Probing superfast quarks in nuclei through dijet production at the LHC

Adam J. Freese,¹ Misak M. Sargsian,¹ and Mark I. Strikman²
¹Florida International University, Miami, FL 33199
²Pennsylvania State University, University Park, PA 16802
(Dated: November 25, 2014)

Abstract

We investigate dijet production from proton-nucleus collisions at the Large Hadron Collider (LHC) as a means for observing superfast quarks in nuclei with Bjorken $x > 1$. Kinematically, superfast quarks can be identified through directly measurable jet kinematics. Dynamically, their description requires understanding several elusive properties of nuclear QCD, such as nuclear forces at very short distances, as well as medium modification of parton distributions in nuclei. In the present work, we develop a model for nuclear parton distributions at large $x$ in which the nuclear distance at short distance scales is described by two- and three-nucleon short range correlations (SRCs). Nuclear modifications are accounted for by the color screening model, which has previously successfully described the available lepton deep inelastic scattering data. We apply QCD evolution at the leading order to obtain nuclear parton distributions in the kinematic regime of the LHC, and based on the obtained distributions we calculate the cross section for dijet production. We find not only that superfast quarks can be observed at the LHC, but that superfast quarks also provide sensitivity to the practically unexplored three-nucleon SRCs in nuclei.
I. INTRODUCTION

The dynamics of quantum chromodynamics (QCD) in the nuclear medium is one of the most interesting areas of modern nuclear physics. Many aspects of this are currently being investigated, including the formation of collective quark-gluon states such as quark-gluon plasma, the shadowing of the small $x$ parton densities in nuclei, the hadronization of quarks and gluons within the nucleus, and the medium modification of partonic distribution functions in nuclei.

Another interesting aspect of nuclear QCD is the possibility for partons to carry a light cone momentum fraction higher than that of a free nucleon at rest. Deep inelastic scattering (DIS) from such a parton will correspond to Bjorken $x > 1$, and we will hereafter refer to partons with such a light cone momentum fraction as superfast partons.

Due to the short-range nature of strong interactions, detecting a superfast quark in a nucleus requires probing the nucleus at extremely short distance scales. The characteristic space-time distances in nuclei become shorter with an increase in $Q^2$ for fixed $x$ owing to a property of QCD evolution, namely that a probed parton at high $Q^2$ will have come from a parent quark with a higher light cone momentum fraction. Thus, the theoretical expectation is that superfast quarks at large $Q^2$ will allow one to probe unprecedented small space-time distances in nuclei\[1\],\[2\].

Our current understanding of the dynamics of nuclei at short distances is extremely limited. Due to the short range nature of the strong interaction, one expects that short distances will be dominated by multi-nucleon short range correlations (SRCs)[3], which includes non-nucleonic degrees of freedom (such as $\Delta\Delta$ and $NN^*$ contributions from the SRCs), followed by the transition from baryonic to quark-gluon degrees of freedom.

There has been considerable progress recently made in studies of two-nucleon SRCs in inclusive and semi-inclusive nuclear processes, which have been dominated by quasi-elastic scattering of a high-energy probe (either an electron or a proton) off of a nucleon in the SRC[4–16]. These studies unambiguously established the existence of two-nucleon SRCs and measured the probability of a nucleon existing in one for the given nuclei. Due to the quasi-elastic nature of the scattering processes, however, it is much more difficult to reach kinematics dominated by three-nucleon SRCs[7]. Moreover, probing the transition to quark and gluon degrees of freedom has so far remained elusive.

The possibility of probing superfast quarks in hard nuclear processes provides a new venue in studies of SRCs. From the theoretical point of view, as was mentioned above, the QCD evolution of quark distributions gives an additional handle on the ability to probe exceedingly small space-time distances through an increase in $Q^2$ (at the scale of Ioffe time $1/m_N$). From the experimental point of view, deep inelastic processes do not restrict the phase space of the final states (as is a case for quasielastic processes), and hence one can use closure to express the cross section through the parton distribution function of the nucleus, which in turn can be expressed through light cone density matrix of the nucleus. Measurement of the parton distribution at $x \geq 1$ allows one to explore the transition from baryonic to quark degrees of freedom. For $x \geq 1.5$, additionally, the parton distribution is very sensitive to the three-nucleon SRCs.

One of the conventional methods for probing superfast quarks in nuclei is deep inelastic scattering (DIS) from nuclei at Bjorken $x > 1$. A number of attempts have been taken over the years. The attempts to measure quarks at $x > 1$ were undertaken at CERN using a muon beam[17] and at FNAL using a neutrino beam[18], with contradictory results.
Measurements using electron beams were taken for $x \sim 1$ at SLAC\cite{19} for $Q^2 \leq 10$ GeV$^2$ and at Jefferson Lab for $Q^2 = 7$ GeV$^2$ \cite{20}. However, this $x$ range gets a significant contribution from higher-twist quasieleastic scattering up to fairly large $Q^2$ ($\sim 15$ GeV$^2$). For example, the nucleus/deuteron cross section ratio is reduced for $x = 1$ and $Q^2 = 10$ GeV$^2$ by a factor of 2 to 2.5\cite{4}. The only way to avoid this is to probe larger values of $Q^2 \geq 30$ GeV$^2$. The only way to avoid this is to probe larger values of $Q^2 \geq 30$ GeV$^2$, for which the quasieleastic contribution will be a small correction\cite{2}. Such experiments are currently included in the physics program for the 12 GeV upgrade of Jefferson Lab\cite{21}, and the first experimental data will be available within the next few years\cite{22}.

In this work, we propose a new approach to probing superfast quarks using data from the recent proton-nucleus run at the LHC. This method is based on the standard method of studying dijet production in hadron-hadron collisions, which allows for the light cone momentum fractions of the initial partons to be given in terms of momenta of the dijets. This gives the possibility of selecting transverse jet momenta and pseudo-rapidities to isolate scattering off the superfast quarks within the nucleus. We develop a theoretical framework for calculating this reaction, which requires addressing several issues, such as modeling the high-momentum (short range) properties of the nuclear wave function, and calculating the medium modification of parton distributions within the nucleus and evolving this modification. Within the framework we develop, we calculate the absolute cross section for the reaction and study its sensitivity to two- and three-nucleon SRCs.

The article is organized as follows: In Sec. II we review the formalism of the dijet production reaction, including its kinematics (and the meaning of $x > 1$) and the cross section formula. The cross section for dijet production depends on the nuclear parton distribution function (PDF), which is discussed in Sec. III. Sec. III also discusses short range correlations and medium modifications, and how they factor into and affect the cross section. Sec. IV discusses the hard subprocesses that contribute to dijet production. In Sec. V, we present numerical estimates for the cross section. Conclusions and outlook are given in Sec. VI. The Appendix gives detailed derivations of the factorization formula and the SRC parts of the nuclear light cone momentum distributions.

II. BASIC FORMALISM

The reaction we consider is the production of a dijet from a proton-nucleus collision,

$$p + A \rightarrow \text{dijet} + X,$$

where the kinematics of the two outgoing jets are fully determined. The reaction is treated at the leading order (LO) in perturbative QCD (pQCD), meaning that the jets are produced by a two-body parton-parton scattering process, in which one parton originates in the proton and the other in the nucleus. In our derivation, we restrict ourselves by considering only nucleonic degrees of freedom in the nucleus, and thus consider the nuclear parton as originating from one of its bound nucleons. This scenario is presented in Fig. 1, where we also define the kinematic variables of Eq. (2).
A. Jet Kinematics

We consider a scenario with a proton moving in the $+z$ direction and a heavy nucleus, with charge number $Z$ and mass number $A$, moving in the $-z$ direction. We describe their four-momenta using light cone coordinates, namely

$$ p^\mu \equiv (p^+, p^-, p_T), $$

where $p^\pm = E \pm p_z$ and $p_T$ is the two-component transverse momentum. Using the on-mass shell condition and the fact that the energies of the proton and nucleus greatly exceed their masses, in the collider reference frame one has

$$ p_p^\mu = \left( \frac{m_p^2}{p_p^+}, 0, 0 \right) = \left( 2E_0, 0, 0 \right) = \left( \frac{\sqrt{A_{NN}^{\text{avg}}}}{Z}, 0, 0 \right), $$

$$ p_A^\mu = \left( \frac{M_A^2}{p_A^+}, P_A, 0 \right) = \left( 0, 2ZE_0, 0 \right) = \left( 0, \sqrt{AZ_{NN}^{\text{avg}}}, 0 \right), $$

where $E_0$ is the beam energy per proton in the collider reference frame, and $s_{NN}^{\text{avg}} = 4\frac{Z}{A}E_0^2$ is the square of the average center-of-mass energy per nucleon. As an example, a lead-proton collision with a beam energy of 4 TeV per proton would have an average center-of-mass energy per nucleon of $\sqrt{s_{NN}^{\text{avg}}} \approx 5.02$ TeV. However, since the motion of the bound nucleon inside the nucleus will in general be variable, the actual center-of-mass energy per nucleon $\sqrt{s_{NN}}$ is not a fixed parameter, though it will be equal to $\sqrt{s_{NN}^{\text{avg}}}$ in the limit where the nucleons do not interact and all move forward with equal momenta.

At leading order, the collision results in an interaction between two partons, one each from the proton and the nucleus. Their respective four-momenta are labeled $p_1$ and $p_2$. We use a collinear approximation, in which the initial partons are treated as having zero transverse momentum. Moreover, they are treated as massless and on-shell, so

$$ p_1 = (p_1^+, 0; 0) $$

$$ p_2 = (0, p_2^-; 0). $$

FIG. 1: Diagram of reaction
The light cone momentum fractions are defined for each parton as

\[ x_p = \frac{p_1^+}{p_p} = \sqrt{Z \frac{p_1^+}{A \sqrt{s_{NN}}} \frac{1}{\sqrt{s_{avg}}}} \]  

(5)

\[ x_A = \frac{A p_2^-}{p_A} = \sqrt{\frac{A}{Z} \frac{p_2^-}{\sqrt{s_{avg}}}} \]  

(6)

Note that \( x_A \) is scaled by a factor of \( A \). The rationale behind this is that the parton from the nucleus is found within one of its nucleons, and in a limiting case where the nucleons are all non-interacting and carry equal momenta, their light cone momentum is equal to \( p_{N, avg} = \frac{p_N}{A} \), meaning \( x_A \leq 1 \) in this case.

In reality, however, the bound nucleons do interact and it is possible for \( p_2^- > p_{N, avg} \). In this situation, the nuclear parton originates from a nucleon which has larger-than-average momentum (\( p_N > p_{N, avg} \)). This indicates that finding an exceedingly large \( x_A > 1 \) will identify a bound nucleon with momentum significantly larger than average.

The parton momentum fractions cannot be directly measured. However, they can be related to the kinematic parameters of the jets. The jets from the proton and the nucleus are respectively ascribed four-momenta \( p_3 \) and \( p_4 \). At leading order, they come from the fragmentation of two partons (with the same momenta), which are treated as massless and on-shell. From energy-momentum conservation, it follows that

\[ p_1 + p_2 = p_3 + p_4. \]  

(7)

and due to the assumed collinear approximation, \( p_{3T} = -p_{4T} \equiv p_T \). Using this relation, and neglecting the masses of the produced jets, we obtain:

\[ p_3^+ p_3^- = p_4^+ p_4^- = p_T^2. \]  

(8)

To proceed, we define the rapidity \( \eta \) as

\[ \eta = \frac{1}{2} \log \left( \frac{p^+}{p^-} \right), \]  

(9)

and use this, with the massless limit (in which \( p^+ p^- = p_T^2 \)) to obtain

\[ p^\pm = p_T e^{\pm \eta}. \]  

(10)

Applied to the jets of reaction (1), this results in \( p_3^\pm = p_T e^{\pm \eta_3} \) and \( p_4^\pm = p_T e^{\pm \eta_4} \). Using these relations, and energy momentum conservation, viz. Eq. (7), the light cone momenta of the initial partons can be expressed through jet kinematics in the following form:

\[ p_1^+ = p_3^+ + p_4^+ = p_T (e^{\eta_3} + e^{\eta_4}) \]

\[ p_2^- = p_3^- + p_4^- = p_T (e^{-\eta_3} + e^{-\eta_4}). \]

These relations can be used to express the momentum fractions \( x_p \) and \( x_A \) in terms of jet observables, namely

\[ x_p = \sqrt{Z \frac{p_T}{A \sqrt{s_{NN}}} (e^{\eta_3} + e^{\eta_4})} \]  

(11)

\[ x_A = \sqrt{\frac{A}{Z} \frac{p_T}{\sqrt{s_{avg}}}} (e^{-\eta_3} + e^{-\eta_4}). \]  

(12)
The main question that concerns us is whether partons with \( x_A > 1 \) can be seen in lead-proton collisions at the LHC. Eq. (12) suggests that we should look for three conditions: large \( p_T \), small (or highly negative) \( \eta_3 \), and small (or highly negative) \( \eta_4 \). The first two conditions are expected to lead to a drop in the cross section, however, so a compromise must be found between the conditions that lead to large \( x_A \) and those that lead to an appreciable cross section. We anticipate that the best scenario is one with a forward jet from the nucleus (highly negative \( \eta_4 \)) and a central jet from the proton (small \( \eta_3 \)). A plot of the dependence of \( x_A \) on \( \eta_4 \) is given in Fig. 2.

In addition to the momentum fractions, we wish to express the characteristic “hardness” scale \( Q^2 \), to be used as both the factorization scale and renormalization scale, in terms of jet observables. It is natural to use \( Q^2 = p_T^2 \), since the Mandelstam variable \(-t\) for the parton-level process is

\[
-t = -(p_1 - p_3)^2 = 2(p_1 \cdot p_3) = p_1^+ p_3^- = p_T^2 \left( 1 + e^{-\eta_3 + \eta_4} \right) \approx p_T^2.
\]

B. Dijet cross section

The factorization formula for the hadronic cross section can be presented in terms of the partonic cross section as follows:

\[
\sigma_{\text{had.}} = \sum_{ij} \int_0^1 dx_p \int_0^A dx_A f_{i/p}(x_p, Q^2) f_{j/A}(x_A, Q^2) \sigma_{\text{part.}},
\]

where \( f_{i/p}(x_p, Q^2) \) and \( f_{j/A}(x_A, Q^2) \) are parton distribution functions (PDFs) for the proton and nucleus, respectively. This is similar in form to the standard factorization formula for \( pp \) collisions\[23, 24\], and in fact reduces to it in the free nucleon limit. A derivation of Eq. (13) can be found in Appendix A.
Since, at leading order, the momenta of the outgoing partons coincide with the jet momenta, we can integrate out the transverse momentum of one of the jets and express differentials through the rapidities of the outgoing jets. For this purpose, we present the differential form of Eq. (13) as

\[
d\sigma_{pA} = \sum_{ijkl} \frac{1}{16\pi p^+_p p^-_A} \frac{1}{2(p_1 \cdot p_2)} f_{ij/p}(x_p, Q^2) f_{j/A}(x_A, Q^2) \frac{1}{1 + \delta_{kl}} |M_{ij\rightarrow kl}|^2 \frac{2\delta^{(1)}(p^+_1 - p^+_3 - p^+_4)}{2(2\pi)^4} \times \delta^{(1)}(p^-_2 - p^-_3 - p^-_4) \delta^{(2)}(p_{3T} + p_{4T}) \frac{d^3 p_3}{2E_3(2\pi)^3} \frac{d^3 p_4}{2E_4(2\pi)^3} dx_p dx_A,
\]

where we have expressed four-momentum conservation through the light cone momenta. Here, the indices \( i, j, k, \) and \( l \) indicate parton types, and \( M_{ij\rightarrow kl} \) is the invariant Feynman amplitude for the hard partonic scattering process.

