SciPost Physics

Submission

All About the Neutron from Lattice QCD

Rajan Gupta[1*]

1 Los Alamos National Laboratory, Theoretical Division T-2, Los Alamos, New Mexico 87545, USA
* CorrespondingAuthor rajan@lanl.gov

December 22, 2022

Abstract

I describe how simulations of lattice QCD using the path integral formulation provide the two basic quantum mechanical properties of QCD, its ground state in which correlation functions are calculated, and Fock state wavefunctions between which matrix elements of operators are calculated. Both constructs are stochastic, so unfortunately one gets no intuitive picture or even qualitative understanding of what they look like, nevertheless they contain and display all the subtleties of the quantum field theory. Today, these simulations provide many quantities that are impacting phenomenology and experiments. I illustrate the methods and the steps in the analysis using, as examples, three observables: the isovector charges of the nucleon, the contribution of the quark's intrinsic spin to the nucleon spin, and the pion-nucleon sigma term.

1 Lattice QCD

The field of lattice QCD (LQCD)—the theoretical rigorous path integral formulation of QCD discretized on a 4D hypercubic grid by Wilson, and formulated to provide non-perturbative predictions [1]—has come of age. This formalism converts quantum field theories into statistical mechanics systems, for example, the 3+1 dimensional QCD in Minkowski time becomes, after a Wick rotation to Euclidean time, a classical system of gluon and quark fields on a 3+1 dimensional lattice in Euclidean time (see Fig. 1). Numerical simulations of it [2] are providing first principle results with control over all systematic uncertainties for a large number of physical observables that elucidate the standard model and probe physics beyond it. The Flavor Lattice Averaging Group (FLAG [3]) provides a community based evaluation of quantities that are considered robust [4, 5][1] With improvements in numerical algorithms and increasing computing resources, the errors on these quantities are being reduced steadily, and many more quantities are being added to the list. At the same time, the need for new ideas for a big leap forward is also evident. In this writeup, I will present an ideosyncratic mixture of topics, starting with explaining what simulations of LQCD give us, and then highlight successes and, at the same time, the need for new ideas for subpercent precision predictions of the properties of nucleons.
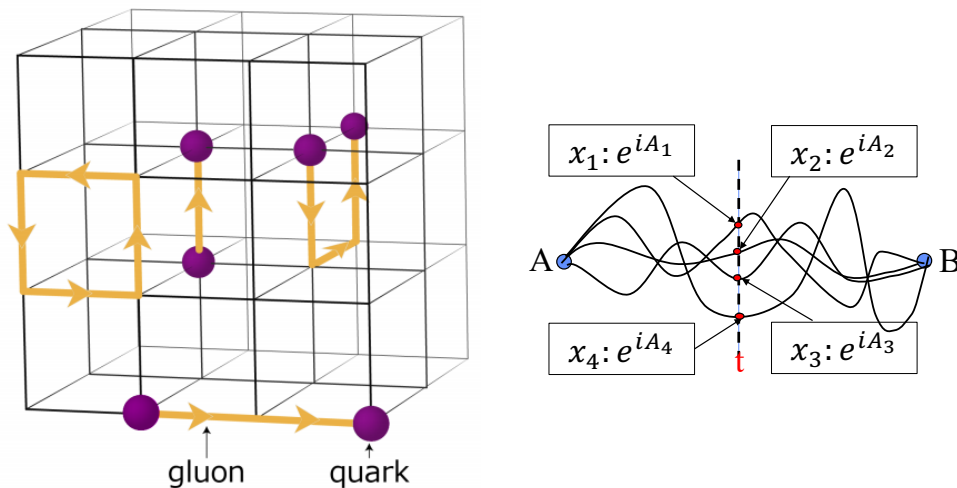
Figure 1: (Left) Discretization of QCD on a hypercubic lattice with quark fields placed on sites and the gluon fields $A_\mu(x)$ on directed gauge links via SU3 matrices $\mathcal{U}_{x,\mu} = e^{iagA_\mu^a\lambda^a}$ where g is the gauge coupling and λ^a are the Gell-Mann matrices. The lattice theory preserves the gauge invariance of the continuum theory. (Right) Illustration of the path intergral formulation of quantum mechanics of a particle moving between points A and B. Each path i has coordinate x_i at time t and is weighted by e^{iA_i}, where A_i is the action. All possible paths connecting A and B contribute.

Let me begin with a brief recap of the path integral formulation of a particle moving in time. Figure 1 (right) illustrates some of the paths that contribute to the quantum mechanical amplitude for the particle to go from point A to point B. In fact all paths one can draw between those two points contribute. Each path has a weight e^{iA}, where A is the action along that path. The paths interfere and, typically, the one with the smallest action gives the largest contribution. In the classical limit, this path converges to that predicted by Newton's equations. The mean value of an observable, say the position x at a prescribed time t is given by

---

[1]The chapter on Nucleon Matrix Eelements in these reports provides an introduction to the many issues relevant to the calculation of nucleon matrix elements discussed below and contains an extensive list of references for the interested reader. Very often I will just refer to the FLAG reports with the assumption that a list of pertinent references are already collected there.

the expectation value $\sum_n x_n e^{iA_n}/\sum_i e^{iA_n}$ where x_n are the values of the position at time t on each path n. Using this example, I now motivate how simulations of lattice QCD give us the analogues of the "paths", the correlation functions corresponding to observables such as x, and the wavefunctions within which matrix elements of observables can be calculated.

In simulations of lattice QCD (LQCD), and of gauge field theories in Euclidean time in general, the analogue of the paths are gauge configurations. Each LQCD configuration $\mathcal{C}_i$ is a specification of the 12 independent entries in each SU(3) matrix, $\mathcal{U}_{x,\mu}$, assigned to each link of the lattice (see Fig. 1. The $\mathcal{C}_i$ have a weight e^{-A_i}, where A_i is the Euclidean QCD action calculated on that configuration. It is a functional of all the $\mathcal{U}_{x,\mu}$. Note, the quark fields are formally integrated out as discussed below, leaving only gauge fields as dynamical variables. Configurations are generated using Markov Chain Monte Carlo Methods with importance sampling and the Metropolis accept/reject step [2]. Conceptually, this algorithm for the generation of the $\mathcal{C}_i$ is the same as used in the classical simulations of spin models, however, simulations are computationally expensive because on a 100^4 lattice, there are $4 \times 12 \times 10^8$ independent variables that specify a $\mathcal{C}_i$ (entries in all the SU(3) link matrices) and evaluating the action A is expensive. The full set of the $\mathcal{C}_i$ (called an ensemble) and their associated A_i provides us the fully quantum mechanical ground state of the field theory, *albeit* stochastically since only a finite $\mathcal{C}_i$ are sampled in practice. The action A is characterized by the input parameters of the simulations: quark masses m_i with $i \in \{u,d,s,c\}$ flavors, lattice spacing a (equivalently the gauge coupling via dimensional transmutation) [2], and the lattice volume $L^3 \times T$.

