Nucleon Structure from Lattice QCD Calculations

Jian-Wei Qiu

1Theory Center, Jefferson Lab, 12000 Jefferson Avenue, Newport News, VA 23606, U.S.A.
E-mail: jqiu@jlab.org

(Received February 15, 2019)

Parton distribution and correlation functions describe the relation between a hadron and the quarks and gluons (or collectively, the partons) within it, and carry rich information on hadron’s partonic structure that cannot be calculated by QCD perturbation theory. In this talk, I will review what lattice QCD can and cannot do for calculating the parton distribution and correlation functions, and the new ideas and efforts around the world to explore nucleon structure from lattice QCD calculations by combining the strength of both lattice QCD and perturbative QCD in such a way that is complementary to our on-going effort to extract these fundamental functions from experimental data.

KEYWORDS: Nucleon Structure, Lattice QCD, Perturbative QCD, Factorization

1. Introduction

The proton and neutron (or collectively, known as the nucleon) are the fundamental building blocks of all atomic nuclei and make up essentially all the visible matter in the universe, including the stars, the planets, and us. The nucleon emerges as a strongly interacting, relativistic bound state of quarks and gluons in Quantum Chromodynamics (QCD), and has a complex internal structure only beginning to be revealed in modern experiments and lattice QCD calculations. Both theory and technology have now reached a point where human is capable of exploring the inner dynamics and structure of nucleons and nuclei at the sub-femtometer distance, which provides new opportunities for us to understand ourselves and the visible world around us in a deeper and more fundamental level and is expected to lead to a new emerging science of nuclear femtography.

QCD, a dynamical theory of quarks and gluons, is believed to be the theory of the strong interaction physics and is responsible for the known macroscopic properties of nucleons, such as mass and spin, as well as their rich, mysterious and largely unknown microscopic internal structure in terms of the properties and interactions of quarks and gluons in QCD. With its color confinement - the defining property of QCD, no modern detector has ever seen quarks and gluons in isolation, and it has been an unprecedented intellectual challenge to explore and to quantify the nucleon’s internal structure without being able to see quarks and gluons directly.

Unlike the well-known atomic, molecular or crystal structure, which is often presented in terms of a “still picture” showing the relative location of various nuclei, the internal structure of nucleon cannot be described by any “still picture”. This is because nuclei are so small comparing to the size of atom, and so heavy that their motion is so much slower than the speed of light, while the motion of quarks and gluons is fully relativistic, their numbers are constantly changing due to quantum fluctuations inside the hadron, and their color is quantum mechanically entangled. Consequently, we quantify the nucleon’s partonic structure in QCD in terms of “quantum probabilities” to find quarks, gluons and their correlations, which are defined in terms of hadron matrix elements of quark ($\bar{q}, q$) and gluon ($A^\mu$) operators: $\langle P, S|O(\bar{q}, q, A)|P, S\rangle$ with hadron momentum $P$ and spin $S$. Although these “quantum probabilities” or hadron matrix elements are well defined in QCD, none of them is a direct physical observable, exactly due to the color confinement and the fact that no quarks and gluons could
be observed directly in a physical detector. That is, we have to identify “probe(s)” that can “see” the
quarks and gluons inside the nucleon in order to measure these quantum probabilities to quantify the
nucleon’s partonic structure or internal landscape.

Fortunately, QCD has another equally important and fundamental defining property, the asymptotic
freedom - the strong force effectively is weak if it is probed at a sufficiently short-distance. It is
the asymptotic freedom that makes it possible for us to develop the powerful theoretical formalism,
known as QCD factorization [1] that factorizes a physical cross section with measured hadron(s) in
high energy collision into a product of a calculable partonic scattering of quark(s) and/or gluon(s) at a
sub-femtometer scale and a set of fundamental and universal matching functions to link the scattering
quark(s) and/or gluon(s) to the identified hadron(s) within well-defined approximations. For example,
inclusive cross section of electron (e) - proton (p) deeply inelastic scattering (DIS) has one identified
hadron - the proton of momentum $p$ and could be factorized as,

