Scaling and Parallel Performance

The most time consuming step in our codes is the `dslash` routine that computes $\psi \to D_W \psi$. Moreover, this is the routine that relies on interprocess communication the most, since most other vector routines are local. Consequently, our optimization efforts focused primarily on implementing it efficiently. We are primarily interested in the double precision implementation since the eigensolvers employed are very sensitive to roundoff errors. The single precision version of this code performs twice as fast.

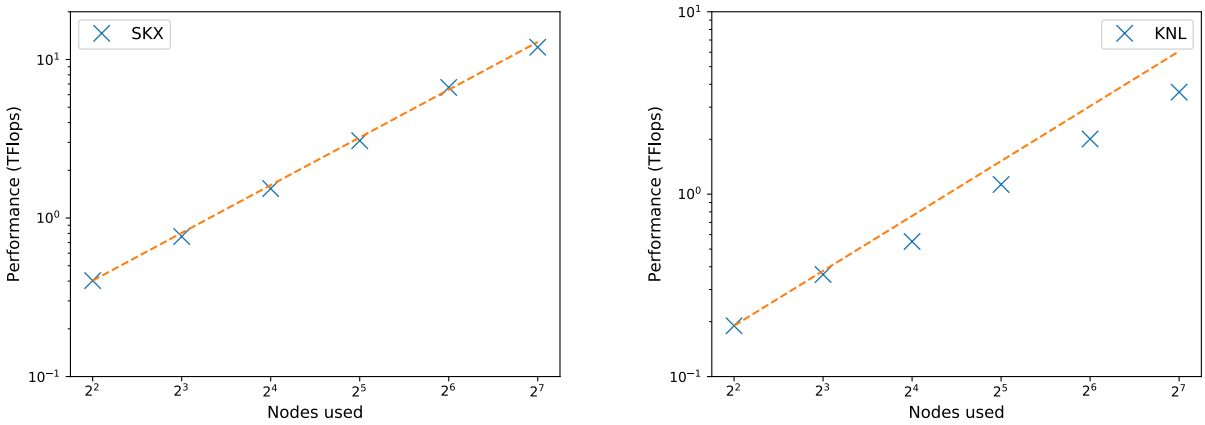

Figure 1: The log-log plots of strong scaling for `dslash` routine on a $48^3 \times 96$ lattice configuration on SKX (left panel) and on KNL (right panel) from 4 up to 128 nodes. The straight line marks a perfect scaling of performance based on the results at 4 nodes which is 0.101 TF/node for SKX and 0.047 TF/node for KNL.

For lattice QCD, a parallel code is efficient when the aggregate performance is proportional to the number of processes. We present here the results for *strong* scaling, i.e. performance of our codes for a lattice of fixed size that gets divided into smaller and smaller pieces as the number of processors is increased. As long as the bulk `dslash` kernel takes more time than communication, the scaling will be almost perfect. Eventually the communication time will dominate and scaling will suffer.

In Fig. 1 we show the log-log plots of the performance of `dslash` using SKX (left panel) and KNL (right panel) nodes on Stampede2 for the lattice size of $48^3 \times 96$. The results are obtained by taking the average of the repeated test over 100 times. As a reference, we plot the straight line which marks a perfect scaling of performance based on the result of 4 nodes which is is 0.101 TF/node for SKX and 0.047 TF/node for KNL. We note that on the $48^3 \times 96$ lattice, the code scales almost perfectly from 4 to 128 nodes on SKX. On KNL, the `dslash` performance degrades after 8 nodes , a clear indication that the communication time is becoming dominant.

Based on the scaling test, we shall plan to run the production of the on 86 SKX nodes for the $48^3 \times 96$ lattice and 43 nodes for the smaller $24^3 \times 48$ and $32^3 \times 64$ lattices.