The full set of configurations are the same independent of the action, $A(m_i, a)$ and depend only on the number of gauge links and the values they can take, *i.e.* on the lattice volume. There are ∞^2 of them: infinite number of variables in the limit the volume $L^3 \times T \to \infty$, and each variable is continuous valued between $\{-1, 1\}$. If even a significant subset of these were needed to calculate observables, precision would not be achieved.

2 Correlation Functions and Observables

Observables O are obtained from ensemble averages of correlation functions Γ_α measured on the $\mathcal{C}_i$. These are again given by the expectation values $\sum_i \Gamma_\alpha e^{-A_i}/\sum_i e^{-A_i}$. What saves us from not having to consider the ∞^2 configurations is that the weight e^{-A} is so very highly peaked about the minimum of A that $10^3 - 10^7$ (depending on the observable) importance sampled and statistically independent configurations suffice to yield expectation values with sufficient precision. The location of the peak of the distribution (minimum of A) changes with the input parameters, i.e., with $A(m_i, a, L)$.

In practice, data (expectation values of correlation functions) are obtained on many ensembles with different $\{m_i, a, L\}$ so that the limits $a \to 0$, m_i to their physical values set by the experimental values of the masses M_π, M_N, M_Ω and M_D, and $L \to \infty$ can be taken to obtain physical results. Typical of current simulations, m_s and m_c, being sufficiently heavy, are already tuned to their physical values before starting production, and only $m_{u,d}$ are varied. In the isospin symmetric limit, $m_u = m_d$, it is typical to represent the common light quark mass $\overline{m}_{ud}$ by the corresponding value of M_π, again tuned before starting production runs. Today, we can perform simulations at $M_\pi = 135$ MeV, but very often data are also obtained at a number of heavier values of $\overline{m}_{ud}$, (equivalently M_π) and then extrapolated to $M_\pi = 135$ MeV using ansatz motived by chiral perturbation theory. These ideas will be illustrated by the calculations/results reviewed later.

An essential simplification, in fact one that allows simulations of QCD on classical computers in the first place, is that the fermion action,

$$A_F = \bar{\psi} D \psi = \bar{\psi}(\gamma_\mu A_\mu + m_f)\psi \to \bar{\psi}(x + a\hat{\mu})(\gamma_\mu U_{x,\mu})\psi(x) + m_f \bar{\psi}(x)\psi(x), \qquad (1)$$

is bilinear in the quark fields where D is the Dirac operator. They can be integrated out exactly from the path integral but impact the $A(m_i, a, L)$, thus the position of, and fluctuations about, the peak of the distribution specifying the ground state. The location of the peak in the distribution e^{-A_i} includes the contribution of the determinant of the Dirac action, $\text{Det}\mathcal{D}_f$, for each flavor f (which requires only the $U_{x,\mu}$ for calculation), i.e., the Boltzmann weight becomes $\prod_f (\text{Det}\mathcal{D}_f) e^{-A_G} = e^{-A_G + \sum_f \text{LnDet}\mathcal{D}_f}$ where A_G is the gauge action. Including this highly non-local determinant in the simulations makes them expensive but does not pose a formal obstruction.

I now illustrate how correlation functions of operators composed of quark fields are calculated even though we have integrated them out formally. Consider 2- and 3-point correlation functions involving time-ordered product $\mathcal{T}$ of the pion and the axial current:

$$\Gamma_\pi^2 = \langle\, \mathcal{T}(\bar{d}\gamma_5 u|_\tau\ \bar{u}\gamma_5 d|_0)\,\rangle; \quad \Gamma_\pi^3 = \langle\, \mathcal{T}(\bar{d}\gamma_5 d|_\tau\ \bar{d}\gamma_\mu\gamma_5 u|_t\ \bar{u}\gamma_5 d|_0)\rangle, \tag{2}$$

with the assumption that all three operators have been projected to zero momentum for simplicity, thus leaving only the time index. The notation $\langle \cdots \rangle$ implies ensemble average over the $\mathcal{C}_i$. At the same time as integrating out the quarks, one can perform a Wick contraction of the fields to get

$$\Gamma_\pi^2 = \langle\, S_F(0,\tau)\gamma_5 S_F(\tau,0)\gamma_5\,\rangle; \quad \Gamma_\pi^3 = \langle\, S_F(0,\tau)\gamma_5 S_F(\tau,t)\gamma_\mu\gamma_5 S_F(t,0)\gamma_5\rangle \tag{3}$$

where $S_F = \mathcal{D}^{-1}$ is the Feynman propagator given by the inverse of the Dirac matrix on that configuration. (This inverse is calculated using iterative Krylov solvers). Now using the hermiticity property of the Dirac action and its inverse, $S_F(0,\tau) = \gamma_5 S_F^\dagger(\tau,0)\gamma_5$, we get

$$\Gamma_\pi^2 = \langle S_F(0,\tau) S_F^\dagger(0,\tau)\rangle; \quad \Gamma_\pi^3 = \langle S_F(0,\tau) S_F^\dagger(\tau,t)\gamma_\mu\gamma_5 S_F^\dagger(0,t)\rangle \tag{4}$$