$$\sigma_{ep-e'p}^{DIS}(x_B, Q^2) = \sum_{i=q,\bar{q},g} \int_{x_B}^{1} \frac{dx}{x} \hat{\sigma}_{e^+e^- \rightarrow X}(\frac{x_B}{x}, \frac{Q^2}{\mu^2}; \alpha_s(\mu^2)) f_i/p(x, \mu^2) + O(\frac{1}{(Q\mu)^2}), \quad (1)$$

where $x_B = Q^2/2p \cdot q$ and $Q^2 = -q^2$ with the $q$ as the momentum transfer between the scattering
electron and proton, and $R \sim 1/A_{QCD} \sim 1$ fm is the hadron radius representing the order of
nonperturbative hadronic scale. The factorization formalism in Eq. (1) presents an indirect and approxi-
mate way to probe quarks and gluons in electron-proton scattering when the momentum transfer
$Q^2 >> 1/R^2$, in terms of the “probe”: $\hat{\sigma}_{e^+e^- \rightarrow X}$ - a perturbatively calculable electron scattering off a
parton of flavor $i = q, \bar{q}, g$ (quark, antiquark, gluon, respectively), the universal matching functions:
$f_i/p(x, \mu^2)$ - the probability distributions to find the parton of flavor $i$ in the proton $p$ carrying its
momentum fraction between $x$ and $x + dx$, probed at the factorization scale $\mu$, and a controlled “cor-
rection” suppressed by the power of the large momentum transfer $Q$. It is the precision, for which we
can calculate the $\hat{\sigma}_{e^+e^- \rightarrow X}$ in QCD perturbation theory at the scale $Q$ - a consequence of the
asymptotic freedom, that provides us the well-controlled short-distance “probe(s)” to effectively “see” the
quark(s) and/or gluon(s) at a sub-femtometer distance ($\sim 1/Q$); it is the matching functions that link the
quark and/or gluon at such a short distance inside the observed hadron to provide the rich informa-
tion on the hadron’s partonic structure; and it is the universality of these matching functions, extracted
from one or few experimental measurements and then used to predict and to be tested in many other
experiments, that ensures the predictive power of QCD. With the factorization formalisms available
for various observables, QCD has been extremely successful in interpreting all available data from
high energy scatterings with probing distance less than 0.1 fm (or equivalently, with the momentum
transfer in the scattering larger than 2 GeV), which has provided us the confidence and the tools to
discover the Higgs particle and to explore new physics beyond the Standard Model of particle physics
in high energy hadronic collisions [2].

The predictive power of QCD factorization formalisms, like the one in Eq. (1), relies on the pre-
cision of these universal matching functions, which are defined in terms of hadron matrix el-
ments of quark-gluon operators and encoded with rich information on hadron’s partonic structure.
The $f_i/p(x, \mu^2)$ in Eq. (1), known as the parton distribution functions (PDFs), are defined as

$$f_{q/p}(x, \mu^2) = \int \frac{dp^+ d\xi^-}{2\pi} e^{-i xp^+ \xi^-} \langle p|\bar{q}(\xi^-)|\psi_q(0)p\rangle \psi_{q}(0)|p\rangle \exp \left\{ -ig \int_0^{\xi^-} d\eta^- A^+(\eta^-) \right\} \quad (2)$$

where the light-cone components: $\nu^b \equiv (v^0 \pm v^3)/\sqrt{2}$ are defined for any four-vector $\nu^\mu = (v^0, v_L, v^3)$,
and might be the most fundamental and the best studied matching functions. In the framework of
QCD factorization, enormous efforts have been devoted to determine PDFs and their uncertainties in
terms QCD global analysis of all existing high energy scattering data by many collaborations around
the world, including MMHT [3], CT [4], NNPDF [5], HERAPDF [6], and JAM [7]. Although we
have improved our knowledge of PDFs tremendously, they still have large uncertainties in both large and small $x$ region, and sometimes, the uncertainties of PDFs become the main sources of systematic errors for comparison between the best theoretical calculations available and the better and more precise experimental data from the LHC. With the hadronic matrix elements well-defined in QCD, it is not only very natural, but also critically important to ask if we can calculate these PDFs and other partonic correlation functions directly in QCD, or at least in Lattice QCD (LQCD), which is by far the most reliable and theoretically justified approach to calculate the nonperturbative quantities in QCD. Unfortunately, it is, if not impossible, very difficult to calculate PDFs and other partonic correlation functions directly in LQCD because the partonic operators of these hadron matrix elements live on the light-cone, as shown in Eq. (2), and are time-dependent with the time defined in Minkowski space, while LQCD calculations are performed in Euclidean space with the Minkowski time $t$ analytically continued to a Euclidean time, $\tau = i t$.