Using the definitions of the light cone momenta fractions in Eqs. (5,6), we can express the delta functions for \( p^\pm \) conservation as:

\[
2\delta^{(1)}(p^+_1 - p^+_3 - p^+_4)\delta^{(1)}(p^-_2 - p^-_3 - p^-_4) = \frac{2A}{p^+_p p^-_A} \delta^{(1)} \left( x_p - \frac{p^+_3 + p^+_4}{p^+_p} \right) \delta^{(1)} \left( x_A - A \frac{p^-_3 + p^-_4}{p^-_A} \right).
\]

Inserting this relation into Eq. (14), one can integrate out \( p_{4T}, x_p \) and \( x_A \), resulting in

\[
d\sigma_{pA} = \sum_{ijkl} \frac{A}{16\pi p^+_p p^-_A} \frac{1}{2(p_1 \cdot p_2)} f_{ij/p}(x_p, Q^2) f_{j/A}(x_A, Q^2) \frac{|M_{ij\rightarrow kl}|^2}{1 + \delta_{kl}} \frac{dp_{3z} dp_{4z}}{E_3 E_4} \frac{dp_T}{E_T}.
\]

The \( p_z \) elements can be rewritten as rapidities, since \( d\eta = \frac{dp_z}{E} \). In addition, as \( 2(p_1 \cdot p_2) = \frac{x_p x_A}{A} p^+_p p^-_A \), we have

\[
d\sigma_{pA} = \sum_{ijkl} \frac{1}{16\pi} \left( \frac{A}{p^+_p p^-_A} \right)^2 f_{ij/p}(x_p, Q^2) f_{j/A}(x_A, Q^2) \frac{|M_{ij\rightarrow kl}|^2}{1 + \delta_{kl}} d\eta_3 d\eta_4 dp_T^2.
\]

Since the square of the average center-of-mass energy per nucleon is

\[
s_{NN}^{\text{avg.}} = \frac{p^+_p p^-_A}{A},
\]

the differential cross section can be written in the standard form:

\[
\frac{d^3\sigma}{d\eta_3 d\eta_4 dp_T^2} = \frac{1}{16\pi (s_{NN}^{\text{avg.}})^2} \frac{f_{ij/p}(x_p, Q^2) f_{j/A}(x_A, Q^2) |M_{ij\rightarrow kl}|^2}{x_p x_A} \frac{1}{1 + \delta_{kl}}.
\]

III. NUCLEAR PDFS

The main theoretical issue to be worked out is the nuclear PDF \( f_{j/A}(x_A, Q^2) \). Currently, there is a lack of strong experimental constraint on nuclear PDFs. Phenomenological parametrizations exist for several nuclei, and are based on experimental measurements of
the nuclear structure function $F_2^{(A)}(x, Q^2)$ in deep inelastic scattering (DIS) experiments in a restricted range of $x$ and $Q^2$. (cf. e.g. Refs. [25–30].) Most treatments of the nuclear PDF parametrize the ratio between the nucleus and the nucleon, which only makes sense for treating $x_A$ up to 1. Ref. [26] instead relates the nuclear and nucleonic PDFs through a Mellin convolution, and parametrizes the function that is convoluted with the nucleonic PDF. This approach has the possibility of accounting for $x_A > 1$, but owing to a lack of experimental constraint, Ref. [26] did not treat this region.

In order to describe $x_A > 1$ in the absence of robust empirical parametrizations of the nuclear PDF, we must account for the nuclear PDF in terms of the nucleonic PDF theoretically. To this end, we must look at the theoretical relationship between the nuclear and nucleonic parton distributions.

Theoretical and experimental studies of DIS reactions have demonstrated that, in a limit where only nucleonic degrees of freedom are accounted for in the nucleus, the nuclear structure function can be expressed as a convolution between the bound nucleons’ structure functions and the nuclear light cone density matrix[1, 31, 32]. In the Bjorken limit, where the photon virtuality $Q^2$ and energy $q_0$ both go to infinity at a fixed $x = \frac{AQ^2}{2M_A q_0}$, the convolution formula takes an especially simple form:

$$F_2^{(A)}(x, Q^2) = \sum_N \int_x^A d\alpha \rho_{N/A}(\alpha) F_2^{(N, \text{eff})}\left(\frac{x}{\alpha}, Q^2\right).$$  \hfill (18)

Here, $\rho_{N/A}(\alpha)$ is the light cone density matrix of a nucleon $N$ in the nucleus $A$, described in Sec. III A. The “eff” in the superscript of $F_2^{(N, \text{eff})}\left(\frac{x}{\alpha}, Q^2\right)$ signifies that this is an effective nucleonic structure function, which differs from the free nucleon structure function owing to modifications from the nuclear medium. When the factorization scale $\mu_F$ is equal to the characteristic hardness scale $Q$, we have

$$F_2(x, Q^2) = \sum_i e_i^2 x f_i(x, Q^2).$$  \hfill (19)

Using this relation and Eq. (18) allows us to obtain a convolution relation between nuclear and bound nucleon PDFs in the form:

$$f_{i/A}(x, Q^2) = \sum_N \int_x^A \frac{d\alpha}{\alpha} \rho_{N/A}(\alpha) f^{(\text{eff})}_{i/N}\left(\frac{x}{\alpha}, Q^2\right).$$  \hfill (20)

This relation indicates that to construct nuclear PDF’s one need to address two theoretical issues: the light cone density matrix of $\rho_{N/A}(\alpha)$, and possible medium modification effects on the bound nucleon PDFs.
A. Light cone density matrix and SRCs

We formally define the light cone density matrix in terms of the nuclear wave function as:

$$\rho_A(\alpha, p_T) = \int \prod_{j=1}^A \frac{d\alpha d^2p_T}{\alpha} \psi^\dagger(\alpha_1, p_{T1}, \ldots, \alpha_A, p_{TA}) \psi(\alpha_1, p_{T1}, \ldots, \alpha_A, p_{TA})$$

$$\times \delta^{(1)} \left(1 - \sum_{i=1}^A \alpha_i \right) \delta^{(2)} \left( \sum_{i=1}^A p_{iT} \right) \left\{ \sum_{i=1}^A \delta^{(1)}(\alpha - \alpha_i) \delta^{(2)}(p_T - p_{iT}) \right\}$$  \hspace{1cm} (21)

where the (scaled) light cone momentum fraction $\alpha$ given by

$$\alpha = \frac{A^+ p_N}{p_A},$$  \hspace{1cm} (22)

and $p_T$ is the transverse momentum of the nucleon. The light cone density matrix is normalized to obey two sum rules, namely baryon number conservation and momentum conservation:

$$\int_0^A d\alpha \int d^2p_T \rho_{N/A}(\alpha, p_T) = 1$$  \hspace{1cm} (23)

$$\int_0^A d\alpha \int d^2p_T \alpha \rho_{N/A}(\alpha, p_T) = 1.$$  \hspace{1cm} (24)

For our calculations, we decompose the light cone density matrix as a sum of contributions from the nuclear mean field and short range correlations (SRCs). The decomposition takes the form:

$$\rho_{N/A}(\alpha, p_T) = \rho_{MF}^{(N)}(\alpha, p_T) + \sum_{j=2}^A \rho_j^{(N)}(\alpha, p_T),$$  \hspace{1cm} (25)

where $\rho_{MF}^{(N)}(\alpha, p_T)$ is the mean field part, and $\rho_j^{(N)}(\alpha, p_T)$ is the density of $j$-nucleon SRCs.

1. Mean field density matrix

The mean field part of the light cone density matrix describes how the nucleons in the nucleus would be distributed if they were only acted upon the mean field generated by the bulk of the nucleus. The mean field density can be related to the wave function of the nucleus, which is calculated in the non-relativistic limit since the relevant momenta are smaller than the typical Fermi momentum of heavy nuclei. The crudest mean field model considers the heavy nucleus as a degenerate Fermi gas, and the typical Fermi momentum is around 250 MeV[33]. More sophisticated momentum distributions (calculated based on, for example, Hartree-Fock approximations) still fall off sharply after the Fermi momentum, so the leading order relativistic corrections are at most on the order of the mean field momentum distribution strength above $k_F$, which is at most a few % in magnitude.

The mean field density is related to the non-relativistic momentum-space wave function using the sum rule of Eq. (23). Short range correlations are briefly neglected, so that the
mean field density matrix obeys Eq. (23) by itself. The momentum-space wave function 
\( \Psi_{MF}^{(N)}(p) \) (which carries an index of \((N)\) since the wave function will in general be different for protons and neutrons) is likewise normalized to unity, so we equate 
\[
\left| \Psi_{MF}^{(N)}(p) \right|^2 d^3p = \rho_{N/A}(\alpha, \mathbf{p}_T) d\alpha d^2\mathbf{p}_T.
\]
In the non-relativistic limit, we can write, in the nuclear center of mass frame,
\[
\alpha = A \frac{E + p_z}{m_A} \approx A \frac{m_N + p_z}{m_A},
\]
\[
p_z \approx \frac{m_A}{A} \alpha - m_N
\]
\[
dp_z \approx \frac{m_A}{A} d\alpha.
\]
Consequently, we can identify
\[
\rho_{MF}^{(N)}(\alpha, \mathbf{p}_T) = \frac{m_A}{A} \left| \Psi_{MF}^{(N)}(p) \right|^2.
\]
One modification must be made when accounting for SRCs. It is not the mean field part of the light cone density matrix, but the light cone density matrix taken as a whole that is normalized to satisfy the sum rules of Eqs. (23, 24). Therefore, \( \rho_{MF}^{(N)}(\alpha, \mathbf{p}_T) \) must be scaled down by a factor of \( a_1^{(N)}(A) \), which is defined as
\[
a_1^{(N)}(A) = 1 - \sum_j \int d\alpha d^2\mathbf{p}_T \rho_j^{(N)}(\alpha, \mathbf{p}_T),
\]
i.e. it subtracts off the probability that the nucleon is in a short range correlation. For instance, if a nucleon in is in a short range correlation 25% of the time, then \( a_1^{(N)}(A) \) will be 0.75. In general, however, \( a_1^{(N)} \) will be different for protons and neutrons.

For numerical estimates, we will use the momentum distributions calculated in Ref. \[34\] for the mean field.

2. Two-nucleon correlations

Around 25% of the time, a nucleon in a heavy nucleus is in a two-nucleon short range correlation\[4\]. A short range correlation occurs when two nucleons are separated by a short distance on the order of 1 fm and consequently have a relative momentum larger than \( k_F \). The analysis of recent experiments indicates\[12\] that in the momentum range from \( k_F \) to \(~ 650 \text{ MeV} \), the nucleus is well-described by two-nucleon SRCs with negligible center of mass motion. Due to the short distance between the nucleons, the dynamics of the nucleons are primarily influenced by their mutual interaction rather than by the mean field. The most immediate consequence of two-nucleon SRCs is a large tail in the momentum distribution above the Fermi momentum, which is not reproduced by mean field models\[14\].

Within the last decade, there has been considerable experimental and theoretical effort put into studying two-nucleon SRCs and their properties. There are a fair number of triple-coincidence experiments demonstrating their existence\[10, 11, 16\] and demonstrating that
two-nucleon SRCs primarily form between a proton and a neutron[9–11]. The fact that most two-nucleon SRCs are predominantly \( pn \) pairs has an important implication for momentum distributions (and thus for the 2N light cone density): for neutron rich nuclei, a given proton is more likely to be in a short range correlation than a given neutron[35–37]. In fact, it means \( \rho_{2}^{(p)}(\alpha, p_{T}) > \rho_{2}^{(n)}(\alpha, p_{T}) \), and consequently (for the sum rules Eqs. (23,24) to be satisfied), \( a_{1}^{(p)} < a_{1}^{(n)} \).

Another important aspect of two-nucleon SRCs is the universality of their behavior. The high-momentum tail that they introduce to the momentum distribution is universal in its form among nuclei[36, 38]. Most significantly, because of the dominance of spin-one, isosinglet SRCs, the high-momentum tail behaves like a scaled version of the tail in the deuteron momentum distribution.

Based on the arguments discussed above, a model of the two-nucleon SRC contribution to the nuclear light cone density matrix should incorporate two main properties: first, the universality of the form of the high momentum tail, and second, that the SRC density matrix for the proton and the neutron should be inversely proportional to its respective relative fraction. For the first property we use the light cone approximation of Ref. [1, 39], which uses the requirement of rotational invariance to relate the light cone deuteron wave function to the non-relativistic wave function through the effective \( pn \) relative momentum:

\[
k = \sqrt{\frac{m^2(\alpha - 1)^2 + p_{T}^2}{\alpha(2 - \alpha)}}.
\]

(28)

For the second property, we use the model of Ref. [40], in which the high momentum distribution of nuclei is inversely proportional to the relative fraction to the nucleon type. The model additionally includes a scaling factor \( a_2(A) \), which is extracted from SRC studies in inclusive \( eA \) processes.

Combining these two properties produces the following for the two-nucleon SRC density matrix:

\[
\rho_{2}^{(N)}(\alpha, p_{T}) = \frac{a_2(A) |\psi_d(k)|^2}{2\chi_N \alpha(2 - \alpha)} \Theta(k - k_F),
\]

(29)

where the extra factor of \( \alpha(2 - \alpha) \) is from the phase space of the struck and spectator nucleons. The details of the derivation of Eq. (29), can be found in Appendix B 1. Here, \( \psi_d(k) \) is the relativistic, light-cone deuteron wave function, related to the non-relativistic wave function by

\[
\psi_d(k) = \sqrt{m^2 + k^2}\psi_{NR}(k),
\]

(30) and \( \chi_N \) is the relative abundance of nucleons of type \( N \) in the nucleus, i.e. \( \chi_p = \frac{Z}{A} \) and \( \chi_n = \frac{A-Z}{A} \). The factor of \( \chi_N \) is present in Eq. (29) results in the equality

\[
\sum_p \rho_{2}^{(p)}(\alpha, p_{T}) = \sum_n \rho_{2}^{(n)}(\alpha, p_{T}).
\]

This relation is confirmed by variational Monte-Carlo calculations for light nuclei up to \( A = 12 \), and for medium to heavy nuclei it was confirmed using a correlated basis calculation of the nuclear momentum distributions in a non-relativistic approach[41].

In the present model we neglect by center of mass motion of the SRC. As a result, the two-nucleon SRC to the nuclear light cone density matrix sets in as soon as \( k > k_F \), which is signified by the step function \( \Theta(k - k_F) \) in Eq. (29).
The remaining factors in Eq. (29) need to be accounted for. $a_2(A)$ is a scaling factor that describes how big the high-momentum tail of the nucleus is relative to the high-momentum tail of the deuteron. It is determined experimentally by examining ratios of quasielastic inclusive cross sections of $A(e,e'X)$ and $^2\text{H}(e,e'X)$ reactions:

$$a_2(A) = \frac{2\sigma_{eA}}{A\sigma_{ed}} : \quad x \gtrsim 1.4,$$

where a roughly flat plateau in the ratio of the quasielastic cross sections is known to occur\cite{7,8,13,15}. The values of $a_2(A)$ for various nuclei are well-constrained by experiment\cite{8} and it is equal to roughly 5.6 for heavy nuclei such as iron-56. In this work, we will also use $a_2(^{208}\text{Pb}) = 5.6$, as no experimental data for this quantity exist for lead.