Thus the Wick contraction has replaced quark fields in terms of $S_F = \mathcal{D}^{-1}$, which depend only on the gauge links. *In short, both in the generation of the configurations and in the calculation of correlation functions, the quark fields have been removed exactly.* The expressions in Eq. 4 for the pion are the quark line diagrams shown in Fig. 2, whose expectation value gives the desired non-perturbative correlation functions.
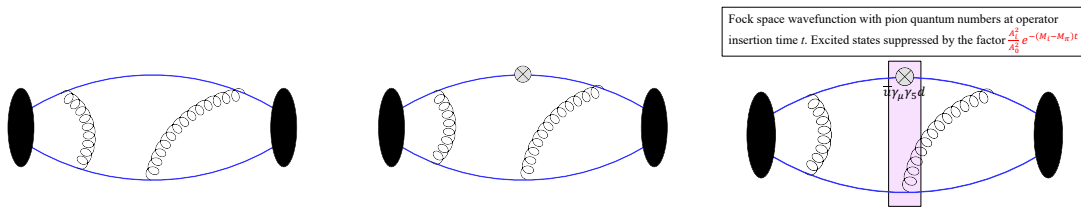
Figure 2: Illustration of quark-line diagrams for 2-point (left) and 3-point for the pion (middle and right). The gluon lines are just for illustration and to remind the reader that all orders of gluon exchanges are implicit in these diagrams. (Right) The axial current, shown by $\otimes$, is inserted at intermediate Euclidean time t and with momentum $\vec{q}$. The ensemble average in LQCD simulations creates the stocastic Fock state wavefunction, indicated by the pink band at time t, and the operator causes transitions between the various "pion" states present, i.e., gives us the matrix elements. Using Eq. (5), we extract the pion's axial form factors from this correlation function.

The question I hope you are dying to ask is how does the quark propagator, S_F, calculated on a given configuration and combined to form the quark-line diagrams shown in Fig. 2, know anything about the non-perturbative propagation of the pion or any of the thousands of possible states of QCD, the analogous quark-line diagrams for which are obtained by simply

changing the interpolating operators. As already explained, the interpolating operators create states with given quantum numbers, for example $\bar{d}\gamma_5 u$ or $\bar{d}\gamma_0\gamma_5 u$ for the pion, that are propagated in time by the transfer matrix. The non-perturbative properties of the state and its dynamics arise from the coherent addition of those gauge fluctuations on each configuration that correspond to a pion propagating. The miracle (actually a dictate of quantum field theory) is that these fluctuations get picked up and add in the ensemble average. In short, the frothing vacuum has all possible fluctuations present, and the ensemble average picks up those that conform to the quantum numbers of the created state.

An interesting aside is that $\Gamma_\pi^2 = \langle|S_F(0,\tau)|^2\rangle$ in Eq. 4 is a positive definite quantity. So you may ask–how can averaging over only the small "working ensemble" give a precise unbiased result? The answer lies in the fact that configurations importance sampled according to the Boltzmann weight e^{-A} do an excellent job of approximating the full ensemble (path integral). Clearly, to improve statistical precision, one needs to enlarge the "working ensemble".

Now we come to the last part of the introduction to LQCD–how does one get physics from correlation functions such as those in Eq. 4? For this we invoke the spectral decomposition of Γ_π^2 and Γ_π^3, i.e., inserting a complete set of states at each intermediate time step, with the evolution between steps given by the transfer matrix. The result is

$$\Gamma_\pi^2 = \sum_i |\langle 0|\hat{\pi}_i|\pi\rangle|^2 \; e^{-E_i\tau}; \quad \Gamma_\pi^3 = \sum_{i,j} \langle 0|\hat{\pi}|\pi_i\rangle^* \; e^{-E_i(\tau-t)}\langle\pi_i|\hat{A}_\mu|\pi_j\rangle e^{-E_i(t-0)}\langle\pi_j|\hat{\pi}|0\rangle, \quad (5)$$

where $\hat{\pi}$ is the pion interpolating operator, and the sum over $\{i,j\}$ is over all the states of the Transfer matrix with the quantum numbers of the pion. Such decompositions of Γ^n hold for all states. Simply replace the symbol π by the state of interest. By fitting Γ_π^2 versus τ, we can extract the amplitudes, $|\langle 0|\hat{\pi}|\pi\rangle|^2$ and the energies E_i for all the "pion" states that couple to $\hat{\pi}$. In the limit $\tau \to \infty$, only the ground (lowest) state contributes, and for $\hat{\pi} = \bar{d}\gamma_0\gamma_5 u$, one gets from Γ^2 the pion decay constant F_π since $|\langle 0|\hat{\pi}|\pi\rangle|^2 = F_\pi^2$, and its energy E_π. Thus 2-point functions give us the amplitudes for creating the sate and the spectrum of the theory (actually of the Transfer Matrix in finite volume).

Next, consider Γ_π^3. It has an additional operator, $\hat{A}_\mu$, sandwiched between the pion creation and annihilation operators, i.e., the current interacts with a propagating pion at time t and the strength of this interaction is given by the matrix element, $\langle\pi_i|\hat{A}_\mu|\pi_j\rangle$. This matrix element of the axial current between pion states can be isolated from the fit to $\Gamma_\pi^3|_{\tau\to\infty}$ since all the other terms, can in principle, be determined from the fit to Γ_π^2. It is easy to check that in the limit that only the ground state contributes (large τ), the matrix element is given by the ratio Γ_3/Γ_2.

If any two operators in Eq. 4 are projected to $\vec{p} = 0$ (LQCD conserves momentum), then we get the axial charge of the pion. If the axial operator inserts momentum $\vec{p}$ and one of the pion interpolating operator removes it, we get the axial form factor describing the semileptonic decay of pions.

Once such data (amplitudes $\to$ decay constants, energies E_i and matrix elements) are generated at a number of values of $\{m_i, a, L\}$, physical results for that observable are obtained by extrapolation: taking the limits $M_\pi \to 135$ MeV, $a \to 0$, and $L \to \infty$ using physics motivated ansätz. This extrapolation is common to all LQCD calculations. I present examples below.

3 Renormalization of operators

We can write down a number of lattice operators, however, results for observables in the continuum limit should be the same. At finite a they have different discretization errors and give different values over and above the difference in the amplitudes, such as $\langle\Omega|\hat{O}_i|\pi\rangle$. Lattice renormalization factors, Z_O^i, relate the different O^i at a given a, and their scaling behavior as $a \to 0$. Results using renormalized operators, say $Z_A^i A_\mu^i$, should agree in the continuum limit.