LQCD calculation of hadronic matrix elements is based on the QCD path integral formalism evaluated in a discretized finite volume of Euclidean space-time with the integration of all field configurations done by the Monte Carlo method [8]. It requires a minimum input to define the QCD action, such as the overall hadronic scale and quark masses, which could be fixed by a few known baryon and meson masses. Due to the computation resources and technical challenges, early LQCD efforts for determining hadron’s partonic structure were limited to calculations of local matrix elements. With the improvement of computation power and remarkable new ideas in recent years, there has been a tremendous growth of LQCD efforts for calculating nonlocal matrix elements, from which parton distribution and correlation functions could be extracted [8, 9]. In the rest of this writeup for my talk, I will review briefly what lattice QCD can and cannot do for calculating the parton distribution and correlation functions that provide rich information on hadron’s partonic structure. I will first briefly summarize the early LQCD effort to study hadron structure in terms of the calculation of local matrix elements in the next section. In the Sec. 3, I will focus on the fast developing world effort for extracting the parton distribution and correlation functions by calculating non-local matrix elements in LQCD, and finally, I will provide the summary and outlook in the Sec. 4.

### 2. Hadron structure in terms of local matrix elements

Hadron structure with relativistic quarks and gluons is dynamical and nonlocal. Parton distribution and correlation functions, which carry the structure information, cannot be directly calculated in LQCD defined with a Euclidean space-time. Historically, LQCD has been used to calculate hadron form factors of local currents [10] and Mellin moments of PDFs, which are given in terms of hadronic matrix elements of local twist-2 operators of quark and gluon fields [8].

In principle, we could determine the nonlocal correlation of quark and gluon fields and reconstruct PDFs if we are able to calculate a sufficient number of the moments in LQCD. The Mellin moments of PDFs are defined as

$$\langle x^n \rangle_{\mu^2} \equiv \int_0^1 dx x^n f(x, \mu^2), \quad f = \{u, \bar{u}, d, \bar{d}, s, \bar{s}, \ldots; g\}, \quad (3)$$

with $q^k = q = \bar{q}$ for quark flavors and $\Delta f$ for helicity distribution functions. LQCD calculation has done a good job in determining the lower moments of PDFs. In Fig. 1, we compare the lower moments of PDFs determined by QCD global fits and LQCD calculations. In the left plot, we show the first moment of unpolarized PDFs, which is equal to the momentum fraction carried by the quark and gluon; and in the right plot, we present the zeroth moment of quark helicity distributions, which is equal to the helicity fraction of a fast moving proton carried by the quarks. With the precise date from experiments and a wide range of collision energies, QCD global fits have provided much accurate determination of the momentum fractions carried by quarks and gluons, which effectively set the benchmarks for LQCD calculations. On the other hand, LQCD determination of quark helicity
contribution to the proton spin is comparable to what have been determined by QCD global fits. With less data available for providing information on quark transversity distribution, precise determination of quark tensor charge from LQCD calculations has helped to better determine the x-dependence of quark transversity distribution [11].

However, due to the power divergent mixing between twist-2 operators on the finite lattice, it is extremely difficult and effectively unpractical to determine the higher moments of PDFs. In practice, LQCD calculations of the Mellin moments of unpolarized PDFs are limited to the lowest three moments with the first moment results in Fig. 1, and the other two moments in Refs. [12,13], respectively. With the limited information on the lowest three moments, it is not realistic to determine the x-dependent PDFs to be comparable with those from QCD global fits [14].