For numerical estimates, we use the Paris potential for parameterizing the non-relativistic deuteron wave function\cite{42} in Eq. (30).

3. Three-nucleon correlations

When the light cone momentum fraction $\alpha > 2$, one expects the nuclear light cone density matrix to be dominated by three-nucleon short range correlations. As with two-nucleons SRCs, we consider a three-nucleon SRC to occur when the center of mass momentum is small, and when the struck nucleon has large ($> k_F$) momentum relative to the other two nucleons. Such a configuration can naturally generate a light cone momentum fraction exceeding 2.

There are two distinct mechanisms that could generate three-nucleon correlations. The first is through a sequence of two-nucleon interactions, and the second an irreducible three-nucleon interaction. The second contributes primarily to the large removal energy part of the nuclear spectral function, and since the light cone momentum distribution is an integral over removal energies, one expects it to be dominated by the first mechanism. Therefore, we develop a model of three-nucleon SRCs where the struck nucleon obtains its momentum from a sequence of two-nucleon interactions. This model is based on the collinear approximation of Ref. [3], where the three nucleons in the SRC are moving collinearly prior to the SRC generation mechanism.

In the present work, we develop the model further using the recently observed dominance of $pn$ pairwise correlations. This allows us to express the three-nucleon SRC as occurring through a sequence of $pn$ interactions, and thus the 3N SRC part of the nuclear light cone density matrix as a convolution of two 2N SRC light cone momentum distributions. A full derivation of the three-nucleon SRC density matrix can be found in Appendix B 2, with the final result:

$$\rho_3(\alpha, P_T) = \{a_2(A)\}^2 \frac{1}{\alpha} \int \frac{d\alpha_3 d^2P_3 T}{\alpha_3(3 - \alpha_3)} \left\{ \frac{3 - \alpha_3}{2(2 - \alpha_3)} \right\}^2 \left| \psi_d(k_{12}) \right|^2 \Theta(k_{12} - k_F) \left| \psi_d(k_{23}) \right|^2 \Theta(k_{23} - k_F). \quad (31)$$

Here, $k_{12}$ and $k_{13}$ are relative light cone momenta between pairs of nucleons in the three-nucleon SRC. Since the three-nucleon SRC is effected through a sequence of two short-range $pn$ interactions, the three-nucleon SRC can only occur if the conditions under which the individual $pn$ interactions would occur are satisfied. For this reason, the threshold for
three-nucleon SRCs to appear is that the relative light cone momenta of the pairs should each satisfy the threshold condition for which short range two-nucleon interactions occur, namely they should both be above the Fermi momentum $k_F$. It is worth noting that $k_{12}$ and $k_{23}$ are both functions of the light cone momentum fraction $\alpha_3$ of a spectator nucleon, which is integrated over. (c.f. Appendix B2 for details.) Because of this, despite the jump discontinuity in the step function, the discontinuity is smeared out and $\rho_3(\alpha, \mathbf{p}_T)$ is itself a smooth function. The form of $\rho_3(\alpha, \mathbf{p}_T)$ is given in Fig. 3.

The factor $\{a_2(A)\}^2$ appearing in Eq. (31) is a consequence of the fact that three-nucleon SRCs arise from a sequence of short-range $pn$ interactions. The probability of having a three-nucleon correlation should be proportional to the square of the probability of having a two-nucleon correlation. We have taken them to be equal, i.e. for the proportionality factor to be 1. In principle, the proportionality factor may be less than 1, owing to restrictions such as selection rules. However, in the present work, we will assume the maximal reasonable coefficient, which is $\{a_2(A)\}^2$.

Eq. (31) may also be written in terms of a scaling factor $a_3(A)$, similar to $a_2(A)$ the scaling factor for two-nucleon SRCs. If it is assumed that three-nucleon SRCs are universal in their behavior, the parameter $a_3(A)$ would be possible to extract in a similar way to $a_2(A)$ in the case of 2N SRCs. In particular, as is the case for $a_2(A)$, it should be possible to extract $a_3(A)$ through the ratio of quasielastic cross sections for $eA$ and $e^3\text{He}$ processes at high $x > 2$, viz.

$$a_3(A) = \frac{3\sigma_{eA}}{A\sigma_{e^3\text{He}}} : x \gtrsim 2.4,$$

provided that another plateau is actually observed in this ratio. In principle, the observation of $a_3(A)$ may provide an experimental test of the three-nucleon SRC model presented in this work. By definition, $a_3(3\text{He}) = 1$, and for Eq. (31) and Eq. (32) to be consistent would require that $a_3(A) \propto \left(\frac{a_2(A)}{a_2(3\text{He})}\right)^2$, where we take the proportionality constant to be 1 in this
FIG. 4: (Color online.) The nuclear light cone density matrix for $^{208}$Pb.

work. Consequently, we have

$$\rho_3(\alpha, p_T) = \left\{a_2(^3\text{He})\right\}^2 \frac{a_3(A)}{\alpha} \int \frac{d\alpha_3 d^2P_T^3}{\alpha_3(3 - \alpha - \alpha_3)} \left\{ \frac{3 - \alpha_3}{2(2 - \alpha_3)} \right\}^2 \frac{|\psi_d(k_{12})|^2 \Theta(k_{12} - k_F)}{|\psi_d(k_{23})|^2 \Theta(k_{23} - k_F)}. \quad (33)$$

The experimental status of a three-nucleon scaling plateau is, however, ambiguous. While observation of an $x > 2$ plateau was reported in Ref. [8], a later experiment at similar kinematics and better resolution did not observe the plateau[13]. The discrepancy may be due to a resolution issue in the first experiment[43]. It is not expected that a plateau should be seen at the kinematics probed in these experiments, since the momentum transfer was too low to probe three-nucleon SRCs and thus to observe the scaling plateau in the ratio of Eq. (32)[44].

In lieu of experimental data for $a_3(A)$, we calculate $\rho_3(\alpha, p_T)$ in this work using Eq. (31), which requires knowing only the experimentally well-known quantity $a_2(A)$. Fig. 3 contains a parametrization of just the 3N part of the nuclear light cone density calculated using Eq. (31) with $a_2(A) = 5.6$. In Fig. 4, we present the nuclear light cone density matrix for lead. In Fig. 4a, we compare the light cone density using Eq. (25), both considering only the mean field, and considering two-nucleon and three-nucleon correlations in addition. Despite most of the three-nucleon light cone density being concentrated at $\alpha < 2$, it makes a negligible contribution on top of the mean field and 2N density until around $\alpha \gtrsim 1.7$.

In Fig. 4b, we have compared our model for the nuclear light cone density to a model explicated in Chapter 5 of Ref. [3]—in particular, in its Eq. (5.11)—which we call the FS81 model. This model was constructed to fit $p + A \rightarrow p + X$ data with 400 GeV protons[45]. The model implicitly considers four-nucleon, five-nucleon, etc. contributions to exist in the nuclear light cone density matrix, and accordingly is expected to exceed our model, as can be seen in Fig. 4b.
B. Medium modifications

If a nucleon within the nucleus were identical to a nucleon in free space, then with both the light cone density matrix described in Sec. III A and an empirical parametrization of the proton PDF, we could now use Eq. (20) to obtain the nuclear PDF and proceed to calculate the dijet cross section. However, the nuclear medium is strongly interacting, and each nucleon has a high probability (20-30%) of being in a short-range correlation where the nucleons themselves overlap. As a result, it should be expected that the parton distributions for nucleons immersed in the nuclear medium should be modified in some way.

Modification of nucleons within the nuclear medium was first observed by the European Muon Collaboration\[46\] in measurements of the ratio of cross sections for deeply inelastic muon scattering for nuclear and deuteron targets, i.e.

\[
\mathcal{R}_{EMC}(x,Q^2) = \frac{2\sigma_{eA}}{A\sigma_{ed}} \approx \frac{2F_2^{(A)}(x,Q^2)}{AF_2^{(d)}(x,Q^2)}.
\] (34)

The small angle of the scattering that was measured allowed the cross sections to be related to the structure functions. It was originally expected that, except for smearing from Fermi motion at high \(x\), the structure functions should just be the structure functions for free nucleons, and thus the ratio should be 1. However, there is a dip below 1 in the range 0.3 < \(x\) < 0.7 that cannot be reproduced by nucleonic motion alone\[47, 48\]. The effect is commonly referred to as the EMC effect.

This dip in the ratio \(\mathcal{R}_{EMC}\) is ubiquitous throughout nuclei, and is roughly proportional to the nuclear mass number\[49\]. Additionally, it is proportional to the local density of the nucleus considered\[50\], suggesting that it is an effect of the nuclear medium. Since the structure function \(F_2(x,Q^2)\) is directly related to the parton distribution functions, the EMC effect is concluded to be due to a modification of the PDF of a bound nucleon in the nuclear medium. Even though there is a consensus about this, there is a plethora of theoretical models based on different aspects of nuclear medium effects (for reviews on the EMC effect see \[51, 52\]). Despite differences in the approaches taken by these models, one thing they have in common is that in order to satisfy locality, the size of the effect should be proportional to the virtuality of the bound nucleon. This indicates that more modification should occur in the high momentum part of the nuclear wave function.

Since our goal is to study the effects of SRCs on the reaction (1), and since SRCs dominate the high momentum part of the nuclear wave function, it will be necessary to account for the medium modifications that produce the EMC effect when studying this reaction. In particular, the medium modification effects will be prominent in the nuclear parton distribution \(f_{i/A}(x,Q^2)\) present in Eq. (17). Numerical estimates of the EMC effect will be based on the color screening model\[1, 53\], which satisfactorily describes the phenomenology of EMC effects observed in inclusive DIS reactions.

Before we proceed to describe medium modifications, we present estimates of the ratio given in Eq. (34) in the absence of medium modifications. The quantity \(F_2^{(A)}(x,Q^2)\) is calculated using the convolution formula Eq. (18), and the formalism of Sec. III A is applied for computing the light cone density matrix. These estimates are presented in Fig. 5, where a significant effect from SRCs can be seen for \(x > 0.6\). The discrepancy between theory and experiment becomes worse in this region with SRCs present, requiring that SRCs are more strongly modified by the nuclear medium than nucleons in the mean field. This is consistent
FIG. 5: (Color online.) EMC ratios at $Q^2 = 10$ GeV$^2$, without medium modifications. Data are from SLAC experiments: Stars (blue) are from [49]; circles (red) are from [54]. The nucleon structure functions are computed with the Bodek-Ritchie parametrization [55–57].

with our expectation that modification should increase with nucleon virtuality, since SRCs occur predominantly at especially large momenta.

1. Correctly defining $x$

Up until now, data from DIS experiments have been plotted against the Bjorken scaling variable $\hat{x} = \frac{Q^2}{2m_pv}$, where $m_p$ is the mass of the proton. However, the Bjorken scaling variable that enters into the dynamics of QCD, including the convolution formulas of Eqs. (18,20), is instead $x = \frac{AQ^2}{2MA_v}$. Thus, it is as an argument of $x$ rather than $\hat{x}$ that the ratio of Eq. (34) should be presented in order to see dynamical QCD effects. By contrast, presenting data as a function of $\hat{x}$ instead of $x$ artificially shifts the arguments of the nuclear and deuteron structure functions in Eq. (34) by different amounts, since $x_A = \frac{Am_p}{MA_v} \hat{x}$ and $x_d = \frac{2m_p}{Md} \hat{x}$. Since the deuteron is a loosely bound system with little binding energy, the arguments for the nuclear and deuteron structure functions are shifted by substantially different amounts when the nucleus has a high binding energy.

It has been observed [30, 58] that presenting structure function ratios as an argument of $\hat{x}$ rather than $x$ introduces an artificial dip in the ratio for $x \gtrsim 0.5$. In Fig. 6, we present theoretical calculations for $R_{EMC}(x)$ (as a function of the dynamical QCD scaling variable $x$) and $R_{EMC}(\hat{x})$ (as a function of the kinematical Bjorken scaling variable $\hat{x}$) for both $^{56}$Fe and $^{208}$Pb. In these plots, the use of $\hat{x}$ as a variable produces an artificial dip, which partially explains the EMC effect up to $x \sim 0.5$. However, there continues to be a discrepancy between theory and data, which requires medium modification to explain.

2. Color screening model of the EMC effect

Any nucleon can be expected to spend some of its time in a point-like configuration (PLC), in which its constituent quarks are compressed into a tiny volume compared to the average
nucleon radius. A PLC is largely invisible to the color force, in analogy to neutrally charged atoms in a gas whose van der Waals forces become weaker if the atoms are compressed. The nucleon, after all, has an overall neutral color charge, so any color interaction between nucleons owes to higher moments (dipole, quadrupole, etc.), which decrease with distance between the color-charged constituents. Moreover, it can be shown by the renormalizability of QCD that meson exchange between nucleons, one of which is in a PLC, is suppressed\(^{53}\).

Since nucleons in an average-sized configuration (ASC) and a PLC will interact differently, the probability that the nucleon can be found in either configuration should be modified by the immersion of a nucleon in the nuclear medium. In particular, PLCs are expected to be suppressed compared to ASCs since the bound nucleon will assume a configuration that maximizes the binding energy and brings the nucleus to a lower-energy ground state. The change in probability can be estimated using non-relativistic perturbation theory, as has been done in Refs. [1, 53]. What is found is that the light cone density matrix should be modified by a factor \(\delta_A(k^2, x)\), which depends on the nucleon momentum (or virtuality) as

\[
\delta_A(k^2, x \geq 0.6) = \frac{1}{(1 + z)^2}
\]

\[
z = \frac{k^2}{m_p} + 2\epsilon_A \Delta E_A.
\]

In analogy with electric charge screening, this is called the color screening model of the EMC effect. We shall use it as an example of accounting for medium modifications when calculating dijet cross sections. For small \(x \lesssim 0.45\), there should be little or no color screening, and the region between 0.45 < \(x\) < 0.6 is smoothly interpolated with a linear damping factor:

\[
\delta_A(k^2, 0.45 < x < 0.6) = 1 + \frac{0.45 + x}{0.15} \left\{ \delta_A(k^2, x \geq 0.6) - 1 \right\}.
\]
Since the suppression factor depends on the total nucleon momentum rather than just the light cone momentum fraction $\alpha$, it is necessary to use the three-dimensional light cone density $\rho(\alpha, \mathbf{p}_T)$ when applying the color screening model. Moreover, since the suppression of PLCs depends on inter-nucleon dynamics, it is expected not just that the parameters of $\delta_A(k^2, x)$ should vary with the nucleus considered, but with whether the nucleons are moving in the mean field or are in a SRC. For a nucleon in the mean field of a heavy nucleus, we expect the excitation energy $\Delta E_A$ to be in the range $300 - 500$ MeV, namely between the excitation energies of a $\Delta$ and an $N^*$ resonance. The best bit to data appears to be with the $N^*$ excitation energy $\Delta E_A \approx 500$ MeV. However, for the deuteron, as well as for a nucleon in a 2N SRC, we expect the lowest excited state to be a $\Delta\Delta$ configuration, giving $\Delta E_d \approx 600$ MeV. We take $\epsilon_A$ to be the binding energy per nucleon of the nucleus under consideration when in the mean field, and estimate $\epsilon_A = \epsilon_d \approx 2.2$ MeV for a nucleon in a two-nucleon SRC, and since a three-nucleon SRC consists of a sequence of two two-nucleon SRCs. We expect the characteristic excitation energy for a nucleon in a 3N correlation to be around $600$ MeV, since the 3N correlation is effected through a sequence of 2N correlations, and likewise we take $\epsilon_A = 2\epsilon_d \approx 4.4$ MeV for a nucleon in a three-nucleon SRC. As can be seen in Fig. 7, this results in a fairly good account of the EMC effect.

