19] Accardi, A., et al. *Phys. Rev. D* 93, 114017, 2016
https://link.aps.org/doi/10.1103/PhysRevD.93.114017

[20] Close, F. *Physics Letters B* 43, 422 , 1973. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269373903894

[21] Farrar, G.R. and Jackson, D.R. *Phys. Rev. Lett.* 35, 1416, 1975
https://link.aps.org/doi/10.1103/PhysRevLett.35.1416

[22] Brodsky, S.J., Burkardt, M., and Schmidt, I. *Nuclear Physics B* 441, 197 , 1995. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/055032139500009H

[23] Aubert, J., et al. *Physics Letters B* 105, 315 , 1981. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269381908960

[24] Seely, J., et al. *Phys. Rev. Lett.* 103, 202301, 2009
https://link.aps.org/doi/10.1103/PhysRevLett.103.202301

[25] Gomez, J., et al. *Phys. Rev. D* 49, 4348, 1994
https://link.aps.org/doi/10.1103/PhysRevD.49.4348

[26] Amaudruz, P., et al. *Nuclear Physics B* 441, 3 , 1995. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/0550321394000239

[27] Arneodo, M., et al. *Nuclear Physics B* 481, 3 , 1996. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/S0550321396901170

[28] Arneodo, M., et al. *Nuclear Physics B* 333, 1 , 1990. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/055032139090221X

[29] Bodek, A., et al. *Phys. Rev. Lett.* 51, 534, 1983
https://link.aps.org/doi/10.1103/PhysRevLett.51.534

[30] Bodek, A., et al. *Phys. Rev. Lett.* 50, 1431, 1983
https://link.aps.org/doi/10.1103/PhysRevLett.50.1431

[31] Benvenuti, A., et al. *Physics Letters B* 189, 483 , 1987. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269387906642

[32] Malace, S., et al. *International Journal of Modern Physics E* 23, 1430013, 2014
https://doi.org/10.1142/S0218301314300136

[33] West, G. *Physics Letters B* 37, 509 , 1971. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269371903583

[34] Frankfurt, L. and Strikman, M. *Physics Reports* 76, 215 , 1981. ISSN 0370-1573
http://www.sciencedirect.com/science/article/pii/0370157381901290

[35] Bodek, A., et al. *Phys. Rev. D* 20, 1471, 1979
https://link.aps.org/doi/10.1103/PhysRevD.20.1471

[36] Frankfurt, L. and Strikman, M. *Physics Reports* 160, 235 , 1988. ISSN 0370-1573
http://www.sciencedirect.com/science/article/pii/0370157388901792

[37] Melnitchouk, W., Schreiber, A., and Thomas, A. *Physics Letters B* 335, 11 , 1994. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269394915504

[38] The MARATHON collaboration. *Jefferson Lab Proposal, PAC 36* 2010
https://www.jlab.org/exp_prog/PACpage/PAC36/Proposals/conditionally%20approved/12-06-118.pdf

[39] Afnan, I., et al. *Physics Letters B* 493, 36 , 2000. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/S0370269300010121

[40] Airapetian, A., et al. *Physics Letters B* 567, 339 , 2003. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/S0370269303009444

[41] *JLab Tritium Target Wiki* 2018
https://wiki.jlab.org/jlab_tritium_target_wiki/index.php/Main_Page

[42] Alcorn, J., et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 522, 294 , 2004. ISSN 0168-9002
http://www.sciencedirect.com/science/article/pii/S0168900203033977

[43] Liyanage, N. *JLAB-TN-02-012* 2002
https://hallaweb.jlab.org/publications/Technotes/files/2002/02-012.pdf

[44] Fissum, K., et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 474, 108 , 2001. ISSN 0168-9002
http://www.sciencedirect.com/science/article/pii/S0168900201008750

[45] Iodice, M., et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 411, 223 , 1998. ISSN 0168-9002
http://www.sciencedirect.com/science/article/pii/S0168900298002356

[46] Tsai, Y.S. *SLAC-PUB-848* 1971
https://www.slac.stanford.edu/cgi-wrap/getdoc/slac-pub-0848.pdf

[47] Tvaskis, V. *Ph.D. thesis, University of Vrije* 2004
http://inspirehep.net/record/665779/files/tvaskis.04-065.thesis.pdf

[48] MO, L.W. and TSAI, Y.S. *Rev. Mod. Phys.* 41, 205, 1969
https://link.aps.org/doi/10.1103/RevModPhys.41.205

[49] Arneodo, M., et al. *Physics Letters B* 364, 107 , 1995. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/0370269395013189

[50] Abe, K., et al. *Physics Letters B* 452, 194 , 1999. ISSN 0370-2693
http://www.sciencedirect.com/science/article/pii/S0370269399002440

[51] Kulagin, S. and Petti, R. *Nuclear Physics A* 765, 126 , 2006. ISSN 0375-9474
http://www.sciencedirect.com/science/article/pii/S0375947405011838

[52] Kulagin, S.A. and Petti, R. *Phys. Rev. C* 82, 054614, 2010
https://link.aps.org/doi/10.1103/PhysRevC.82.054614

[53] Kulagin, S.A. and Petti, R. *Private communication* 2018

[54] Amaudruz, P., et al. *Nuclear Physics B* 371, 3 , 1992. ISSN 0550-3213
http://www.sciencedirect.com/science/article/pii/0550321392902273