3. Hadron structure from non-local matrix elements

Over the years, various ideas have been proposed to extract x-dependent PDFs from LQCD calculations, such as the path-integral formulation of the deep-inelastic scattering hadronic tensor [15, 16] and the inversion method [17]. Recently, Ji [18] introduced a set of LQCD-calculable quasi-PDFs,

\[ \tilde{q}(\vec{x}, \mu^2, P_z) = \int \frac{dP_z}{2\pi} e^{-i\vec{P}_z \vec{x}} \left\langle \frac{P}{P_z} \right\rangle \frac{\gamma_z}{2} \exp \left\{ -ig \int_0^{\xi_z} dy \langle \tilde{z} \rangle \right\} \psi_q \left( \frac{-\xi_z}{2} \right) |P \rangle \]  

(4)

for quasi-quark distribution, and argued that the quasi-PDFs of hadron momentum \( P_z \) become corresponding PDFs when \( P_z \) is boosted to infinity. Recognizing the fact that a separation in \( \xi_z \) at an equal time becomes a separation in \( \xi^- \) while \( \xi^+ = 0 \) when \( P_z \rightarrow \infty \), and the idea to calculate hadron matrix elements of parton correlation along the z-direction in LQCD is novel. Taking the limit, \( P_z \rightarrow \infty \) is difficult to achieve in LQCD calculation due to the limited size of lattice spacing. A perturbative matching of quasi-PDFs at a finite \( P_z \) to the standard light-cone PDFs was proposed [18],

\[ \tilde{q}(\vec{x}, \mu^2, P_z) = \int_\xi^1 dy Z \left( \frac{\vec{x}}{y}, \frac{\mu}{P_z} \right) q(y, \mu^2) + O \left( \frac{\Lambda^2}{P_z^2}, \frac{M^2}{P_z^2} \right) \]  

(5)

with \( q(y, \mu^2) \) being the standard light-cone quark distribution and perturbatively calculable matching coefficient, \( Z \) and power corrections in \( 1/P_z \). It was noted recently that instead of \( 1/P_z^2 \), the power corrections could be proportional to \( 1/[x^2(1 - x)P_z^2] \), which could be significantly enhanced at a large \( x \) or small \( x \) [19]. Comparing Eqs. (2) and (4), one recognizes that the operator defining the quasi-quark distribution is not a twist-2 operator, and consequently, unlike the quark distribution,
quasi-quark distribution is not boost invariant and its dimension-three operator leads to a ultraviolet (UV) linear power divergence, which makes the renormalization of the LQCD calculation much more challenging [20, 21].

Regardless the difficulties, tremendous efforts have been investigated by both LQCD and perturbative QCD (PQCD) communities to study the quasi-PDFs in recent years, and important progresses have been made in our understanding of the relationship between the PDFs and quasi-PDFs, and their determination from LQCD [8, 9, 22]. The power UV divergence of quasi-quark distributions were proved to be multiplicative renormalizable by three groups [23–25]. While the operator defining the quasi-gluon distribution has a dimension of four, QCD color gauge invariance ensures that its maximum power of UV divergence is linear, which were proved to be multiplicative renormalizable by two groups with completely different methods [26, 27]. With the proof of QCD collinear factorization of hadronic matrix elements of operators defining quasi-PDFs in terms of PDFs [28], extracting PDFs from LQCD calculated hadronic matrix elements of quasi-parton operators has a solid theoretical foundation, while technical challenges, such as nonperturbative renormalization of power divergence and large hadron momentum in LQCD calculations, still exist.