The color screening model is now applied to PDFs rather than structure functions. We assume that the suppression factor $\delta_A(k^2, x)$ that appears in the convolution formula for $F_2^A$ similarly appears in the formula for nuclear PDFs, namely

$$f_{i/A}(x, Q^2) = \sum_N \int_x^A \frac{d\alpha}{\alpha} d^2p_T \rho_{N/A}(\alpha, \mathbf{p}_T) f_{i/N} \left( \frac{x}{\alpha}, Q^2 \right) \delta_A(k^2, x).$$

(38)

The form of $k^2$, as well as the values of the parameters in $\delta_A(k^2, x)$ depend on whether the nucleon is moving in the mean field or in an SRC, as before. There is one ambiguity in applying the color screening model to PDFs, however: the correct modification appears in $F_2^{(A)}(x, Q^2)$ by modifying quarks and anti-quarks according to Eq. (38); the modification of gluons must still be accounted for. Since medium modifications occur predominantly at
large $x$, which is the region where valence quarks predominate, gluons in this region should originate as radiation from quarks, and should therefore inherit medium modifications from the quarks. For this reason, we apply Eq. (38) to gluons as well as quarks and anti-quarks. The result of applying color screening to the nuclear PDF is demonstrated in Fig. 8.

3. Evolution of medium modifications

The EMC effect has so far been observed in a restricted range of $Q^2$, around $Q^2 \sim 10 \text{ GeV}^2$. One would expect that medium modification effects will be present at larger $Q^2$ values as well, such as at the $Q^2$ values characteristic of dijet production at the LHC, which can reach tens of thousands of GeV$^2$. Therefore, it is necessary to predict how the EMC effect will manifest at $Q^2$ well above the domain where the effect has been observed. Assuming that no higher twist effects are involved in the EMC effect, the modified nuclear PDFs at high and low $Q^2$ should be related by QCD evolution equations. For this reason, we evolve the nuclear PDFs obtained through medium modification at $Q^2 = 10 \text{ GeV}^2$ using DGLAP evolution\textsuperscript{[60–62]}. The standard DGLAP formula is modified to account for the possibility of $x > 1$ inherent in dealing with nuclei; in particular, we have

$$\frac{\partial f_{i/A}(x, Q^2)}{\partial \log(Q^2)} = \frac{\alpha_s(Q^2)}{2\pi} \sum_{ij} \int_x^A \frac{dy}{y} P_{ij} \left( \frac{x}{y} \right) f_{j/A}(y, Q^2). \tag{39}$$

Here, $P_{ij}$ are the Alterelli-Parisi splitting functions, found at leading order in Ref. [62]. In our calculation we solved Eq. (39) numerically, using the nuclear PDFs obtained in the previous sections as input.

The results of this evolution are shown in Fig. 9, which compares the effects of evolution on nuclear PDFs with and without medium modification accounted for. In both cases, the
FIG. 9: (Color online.) Ratio of $u + d$ quark distribution for $^{208}$Pb to that of the proton. Unmodified lines are at $Q^2 = 10$ GeV$^2$, and evolved at $Q^2 = 10000$ GeV$^2$. The $Q^2 = 10$GeV$^2$ PDFs are computed using the CT10 parametrization\cite{59}, as well as the convolution formula Eq. (20).

ratio of the $u + d$ quark distribution for the nucleus to the proton is used, since the effects of evolution are most easily seen in ratios, and since adding the $u$ and $d$ distributions removes peculiarities of isospin asymmetry.

IV. HARD SUBPROCESSES

We have developed a model of nuclear PDFs that accounts for both short range correlations and effects from medium modification. In addition to a model for PDFs, however, we also need to calculate the squares of the invariant matrix elements $|\mathcal{M}_{ij}|^2$ for the hard partonic subprocesses that contribute to Eq. (17). We consider two partons, labeled $i$ and $j$, to exist in the initial state of the reaction—one from each of the proton and the heavy nucleus participating in the reaction (1). At leading order, a dijet is associated with two partons in the final state, which we label $k$ and $l$. Each parton can in principle be a quark, an anti-quark, or a gluon, although what partonic subprocesses are possible is limited by quark flavor conservation. A list of the possible two-parton to two-parton subprocesses and their squared matrix elements at leading order is given in Table I.

At next-to-leading order, a dijet can be associated either with two or with three partons in the final state, provided two of the three partons have small differences in both their rapidities and their azimuthal angles. This is possible because jets are identified as sprays of hadrons that fall within a cone of radius $R$ in $\eta$-$\phi$ space, i.e. for which $\Delta \eta^2 + \Delta \phi^2 < R^2$. If $\Delta \eta$ and $\Delta \phi$ between two partons are both small, the sprays of hadrons produced by each will overlap and will be detected as a single jet\cite{64}. In addition to the concern of overlapping sprays of hadrons, it is theoretically necessary at next-to-leading order to account for final states with three partons, as the next-to-leading order matrix elements for two-parton to two-parton processes contain infrared divergences that are only canceled when two-parton to three-parton processes are also accounted for\cite{65}.
TABLE I: Squared matrix element for hard partonic processes at leading order, summed over final spins and averaged over initial spins, and divided by the square of the coupling strength $g$. $j \neq k$ in subprocesses where both indices are present. The table is from Ref. [63].

| Subprocess | $|\mathcal{M}|^2 / g^2$ |
|------------|--------------------------|
| $q_j + q_k \rightarrow q_j + q_k$ | $\frac{4}{9} \frac{s^2+u^2}{t^2}$ |
| $q_j + q_j \rightarrow q_j + q_j$ | $\frac{4}{9} \left( \frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} \right) - \frac{8}{27} \frac{s^2}{ut}$ |
| $q_j + \bar{q}_j \rightarrow q_k + \bar{q}_k$ | $\frac{4}{9} \frac{t^2+u^2}{s^2}$ |
| $q_j + \bar{q}_j \rightarrow q_j + \bar{q}_j$ | $\frac{4}{9} \left( \frac{s^2+u^2}{t^2} + \frac{t^2+u^2}{s^2} \right) - \frac{8}{27} \frac{u^2}{ut}$ |
| $q_j + g \rightarrow q_j + g$ | $\frac{32}{27} \frac{u^2+t^2}{ut} - \frac{8}{3} \frac{u^2+t^2}{s^2}$ |
| $g + g \rightarrow q_j + q_j$ | $\frac{1}{6} \frac{u^2+t^2}{ut} - \frac{3}{8} \frac{u^2+t^2}{s^2}$ |
| $q_j + g \rightarrow q_j + g$ | $-\frac{4}{9} \frac{u^2+s^2}{us} + \frac{8}{3} \frac{u^2+s^2}{t^2}$ |
| $g + g \rightarrow g + g$ | $\frac{9}{2} \left( 3 - \frac{4u}{s} - \frac{4u}{t} - \frac{4t}{u} \right)$ |

Since, at next-to-leading order, the relevant dijet processes include final states where two partons correspond to a single jet, jets will not in general be massless at this order. However, the kinematical formula Eq. (6) for determining $x_A$ in terms of jet kinematics requires that the two jets are associated with two on-shell, massless partons in the final state. This means we would like to consider two-parton to two-parton subprocesses only, which requires treating the problem at leading order.

The question is whether the leading order in pQCD is sufficient to calculate inclusive dijet production in the kinematical regime we are investigating. To conclude that leading order is sufficient, we perform a leading order calculation of the reaction (1) for the case of $A = 1$, i.e. for dijet production from proton-proton collisions, for which a fair amount of experimental data in the kinematical regime of interest exist. We compare the calculation within the leading order to experimental data from Ref. [66]. The data, however, are given in terms of a two-fold differential cross section $d^2\sigma/dM_{jj}d|y^*|$. We will briefly describe the variables used in Ref. [66].

At next-to-leading order, a jet may be massive, which requires distinguishing between rapidity ($y$) and pseudo-rapidity ($\eta$). At leading order, with massless jets, $\eta = y$, so for our calculations we will henceforth denote rapidity as $\eta$. With jet rapidities $\eta_3$ and $\eta_4$, $\bar{\eta}$ is the rapidity of the dijet as a whole, and $\eta^*$ is the rapidity of an individual jet in the dijet...
FIG. 10: (Color online) Comparison of leading-order (LO) calculation of the two-fold differential cross section using Eq. (43) to experimental data from Ref. [66]. Proton PDF used is CT10[59]

center-of-mass frame, where

$$\bar{\eta} = \frac{\eta_3 + \eta_4}{2}, \quad (40)$$

$$\eta^* = \frac{\eta_3 - \eta_4}{2}, \quad (41)$$

Note that the Jacobian for the transformation from ($\eta_3, \eta_4$) to ($\bar{\eta}, \eta^*$) gives us $d\eta_3 d\eta_4 = 2 d\bar{\eta} d\eta^*$. In addition to the rapidity in the center-of-mass frame $\eta^*$, Ref. [66] uses the dijet mass $M_{JJ}$, which is given by

$$M_{JJ}^2 = (p_3 + p_4)^2 = 4p_T^2 \cosh^2(\eta^*), \quad (42)$$

or $M_{JJ} = 2p_T \cosh(\eta^*)$. We note that the overall Jacobian is given by $dp_T^2 d\eta_3 d\eta_4 = \frac{2p_T}{\cosh(\eta^*)} dM_{JJ} d\bar{\eta} d\eta^*$, and thus the two-fold differential cross section given in Ref. [66] can be written as

$$\frac{2d^2\sigma}{dM_{JJ} d\eta^*} = \frac{4p_T}{\cosh(\eta^*)} \int d\bar{\eta} \frac{d^2\sigma}{d\eta_3 d\eta_4 dp_T^2}. \quad (43)$$

The three-fold differential cross section is evaluated using Eq. (17), but with $f_A = f_p$ and $Z = A = 1$ to compare with proton-proton collision data. The data of Ref. [66] are given at $\sqrt{s_{NN}} = 7$ TeV, and are given in bins of $|\eta^*|$. Since bins are of $|\eta^*|$ rather than $\eta^*$, it is necessary to add the cross sections at $\eta^* = |\eta^*|$ and $\eta^* = -|\eta^*|$ together in the theoretical calculation. We have performed the calculation at leading order in $\alpha_s$, and present the calculation against experimental data for two $|\eta^*|$ bins in Fig. 10. As can be seen in the figure, the leading order is adequate for describing the data.

V. NUMERICAL ESTIMATES

In the previous sections, we have supplied all the ingredients necessary to calculate the three-fold differential cross section of Eq. (17) for the dijet production reaction (1). It is
our intention to use these ingredients to determine the sensitivity of the dijet production reaction (1) to the presence of short range correlations in the nucleus, as well as to medium modification effects. Additionally, we will look at nuclear partons with $x_A > 1$ in particular. Since $x_A$ can be related to measurable jet kinematics via Eq. (6), it is necessary to elaborate on the kinematics we will consider. Once the kinematics considered are explicated, we will present numerical estimates of the three-fold differential cross section of Eq. (17) at these kinematics, and additionally we will estimate the partially and fully integrated cross section to demonstrate that events corresponding to $x_A > 1$ are frequent enough to be measurable at the LHC.

A. Kinematics considered

Since it is our goal to probe $x_A > 1$, we consider kinematics that maximize $x_A$. According to Eq. (6), three variables are related to $x_A$. The jet rapidities $\eta_3$ and $\eta_4$ should be as “small” (as highly negative) as possible in order to maximize $x_A$. Additionally, the transverse jet momentum $p_T$ should be as large as possible. On the other hand, non-central jets are more difficult to detect and identify, and the jet cross section for large $p_T$ is known to drop rapidly. It is necessary to find a balance between kinematics that will increase $x_A$ and those that will produce an appreciable yield.

Since central jets are easier to detect, and since we are interested in superfast partons in the nucleus, we consider jets that are central from the proton, namely $-2.5 < \eta_3 < 2.5$, and we consider jets that are moving forward in the direction of the nucleus beam, namely $3 < -\eta_4 < 5$. We also consider a $p_T$ range from 40 GeV to 200 GeV.

B. Three-fold differential cross section

Eq. (17) can be used to obtain numerical estimates of the three-fold differential cross section $d^3\sigma/d\eta_3 d\eta_4 dp_T^2$ for the dijet production reaction (1). Numerical estimates are presented in Fig. 11, where estimates with and without short range correlations are presented for comparison, and likewise estimates with and without medium modification are presented. As can be seen in Figs. 11a and 11b, the presence of SRCs increases the cross section for superfast ($x_A > 1$) quarks considerably. However, medium modification decreases the cross section, especially at large $x$ where highly-modified SRCs dominate the cross section, as can be seen in Figs. 11c and 11d.

In addition to comparisons of the effects of short range correlations and medium modifications within the model for the nuclear light cone density explicated in the present work, we compare in Fig. 12 the prediction for the three-fold differential cross section using our model to the prediction made using the FS81 model[3]. The FS81 model uses exponential decay to model the large $\alpha$ part of the nuclear light cone density, and likewise the nuclear PDF calculated using this model exhibits exponential decay for large $x_A$. It is for this reason the curve for the FS81 model appears as a straight line in the log plot of Fig. 12. Qualitatively, the model of this work can be distinguished experimentally from the FS81 model by determining whether the three-fold differential cross section of Eq. (17) falls exponentially or faster-than-exponentially.
FIG. 11: (Color online.) Three-fold differential cross section for $p+^{208}$Pb→dijet + $X$ according to Eq. (17). $\eta_4$ is constrained to the range $(-5, -3)$, and $\eta_3 = 0$. Proton PDF used is CT10[59].

C. Partially integrated cross section

We will continue to see a large sensitivity to the presence of SRCs if we integrate over the $\eta_3$ and $\eta_4$ ranges considered in order to find the one-fold differential cross section $\frac{d\sigma}{dp_T}$. However, simply integrating over these ranges will count both $x_A > 1$ and $x_A < 1$ events for a fixed $p_T$ together, which is potentially a problem if we want to look at $x_A > 1$ events in particular. However, since it is possible to determine $x_A$ from jet kinematics via Eq. (6), we can simply define two one-fold differential cross sections, one of which only counts $x_A > 1$ events. Namely, we define

$$\frac{d\sigma}{dp_T} = \int_{-2.5}^{2.5} d\eta_3 \int_{-5}^{-3} d\eta_4 \frac{2p_T d^3\sigma}{d\eta_3 d\eta_4 dp_T^2}.$$  (44)

$$\frac{d\sigma(x_A > 1)}{dp_T} = \int_{-2.5}^{2.5} d\eta_3 \int_{-5}^{-3} d\eta_4 \frac{2p_T d^3\sigma}{d\eta_3 d\eta_4 dp_T^2} \Theta(x_A - 1).$$  (45)

The differential cross section of Eq. (44) is simply the differential cross section for the full kinematical range considered, whereas the differential cross section in Eq. (45) throws out
FIG. 12: No medium modifications. Calculation within our model considers 2N and 3N SRCs to contribute.

(a) Without medium modifications.  
(b) With and without color screening.

FIG. 13: Differential cross section for $p + ^{208}$Pb → dijet + X, according to Eqs. (44,45). Computed using CT10 for proton PDF\[59\].

events for which $x_A < 1$. By comparing these quantities, it can be seen how often these kinematics correspond to superfast partons with $x_A > 1$.