The experimental results presented by phenomenologists typically use a scheme such as $\overline{MS}$ and a convenient scale such as 2 GeV above which perturbation theory is considered reliable. To translate the lattice result to the $\overline{MS}$ scheme at, say, 2 GeV is a two step process. First one calculates the lattice factors Z_O^i in some scheme (currently two popular ones are the regularization independent [symmetric] momentum schemes labeled RI-MOM and RI-sMOM [6,7]), and a second calculation that relates them to $\overline{MS}$ that is typically done in the continuum using perturbation theory, as is the factor for running in the continuum to a specified scale.

For the calculations described here, the renormalization factors are well-determined. For many other operators, such as the Weinberg and quark chromo EDM operators of dimension 6 and 5, respectively, there have divergent mixing with lower dimension operators, and constructing the finite renormalized operators to use in simulations is very non-trivial [8]. For these two operators, it is still an open problem.

4 Nucleon Correlation Functions

The quark line diagrams for the nucleon 2- and 3-point functions are shown in Fig. 3. Formally, the mechanics of the lattice calculation is very similar to that for the pion, however there are two very important differences:

- The signal to noise ratio falls exponentially as $\sim e^{-(E_N - 1.5 M_\pi)}$ in all nucleon correlation functions. Typical data show that for Γ_2 a good signal extends to about 2 fm and for Γ_3 to about 1.5 fm with $O(10^6)$ measurements [9].

- The spectrum of the towers of multi-hadron excited states, $N\pi$, $N\pi\pi$, ..., labeled by relative momentum $\vec{p}$, begin at about 1200 MeV. For a number of matrix elements these excited states give enhanced contributions that are still large at 1.5 fm. Removing these excited-state contributions (ESC) is still a work under progress for many observables.

To determine various quantities, we use appropriate probes. Changing the operator to a scalar, $\hat{S} = \bar{d}u$ or tensor, $\hat{T} = \bar{d}\sigma_{\mu\nu}u$, gives us nucleon's scalar and tensor charges that are also probed in precision measurements of neutron decay distributions [10]. One link operators give us the momentum fraction, helicity and transversity moments [11]. And the list continues.

Having laid out, hopefully, an intuitive foundation, I now discuss three calculations in order of increasing complexity.
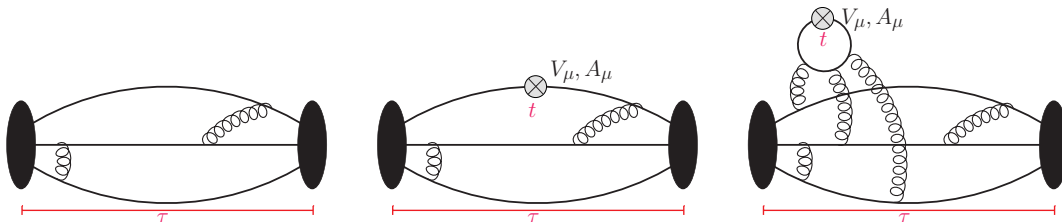
Figure 3: Illustration of quark-line diagrams for 2-point (left), connected 3-point for vector and axial operators $\bar{u}\gamma_\mu d$ and $\bar{u}\gamma_5\gamma_\mu d$ (middle), and the additional disconnected contributions to matrix elements of flavor diagonal axial and vector operators $\bar{q}\gamma_\mu q$ and $\bar{q}\gamma_5\gamma_\mu q$ (right). The operator, axial/vector, current, is inserted at intermediate Euclidean time t and with momentum insertion $\vec{q}$. From this correlation function we get the axial/vector form factors of the nucleon.

5   Isovector charges of the Nucleon

The iso-vector axial, scalar, tensor charges of the nucleon, g_A^{u-d}, g_S^{u-d}, and g_T^{u-d}, probed in the $N \to P$ decay, are extracted from $\Gamma_3(\vec{p} = 0)$, i.e., from the forward matrix element

$$\langle P(p = 0, s') | Z_O \, \bar{u} X_O d |_{\vec{q}=0} | N(p = 0, s) \rangle = g_O \, \bar{u}_P(0, s') X_O u_N(0, s). \qquad (6)$$

with $X_O = \gamma_\mu \gamma_5$, $\mathbb{1}$, $\sigma_{\mu\nu}$ specifying the insertion of the axial, scalar and tensor operators at zero momentum transfer. For iso-vector charges in the isospin symmetric limit, only the connected quark line diagram in the middle panel in Fig 3 contributes. (Even for $\langle N | O | N \rangle$, the disconnected contribution cancels between the insertions on the u and d quarks.)

The data in Fig. 4 for g_A^{u-d} from a $\{a = 0.071\text{fm}, M_\pi = 170\text{MeV}\}$ ensemble (see Ref. [9]) illustrate what the presence of ESC does. The data (same in the two panels) display the following features of the ESC and our goal is to understand and reliably remove them.

- The variation of the data with t and τ is a result (demonstration) of ESC. In the limit $\tau - t$ and $\tau \to \infty$, the data should be flat and lie on top of each other, i.e., independent of τ and t, particularly near $t - \tau/2$, i.e., away from the source/sink.

- The data should be symmetric about $t - \tau/2$ because Γ^3 is. The statistical quality of the data for $\tau = 19$ (=1.35 fm) is already borderline in this respect.

- The convergence of the data with τ for fixed t is monotonic and from below. This shows that ESC causes g_A^{u-d} to be underestimated.

- The ESC is removed by fitting the data for the three largest values of τ using the spectral decomposition given in Eq. 5. The fit shown includes 3 (ground plus two excited) states. The value for the ground state matrix element, given by the fit, is shown by the grey band.

- The data in the two panels are the same. The fits differ in the energy of the first excited state used. In the left panel it is the output of the fit to Γ^2 while in the right, the energy of the lowest $N\pi$ state with relative momentum $(0, 0, 1)$, is input using a narrow prior since it contributes at one loop in χPT.

- The ground state value given by the two fits is different but the augmented χ^2/dof of the two fits are comparable. This shows that the current data are not at sufficiently large τ (or precise enough) to allow us to choose between the two fits.

- The two first excited-state energies, E_1, selected are, physics wise, reasonable options: the first is given by the fit to Γ^2 and lies close to the N(1440), while in the second fit we input, $N(0, 0, 1)\pi(0, 0, -1)$. But so does $N(0, 1, 1)\pi(0, -1, -1)$ and the rest of the tower contribute. In fact, all states with the same quantum numbers contribute! What we do not know, a priori, are the amplitudes, and thus the size of the contribution of each possible excited state. In short, the statistical precision of the current data allow fits with three states, however, these fits show thatthere are large regions in E_1 and E_2 that give similar χ^2/dof but give significantly different extrapolated values.