Meanwhile, a number of other approaches to extract PDFs from LQCD calculations were also proposed, such as the good “lattice cross section” - a QCD factorization based general approach to calculate PDFs in LQCD indirectly [28, 29], the pseudo-PDFs [30, 31], and the “OPE without OPE” approach [32]. The good “lattice cross section” (LCS) is defined as single hadron matrix elements that satisfy (1) calculable in Euclidean-space LQCD, (2) renormalizable for UV divergences to ensure a reliable continue limit, and (3) factorizable to PDFs with infrared-safe matching coefficients. It is the (3)-factorization that relates the desired PDFs to the LQCD calculable LCSs, just like how PDFs are related to the factorizable hadronic cross sections. To overcome the power UV divergence of the operator defining the quasi-PDFs, a set of current-current correlators were introduced for good LCSs [29]. With renormalized or conserved currents, the UV renormalization of the good LCSs is trivial; and with many choices and combinations of currents, this approach could help construct many more LQCD calculable quantities to provide more information and constraints to determine the hadron’s partonic structure. Similar to the extraction of PDFs from experimental data of factorizable and measurable hadronic cross sections, PDFs could be extracted from QCD global analysis of data generated by LQCD calculations of good LCSs.

![Fig. 2. Left: Scale dependence of Pseudo-quark distribution function [31]. Right: Effect of one-loop matching between quasi-quark and quark distribution function [35].](image)

Pseudo-PDFs approach was introduced to better control the power UV divergence of quasi-PDFs in LQCD calculation, and to avoid the operator mixing of the quasi-quark distribution under the renormalization [31]. A pseudo-quark distribution is defined as,

$$\mathcal{P}(x, \xi^2) \equiv \int \frac{d\omega}{2\pi} e^{-i\omega x} \frac{M_{p=p^0}(\omega, \xi^2)}{M_{p=p^0}(0, \xi^2)} \quad (6)$$
where the hadronic matrix element $M_p(\omega, \xi^2)$ with the Ioffe time $\omega = p \cdot \xi$ is defined as

$$M^Q(\omega, \xi^2) \equiv \langle p|\bar{u}(\xi)\gamma^\mu \exp\left\{-ig \int_0^\xi d\lambda \nu \cdot A(\lambda\nu)\right\}\psi(0)|p \rangle \approx 2p^\mu M_p(\omega, \xi^2).$$

With the ratio in Eq. (6), the power UV divergence of the matrix elements cancels since the UV divergence is multiplicatively renormalizable and is insensitive to the details of the long-distance hadron states. The $\xi^2$ is the hard scale at which the distribution is probed and its dependence should obey the DGLAP evolution, which is clearly demonstrated in the left plot in Fig. 2.

**Fig. 3.** LQCD determination of $u(x) - d(x)$ (left [35]) and $\Delta u(x) - \Delta d(x)$ (right [37]) by the LP3 collaboration.

Significant achievements in the LQCD calculations of quasi-PDFs have been made in recent years [33–37]. At a finite hadron momentum $P_z$, as shown in the right plot in Fig. 2, the one-loop matching coefficient plays a very significant role [35]. The latest and the state of art extraction of unpolarized distribution $u(x) - d(x)$, as well as polarized helicity distribution, $\Delta u(x) - \Delta d(x)$, from LP3 collaboration are presented in Fig. 3, while the same combination of distributions from LQCD calculations of ETMC collaboration are shown in Fig. 4. References to many additional LQCD calculations, the current status and challenges for extracting reliable PDFs and other parton correlation functions from such LQCD calculations can be found in Refs. [8, 9].

**Fig. 4.** LQCD determination of $u(x) - d(x)$ (left) and $\Delta u(x) - \Delta d(x)$ (right) by the LP3 collaboration [34].

The good “lattice cross section” approach with current-current operators trades the speed of LQCD calculations to a better control of UV renormalization and many more calculable and factorizable observables, by calculating four-point functions in LQCD instead of the three-point functions in the case of calculating quasi- or pseudo-PDFs [29]. With different currents and QCD symmetries, we could construct various combinations of these currents to have access to different distribution and correlation functions. For example, under the parity and time-reversal invariance of QCD, pion matrix element of the following combination of a vector current ($V$) and an axial-vector current ($A$) has to be antisymmetric in the Lorentz indices of these two currents,