In Fig. 13 we present estimates of the differential cross sections of Eqs. (44,45). Fig. 13a compares the differential cross section, in the absence of medium modifications, without and with short-range correlations accounted for. Most dijets with transverse momentum up to about 100 GeV can be attributed to nucleons in the mean field. In contrast, dijets with high $p_T \gtrsim 150$ GeV can be attributed predominantly to short range correlations. In addition, selecting only dijets with kinematics corresponding to an initial state nuclear parton with $x_A > 1$ produces a differential cross section that is at least an order of magnitude larger when SRCs are considered than for the mean field by itself. This confirms that $x_A > 1$ events can be attributed predominantly to the presence of SRCs in the nucleus, and also
indicates that finding a large yield for \( x_A > 1 \) events can demonstrate the presence of SRCs. In the absence of medium modification effects, the value of the cross section could also be used to probe their strength.

On the other hand, since SRCs are highly modified in the nuclear medium, \( x_A > 1 \) events in particular are suppressed by the EMC effect, as can be seen in Fig. 13b. This means the value of the one-fold differential cross section for \( x_A > 1 \) dijet events suffers two effects that compete with one-another: it is increased by the presence of SRCs, but it is at the same time suppressed by medium modifications, which most strongly affect SRCs. It would be difficult to disentangle these competing effects, so it may not be possible to find precision values for \( a_2(A) \) and \( a_3(A) \) using the differential cross section of Eq. (45).

In addition to one-fold differential cross sections, we also present estimates for the total integrated cross section in order to show that \( x_A > 1 \) events should produce a measurable yield. Since jet identification at low \( p_T \) is difficult, we consider the total cross section integrated from a minimum \( p_T \) value, and in particular consider several minimum values of \( p_T \). In Table II we present numerical estimates for the total cross section with different minimum \( p_T \) values as indicated in the table. Table II considers the integrated cross section both in the case that all events are counted, and in the case that only \( x_A > 1 \) events are counted. Since larger \( p_T \) correspond to a greater proportion of \( x_A > 1 \) events, we also include in the table estimates of the integrated cross section when \( p_T \) is integrated from a minimum value of 150 GeV.

As can be seen in Table II, medium modifications suppress the \( x_A > 1 \) and high-\( p_T \) integrated cross section just as they suppress the one-fold differential cross section. This is consistent with the expectation that medium modifications affect SRCs especially strongly, since these kinematics are dominated by SRCs. However, despite the suppression by medium modification, the cross section remains large enough to be observable.

### VI. CONCLUSIONS

In this work, we have shown that superfast partons inside a heavy nucleus with \( x > 1 \) are experimentally accessible at the LHC, and can be probed in dijet production reactions from lead-proton collisions. The value of \( x \) can be related to jet kinematics in the LHC rest frame in a straightforward way, so it is possible to select events that correspond to \( x > 1 \). Partons with \( x > 1 \) can predominantly be attributed to the presence of short range correlations between nucleons within the lead nucleus. Medium modifications are necessary to account for because they suppress the cross section appreciably, but the total \( x > 1 \) cross section

| \( p_T,\text{min} \) | \( \sigma_{\text{unmod}} \) | \( \sigma_{\text{clrcn}} \) | \( \sigma_{\text{unmod}} \) | \( \sigma_{\text{clrcn}} \) |
|-----------------|----------------|----------------|----------------|----------------|
| 50 GeV          | 7.8 mb         | 6.9 mb         | 12 nb          | 5.7 nb         |
| 60 GeV          | 2 mb           | 1.7 mb         | 6.5 nb         | 3 nb           |
| 70 GeV          | 0.57 mb        | 0.45 mb        | 3.7 nb         | 1.7 nb         |
| 150 GeV         | 27 pb          | 11 pb          | 26 pb          | 11 pb          |

TABLE II: Estimates of integrated cross sections, at different values of the minimum \( p_T \) considered. Calculations include two- and three-nucleon SRCs.
remains on the order of a few nanobarns, making measurement feasible.

The observability of superfast partons in the nucleus owes itself to the presence of short range correlations. This fact tells us that it is inadequate to treat the nucleus as a collection of free nucleons, moving forward with equal light cone momenta. The internal structure of the nucleus, and the relative momenta of its nucleons, are necessary to account for. Additionally, accounting for the EMC effect using the color screening model resulted in an appreciable suppression of the dijet cross section at moderate values of \( x \). Nuclear effects in general cannot be ignored in collisions involving heavy ions. This fact may be especially relevant to nucleus-nucleus collisions, in which jet quenching is used to study the existence and properties of quark-gluon plasma. The EMC effect by itself will suppress the cross section, while at large \( x \) (i.e. at high \( p_T \) prior to quenching, and high jet rapidity) the presence of SRCs will increase the cross section. These effects may be necessary to account for in order to correctly predict the expected jet yields in the absence of jet quenching, and thus to correctly gauge the magnitude of quenching.

We expect the formalism to have applications beyond dijet production in proton-nucleus collisions. In particular, it will prove necessary to account for multi-nucleon short range correlations to correctly predict the high-\( x \) cross section in inelastic nuclear DIS experiments at high energy. The Electron Ion Collider\[67] in particular will be able to carry out such experiments, where the formalism explicated in this work will prove necessary. In particular, the nuclear PDF at high \( Q^2 \) will need to be obtained through DGLAP evolution, since the phenomenological models that exist to account for nuclear modifications are constructed using low-\( Q^2 \) experimental data.

ACKNOWLEDGMENTS

We are thankful to Drs. Alberto Accardi, Wim Cosyn, Leonid Frankfurt, Shunzo Kumano, and Gerald Miller for many useful discussions. This work is supported by US Department of Energy grants under contracts DE-FG02-01ER-41172 and DE-FG02-93ER40771.

Appendix A: Derivation of factorization formula

The factorization formula is here derived using a Feynman diagram formalism. To be sure, QCD factorization is not rigorously proved here: we merely present a demonstration that, assuming the usual factorization theorem holds, Eq. (13) is correct for the nuclear PDF as we have defined it.

This derivation is performed by treating hadrons as consisting of two parts: a parton that is probed, and the remainder of the hadron, which is a spectator to the probing reaction. Likewise, the nucleus will be treated the same way: it consists of a nucleon \( N \) that is probed, and a spectator consisting of the other \( (A - 1) \) nucleons. Subscripts with the letter \( s \) will be used to signify spectators to the reactions. e.g. \( p_{sp} \) is the four-momentum of the part of the proton left over when \( i \) leaves it, which is a spectator to the parton-level interaction.

We make the assumption that the parton coming from the nucleus was found within a particular nucleon. The wave function of the nucleon itself may be modified by the nuclear medium, however.

For notational simplicity, we will use the symbol \( u \) to signify the part of a particle’s wave function that transforms under a representation of the Lorentz group, even if it is
not a spinor. For instance, \( u \) will be a spinor for fermions, but for vector bosons will be a polarization vector.

With this in mind, the hadron-level Feynman amplitude is given by

\[
i\mathcal{M}_{\text{had.}} = \sum_{ij} \bar{u}_k^{(\lambda_3)}(p_3)\bar{u}_l^{(\lambda_4)}(p_4)i\Gamma_{ij\rightarrow kl}(i\Delta_F^{(i)}(p_1))(i\Delta_F^{(j)}(p_2))\bar{u}_{sp}^{(\lambda_{sp})}(p_{sp})\bar{u}_s^{(\lambda_N)}(p_{sN})
\]

\[\times (i\Gamma_{i/p})(i\Gamma_{j/N})u_p^{(\lambda_p)}(p_p)i\frac{\not{\!p_N} + m_N}{p_N^2 - m_N^2 + i\epsilon}\bar{u}^{(\lambda_A)}(p_A)i\Gamma_{N/A}u_A^{(\lambda_A)}(p_A),\]

where \( \Delta_F^{(i)}(p) \) is the Feynman propagator of a particle of type \( i \) with four-momentum \( p \). For both types of particles, the completeness relation takes the form

\[
\Delta_F^{(i)}(p) = \sum_{\lambda} \frac{\bar{u}_i^{(\lambda)}(p)\bar{u}_i^{(\lambda)}(p)}{p^2 - m_i^2 + i\epsilon},
\]

so applying (approximate) completeness relations to each of the intermediate particles gives

\[
i\mathcal{M}_{\text{had.}} = -i \sum_{ij} \sum_{\text{spins}} \left( \bar{u}_k^{(\lambda_3)}(p_3)\bar{u}_l^{(\lambda_4)}(p_4)i\Gamma_{ij\rightarrow kl}u_i^{(\lambda_1)}(p_1)u_j^{(\lambda_2)}(p_2) \right)
\]

\[\times \left( \frac{\bar{u}_i^{(\lambda_1)}(p_1)\bar{u}_{sp}^{(\lambda_{sp})}(p_{sp})\Gamma_{i/p}u_p^{(\lambda_p)}(p_p)}{p_1^2 + i\epsilon} \right) \left( \frac{\bar{u}_j^{(\lambda_2)}(p_2)\bar{u}_s^{(\lambda_N)}(p_{sN})\Gamma_{j/N}u_A^{(\lambda_A)}(p_A)}{p_2^2 + i\epsilon} \right)
\]

\[\times \left( \frac{\bar{u}_N^{(\lambda_N)}(p_N)\bar{u}_{sA}^{(\lambda_A)}(p_{sA})\Gamma_{N/A}u_A^{(\lambda_A)}(p_A)}{p_N^2 - m_N^2 + i\epsilon} \right).
\]

The terms have been grouped for easy identification. The parton-level Feynman amplitude can be read off easily. There are also two partonic wave functions, for partons \( i \) and \( j \), which can be observed using the definition\[68]\n
\[
\psi^{(\lambda;i;\lambda_N)}_{i/N} = -x_N \frac{\bar{u}_i^{(\lambda_i)}(p_i)\bar{u}_s^{(\lambda_s)}(p_s)\Gamma_{i/N}u^{(\lambda_N)}(p_N)}{p_i^2 + i\epsilon}. \tag{A1}
\]

There is a light cone wave function for the nucleon \( N \) within the nucleus as well. In this work we shall use the definition

\[
\psi^{(\lambda_N;\lambda_A)} = -\alpha \frac{1}{\sqrt{2A(2\pi)^3}} \bar{u}_N^{(\lambda_N)}(p_N)\bar{u}_s^{(\lambda_s)}(p_s)\Gamma_{N/A}u_A^{(\lambda_A)}(p_A), \tag{A2}
\]

a definition which ensures the normalization condition

\[
\sum_{\lambda_N} \int \frac{d\alpha}{\alpha(A - \alpha)} d^2 p_T \left| \psi^{(\lambda_N;\lambda_A)} \right|^2 = 1. \tag{A3}
\]

In light of these identities, the hadronic Feynman amplitude can be written

\[
\mathcal{M}_{\text{had.}} = \sum_{ij} \sum_{\text{spins}} \mathcal{M}_{ij\rightarrow kl} \psi^{(\lambda_1;\lambda_p)}(p_1) \psi^{(\lambda_2;\lambda_N)}(p_2) \sqrt{2A(2\pi)^3} \frac{x_p}{x_N} \frac{\alpha}{\alpha} \psi^{(\lambda_3;\lambda_A)}(p_A).
\]
Due to flavor and spin conservation, the spin and flavor of intermediate states can be identified using the spin and flavor of jets and spectators present in the final state. Consequently, the sums over these things are incoherent, and can be factored out of the square of the amplitude. Now taking the square, and then summing over final states and averaging over initial states, gives making the sum-and-average of the squares

\[
|\mathcal{M}_{\text{had.}}|^2 = \frac{2A(2\pi)^3}{x_p^2 x_N^2 \alpha^2} \sum_N \sum_{ijkl} |\mathcal{M}_{ijkl}|^2 \left| \psi_{i/p}(p_1, p_{sp}; p_p) \right|^2 \left| \psi_{j/N}(p_2, p_{sN}; p_N) \right|^2 \left| \psi_{N/A}(p_N, p_{sA}; p_A) \right|^2,
\]

where spin indices have been suppressed since they are now averaged over. The hadronic and partonic fluxes are related as

\[
\Phi_{\text{had.}} = 4(p_A \cdot p_p) = 2p_A \bar{p}_p^+ \\
\Phi_{\text{part.}} = 4(p_1 \cdot p_2) = 2p_1^+ p_2^+ = 2\frac{\alpha}{A} x_p x_N p_A \bar{p}_p^+ = \frac{\alpha x_p x_N}{A} \Phi_{\text{had.}},
\]

and the phase space elements are related as

\[
dQ_p = (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) \frac{dp_1^+ dp_2^+ d^3p_{ST} d\bar{p}_4^+ d^3p_{AT}}{2p_3^+(2\pi)^3 2p_4^-(2\pi)^3} \\
dQ_h = (2\pi)^4 \delta^{(4)}(p_A + p_p - p_3 - p_4 - p_{sp} - p_{sN} - p_{sA}) \frac{dp_1^+ dp_2^+ d^3p_{ST} d\bar{p}_4^+ d^3p_{AT}}{2p_3^+(2\pi)^3 2p_4^-(2\pi)^3} \\
\times \frac{dp_3^- dp_4^+ d^3p_{sp,T} d\bar{p}_s^N d^3p_{sN,T} d\bar{p}_s^A d^3p_{sA,T}}{2p_3^- (2\pi)^3 2p_s^N (2\pi)^3 2p_s^A (2\pi)^3} \\
= dQ_p \frac{dp_1^+ dp_2^+ d^3p_{sp,T} d\bar{p}_s^N d^3p_{sN,T} d\bar{p}_s^A d^3p_{sA,T}}{2p_3^+(2\pi)^3 2p_4^-(2\pi)^3 2p_s^N (2\pi)^3 2p_s^A (2\pi)^3}.
\]

Thus the overall hadronic cross section element is

\[
d\sigma_{\text{had.}} = \sum_N \sum_{ijkl} \frac{\alpha x_p x_N}{A \Phi_{\text{part.}}} \frac{2A(2\pi)^3}{x_p^2 x_N^2 \alpha^2} |\mathcal{M}_{ijkl}|^2 \left| \psi_{i/p}(p_1, p_{sp}; p_p) \right|^2 \left| \psi_{j/N}(p_2, p_{sN}; p_N) \right|^2 \left| \psi_{N/A}(p_N, p_{sA}; p_A) \right|^2 \\
\times \left| \psi_{j/N}(p_2, p_{sN}; p_N) \right|^2 \left| \psi_{N/A}(p_N, p_{sA}; p_A) \right|^2 \left( \frac{dp_1^+ dp_2^+ d^3p_{sp,T} d\bar{p}_s^N d^3p_{sN,T} d\bar{p}_s^A d^3p_{sA,T}}{2p_3^+(2\pi)^3 2p_4^-(2\pi)^3 2p_s^N (2\pi)^3 2p_s^A (2\pi)^3} \right)
\]

\[
= \sum_N \sum_{ijkl} \frac{1}{x_p x_N \alpha} d\sigma_{ij \rightarrow kl} \left| \psi_{i/p}(p_1, p_{sp}; p_p) \right|^2 \frac{dx_{sp} d^3p_{sp,T}}{2x_{sp} (2\pi)^3} \\
\times \left( \frac{dx_{N}}{2x_{N} (2\pi)^3} \right) \left( \frac{dx_{sN}}{2x_{sN} (2\pi)^3} \right) \left( \frac{d\alpha_s d^3p_{sA,T}}{\alpha_s} \right). \tag{A7}
\]

From here, we can introduce several identities that are consequences of momentum conservation. For instance, \(x_N = \frac{4x_p}{\alpha} \). In addition, \(1 = x_p + x_{sp}\) implies \(x_{sp}\) integration can be replaced by \(x_p\) integration. Similarly, \(x_{sN}\) integration can be replaced by integration over \(dx_N = \frac{1}{\alpha} dx_A\), and integration over \(\alpha_s = A - \alpha\) can be replaced by integration over \(\alpha\), and
all transverse momentum integrations can be transformed similarly. Thus we have

\[ \sigma_{\text{had.}} = \sum_N \sum_{ijkl} \int \frac{dx_p d^2 p_{1T}}{2x_p (1-x_p) (2\pi)^3} |\psi_{i/p}|^2 
\times \int \frac{d\alpha}{\alpha} d^2 p_N \frac{|\psi_{N/A}|^2}{\alpha (A - \alpha)} \int \frac{d\alpha_A d^2 p_{2T}}{2z_A (1 - \frac{z_A}{\alpha}) (2\pi)^3} |\psi_{j/N}|^2 \sigma_{ij\rightarrow kl} \]

\[ = \sum_N \sum_{ijkl} \int_0^1 dx_p \int_0^A dx_A f_p(x_p) \int_0^A \frac{d\alpha}{\alpha} \rho(\alpha) f_{j/N}^{(eff)}(x_A/\alpha) \sigma_{ij\rightarrow kl}. \]

With the convolution formula Eq. (20), this gives us the factorization formula Eq. (13).