Quark and Glue Spins and Angular Momenta of the Nucleon from Lattice QCD

χQCD Collaboration

Senior Researchers

Keh-Fei Liu (PI), Terrence Draper (Co-PI)	University of Kentucky
Andrei Alexandru (Co-PI), Frank Lee (Co-PI)	George Washington University
Ying Chen, Ming Gong, Zhaofeng Liu	Inst. of High Energy Physics, Beijing
Mridupawan Deka	JINR, Dubna, Russia
Takumi Doi	RIKEN
Shao-Jing Dong, Ivan Horvath	University of Kentucky
Balint Joo	Jefferson Laboratory
Huey-Wen Lin	Michigan State University
Nilmani Mathur	Tata Institute, Mumbai
Hank Thacker	University of Virginia
Jianbo Zhang	Zhejiang University

Postdoctoral Researchers

Jian Liang, Jacob Fallica	University of Kentucky
Raza Sabbir Sufian	JLab
Yi-Bo Yang	Michigan State University

Graduate Students

Chris Culver, Dehua Guo, Hossein Niyazi	George Washington University
Gen Wang	University of Kentucky
Yujiang Bi, Wei Sun	Inst. of High Energy Physics, Beijing

Abstract

We propose to carry out the Lattice Quantum Chromodynamics (LQCD) calculation of the quark and glue spins and their angular momenta of the nucleon with overlap quark propagators on 2+1 flavor domain-wall fermion configurations on three lattices, i.e. $48^3 \times 96$ lattice at lattice spacing $a = 0.114$ fm, $24^3 \times 48$ and $32^3 \times 64$ lattices at $a = 0.2$ fm. All three of them have the light sea quark masses which correspond to the physical pion mass of 139 MeV and they are produced by the RBC and UKQCD Collaborations. Together with results from smaller lattices and different lattice spacings from the RBC and UKQCD Collaborations, we shall perform a global fit in order to study the valence and sea quark mass dependence and control the systematic errors in regard to the continuum limit and infinite volume corrections of the physical observables that we propose to calculate.

We aim to do production of the quark propagators with eigenmode deflation. For the valence quarks in the nucleon correlator, we shall use the grid-smeared noise source with low-mode substitution which is tested and shown to be very efficient. The quark loops will be calculated with low-mode averaging and noise estimation of the high modes. Since the low modes dominate the quark loops for the pseudoscalar and scalar densities, this low-mode-average approach is effective in variance reduction in the calculation of the quark condensate and quark spin via the anomalous Ward identity.

The nucleon structure that we plan to calculate includes the quark spin from chiral axial-vector current, the glue spin, and quark and glue momenta and angular momenta so that we can completely apportion the proton spin into quark spin, quark orbital angular momentum, and glue angular momentum.

Our overall request is 940,000 SUs on Stampede2 at TACC. We also request 392 TB of storage space at TACC.

1 Project Description

1.1 Physics Goals and Significance of Research

An important task in nuclear physics, as spelled out in the Long Range Plan prepared by the 2015 Nuclear Science Advisory Committee (NSAC) [1] is to delineate hadron structure and spectroscopy in terms of quarks and gluons from Quantum Chromodynamics (QCD). It is emphasized in the Plan [1] that "understanding the structure of hadrons in terms of QCD's quarks and gluons is one of the central goals of modern nuclear physics." Over the last three decades, lattice QCD has developed into a powerful tool for *ab initio* calculations of strong-interaction physics. It is presently the only theoretical approach to solving QCD with controlled statistical and systematic errors.

Determining the contributions of the quarks and gluons to the nucleon spin is one of the most challenging issues in QCD both experimentally and theoretically. It was an utter surprise 29 years ago when the deep inelastic scattering experiments revealed that the quark spin contributes only ∼30% to the total proton spin [2]. This is contrary to the long-held conventional wisdom from the quark model that 100% of the proton spin is due to the quark spin. This result has plagued the physics community. While thousands of papers have been written on the subject with much speculation, there is no proof which one is correct. As such, the puzzle has been dubbed the 'proton spin crisis.' The experimental search for the remaining spin components from the gluon helicity distribution $\Delta G(x)/G(x)$ from both COMPASS and STAR experiments has found it to be small and close to zero [3]. A recent analysis [4] of high-statistics 2009 STAR [5] and PHENIX [6] data showed evidence of non-zero glue helicity in the proton. For $Q^2 = 10\,\mathrm{GeV}^2$, they found gluon helicity distribution $\Delta g(x, Q^2)$ positive and away from zero in the momentum fraction region $0.05 \le x \le 0.2$. However, the result presented in [4] has very large uncertainty in the small-x region. There are also experimental attempts to measure the quark orbital angular momentum from the generalized parton distribution (GPD) in deeply virtual Compton scattering experiments and from the transverse momentum dependent (TMD) parton distribution to search for the missing proton spin components. This is an active research area both experimentally and theoretically and is one of the major goals of the planned electron-ion collider (EIC) experimental facility.