[55] Lucas, L. and Unterweger, M.P. *Journal of Research of the National Institute of Standards and Technology* 105, 541, 2000

[56] Abazov, V.M. and Anonymous. *Phys. Rev. Lett.* 114, 049901, 2015
https://link.aps.org/doi/10.1103/PhysRevLett.114.049901

[57] Abazov, V.M., et al. *Phys. Rev. Lett.* 112, 151803, 2014
https://link.aps.org/doi/10.1103/PhysRevLett.112.151803

[58] Tropiano, A.J., et al. *Phys. Rev. C* 99, 035201, 2019
https://link.aps.org/doi/10.1103/PhysRevC.99.035201

[59] Kulagin, S.A. *Private communication* 2019

Appendix A

Data tables

This appendix contains the measured (Table A.1) and renormalized (Table A.2) MARATHON per-nucleon cross section ratios. As the D/p ratio is not renormalized, it only appears in Table A.1. The δ columns indicate the absolute uncertainty for each ratio, including statistical and systematic uncertainty added in quadrature. The systematic uncertainty does not include the normalization uncertainty. See text for details on the normalization uncertainty. Note that all systematic uncertainties are currently being finalized for the official MARATHON publication, and are subject to change slightly.

x	D/p	δ(D/p)	^{3}H/^{3}He	$\delta(^3$H/^{3}He)	^{3}H/D	$\delta(^3$H/D)	^{3}He/D	$\delta(^3$He/D)
0.195	0.862	0.0060	0.924	0.0061	0.962	0.0071	1.045	0.0066
0.225	0.849	0.0059	0.914	0.0060	0.952	0.0071	1.044	0.0065
0.255	0.837	0.0064	0.900	0.0063	0.939	0.0073	1.042	0.0071
0.285	0.828	0.0065	0.885	0.0065	0.930	0.0073	1.050	0.0074
0.315	0.815	0.0066	0.890	0.0070	0.921	0.0077	1.037	0.0075
0.345	0.794	0.0070	0.877	0.0073	0.920	0.0085	1.051	0.0081
0.375	0.772	0.0085	0.873	0.0081	0.926	0.0098	1.064	0.0096
0.405			0.843	0.0085	0.890	0.0103	1.055	0.0106
0.435			0.851	0.0089	0.884	0.0104	1.040	0.0107
0.465			0.838	0.0089	0.885	0.0104	1.054	0.0109
0.495			0.831	0.0086	0.893	0.0102	1.076	0.0107
0.525			0.815	0.0090	0.880	0.0108	1.079	0.0116
0.555			0.811	0.0089	0.862	0.0105	1.063	0.0113
0.585			0.810	0.0092	0.857	0.0107	1.058	0.0116
0.615			0.810	0.0095	0.856	0.0109	1.056	0.0120
0.645			0.795	0.0095	0.858	0.0111	1.079	0.0125
0.675			0.797	0.0101	0.841	0.0115	1.053	0.0130
0.705			0.784	0.0096	0.859	0.0114	1.093	0.0130
0.735			0.789	0.0094	0.861	0.0113	1.088	0.0127
0.765			0.781	0.0094	0.861	0.0114	1.099	0.0129
0.795			0.783	0.0102	0.866	0.0122	1.102	0.0142
0.825			0.789	0.0122	0.885	0.0144	1.119	0.0171

Table A.1: Measured MARATHON per-nucleon cross section ratios.

x	^{3}H/^{3}He	$\delta(^3$H/^{3}He)	^{3}H/D	$\delta(^3$H/D)	^{3}He/D	$\delta(^3$He/D)
0.195	0.904	0.0061	0.960	0.0071	1.064	0.0066
0.225	0.895	0.0060	0.950	0.0071	1.062	0.0065
0.255	0.881	0.0063	0.936	0.0073	1.060	0.0071
0.285	0.866	0.0065	0.927	0.0073	1.068	0.0074
0.315	0.871	0.0070	0.918	0.0077	1.055	0.0075
0.345	0.858	0.0073	0.918	0.0085	1.069	0.0081
0.375	0.854	0.0081	0.924	0.0098	1.083	0.0096
0.405	0.825	0.0085	0.888	0.0103	1.073	0.0106
0.435	0.833	0.0089	0.881	0.0104	1.058	0.0107
0.465	0.820	0.0089	0.883	0.0104	1.073	0.0109
0.495	0.813	0.0086	0.890	0.0102	1.095	0.0107
0.525	0.797	0.0090	0.878	0.0108	1.098	0.0116
0.555	0.793	0.0089	0.859	0.0105	1.081	0.0113
0.585	0.793	0.0092	0.855	0.0107	1.076	0.0116
0.615	0.792	0.0095	0.854	0.0109	1.075	0.0120
0.645	0.778	0.0095	0.856	0.0111	1.098	0.0125
0.675	0.780	0.0101	0.839	0.0115	1.072	0.0130
0.705	0.768	0.0096	0.857	0.0114	1.112	0.0130
0.735	0.772	0.0094	0.859	0.0113	1.107	0.0127
0.765	0.764	0.0094	0.859	0.0114	1.118	0.0129
0.795	0.766	0.0102	0.864	0.0122	1.121	0.0142
0.825	0.772	0.0122	0.883	0.0144	1.139	0.0171

Table A.2: Renormalized MARATHON per-nucleon cross section ratios.