**Appendix B: Light cone densities for short range correlations**

In this Appendix, we will derive the functional forms of the light cone densities for two-nucleon and three-nucleon short range correlations, as well as the functional forms of the light cone momenta that they employ as arguments. Appendix B1 will focus on 2N correlations, while Appendix B2 will focus on 3N correlations.

**1. 2N correlations**

In this part of the Appendix, we will derive Eqs. (29,28), and additionally relate the light cone deuteron wave function \( \psi_d(k) \) to the non-relativistic wave function in order to use a numerical parametrization.

To begin, it is known that two-nucleon SRCs dominate when the momenta are above the Fermi momentum \( k_F \). It is unknown how exactly the light cone density matrix behaves in the transition region around \( k_F \), so a cutoff \( \Theta(k - k_F) \) is imposed: 2N correlations will contribute only at momenta above the Fermi momentum. Moreover, since the high-momentum tail of nuclei are known to scale as the high-momentum tail of the deuteron, we will specifically have

\[ \rho_2^{(N)}(\alpha, p_T) = C^{(N)}(A) \rho_d(\alpha, p_T) \Theta(k - k_F), \] (B1)

where \( \rho_d(\alpha, p_T) \) is the light cone density matrix of the deuteron. The scaling constants \( C^{(N)}(A) \) will be different for protons and neutrons (except in the case of symmetric nuclei), and in particular must satisfy

\[ ZC^{(p)}(A) = (A - Z)C^{(n)}(A), \] (B2)

since a proton can only be in a 2N correlation by being paired with a neutron (within the approximation we use, which neglects \( pp \) and \( nn \) correlations), and vice-versa. These can additionally be related to the scaling constants \( a_2(A) \) through the scaling relation

\[ a_2(A) = \frac{2}{A} \sum_N \frac{\rho_2^{(N)}(\alpha, p_T)}{2\rho_d(\alpha, p_T)} : k \gg k_F, \] (B3)

which tells us that

\[ a_2(A) = \frac{ZC^{(p)}(A) + (A - Z)C^{(n)}(A)}{A}. \] (B4)
Combining Eqs. (B2,B4) gives us

\[ C^{(N)}(A) = \frac{a_2(A)}{2\chi_N}, \]  

where \( \chi_N \) is the relative fraction of nucleons of type \( N \) in the nucleus, i.e. \( \chi_p = \frac{Z}{A} \) and \( \chi_n = \frac{A-Z}{A} \).

We have most of the factors of Eq. (29), and now need to derive the form of \( \rho_d(\alpha, \mathbf{p}_T) \). To do this, we will evaluate the Feynman diagram in Fig. 14, which describes the deuteron interacting with an external probe in the limit of zero momentum transfer, and apply the optical theorem to the result. The Feynman amplitude for this process is

\[ i\mathcal{M}_{Xd} = 2 \int \frac{d^4p_2}{(2\pi)^4} \left[ \bar{u}^{\sigma_1} u^{\sigma_2} (i\Gamma_{p_1}) \left( \frac{\not{p}_1 + m}{p_1^2 - m^2 + i\epsilon} \right) \right. \\
\left. \times \left( i\Gamma_{\chi_N} \right) \left( \frac{i\not{p}_1 + m}{p_1^2 - m^2 + i\epsilon} \right) \left( i\Gamma_{p_1} \right) u^{\sigma_1} u^{\sigma_2} \right], \]  

with the factor of 2 owing to the possibility of scattering occurring from either nucleon. Here, momentum arguments have been left out for the initial and final state spinors since these do not affect the calculation.

From here, the amplitude is evaluated within the light cone formalism. For instance, we write out the integration element as

\[ d^4p_2 = \frac{1}{2}dp_2^+ dp_2^- d^2\mathbf{p}_2T. \]

We rewrite this, additionally, in terms of other variables. Since the light cone momentum \( p_d^+ = p_1^+ + p_2^+ \) of the deuteron is a constant, we can substitute \( dp_2^+ \) with \( dp_1^+ \) in the integration element. Similarly, we can substitute \( d^2\mathbf{p}_2T \) with \( d^2\mathbf{p}_1T \). Additionally, as \( p_1^+ = \frac{3}{2}p_d^+ \),

\[ d^4p_2 = \frac{p_d^+}{4} d\alpha d^2\mathbf{p}_{1T} dp_2. \]

The next issue concerns the denominators of the propagators. In general, we have

\[ p^2 - m^2 = p^+ p^- - \mathbf{p}_T^2 - m^2 = p^+ \left( p^- - \frac{m^2 + \mathbf{p}_T^2}{p^+} \right), \]

where for an on-shell particle, since \( p^2 = m^2 \),

\[ p_{on}^- = \frac{m^2 + \mathbf{p}_T^2}{p^+}. \]

FIG. 14: Coherent scattering diagram for \( pn \) cluster.
The propagator for $p_2$ in particular can be rewritten
\[
\frac{1}{p_2^2 - m^2 + i\epsilon} = \frac{1}{p_2^+ p_2^-} - \frac{1}{(p_2^+ - p_{2\text{ on}}^- - i\epsilon)}.
\]

The integration over $p_2$ is performed as a contour integration in the complex $p_2^-$ plane, with the semicircular arc encircling the lower complex plane. The pole at $p_2^- = p_{2\text{ on}}^- - i\epsilon$ is picked up and $\epsilon$ afterwards set to zero, effectively setting $p_2^-$ on its light cone energy shell. A factor of $-2\pi i$ is also picked up from the contour integration. Consequently, the Feynman amplitude becomes
\[
\mathcal{M}_{Xd} = 2 \int \frac{d\alpha d^2 p_{1T}}{(2\pi)^3} \left[ \frac{p_{d}^+}{4} \bar{u}^s_1 \bar{u}^s_2 \Gamma_p \left( \frac{\rho_1 + m}{p_{1T}^2 - m^2 + i\epsilon} \right) \Gamma_{XN} \left( \frac{\rho_1 + m}{p_{1T}^2 - m^2 + i\epsilon} \right) \frac{\rho_2 + m}{p_{2T}^2} \right].
\]

Since nucleon 2 is now on-shell, the completeness relation $\rho_2 + m = \sum \lambda_2 u^{\lambda_2}(p_2)\bar{u}^{\lambda_2}(p_2)$ is exact, so can be substituted in to the expression for $\mathcal{M}_{Xd}$. Additionally, we will use the completeness relation as approximately true for nucleon 1, giving us
\[
\mathcal{M}_{Xd} = 2 \sum_{\lambda_1, \lambda_2} \int \frac{d\alpha d^2 p_{1T}}{(2\pi)^3} \left[ \frac{p_{d}^+}{4} \bar{u}^s_1 \bar{u}^s_2 \Gamma_p \frac{u^{\lambda_1}(p_1)\bar{u}^{\lambda_2}(p_2)}{p_{1T}^2 - m^2 + i\epsilon} \right]
\left( \bar{u}^{\lambda_1}(p_1)\Gamma_{XN} u^{\lambda_1}(p_1) \right) \left( \frac{u^{\lambda_2}(p_1)\bar{u}^{\lambda_2}(p_2)}{p_{2T}^2 - m^2 + i\epsilon} \right). \tag{B7}
\]

The amplitude for a single nucleon to interact with the probe is given by
\[
i \mathcal{M}_{XN} = \bar{u}^{\lambda_1}(p_1)(i \Gamma_{XN}) u^{\lambda_1}(p_1), \tag{B8}
\]
which allows one of the bracketed terms in Eq. (B7) to be rewritten. The remaining bracketed terms should be related to the deuteron wave function, but we have to determine how. To this end, we look at the remaining propagator denominator; namely,
\[
p_{1T}^2 - m^2 = \frac{1}{2} \left( p_{d}^+ p_{d}^- - \frac{m^2 + p_{1T}^2}{\alpha/2} \right)
= \frac{\alpha}{2} \left( p_{d}^+ p_{d}^- - \frac{m^2 + p_{1T}^2}{\alpha/2} \right) = \frac{\alpha}{2} \left( M_d^2 + p_{dT}^2 - \frac{p_{d}^+ p_{d}^-}{(2 - \alpha)/2} - \frac{m^2 + p_{1T}^2}{\alpha/2} \right)
= \frac{\alpha}{2} \left( M_d^2 - \alpha(2 - \alpha) \left[ \frac{\alpha}{2} \left( \frac{2 - \alpha}{2} \right) p_{dT}^2 + \alpha \frac{p_{2T}^2}{2} + \frac{2 - \alpha}{2} \right] \right).
\]

From here, note that $p_{2T} = p_{dT} - p_{1T}$ and look at the terms in the square brackets:
\[
\left[ \right] = -\frac{\alpha}{2} p_{dT}^2 + \frac{\alpha}{4} p_{dT}^2 + \frac{\alpha}{2} p_{dT}^2 - \alpha p_{dT} \cdot p_{1T} + \frac{\alpha}{2} p_{1T}^2 + p_{1T}^2 - \frac{\alpha}{2} p_{1T}^2 + m^2
= \left( p_{1T} - \frac{\alpha}{2} p_{dT} \right)^2 + m^2.
\]

If we now define
\[
k^2 = \frac{\left( p_{1T} - \frac{\alpha}{2} p_{dT} \right)^2 + m^2}{\alpha(2 - \alpha) - m^2}, \tag{B9}
\]

32
without yet committing to an interpretation of $k$ as a momentum, we can write
\[ p_1^2 - m^2 + i\epsilon = \frac{\alpha}{2} \left( M_d^2 - 4 \left[ m^2 + k^2 \right] + i\epsilon \right). \] (B10)

Now, we wish to justify the interpretation of $k$ as a momentum. To this end, let us look at the light cone kinematics of the bound system. The deuteron itself is on its light cone energy shell, so $p_d^+ p_d^- - p_{2T}^2 = M_d^2$. However, since it is a bound system, its constituents are not on-shell. We have chosen for $p^- \rightarrow 0$ to contain the virtuality of the off-shell nucleon, so we can write
\[ p_d^- = p_1^- + p_2^- = p_{1\text{ on}} + p_{2\text{ on}} + V = \frac{m^2 + p_{1T}^2}{p_1^+} + \frac{m^2 + p_{2T}^2}{p_2^+} + V. \]
where $V$ is a term that represents the virtuality of the bound nucleons. Placing this expression for $p_d^-$ into the on-shell condition for the deuteron gives
\[ M_d^2 = \left( -p_{2T}^2 + \frac{m^2 + p_{1T}^2}{\alpha/2} + \frac{m^2 + p_{2T}^2}{(2 - \alpha)/2} + Vp_d^+ \right) = 4(m^2 + k^2) + Vp_d^+. \] (B11)

Consider now the deuteron in its rest frame, so that it consists of two nucleons with equal and opposite momenta $\pm k$. In the absence of binding energy, the rest and kinetic energy would for each particle would be $E_k = \sqrt{m^2 + k^2}$, giving a total squared center of mass energy of $4(m^2 + k^2)$. The expression Eq. (B11) gives exactly this, with $Vp_d^+$ as a correction owing to binding energy. $k$ does correspond to a momentum after all, and Eq. (B9) is equivalent to Eq. (28), in the special case that $p_{2T} = 0$, which is relevant to our purpose since we are neglecting the transverse motion of the 2N correlation.

Now, we combine Eqs. (B7,B8,B10) to obtain
\[ M_{Xd} = 2 \sum_{\lambda_1, \lambda_2} \int \frac{d\alpha d^2 p_{1T}}{(2\pi)^3} \left[ \frac{1}{2(2 - \alpha)\alpha} \left( \bar{u}^{\sigma_1} u^{\sigma_2} \Gamma_{p_1 u^{\lambda_1}(p_1) u^{\lambda_2}(p_2)} \right) \frac{1}{2} \left( M_d^2 - 4(m^2 + k^2) \right) \right] \]
\[ M_{XN} \left( \frac{1}{2} \left( M_d^2 - 4(m^2 + k^2) \right) \Gamma_{p_1 u^{\lambda_1}(p_1) u^{\lambda_2}(p_2)} \right) - \bar{u}^{\lambda_1}(p_1) u^{\lambda_2}(p_2) \Gamma_{p_n u^\sigma_1 u^\sigma_2} \right]. \] (B12)

To this, we apply the optical theorem by taking the imaginary part of Eq. (B12) on both sides. In particular, the optical theorem gives us
\[ \text{Im}(M_{Xd}) = s_{Xd} \sigma_{Xd} = 2s_{Xd} \sigma_{XN} \] (B13)
\[ \text{Im}(M_{XN}) = s_{XN} \sigma_{XN} = \frac{\alpha}{2} s_{Xd} \sigma_{XN}, \] (B14)
and therefore
\[ 1 = \sum_{\lambda_1, \lambda_2} \int \frac{d\alpha d^2 p_{1T}}{2(2\pi)^3} \left[ \frac{1}{2(2 - \alpha)\alpha} \left( \bar{u}^{\sigma_1} u^{\sigma_2} \Gamma_{p_1 u^{\lambda_1}(p_1) u^{\lambda_2}(p_2)} \right) \frac{1}{2} \left( M_d^2 - 4(m^2 + k^2) \right) \right] \]
\[ \left( \frac{1}{2} \left( M_d^2 - 4(m^2 + k^2) \right) \Gamma_{p_n u^\sigma_1 u^\sigma_2} \right) \] (B15)
In its current form, Eq. (B15) allows us to identify \( \rho_d(\alpha, p_{1T}) \) in terms of the given quantities. It also allows us to identify the light cone wave function up to some arbitrary conventions. For instance, we can define the wave function so that a factor of \((2\pi)^{-3}\) must be present in the integration element for its normalization condition, or we may instead absorb this factor into the wave function itself. We will employ the following standard, also employed in Ref. [3]:

\[
\psi_d^{(\lambda_1, \lambda_2)}(k) = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2(2\pi)^3}} \frac{\bar{u}^{\lambda_1}(p_1)u^{\lambda_2}(p_2)}{\frac{1}{2}(M_d^2 - 4(m^2 + k^2))} \Gamma_{\mu
u} u^\mu u^\nu, \tag{B16}
\]

meaning that Eq. (B15) gives

\[
\sum_{\lambda_1, \lambda_2} \int d\alpha d^2 p_{1T} \frac{|\psi_d^{(\lambda_1, \lambda_2)}(k)|^2}{\alpha(2 - \alpha)} = 1 \tag{B17}
\]

\[
\rho_d(\alpha, p_{1T}) = \sum_{\lambda_1, \lambda_2} |\psi_d^{(\lambda_1, \lambda_2)}(k)|^2 \frac{\alpha(2 - \alpha)}{\alpha(2 - \alpha)}. \tag{B18}
\]

If \( \psi_d(k) \) without helicity indices is considered as the spin-averaged deuteron wave function, then this, with Eq. (B1) and the expressions for \( C^{(N)}(A) \), gives us Eq. (29).