Lattice fermion formulation accommodating chiral symmetry has been developed in the last two decades in the form of domain-wall fermion (DWF) [7, 8], overlap fermion [9, 10, 11], and fixed-point fermion [12]. In the past decade or so, the Riken-Brookhaven-Columbia (RBC) collaboration, the UKQCD collaboration of the United Kingdom, and Lattice Hadron Physics Collaboration (LHPC) have used QCDOC machines at BNL, RIKEN, and Edinburgh and IBM Blue Gene P and Q at Brookhaven and Argonne to generate $2+1$ flavor dynamical domain-wall fermion [7, 8] gauge configurations on an ensemble of lattices with spatial sizes ranging from 2.8 fm [13, 14] with pion mass as low as 290 MeV to 4.5 fm with pion mass at 170 MeV [15]. RBC and UKQCD Collaborations have also generated configurations on $48^3 \times 96$ lattice [16] with lattice spacing of $a = 0.114$ fm. The sea u/d quark mass corresponds to a pion mass at the physical 139 MeV and the box size are 5.5 fm which is twice as large as those of the earlier $24^3 \times 64$ and $32^3 \times 64$ lattices. Recently, the same collaborations have produced configurations on two more lattices at physical pion mass of 137 MeV, i.e. $24^3 \times 48$ and $32^3 \times 64$ lattices at $a = 0.201$ fm so that the lattice sizes are 4.8 fm and 6.4 fm respectively. Combined analysis of these lattices have shown a significant improvement in the control of the chiral extrapolation which led to improved continuum predictions for the light hadron properties, such as f_π, f_K, light quark masses and B_K [15, 16].

In view of the fact that the formulation of the overlap fermion [10] can be readily implemented to satisfy the Ginsparg-Wilson relation, which is required for exact chiral symmetry on the lattice, to high precision, we have studied the overlap fermion since 2000 and have developed algorithms to efficiently approximate the matrix sign function in the overlap Dirac matrix [17, 18] and correlators [19]. Since 2008, we have initiated a long-term project to study nucleon form factors, hadron spectroscopy, and decay constants with valence overlap fermion on the above-mentioned $2+1$-flavor dynamical gauge configurations with domain-wall fermions. Undertaking a global fit [20] which combines the chiral extrapolation, the physical quark mass interpolation, and the continuum extrapolation for the charmonium, D_s and D_s^* on the 24^3 and $32^3 \times 64$ lattice with 3 sea quark masses each, we have been able to obtain the charm and strange quark masses, the charmonium hyperfine splitting, the 3 charmonium P-states, and the decay constant, f_{D_s}, all within one sigma from those summarized in PDG Live [21].

These newly generated large lattices with chiral fermion at the physical pion mass represent the state-of-the-art dynamical fermion configurations with good chiral symmetry. These configurations are an extremely valuable asset for the world lattice QCD community; quark propagators with the overlap fermion formalism can be calculated on these to study important physical quantities, particularly the nucleon properties where large volumes are needed to control the finite volume effects. It is an exciting time to use them to study physical observables which are either not readily available or not reliable from the present models or experiments. When the overlap valence quark propagators are calculated, they will represent cutting-edge resources for lattice QCD, since this would be the first time that exact chiral fermion propagators are calculated on the $2+1$ flavor DWF sea with as many as 10 valence quarks which cover the range from charm quark mass down to the light quark masses at the same physical pion mass of 139 MeV as that of the sea u/d quarks. We list in Table 1 the ensemble of these DWF configurations and their lattice spacings, spatial sizes and the pion masses corresponding to light u/d quark masses in the sea.

Table 1: The parameters for the RBC/UKQCD configurations[16]: spatial/temporal lattice sizes, lattice spacings, physical spatial sizes, and the pion masses with the degenerate light sea quark.