Bottom line: Until the data are good enough to distinguish between fits with different number or combinations of plausible excited states, and lacking a theoretical reason for a particular choice, the difference between the extrapolated values with different possible excited states can be regarded as an estimate of the systematic uncertainty due to ESC. In the data shown in Fig. 4, a factor of 10X increase in statistics will allow similarly precise data for $\tau = 23$, at which point we will be able to, I believe, resolve between the two fits.

Our data suggest that the uncertainty due to including the $N(0,0,1)\pi(0,0,-1)$ state or not could be a $\sim 5\%$ effect in g_A^{u-d}, but is much smaller in g_S^{u-d} or g_T^{u-d} [9]. At this point in time, controlling ESC is the key outstanding systematic for achieving subpercent precision in the prediction of the isovector properties of nucleons.
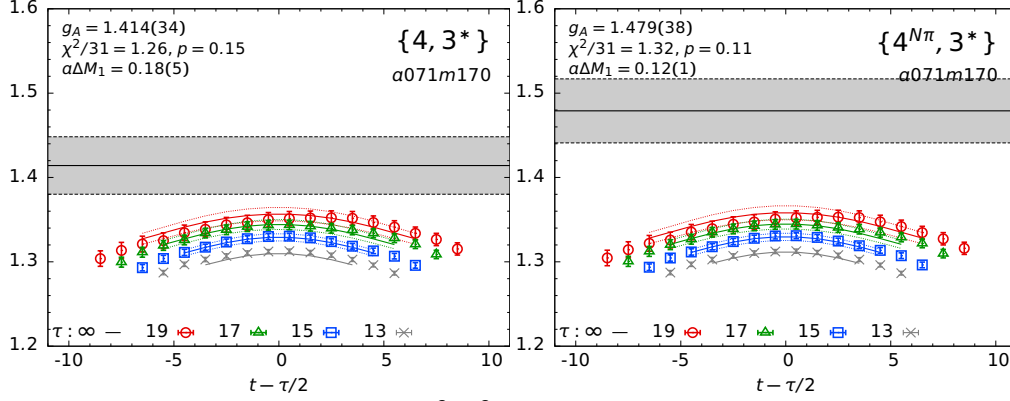
Figure 4: Data for the ratio Γ_A^3/Γ^2 with the insertion of the axial current at $\vec{p}=0$ shown as a function of the displaced operator insertion time $t-\tau/2$. Data for different values of the source-sink separation τ are shown with different colors. Result for the ground state ($\tau \to \infty$ limit) given by a 3-state fit to the spectral decomposition of Γ_A^3 is shown by the grey band, and for different values of τ by lines of the same color as the data.

The illustration of the next step in the analysis using the g_A^{u-d} data is the chiral-continuum-finite-volume extrapolation using a simultaneous fit in $\{M_\pi, a, M_\pi L\}$ shown in Fig. 5. The methodology is the same as described in [9], except we now have data on 13 ensembles with different $\{M_\pi, a, M_\pi L\}$ with $\overline{m}_{ud}$ specified in units of M_π. For the three iso-vector charges, the lowest order corrections in each of these variables is given by [9]

$$g(a, M_\pi, M_\pi L) = c_1 + c_2 a + c_3 M_\pi^2 + c_4 \frac{M_\pi^2 e^{-M_\pi L}}{\sqrt{M_\pi L}} \,. \tag{7}$$

The ansätz depends, in general, on the lattice action and the observable. Each panel in Fig. 5 shows the result of this fit versus a single variable with the other two set to their physical values. For example, the panel plotted versus a takes the simultaneous fit to Eq. 7 to determine the c_i, and then sets $M_\pi = 135$ MeV and $M_\pi L \to \infty$. This gives the pink band. The data plotted are, however, not shifted in the other two variables, which is why the fit does not fall on the data.

Features worth remarking are

- For the 2+1-flavor clover-Wilson action we have used, the value of g_A^{u-d} decreases as $a \to 0$, i.e., the slope with respect to a is positive.

- The value of g_A^{u-d} increases as $M_\pi \to 0$.

- The dependence on a and M_π is largely independent of each other, with opposite slopes.

- There are no significant finite volume corrections observed for $M_\pi L > 4$. This welcome feature has been observed in all calculations involving single nucleons.

Such simultaneous fits are now routine for getting physical values for all observables.

Phenomenologically, the most interesting of the isovector charges is the axial charge, g_A^{u-d}, which has been extracted from experiments with high precision, $g_A^{u-d}/g_V = 1.2754(13)$ [12]. The precision and robustness of lattice results have increased over the last decade, but my

conclusion, in light of possible unresolved ESC of multihadron states such as the $N\pi$ that can cumulatively be as large as $\approx 5\%$, is we need to better quantify and remove the ESC contributions before we can claim sub-percent precision.
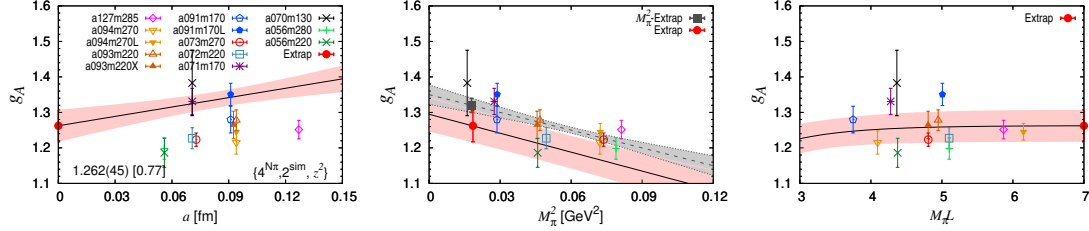
Figure 5: The result of the simultaneous chiral-continuum-finite-volume extrapolation of the g_A^{u-d} data shown by the pink band and plotted versus the lattice spacing a (left), pion mass M_π (middle), and lattice size $M_\pi L$ in units of M_π (right). The grey band in the middle panel shows a simple chiral fit, i.e., with $c_2 = c_4 = 0$ in Eq. 7.