Now all that remains is to relate \( \psi_d(k) \) to the non-relativistic wave function, so that we can employ a numerical parametrization. We will find the relationship in particular through the normalization condition Eq. (B17). First, we find how \( \alpha \) and the longitudinal component \( k_z \) of the momentum are related. We already identify \( k_T = p_{1T} - \frac{\alpha}{2} p_{dT} \), so the Pythagorean condition \( k^2 = k_z^2 + k_T^2 \) gives us

\[
k_z^2 = k^2 - k_T^2 = \frac{(\alpha - 1)^2 m^2 + k_T^2}{\alpha(2 - \alpha)} - k_T^2 = \frac{(\alpha - 1)^2 (m^2 + k_T^2)}{\alpha(2 - \alpha)}
\]

\[
= \frac{(\alpha - 1)^2 (E_k^2 - k_T^2)}{\alpha(2 - \alpha)}
\]

\[
\left(1 + \frac{(\alpha - 1)^2}{\alpha(2 - \alpha)}\right) k_z^2 = \left(\frac{(\alpha - 1)^2}{\alpha(2 - \alpha)}\right) E_k^2
\]

\[
(2\alpha - \alpha^2 + \alpha^2 - 2\alpha + 1) k_z^2 = k_z^2 = (\alpha - 1)^2 E_k^2
\]

\[
k_z = (\alpha - 1) E_k
\]

\[
\alpha = \frac{E_k + k_z}{E_k}. \tag{B19}
\]

The last few lines employed the condition, when the square roots of \( k_z^2 \) and \( E_k^2 \) were taken, that \( k_z \) is positive when \( \alpha > 1 \) and negative when \( \alpha < 1 \), corresponding to the convention that \( \alpha > 1 \) signifies a forward-going nucleon in the deuteron rest frame. Using Eq. (B19), we can obtain the following relations:

\[
\alpha(2 - \alpha) = \frac{E_k + k_z}{E_k} \times \frac{E_k - k_z}{E_k} = \frac{E_k^2 - k_z^2}{E_k^2}
\]

\[
\frac{d\alpha}{dk_z} = \frac{1}{E_k} - \frac{k_z^2}{E_k^3} = \frac{1}{E_k} \times \frac{E_k^2 - k_z^2}{E_k^2}
\]

\[
\frac{d\alpha}{\alpha(2 - \alpha)} = \frac{dk_z}{E_k}.
\]
This means the normalization condition Eq. (B17) can be written
\[ \int |\psi_d(k)|^2 \frac{d^3k}{E_k}, \]
which, in light of the non-relativistic normalization condition \( \int |\psi_{NR}(k)|^2 d^3k = 1 \), tells us
\[ |\psi_d(k)|^2 = |\psi_{NR}(k)|^2 E_k. \] (B20)

Notice that the argument the non-relativistic wave function takes in this relation is the light cone momentum \( k \).

2. 3N correlations

In this section of the Appendix, we derive the functional form of the three-nucleon SRC density, namely Eq. (31). In addition, we derive the form of the momenta \( k_{12} \) and \( k_{23} \) that enter into the wave function as arguments.

As discussed in Sec. III A 3, we consider the 3N correlation as arising from a sequence of two short-range two-body interactions. Since we consider only \( pn \) pairs to interact at short distance, the 3N cluster can itself be either a proton and two neutrons, or a neutron and two protons.

We will calculate the 3N light cone density matrix by considering a coherent scattering process in the limit of zero momentum transfer, just as we did for the 2N density in Appendix B 1. On the face of it, we appear to have several diagrams that can contribute, but in fact we only have one. To see this, suppose (without loss of generality) that the probed nucleon is a proton: then either the cluster is \( pnp \), in which case the diagram for the process is depicted in Fig. 15, or the cluster is \( pnn \), in which case the diagrams for the process are depicted in Fig. 16. In the latter case, both Fig. 16a and Fig. 16b represent means by which the scattering process can occur, and since they are related by the exchange of an identical fermion, their amplitudes are exactly opposite and cancel perfectly \[69]. Thus, we only have to consider the diagram in Fig. 15.
To begin, we write the full Feynman amplitude in all its glory:

\[
i\mathcal{M}_{3N,X} = 2 \int \frac{d^4p_2}{(2\pi)^4} \frac{d^4p_3}{(2\pi)^4} \left[ \bar{u}^{\sigma_1}(k_1) \bar{u}^{\sigma_2}(k_2) \bar{u}^{\sigma_3}(k_3) \left( i\Gamma_{pn}^{(2',3)} \right) \left( \frac{i}{p_2^2 + m - m^2 + i\epsilon} \right) \left( i\Gamma_{pn}^{(1,2)} \right) \left( \frac{i}{p_3^2 + m - m^2 + i\epsilon} \right) \right] + \ldots,
\]

where the factor of 2 is present since the process can occur by scattering off of either of the protons (or either of the neutrons if it’s a npn cluster), and where the \ldots signifies terms not calculated here that are relevant at lower light cone momentum fractions. (These terms are not important for the form of \( \rho_3(\alpha, p_T) \) at the high \( \alpha \) at which three-nucleon SRCs become relevant, but the fact that they are present is important for its normalization.) In contrast to the two-nucleon case, we have explicitly written down the initial and final nucleon momenta as \( k_i \). These will enter into the result, and accordingly we will have to estimate them. In particular, we will use a collinear approximation where all three \( k_i^+ = \frac{1}{3} p_{3N}^+ \), and all three \( k_{iT} = 0 \).

As in the 2N case, we use light cone coordinates, so write, for instance

\[
d^4p_2 = \frac{1}{2} dp_2^+ dp_2^- d^2 p_{2T},
\]

with a similar formula holding for \( p_3 \). Combining this with the \( p_2^2 - m^2 + i\epsilon \) denominator, and performing a contour integration in the complex \( p_2^2 \) plane, will give

\[
\frac{d^4p_2}{(2\pi)^4} \frac{1}{p_2^2 - m^2 + i\epsilon} \rightarrow -i \frac{dp_2^+ d^2 p_{2T}}{2p_2^+ (2\pi)^3} = -i \frac{d\alpha_2 d^2 p_{2T}}{2\alpha_2 (2\pi)^3},
\]

with again a similar formula holding for \( p_3 \). These integrations place nucleons 2 and 3 on-shell, allowing the use of exact completeness relations. For the remaining internal nucleon lines, we employ approximate completeness relations. Consequently, we can write the matrix

FIG. 16: Diagrams for 3N cluster: \( pnn \) if \( p \) is probed, or \( npp \) if \( n \) is probed.
element as

\[ \mathcal{M}_{3N,X} = 2 \sum_{\lambda_1,\lambda_2,\lambda_3,\lambda'_2} \int \frac{d\alpha_2d^2p_{2T}}{2\alpha_2(2\pi)^3} \frac{d\alpha_3d^2p_{3T}}{2\alpha_3(2\pi)^3} \left[ \left( \bar{u}^{\sigma_2}(k_2)\bar{u}^{\sigma_3}(k_3)\Gamma_{pn}^{(2',3)} \frac{u^{\lambda_2}(p'_2)u^{\lambda_3}(p_3)}{(p'_2)^2 - m^2 + i\epsilon} \right) \right. \\
\left. \left( \bar{u}^{\sigma_1}(k_1)\bar{u}^{\lambda_1}(p_1)\Gamma_{pn}^{(1,2)} \frac{u^{\lambda_1}(p_1)u^{\lambda_2}(p_2)}{(p_1^2 - m^2 + i\epsilon)} \right) \right] \left( \bar{u}^{\lambda_2}(p_2)\Gamma_{pn}^{(1,2)} \frac{u^{\sigma_1}(k_1)u^{\lambda_2}(p_2)}{(p_2^2 - m^2 + i\epsilon)} \right) \left( \bar{u}^{\lambda_1}(p_1)\Gamma_{XN}^{(1)} \frac{u^{\lambda_1}(p_1)}{(p_1^2 - m^2 + i\epsilon)} \right) \] \\
+ \ldots \hspace{1cm} (B21)

We can immediately use \( \mathcal{M}_{XN} = \bar{u}^{\lambda_1}(p_1)\Gamma_{XN}^{(1)} \frac{u^{\lambda_1}(p_1)}{(p_1^2 - m^2 + i\epsilon)} \). What must be done next is to deal with the denominator terms \( p_1^2 - m^2 \) and \( (p'_2)^2 - m^2 \) in order to write the amplitude in terms of 2N wave functions, and find their momentum dependence. In particular, since we are to find momenta that enter into a 2N wave function, we must state things in terms of relative momentum fractions of the virtual nucleons with respect to the two-nucleon clusters they make up. In particular, we use

\[ \gamma_1 = \frac{2p_1^+}{p_1^+ + p_2^+} = \frac{2\alpha}{3 - \alpha_3} \]
\[ \gamma'_2 = \frac{2(p'_2)^+}{p'_2^+ + p_3^+} = \frac{2p'_2^+ + p_2^+ - k_1^+}{k_2^+ + k_3^+} = \alpha + \alpha_2 - 1 = 2 - \alpha_3. \]

With this in mind, first consider the \( p_1 \) denominator:

\[ p_1^2 - m^2 = p_1^+p_1 - P_{1T}^2 - m^2 = \frac{\gamma_1}{2} \left( p_{12}^+p_{12}^- - \frac{m^2 + P_{12T}^2}{\gamma_1/2} \right) = \frac{\gamma_1}{2} \left( p_{12}^+p_{12}^- - \frac{M_{12}^2 + P_{12T}^2}{(2 - \gamma_1)/2} - \frac{m^2 + P_{12T}^2}{\gamma_1/2} \right) \]
\[ = \frac{\gamma_1}{2} \left( M_{12}^2 + P_{12T}^2 - \frac{m^2 + P_{12T}^2}{(2 - \gamma_1)/2} - \frac{m^2}{\gamma_1/2} \right) = \frac{\gamma_1}{2} \left( M_{12}^2 - 4 \left[ \frac{m^2 + (P_{1T} - \frac{\gamma_1}{2}P_{12T})^2}{\gamma_1(2 - \gamma_1)} \right] \right) \]
\[ = \frac{\gamma_1}{2} \left( M_{12}^2 - 4(m^2 + k_{12}^2) \right), \]

where we have identified

\[ k_{12}^2 = \frac{(\gamma_1 - 1)^2 m^2 + (P_{1T} - \frac{\gamma_1}{2}P_{12T})^2}{\gamma_1(2 - \gamma_1)} \]

as the light cone momentum for the pair (1,2). For the pair (2',3), we manipulate the \( p'_2 \)
\[(p'_2)^2 - m^2 = (p'_2)^+(p'_2)^- - (p'_{2T})^2 - m^2 = \frac{\gamma'_2}{2} \left( p'_{23}(p'_2)^- - \frac{m^2 + (p'_{2T})^2}{\gamma'_2/2} \right) \]

\[= \frac{\gamma'_2}{2} \left( p'_{23}p'_{23} - p'_{23}p'_2 - \frac{m^2 + (p'_{2T})^2}{\gamma'_2/2} \right) = \frac{\gamma'_2}{2} \left( M'_{23}^2 + p'_{23T}^2 - \frac{m^2 + (p'_{2T})^2}{(2 - \gamma'_2)/2} \right) \]

\[= \frac{\gamma'_2}{2} \left( M'_{23}^2 + p'_{23T}^2 - \frac{m^2 + (p'_{2T})^2}{(2 - \gamma'_2)/2} \right) \]

\[\]

\[= \frac{\gamma'_2}{2} \left( M'_{23}^2 - 4 \left[ \frac{m^2 + (p'_{2T} - \frac{\gamma'_2}{2} p'_{23T})^2}{\gamma'_2(2 - \gamma'_2)} \right] \right) = \frac{\gamma'_2}{2} \left( M'_{23}^2 - 4(m^2 + k_{23}^2) \right), \]

where we have identified

\[k_{23}^2 = \frac{(\gamma'_2 - 1)^2 m^2 + (p'_{2T} - \frac{\gamma'_2}{2} p'_{23T})^2}{\gamma'_2(2 - \gamma'_2)}. \tag{B25} \]

With these identifications, Eq. (B21) can now be written

\[\mathcal{M}_{3N, X} = \frac{2}{\lambda_1\lambda_2\lambda_3\lambda'_2} \sum \int \frac{d\alpha_3 d^2p'_{2T}}{2\alpha_2(2\pi)^3} \frac{d^2p'_{3T}}{2\alpha_3(2\pi)^3} \left[ \frac{1}{(\gamma^2\gamma'\gamma'_2)^2} \left( \hat{u}^{(2)}(k_2)\bar{u}^{(3)}(k_3)G^{(2,3)}_{pm} \frac{u^{(2)}(p_2)}{\gamma_2(M'_{23}^2 - 4(m^2 + k_{23}^2))} \right) \right. \]

\[\left( \hat{u}^{(2)}(k_1)\bar{u}^{(3)}(p_2)G^{(1,2)}_{pm} \frac{u^{(1)}(p_1)u^{(2)}(p_2)}{\gamma_2(M'_{23}^2 - 4(m^2 + k_{23}^2))} \right) \mathcal{M}_{3N} \left( \frac{\hat{u}^{(1)}(p_1)\bar{u}^{(2)}(p_2)}{\gamma_2(M'_{23}^2 - 4(m^2 + k_{23}^2))}G^{(1,2)}_{pm}u^{(3)}(k_3) \right) \right] + \ldots. \tag{B26} \]

Now, we wish to apply Eq. (B16) to Eq. (B26) in order to condense terms into short-range 2N wave functions. However, in order to apply Eq. (B16), we need to investigate the mass terms \(M'_{12}^2\) and \(M'_{23}^2\). These should be close to \(M_0^2 \approx 4m^2\) in order for Eq. (B16) to be applicable. To be sure, a large difference \(\Delta = 4m^2 - M^2\) could be absorbed into \(k^2\) instead, allowing \(\psi_d\left(\sqrt{k^2 + \frac{\Delta}{4}}\right)\) to be used instead. We will elaborate on this below.