$$\frac{1}{2} \left[ \sigma_{\mu V}(\xi, p) + \sigma_{\mu A}(\xi, p) \right] = e^{\mu\nu} \xi_\alpha p_\beta T_1(v, \xi^2) + (p^\mu \xi^\nu - \xi^\mu p^\nu) T_2(v, \xi^2),$$

where

$$T_1(v, \xi^2) = \frac{1}{2} \left\{ \epsilon_1(\xi, \eta), \epsilon_2(\xi, \eta), \epsilon_3(\xi, \eta) \right\},$$

$$T_2(v, \xi^2) = \frac{i}{2} \left\{ \epsilon_1(\xi, \eta), \epsilon_2(\xi, \eta), \epsilon_3(\xi, \eta) \right\},$$

and

$$\epsilon_k(\xi, \eta) = \epsilon_j_k(\xi, \eta) = \epsilon^j_k(\xi, \eta) = \langle 0|J_\mu^k(\xi) J_\nu^\mu(\eta)|0 \rangle,$$
where the LCS, \( \sigma^\mu_{ij} \equiv \xi_1^4(p(p)| \mathbb{P}_i^\mu(\xi/2) \mathbb{P}_j^\mu(-\xi/2)|p(p)) \) with \( i, j = V, A \), the Ioffe time, \( \omega = p \cdot \xi \), and \( T_i \) with \( i = 1, 2 \) are functions of Lorentz scalars, \( \omega \) and \( \xi^2 \), which can be factorized to PDFs as [29, 38],

\[
T_i(\omega, \xi^2) = \sum_{a,q,g} \int_0^1 \frac{dx}{x} f_a(x, \mu^2)C_i^a(x \omega, \xi^2, \mu^2) + O(\xi^2 \Lambda_{QCD}^2)
\]  

with the lowest order (LO) coefficients \( C_1^{(0)}(x \omega, \xi^2) = 2 \bar{\alpha}_s \cos(x \omega) \) and \( C_2^{(0)}(x \omega, \xi^2) = 0 \), which gives,

\[
\bar{T}_i(x, \xi^2) = \int \frac{d\omega}{2\pi} e^{-ix\omega}T_i(\omega, \xi^2) \approx \frac{1}{\pi^2} \{ q(x) - \bar{q}(x) \} = \frac{1}{\pi^2} \bar{q}_i(x).
\]

That is, the LQCD calculated LCS in Eq. (8) is directly proportional to the pion’s valence quark distribution. In the left plot in Fig. 5, LQCD calculated LCS in Eq. (8) is shown for several hadron momenta. On the right, the valence quark momentum distribution \( \bar{x}q_i(x) \), extracted from LQCD calculation with the LO kernel is compared with those extracted from data or calculated in models, and the LO result is encouraging.

Fig. 5. Left: LQCD determination of pion matrix element in Eq. (8) [38]. Right: Comparison of valence quark distribution extracted from LQCD calculation with those obtained with other approaches [38].

4. Summary and outlook

I have briefly reviewed what LQCD can and cannot do for calculating the parton distribution and correlation functions, and the new approaches to explore nucleon structure from LQCD calculations. I presented the recent state-of-art achievements in determine PDFs by the joint effort of LQCD and PQCD communities. LQCD calculations of hadron structure are complementary to experimental measurements, since LQCD can have access to the kinematic regime, such as the large \( x \), where is hard to reach by experimental measurements; and LQCD can study partonic structure of hadrons, such as free neuron, that are hard to do experiments with. The results of initial exploratory LQCD calculations of parton distribution and correlation functions are very encouraging, while more works are obviously needed.

Acknowledgement: This work is supported in part by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Award No. DE-AC05-06OR23177, within the framework of the TMD Topical Collaboration.

References
[1] J. C. Collins, D. E. Soper and G. F. Sterman, Adv. Ser. Direct. High Energy Phys. 5, 1 (1989) [arXiv:hep-ph/0409313].
[2] N. Brambilla et al., Eur. Phys. J. C 74, no. 10, 2981 (2014) [arXiv:1404.3723 [hep-ph]].