6   Contribution of the spin of the quarks to Nucleon Spin

Using Ji's gauge invariant decomposition $1/2 = \sum_{q=u,d,s,c}\left(\Delta q/2 + L_q\right) + J_g$ [13], where L_q is the quark orbital and J_g the gluon total angular momentum, the contribution of the intrinsic spin of a quark with flavor q, $\Delta q/2$, to the proton spin is given by a relation very similar to Eq. 6:

$$\langle P(p,s')|Z_A \bar{q}\gamma_\mu\gamma_5 q|P(p,s)\rangle = g_A^q\, \bar{u}_P(p,s')\gamma_\mu\gamma_5 u_P(p,s). \qquad (8)$$

with $\Delta q = g_A^q$ and u and $\bar{u}$ are the quark spinors. Because the operator is diagonal in flavor, there is now an additional Wick contraction in which the operator forms a closed loop as illustrated in the right panel in Fig. 3. This is called a "disconnected diagram". The full contribution to g_A^q is the sum of the connected (middle) and disconnected (right) quark line diagrams.

The calculation of disconnected diagrams introduces a new layer of computational cost. The straightforward solution to calculate the momentum projected quark loops with operator insertion on all time slices is to calculate the all-to-all propagator. This is impractical as it constitutes a $(12 \cdot 10^8) \times (12 \cdot 10^8)$ complex matrix for a 100^4 lattice. The solution has been to construct a stochastic estimate. This approach works well, is bias-free but introduces additional statistical uncertainty due to the stochastic estimation of the disconnected quark loop whereas the calculation of the connected quark-line diagrams is exact upto matrix inversion precision for S_F. This uncertainty in the measurement of the loop on each configuration gets convoluted with that due to gauge fluctuations in the ensemble average. Methods such as deflation and bias-corrected truncated solver methods have allowed the reduction in errors in the disconnected contributions to be of the same size as in the connected, however, since their central value is smaller they contribute a larger fraction to the overall error.

The steps in the analysis of these data to get physical results are the same as for the isovector case described in Sec. 5. Post this analysis, I show our PNDME results (2018 and 2022 (preliminary)) for g_A^u, g_A^d, g_A^s along with those from different collaborations in the FLAG format [4,5]. In the FLAG review process, results that pass the criteria for control over discretization errors, finite lattice volume, renormalization and ESC, and obtained sufficiently close to physical pion mass (or extrapolated to $M_\pi = 135$ MeV), are then averaged with appropriate consideration given to possible correlations between results and the overall error assigned. These results are presented as FLAG averages [4,5].

Our results, (PNDME 2018 and 2022 (preliminary)), give $\sum_q \Delta q/2 = \sum_q g_A^q/2 = 0.14(3)$ that is in good agreement with the COMPASS result $0.13 < \Delta\Sigma/2 < 0.18$ [14].
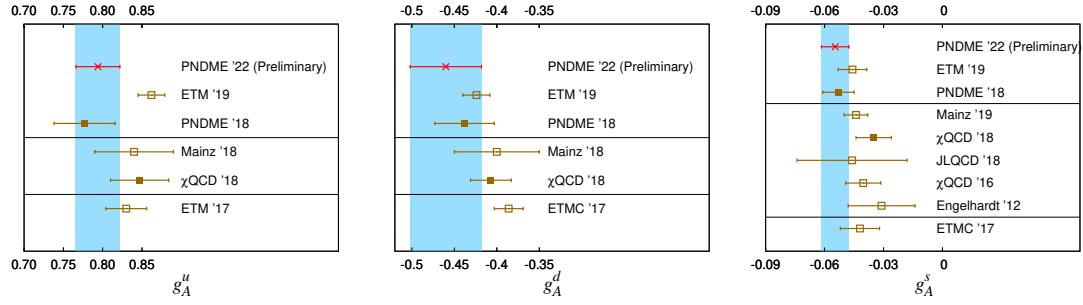
Figure 6: Comparison of results for flavor diagonal axial charges of the nucleon in the FLAG format [4, 5] obtained by different collaborations (PNDME '22 [15] ETM '19 [16], PNDME '18 [17] Mainz '18 [18], χQCD 18 [19], JLQCD '18 [20], ETM '17 [21], χQCD 15A [22], Engelhardt '12 [23]). The points with filled squares meet the FLAG criteria for inclusion in the FLAG average. The contribution to the nucleon spin from quarks with flavor q is $\Delta q/2 = g_A^q/2$.

7 The pion nucleon sigma term

The pion–nucleon σ-term, in the isospin symmetric limit $\overline{m}_{ud} = (m_u + m_d)/2$, is defined as

$$\sigma_{\pi N} \equiv \overline{m}_{ud}\, g_S^{u+d} \equiv m_{ud}\, \langle N(\mathbf{k},s)|\bar{u}u + \bar{d}d|N(\mathbf{k},s)\rangle\,. \tag{9}$$

The calculation of $\overline{m}_{ud}$ is done separately. It is a fundamental parameter of QCD that quantifies the amount of the nucleon mass generated by having u- and d-quarks with non-zero mass. In this direct method, we calculate the scalar charge g_S^q, which is determined from the forward matrix element of the scalar density $\bar{q}q$ between the nucleon state:

$$\bar{u}_N(0,s)g_S^q u_N(0,s) = \langle N(\mathbf{k}=0,s)|Z_S\, \bar{q}q|N(\mathbf{k}=0,s)\rangle, \tag{10}$$

where Z_S is the renormalization constant and the nucleon spinor has unit normalization. The scalar charge, g_S^q, determines the coupling of the nucleon to any scalar mediator with quark content of the coupling given by $\bar{q}q$. It enters in the search for physics beyond the Standard Model (SM), including in direct-detection searches for dark matter scattering off nuclei via a scalar mediator, lepton flavor violation in $\mu \to e$ conversion in nuclei and in electric dipole moments. The calculation of g_S^q (and of g_T^q) is simular to that discussed above for g_A^q, and note that g_A^q give the spin dependent couplings.