Lattice	$L^3 \times T$	a (fm)	L(fm)	m_π (MeV)
24I	$24^3 \times 64$	0.1105(3)	2.65	339
32I	$32^3 \times 64$	0.0828(3)	2.65	302
32ID	$32^3 \times 64$	0.1432(7)	4.58	172
48I	$48^3 \times 96$	0.1141(2)	5.48	139
24IDC	$24^3 \times 48$	0.2013(17)	4.8	137
32IDC	$32^3 \times 64$	0.2013(17)	6.4	137

Unlike as with purely theoretical work, experimental and large numerical-simulation research requires an antecedent sequence of many technical papers with intermediate results which establish and vet the techniques which pave the way for the culminating paper which addresses the important physical results. Over the last decade, we have been continually developing new algorithms to speed up the quark propagator calculation and enhance statistics with greatly improved correlators [19, 67, 65], especially for the cases with the physical pion mass and large volume where the signal to noise ratio for the nucleon two- and three-point functions fall off exponentially in time separation as $e^{-(m_N - 3/2 m_\pi)t}$ [31] where m_N/m_π is the nucleon/pion mass. We will explain these algorithm improvements in the next subsection.

In the intervening years, we have worked on the 24I, 32I, 48I, and 32ID lattices on various physical quantities such as the charmonium spectrum, the charm and strange quark masses and the decay constant f_{D_s} [20], the strange and charm quark spins [32], the πN and strangeness sigma terms [22], and the strange quark magnetic moment [23]. These quantities are intimately related to phenomenology and experiments. The quark spin is related to deep inelastic scattering experiment and our goal is to understand and resolve the "proton spin crisis" where the total quark spin contributes only $\sim 30\%$ of the proton spin. Precise determination of πN and strangeness sigma terms are important in the dark matter search for neutralinos, for πN and KN scattering and kaon condensate in the neutron star. The strange quark magnetic moment $G_M^s(0)$ has been compared to that obtained from the parity-violating e-p scattering experiment. Our latest result on gluon spin [24] is aimed to relate to the ΔG measurement from the STAR and PHENIX experiments at RHIC and COMPASS experiment at CERN and the future experiments on EIC.

In addition to the above mentioned quantities, we propose to continue the calculation of a complete decomposition of the proton spin into quark spin, quark orbital angular momentum and glue angular momentum as has been carried out in the quenched approximation by our group [33]; the orbital angular momentum can be determined from subtracting the spins from the angular momenta. We will also calculate another definition of the orbital angular momenta from the generalized transverse momentum distribution function (GTMD) [25].

One of the main reasons that we adopt overlap fermion in our lattice calculation is because it satisfies chiral symmetry exactly on the lattice and there are many physical quantities which are sensitive to chiral symmetry. For example, due to the explicit chiral symmetry breaking, there is an additive mass renormalization for the Wilson type fermions. As a consequence, there is flavor mixing which has plagued early lattice calculations of strangeness and $\pi N \sigma$ term [26, 27]. This can be ameliorated or corrected by calculating the singlet and octet

mass renormalizations and imposing Ward identities [28]. However, this will require additional work and at the expense of additional systematic errors. The overlap is free of this problem – the quark mass has no additive renormalization and the πN and strange sigma terms are renormalization group invariant. One can simply use the bare quark mass and bare lattice matrix elements to obtain directly the results without renormalization [67, 22]. It is even more important to use overlap fermion to calculate quark spin. As is pointed out, it is imperative to impose the anomalous Ward identity to obtain the normalization constant for the point flavor-singlet axial-vector current [32], analogous to the case of flavor octet axial-vector current with the chiral Ward identity. Without such a normalization, the lattice calculations of the strange quark spin is smaller in magnitude than that obtained with the local axial-vector current [32]. By the same token, the light u/d quark spin in the disconnected insertion is also small. In addition to the above examples which are sensitive to chiral symmetry, it is known that overlap does not have an $\mathcal{O}(a)$ error because of chiral symmetry and its $\mathcal{O}(a^2)$ is small so that it can be extrapolated to the continuum from relatively coarse lattices which have smaller lattice sizes and, thus, less computing time requirement. Its $\mathcal{O}(m^2a^2)$ is also small so that the charm quark and light quarks can be accommodated on the same lattice without having to resort to different actions for heavy and light quarks. The upshot is that physical results with overlap fermion are more accurate at finite lattice spacings with smaller systematic errors compared to other fermion formulations without chiral symmetry.