First, let us calculate \(M'_{23}^2\). Within the collinear approximation we use, we have

\[M'_{23}^2 = p'_{23}p'_{23} - p'_{23T}^2 = (p'_{3N} + k'_1)(p'_{3N} - k'_1) = p'_{3N}^2 - k'_1 p'_{3N} - k'_1 p'_{3N} + k'_1 k'_1 \]

\[= \frac{1}{3} M'_{3N}^2 - \frac{1}{3} M'_{3N}^2 - \frac{1}{1/3} m^2 + m^2 = \frac{2}{3} M'_{3N}^2 - 2m^2 \approx 4m^2, \]

meaning the 2N wave function for the pair \((2',3)\) is fine as it is. However, for the case of \(M'_{12}^2\),

\[M'_{12}^2 = p'_{12}p'_{12} - p'_{12T}^2 = (p'_{3N} - p'_{3})(p'_{3N} - p'_{3}) - p'_{3T}^2 = p'_{3N}p'_{3N} - p'_{3N}p'_{3N} - p'_{3N}p'_{3N} + p'_{3N} - p'_{3N} - p'_{3N}^2 \]

\[= \frac{3}{3} M'_{3N}^2 - \frac{1}{3} M'_{3N}^2 - \frac{1}{3} (m^2 + p'_{3T}^2) + m^2 = \left(1 - \frac{3}{3}\right) M'_{3N}^2 + \left(1 - \frac{3}{3}\right) m^2 - \frac{3p'_{3T}^2}{3} \approx \left(10 - 3 \left(\frac{1}{3}\right)\right) m^2 - \frac{3p'_{3T}^2}{3}. \]
meaning the bound mass-squared of the 2N system is short by

\[ \Delta = 4m^2 - M_{12}^2 = 3\left(\alpha_3 + \frac{1}{\alpha_3} - 2\right)m^2 + \frac{3p^2_{2T}}{\alpha_3}, \]  

(B27)

and the wave function should take \( \hat{k}_{12} \equiv \sqrt{k_{12}^2 + \Delta} \) as its argument rather than \( k_{12} \). However, if \( k_{12} \gg \sqrt{\Delta} \), the correction to the argument of the wave function will be small. At sufficiently short distance scales, the virtuality of the bound nucleons in the pair will be more significant than the virtuality of the pair itself. This is the case at kinematics where three-nucleon SRCs dominate, so we take \( k_{12} = \hat{k}_{12} \) for practical computations.

Now that we have found the 2N correlation masses, we can rewrite Eq. (B26) as

\[ M_{3N,X} = 2 \int \frac{d\alpha_2 d^2 p_{2T} d\alpha_3 d^2 p_{3T}}{\alpha_2 \alpha_3} \left[ \frac{4}{(\gamma_1 \gamma_2')^2} \left| \psi_d(\hat{k}_{12}) \right|^2 \left| \psi_d(k_{23}) \right|^2 M_{XN} \right] + \ldots. \]  

(B28)

Now, we apply the optical theorem to Eq. (B28) by taking the imaginary part of both sides. Note that we have

\[ \text{Im}(M_{3N,X}) = s_{3N,X} \sigma_{3N,X} + C = 2s_{3N,X} \sigma_{XN} + C \]  

(B29)

\[ \text{Im}(M_{XN}) = s_{XN} \sigma_{XN} = \frac{\alpha}{3} s_{3N,X} \sigma_{XN}, \]  

(B30)

where \( C \) corresponds to the part of \( \text{Im}(M_{3N,X}) \) coming from processes we did not consider (i.e. the \( \ldots \) processes), and where we would have \( \sigma_{3N,X} = 2\sigma_{XN} \) if the process depicted in Fig. 15 were the only means by which the probing reaction could occur. These relations give

\[ 1 - C = \frac{4}{3} \int d\alpha_2 d^2 p_{2T} d\alpha_3 d^2 p_{3T} \frac{\alpha_1}{\alpha_2 \alpha_3 (\gamma_1 \gamma_2')^2} \left| \psi_d(\hat{k}_{12}) \right|^2 \left| \psi_d(k_{23}) \right|^2. \]

This tells us that that, sans the threshold condition \( \Theta_3(\alpha, p_T) \) and up to an undetermined constant \( C' \), the 3N light cone density matrix is

\[ \rho_3(\alpha, p_T) = C' a_3(A) \alpha \int d\alpha_3 d^2 p_{3T} \frac{1}{\alpha_3(3 - \alpha - \alpha_3)} \frac{1}{\gamma_1 \gamma_2')^2} \left| \psi_d(\hat{k}_{12}) \right|^2 \left| \psi_d(k_{23}) \right|^2. \]  

(B31)

This should be normalized to \( \frac{3(1-C)}{4} \leq \frac{3}{4} \). The constant \( C \) cannot be determined theoretically within this framework, but numerical evaluation gives the normalization of \( \rho_3(\alpha, p_T) \) to be \( \sim 0.45 \), giving us \( C \approx 0.66 \).

Now, we would like to state the prefactor that exists aside from the squares of wave functions either entirely in terms of the relative fractions \( \gamma_1 \) and \( \gamma_2' \), or in terms of the external variable \( \alpha \) and the integration variable \( \alpha_3 \). To this end, note that

\[ \frac{1}{(\gamma_1 \gamma_2')^2} = \left( \frac{3 - \alpha_3}{2\alpha} \right) \left( \frac{1}{2 - \alpha_3} \right) = \frac{1}{\alpha^2} \left\{ \frac{3 - \alpha_3}{2(2 - \alpha_3)} \right\}^2, \]

and therefore

\[ \rho_3(\alpha, p_T) = C' \frac{1}{\alpha} \int d\alpha_3 d^2 p_{3T} \frac{1}{\alpha_3(3 - \alpha - \alpha_3)} \left\{ \frac{3 - \alpha_3}{2(2 - \alpha_3)} \right\}^2 \left| \psi_d(\hat{k}_{12}) \right|^2 \left| \psi_d(k_{23}) \right|^2. \]  

(B32)
Note that this differs from the result of Ref. [3] for the 3N correlation using a similar model, by the prefactor in curly brackets.

What remains is to determine the constant $C'$. For this, we consider that the probability for two short-range 2N interactions to occur should be, for large nuclei, equal to the square of the probability for one short-range 2N interaction to occur. In other words, the total factor $C'a_3(A) = \{a_2(A)\}^2$. Considering the definition of $a_3(A)$ in Eq. (32), we have

$$a_3(^3\text{He}) = 1,$$

meaning that $C' = \{a_2(^3\text{He})\}^2$. When we consider Eq. (B32) defines $p_3(\alpha, p_T)$ simply without a factor of $a_3(A)$ and without kinematical thresholds, multiplying it by $a_3(A)$ and using $C'a_3(A) = \{a_2(A)\}^2$ gives us Eq. (31).

[1] L. L. Frankfurt and M. I. Strikman, Phys. Rept. 160, 235 (1988).
[2] M. Sargsian, J. Arrington, W. Bertozzi, W. Boeglin, C. Carlson, et al., J. Phys. G29, R1 (2003), arXiv:nucl-th/0210025 [nucl-th].
[3] L. L. Frankfurt and M. I. Strikman, Phys. Rept. 76, 215 (1981).
[4] L. Frankfurt, M. Strikman, D. Day, and M. Sargsian, Phys. Rev. C48, 2451 (1993).
[5] J. L. Aclander, J. Alster, D. Barton, G. Bunce, A. Carroll, et al., Phys. Lett. B453, 211 (1999).
[6] A. Tang, J. Watson, J. L. Aclander, J. Alster, G. Asryan, et al., Phys. Rev. Lett. 90, 042301 (2003), arXiv:nucl-ex/0206003 [nucl-ex].
[7] K. Egiyan et al. (CLAS Collaboration), Phys. Rev. C68, 014313 (2003), arXiv:nucl-ex/0301008 [nucl-ex].
[8] K. Egiyan et al. (CLAS Collaboration), Phys. Rev. Lett. 96, 082501 (2006), arXiv:nucl-ex/0508026 [nucl-ex].
[9] E. Piasetzky, M. Sargsian, L. Frankfurt, M. Strikman, and J. Watson, Phys. Rev. Lett. 97, 162504 (2006), arXiv:nucl-th/0604012 [nucl-th].
[10] R. Shneor et al. (Jefferson Lab Hall A Collaboration), Phys. Rev. Lett. 99, 072501 (2007), arXiv:nucl-ex/0703023 [nucl-ex].
[11] R. Subedi, R. Shneor, P. Monaghan, B. Anderson, K. Aniol, et al., Science 320, 1476 (2008), arXiv:0908.1514 [nucl-ex].
[12] L. Frankfurt, M. Sargsian, and M. Strikman, Int. J. Mod. Phys. A23, 2991 (2008), arXiv:0806.4412 [nucl-th].
[13] N. Fomin, J. Arrington, R. Asaturyan, F. Benmokhtar, W. Boeglin, et al., Phys. Rev. Lett. 108, 092502 (2012), arXiv:1107.3583 [nucl-ex].
[14] J. Arrington, D. Higinbotham, G. Rosner, and M. Sargsian, Prog. Part. Nucl. Phys. 67, 898 (2012), arXiv:1104.1196 [nucl-ex].
[15] J. Arrington, A. Daniel, D. Day, N. Fomin, D. Gaskell, et al., Phys. Rev. C86, 065204 (2012), arXiv:1206.6343 [nucl-ex].
[16] I. Korover, N. Muangma, O. Hen, R. Shneor, V. Sulkosky, et al., Phys. Rev. Lett. 113, 022501 (2014), arXiv:1401.6138 [nucl-ex].
[17] A. Benvenuti et al. (BCDMS Collaboration), Z. Phys. C63, 29 (1994).
[18] M. Vakili et al. (CCFR Collaboration), Phys. Rev. D61, 052003 (2000), arXiv:hep-ex/9905052 [hep-ex].
[19] P. E. Bosted, A. Lung, L. Andivahis, L. Stuart, J. Alster, et al., Phys. Rev. C46, 2505 (1992).
[20] N. Fomin, J. Arrington, D. Day, D. Gaskell, A. Daniel, et al.,
Phys. Rev. Lett. 105, 212502 (2010), arXiv:1008.2713 [nucl-ex].

21] J. Dudek, R. Ent, R. Essig, K. S. Kumar, C. Meyer, et al., Eur. Phys. J. A48, 187 (2012).

22] J. Arrington et al., “Jefferson lab experiment 12-06-105: Inclusive scattering from nuclei at \( x > 1 \) in the quasielastic and deeply inelastic regimes,”
http://www.jlab.org/exp_prog/proposals/06/PR12-06-105.pdf.

23] J. Collins, Foundations of perturbative QCD (2011).

24] R. K. Ellis, W. J. Stirling, and B. R. Webber, QCD and collider physics, Vol. 8 (1996) pp. 1–435.

25] K. Eskola, V. Kolhinen, and C. Salgado, Eur. Phys. J. C9, 61 (1999), arXiv:hep-ph/9807297 [hep-ph].

26] D. de Florian and R. Sassot, Phys. Rev. D69, 074028 (2004), arXiv:hep-ph/0311227 [hep-ph].

27] M. Hirai, S. Kumano, and T.-H. Nagai, Phys. Rev. C76, 065207 (2007), arXiv:0709.3038 [hep-ph].

28] K. Eskola, H. Paukkunen, and C. Salgado, JHEP 0904, 065 (2009), arXiv:0902.4154 [hep-ph].

29] D. de Florian, R. Sassot, P. Zurita, and M. Stratmann, Phys. Rev. D85, 074028 (2012), arXiv:1112.6324 [hep-ph].

30] H. Honkanen, M. Strikman, and V. Guzey, (2013), arXiv:1310.5879 [hep-ph].

31] M. Arneodo, Phys. Rept. 240, 301 (1994).

32] D. F. Goezaman, K. Saito, and A. W. Thomas, Ann. Rev. Nucl. Part. Sci. 45, 337 (1995).

33] E. Moniz, I. Sick, R. Whitney, J. Ficenec, R. D. Kephart, et al., Phys. Rev. Lett. 26, 445 (1971).

34] M. V. Zverev and E. E. Saperstein, Yad. Fiz. 43, 304 (1986).

35] M. McGauley and M. Sargsian, (2011), arXiv:1102.3973 [nucl-th].

36] H. Feldmeier, W. Horiuchi, T. Neff, and Y. Suzuki, Phys. Rev. C84, 054003 (2011), arXiv:1107.4956 [nucl-th].

37] M. M. Sargsian, Phys. Rev. C89, 034305 (2014), arXiv:1210.3280 [nucl-th].

38] M. Alvioli, C. Ciofi Degli Atti, L. P. Kaptari, C. B. Mezzetti, and H. Morita, Int. J. Mod. Phys. E22, 1330021 (2013), arXiv:1306.6235 [nucl-th].

39] L. Frankfurt and M. Strikman, Nucl. Phys. B148, 107 (1979).

40] M. M. Sargsian, J. Phys. Conf. Ser. 496, 012007 (2014), arXiv:1312.2263 [nucl-th].

41] M. Vanhalst, W. Cosyn, and J. Ryckebusch, (2014), arXiv:1405.3814 [nucl-th].

42] M. Lacombe, B. Loiseau, J. Richard, R. Vinh Mau, J. Cote, et al., Phys. Rev. C21, 861 (1980).

43] D. W. Higinbotham and O. Hen, (2014), arXiv:1409.3069 [nucl-ex].

44] D. Day, L. Frankfurt, M. I. Strikman, and M. M. Sargsian, in progress.

45] Y. Bayukov, V. Efremenko, S. Frankel, W. Frati, M. Gazzaly, et al., Phys. Rev. C20, 764 (1979).

46] J. Aubert et al. (European Muon Collaboration), Phys. Lett. B123, 275 (1983).

47] G. A. Miller and J. R. Smith, Phys. Rev. C65, 015211 (2002), arXiv:nucl-th/0107026 [nucl-th].

48] J. Smith and G. A. Miller, Phys. Rev. C65, 055206 (2002), arXiv:nucl-th/0202016 [nucl-th].

49] J. Gomez, R. Arnold, P. E. Bosted, C. Chang, A. Katramatou, et al., Phys. Rev. D49, 4348 (1994).

50] J. Seely, A. Daniel, D. Gaskell, J. Arrington, N. Fomin, et al., Phys. Rev. Lett. 103, 202301 (2009), arXiv:0904.4448 [nucl-ex].

51] P. Norton, Rept. Prog. Phys. 66, 1253 (2003).

52] K. Rith, (2014), arXiv:1402.5000 [hep-ex].
[53] L. Frankfurt and M. Strikman, Nucl. Phys. B250, 143 (1985).
[54] R. Arnold, P. E. Bosted, C. Chang, J. Gomez, A. Katramatou, et al., Phys. Rev. Lett. 52, 727 (1984).
[55] A. Bodek, M. Breidenbach, D. Dubin, J. Elias, J. I. Friedman, et al., Phys.Rev. D20, 1471 (1979).
[56] A. Bodek and J. Ritchie, Phys.Rev. D23, 1070 (1981).
[57] A. Bodek and J. Ritchie, Phys.Rev. D24, 1400 (1981).
[58] L. Frankfurt and M. Strikman, Int.J.Mod.Phys. E21, 1230002 (2012), arXiv:1203.5278 [hep-ph].
[59] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, et al., Phys. Rev. D89, 033009 (2014), arXiv:1302.6246 [hep-ph].
[60] Y. L. Dokshitzer, Sov. Phys. JETP 46, 641 (1977).
[61] V. Gribov and L. Lipatov, Sov. J. Nucl. Phys. 15, 438 (1972).
[62] G. Altarelli and G. Parisi, Nucl. Phys. B126, 298 (1977).
[63] B. L. Combridge, J. Kripfganz, and J. Ranft, Phys. Lett. B70, 234 (1977).
[64] S. D. Ellis, Z. Kunszt, and D. E. Soper, Phys. Rev. Lett. 64, 2121 (1990).
[65] R. K. Ellis and J. Sexton, Nucl. Phys. B269, 445 (1986).
[66] G. Aad et al. (ATLAS Collaboration), Phys. Rev. D86, 014022 (2012), arXiv:1112.6297 [hep-ex].
[67] A. Accardi, J. Albacete, M. Anselmino, N. Armesto, E. Aschenauer, et al., (2012), arXiv:1212.1701 [nucl-ex].
[68] G. P. Lepage, S. J. Brodsky, T. Huang, and P. Mackenzie, in Particles and Fields 2, edited by A. Z. Capri and A. N. Kamal (1982).
[69] O. Artilles and M. Sargsian, private communication.