Figure 7 shows the data for $g_S^u + g_S^d$ from a physical pion-mass ensemble [24]. Again the two fits have very similar χ^2/dof but the one with E_1 from $N\pi$ as the excited state gives almost 50% larger value. Our χPT analysis given in Ref. [24] shows that there are two significant ESC, one from $N\pi$ and the other from $N\pi\pi$. These two states are almost degenerate in our lattice calculation, so they effectively contribute as one, i.e., their amplitudes in the spectral decomposition would add as the exponential factors are very similar. The right panel in Fig. 7 shows pictorially why the $N\pi$ state makes a large disconnected contribution because the scalar current has a large coupling to the quark loop, and the quark-line diagram favors a $N\pi$ configuration.

The two analyses with different values of E_1 led to an interesting conundrum. The standard analysis (with $E_1 \sim 1450$ MeV) gave $\sigma_{\pi N} \approx 40$ MeV consistent with previous lattice analyses [4, 5], whereas the analysis with $E_1 = E_{N\pi} \approx 1230$ MeV gave $\sigma_{\pi N} \approx 60$ MeV, which is consistent with the dispersive analysis starting with the $N\pi$ scattering data [24]. Our preferred solution is the latter based on the χPT analysis. In that case the tension between LQCD and phenomenological estimates is resolved.

Clearly, the 50% difference between the two analyses with similar χ^2/dof calls out for additional LQCD calculations to be done to confirm this exciting result. The key point for future calculations of $\sigma_{\pi N}$, using either the direct method of calculating the charge g_S via the matrix element from Γ^3 or using the Feynman-Hellmann relation (derivative of the nucleon mass with respect to the quark mass), is that they have to be done close to $M_\pi = 135$ as only there the $N\pi$ state becomes much lighter than $N(1440)$ and the ESC is very different and manifest. Data with $M_\pi > 200$ MeV do not give significantly different results between the two kinds of fits. Thus, extrapolation from heavier M_π ensembles will miss this physics.
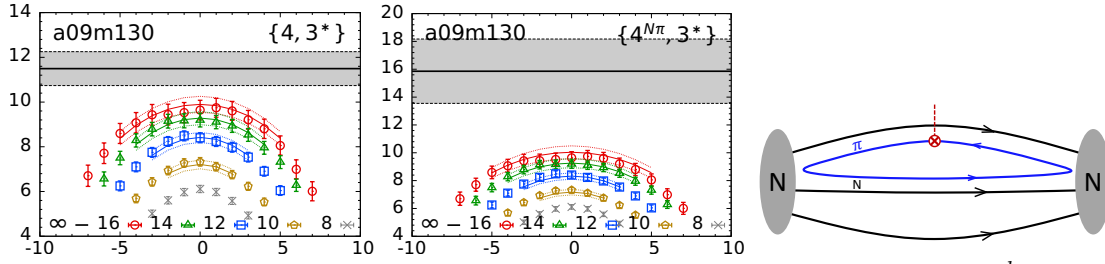
Figure 7: (Left) Fit to get the ground state value for the scalar charge $g_S^u + g_S^d$ using the first excited state energy from Γ^2; (middle) the noninteracting energy of the $N\pi$ state; and (right) the disconnected diagram that is enhanced.

8 Conclusions

Simulations of LQCD provide ensembles of importance sampled configurations whose distribution according to the Boltzmann factor $\mathrm{Det}\mathcal{D}e^{-A_G} = e^{-A_G+\mathrm{Lndet}\mathcal{D}}$ constitutes the nonperturbative ground state of QCD. This construction is exact (bias-free) but stochastic. These ensembles are characterized by the input parameters $\{m_i, a, M_\pi L\}$. Correlation functions of any time-ordered string of operators are given by quark-line diagrams obtained using Wick contraction. Properties of QCD (spectrum, matrix elements, EoS, etc) are extracted from the expectation values, i.e., ensemble averages, of these correlation functions. The full excursion (possible "paths") of the quark propagators (and values of link parameters for non-local and gluon operators) in the quark line diagrams over 3-space but at a fixed intermediate time provides a full Fock space wavefunction, i.e., it includes all states with the same quantum numbers as the interpolating operator. The propagation of each of these states in Euclidean time τ is damped as $e^{-E_n \tau}$, allowing the ground state with energy E_0 to be isolated in the limit $\tau \to \infty$. Once again, this stochastic description of the wavefunction provides no intuition or visualization but allows the calculation of fully quantum mechanical matrix elements of any operator within this state.

The reason LQCD is called a "black box" is because we cannot visualize or represent the vacuum fluctuations or the wavefuctions that get created at intermediate times in the correlation functions Γ_n. Nevertheless, the results obtained are rigorous, display all the subtleties of QCD and confirm quantum field theory as describing quantum phenomenon in nature.

Three kinds of observables have been used to illustrate how LQCD calculations are done and data analyzed. Looking ahead, precision calculations of nucleon correlation functions need to overcome two challenges: the exponential degradation of the signal and how to remove all/most excited state contamination in the wavefunctions to get matrix elements within the nucleon ground state from correlations functions calculated with finite source-sink separation τ. The ESC in current data can be large as illustrated by the pion-nucleon sigma term.

The future is exciting – there is clear demonstration that LQCD has matured and results are having an impact on phenomenology and experiments [4,5]. Methodology and algorithms for many calculations are robust, however, brute force approach to nucleon correlation functions

for sub-percent precision by increasing the statistics is unlikely to succeed in the next 5 years. It is, therefore, time for innovation–new methods and algorithms to reduce systematics and increase statistics efficiently.

Acknowledgements and Funding Information

Many thanks to my collaborators Tanmoy Bhattacharya, Vincenzo Cirigliano, Martin Hoferichter, Yong-Chull Jang, Balint Joo, Huey-Wen Lin, Emanuele Mereghetti, Santanu Mondal, Sungwoo Park, Frank Winter, Junsik Yoo, and Boram Yoon with whom the work presented has been done over the last decade. The calculations used the CHROMA software suite. This research used resources at (i) NERSC, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Award No. DE-AC02-05CH11231; (ii) the Oak Ridge Leadership Computing Facility through ALCC award LGT107 and INCITE award HEP133; (iii) the USQCD Collaboration, funded by DOE HEP; and (iv) Institutional Computing at Los Alamos National Laboratory. The work was partly supported by DOE HEP under Award No. DE-AC52-06NA25396 and by the LANL LDRD program.

References

[1] K. G. Wilson, *Confinement of Quarks*, Phys. Rev. D **10**, 2445 (1974), doi:10.1103/PhysRevD.10.2445.