1.2 Numerical Aspects of Overlap Fermion

The overlap fermion has several theoretical advantages. It satisfies chiral symmetry and its effective quark massless propagator anti-commutes with γ_5, i.e. $\{\gamma_5, D_c^{-1}\} = 0$ and the quark mass is additive in the massive propagator, i.e. $D_c^{-1}(m) = (D_c + m)^{-1}$. These properties are the same as in the continuum. All the continuum current algebra relations are satisfied on the finite lattice. This section will dwell on the numerical issues to show that, despite of the fact that the quark propagator calculation of the overlap fermion is slower than that of the Wilson-type fermions, various algorithmic improvements, some unique to the overlap fermion, make up the speed difference. Therefore, adopting the overlap fermion is not only more accurate with smaller systematic errors, the lattice calculations are at least as precise as with the Wilson-type fermions, given the same resources. The description of the algorithm improvements are technical. Reviewers who are not interested in these numerical issues may wish to skip this section.

On the numerical side, it is well-known that it is computationally intensive to invert the overlap fermion due to its matrix sign function. It is about two orders of magnitude more expensive than for Wilson-type fermions for a straightforward CG inversion without preconditioning or deflation. In view of this, we have implemented eigenvector deflation, to speed up the inversion, and introduced a slew of improvements for the calculation of two-point and three-point correlators over the years.

- Deflating with these eigenmodes and incorporating HYP smearing of gauge links, one can speed up the inversion by a factor from 25 (with 200 pairs of eigenmodes plus zero modes) to 79 (with 300 pairs of eigenmodes plus the zero modes) on different lattice sizes [19]. Except for the one-time overhead of the calculation of the eigenvectors, the inversion itself is now only about an order of magnitude slower than the highly-tuned inverter for clover fermion (e.g. EigCG or multi-grid in Chroma) and twisted mass fermion. More detailed comparison of cost will be given in Sec. 1.2.1 for calculating the nucleon matrix elements. We should point out that with eigenvector deflation, there is no critical slowing down with small quark masses.

- Moreover, these low-energy eigenmodes can be used to construct many-to-all correlators to replace the noise-estimated low-frequency part of the hadron correlators with the exact one. It is found [67] that the smeared-grid noise source with low-mode substitution (LMS) gains in statistics about 60% of the number of grids (8 for smaller lattices and 64 for the 48I lattice) as compared to a single smeared noise source without LMS. We have also calculated quark loops for the scalar, pseudoscalar, and vector currents in the nucleon with low-mode average (LMA) and noise estimate of the high modes [67, 32, 22, 23]. Due to the dominating (more than 90%) low-mode contribution in the scalar and pseudoscalar loops as observed, we have been able to calculate, with the overlap fermion on DWF configurations, the strangeness and charmness contents in the nucleon with unprecedented precision amongst all the current lattice calculations and at a smaller fraction of the cost compared to other calculations [67]. Our most recent work on the strangeness content [22] has

the highest precision among all the 2 + 1-flavor lattice calculations and it has included the physical pion point from the 48I lattice as well as finite volume and continuum corrections. Similarly, a $\sim 50\%$ low-mode contribution in the vector current loop has helped us obtain a more than 4 sigma signal of the strangeness magnetic moment at the physical pion point for the first time whose error is several times smaller than those of experiments [23]. The speed-up in inversion and the improvement in correlator and quark loop calculation are testaments to the fact that low eigenmodes are crucial to all the above improvements.

- Besides the LMS for the nucleon propagator and LMA for the quark loop, both of which are important for disconnected insertion calculations, we have implemented a stochastic sandwich method for the three-point connected insertion (CI) calculations [65]. This involves a multi-grid smeared source with LMS and stochastic sink for the high modes. This saves time as compared to the usual sink sequential method which will need to have multiple inversions at the sink for both u and d, different polarizations, and different momenta. In addition, the stochastic sandwich method can accommodate multi-mass inversions which is important in reducing errors in global fittings with partially quenched data.

- The overlap fermion operator is a normal matrix. As such, one can calculate the pair of complex eigenvalues on the circle from the left-handed and right-handed pair of the eigenmodes of $D^\dagger D$ with the Arnoldi algorithm. This is a one-dimensional search. One could likewise implement deflation with eigenmodes of hermitian Dirac operator for domain-wall or Wilson fermions, D_{dw} and D_W, which entail real eigenvalues, i.e. $\gamma_5 D_{dw}$ or $\gamma_5 D_W$, to gain speedup in inversion. However, these eigenmodes are different for different quark masses and thus they need to be generated for each quark mass.