[2] M. Creutz, *Quarks, gluons and lattices*, Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, UK, ISBN 978-0-521-31535-7 (1985).

[3] The Flavor Lattice Averaging Group (FLAG) (2015).

[4] S. Aoki *et al.*, *FLAG Review 2019: Flavour Lattice Averaging Group (FLAG)*, Eur. Phys. J. C **80**(2), 113 (2020), doi:10.1140/epjc/s10052-019-7354-7, 1902.08191.

[5] Y. Aoki *et al.*, *FLAG Review 2021: Flavour Lattice Averaging Group (FLAG)*, Eur. Phys. J. C **82**(10), 869 (2022), doi:10.1140/epjc/s10052-022-10536-1, 2111.09849.

[6] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa and A. Vladikas, *A General method for nonperturbative renormalization of lattice operators*, Nucl.Phys. **B445**, 81 (1995), doi:10.1016/0550-3213(95)00126-D, hep-lat/9411010.

[7] C. Sturm, Y. Aoki, N. Christ, T. Izubuchi, C. Sachrajda *et al.*, *Renormalization of quark bilinear operators in a momentum-subtraction scheme with a nonexceptional subtraction point*, Phys.Rev. **D80**, 014501 (2009), doi:10.1103/PhysRevD.80.014501, 0901.2599.

[8] T. Bhattacharya, V. Cirigliano, R. Gupta, E. Mereghetti and B. Yoon, *Dimension-5 CP-odd operators: QCD mixing and renormalization*, Phys. Rev. **D92**(11), 114026 (2015), doi:10.1103/PhysRevD.92.114026, 1502.07325.

[9] S. Park, R. Gupta, B. Yoon, S. Mondal, T. Bhattacharya, Y.-C. Jang, B. Joó and F. Winter, *Precision nucleon charges and form factors using (2+1)-flavor lattice QCD*, Phys. Rev. D **105**(5), 054505 (2022), doi:10.1103/PhysRevD.105.054505, 2103.05599.

[10] T. Bhattacharya, V. Cirigliano, S. D. Cohen, A. Filipuzzi, M. Gonzalez-Alonso *et al.*, *Probing Novel Scalar and Tensor Interactions from (Ultra)Cold Neutrons to the LHC*, Phys.Rev. **D85**, 054512 (2012), doi:10.1103/PhysRevD.85.054512, 1110.6448.

[11] S. Mondal, R. Gupta, S. Park, B. Yoon, T. Bhattacharya, B. Joó and F. Winter, *Nucleon momentum fraction, helicity and transversity from 2+1-flavor lattice QCD*, JHEP **21**, 004 (2020), doi:10.1007/JHEP04(2021)044, 2011.12787.

[12] R. L. Workman *et al.*, *Review of Particle Physics*, PTEP **2022**, 083C01 (2022), doi:10.1093/ptep/ptac097.

[13] X.-D. Ji, *Gauge-Invariant Decomposition of Nucleon Spin*, Phys. Rev. Lett. **78**, 610 (1997), doi:10.1103/PhysRevLett.78.610, hep-ph/9603249.

[14] C. Adolph *et al.*, *The spin structure function g_1^p of the proton and a test of the Bjorken sum rule*, Phys. Lett. **B753**, 18 (2016), doi:10.1016/j.physletb.2015.11.064, 1503.08935.

[15] S. Park, T. Bhattacharya, R. Gupta, H.-W. Lin, S. Mondal, B. Yoon and R. Zhang, *Flavor diagonal nucleon charges*, PoS **LATTICE2021**, 558 (2021), doi:10.22323/1.396.0558, 2203.09584.

[16] [ETM 19] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou and A. Vaquero Aviles-Casco, *Nucleon axial, tensor, and scalar charges and σ-terms in lattice QCD*, Phys. Rev. D **102**(5), 054517 (2020), doi:10.1103/PhysRevD.102.054517, 1909.00485.

[17] [PNDME 18A] H.-W. Lin, R. Gupta, B. Yoon, Y.-C. Jang and T. Bhattacharya, *Quark contribution to the proton spin from 2+1+1-flavor lattice QCD*, Phys. Rev. **D98**, 094512 (2018), doi:10.1103/PhysRevD.98.094512, 1806.10604.

[18] D. Djukanovic, H. Meyer, K. Ottnad, G. von Hippel, J. Wilhelm and H. Wittig, *Strange nucleon form factors and isoscalar charges with $N_f = 2+1$ $\mathcal{O}(a)$-improved Wilson fermions*, PoS **LATTICE2019**, 158 (2019), doi:10.22323/1.363.0158, 1911.01177.

[19] [χQCD 18] J. Liang, Y.-B. Yang, T. Draper, M. Gong and K.-F. Liu, *Quark spins and Anomalous Ward Identity*, Phys. Rev. **D98**(7), 074505 (2018), doi:10.1103/PhysRevD.98.074505, 1806.08366.

[20] [JLQCD 18] N. Yamanaka, S. Hashimoto, T. Kaneko and H. Ohki, *Nucleon charges with dynamical overlap fermions*, Phys. Rev. **D98**(5), 054516 (2018), doi:10.1103/PhysRevD.98.054516, 1805.10507.

[21] C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou, A. Vaquero Avilés-Casco and C. Wiese, *Nucleon Spin and Momentum Decomposition Using Lattice QCD Simulations*, Phys. Rev. Lett. **119**(14), 142002 (2017), doi:10.1103/PhysRevLett.119.142002, 1706.02973.

[22] M. Gong, Y.-B. Yang, J. Liang, A. Alexandru, T. Draper and K.-F. Liu, *Strange and charm quark spins from the anomalous Ward identity*, Phys. Rev. D **95**(11), 114509 (2017), doi:10.1103/PhysRevD.95.114509, 1511.03671.

[23] M. Engelhardt, *Strange quark contributions to nucleon mass and spin from lattice QCD*, Phys. Rev. **D86**, 114510 (2012), doi:10.1103/PhysRevD.86.114510, 1210.0025.

[24] R. Gupta, S. Park, M. Hoferichter, E. Mereghetti, B. Yoon and T. Bhattacharya, *Pion–Nucleon Sigma Term from Lattice QCD*, Phys. Rev. Lett. **127**(24), 242002 (2021), doi:10.1103/PhysRevLett.127.242002, 2105.12095.