- One definitive numerical advantage of the overlap fermion is its multi-mass feature with the conjugate gradient algorithm which admits calculating multiple quark masses (e.g. $10 - 20$ which ranges from the physical u/d quark to the charm region) in one stroke with an overhead which is only $\sim 8\%$ of that for calculating the lowest mass alone [34, 35]. This multi-shift algorithm with deflation is predicated on the fact that the eigenmodes can be used for all quark masses to obtain quark propagators for multiple quark masses with very little overhead. Moreover, only half of them (plus the zero modes) need to be saved in disk for future use in constructing hadron correlators and quark loops for any quark mass. The storage is a critical issue for large lattices, such as 48I and 64I, where more than a thousand pairs of eigenvectors need to be saved for analysis. True, hadronic observables are correlated across different quark masses, but not totally. In fact, the error bands of the fitting curve through multiple points are typically substantially smaller than the errors on individual points, so there is demonstrated statistical value in having multiple points in addition to being informed of the functional dependence of observables on quark mass.

These are some of the salient numerical advantages of the overlap fermion.

1.2.1 Cost, Precision and Accuracy

In view of the fact that overlap fermion is known to be more expensive to invert, we made a detailed comparison of the numerical cost of calculating the various nucleon matrix elements with those of the twisted mass fermion [66], the clover fermion [37] and the domain-wall fermion [38]. We have made these comparisons with Kyriakos Hadjiyiannakou of the ETM Collaboration, Boram Yoon of the NME Collaboration, and Sergey Syritsyn of the RBC/LHP Collaboration on the same respective machines that these lattice calculations are carried out. The details of the comparison can be found in the 2016 lattice talk presented by the PI [39]. We considered the ratio of cost for the same statistics in terms of the product of the number of inversions for the total configurations, the inversion time, and the variance of the matrix elements and nucleon mass. The upshot is that despite the fact that the overlap inversion is ~ 10 times slower than that of the twisted mass on GPUs of the same model (K20), and ~ 44 times slower than that of clover on Titan, and ~ 6 times slower than that of the DWF on the Fermilab cluster Pi0, the ratios are close to unity for the comparison with the twisted mass and clover fermions, but ~ 6 times larger for the DWF compared to overlap. When the multi-mass feature is included from partially quenched data, the final error of the $\pi N\sigma$ term, strangeness and strange magnetic moment at the physical point can be reduced by a factor 2 to 3 [22, 23].

We should emphasize that this multi-mass feature of using the same eigenvectors for LMS in the connected insertion and nucleon propagator, and LMA in the quark loop is unique to the overlap fermion. When and if other groups have adopted the LMS and LMA in their calculations with other fermions to speed up their calculations in the future, the multi-mass feature of the overlap alone will offset its higher inversion time so that it can be as precise as other fermion formulations, given the same resources. However, as we stressed in Sec. 1.1, the overlap fermion has chiral symmetry which will make it more accurate in quantities which are more sensitive to chiral symmetry and it has smaller systematic error for the continuum extrapolation due to its smaller $\mathcal{O}(a^2)$ error.

We would like to concentrate on a few projects which we believe will reveal and highlight the desirable features of the chiral fermion nature of the overlap fermion to justify the large effort that we have spent on speeding up the inversion and the large production of valence quark propagators and loops to construct nucleon correlators and three-point functions. The projects include the quark spin and g_A from the chiral current, the glue spin, and the decomposition of the proton spin and mass into its quark and glue components.

1.3 Proposed Projects

In the next year, we would like to concentrate on a few projects which will likely have major impact on phenomenology and will bear direct relevance to experimental and theoretical interests.

1. Quark spin from chiral axial-vector current:

   The quark spin content of the nucleon was found to be much smaller than that expected from the quark model by the polarized deep inelastic lepton-nucleon scattering experiments and the recent global analysis reveals that the total quark spin contributes only $\sim 30\%$ to the proton spin [40, 41].

   In an attempt to understand the smallness of the quark spin contribution from first principles, several lattice QCD calculations [42, 43] have been carried out since 1995 with the quenched approximation or with heavy dynamical fermions [44]. The most challenging part of the lattice calculation is that of the disconnected insertion of the nucleon three-point functions due to the quark loops. Recently, the strange quark spin $\Delta s + \Delta\bar{s}$ has been calculated with the axial-vector current on light dynamical fermion configurations [45, 46, 47, 48, 49, 50] and it is found to be in the range from -0.02 to -0.03. This is about 3 to 4 times smaller in magnitude than that from a global fit of DIS which gives $\Delta s + \Delta\bar{s} \approx -0.11$ [41] and a most recent analysis [51] including the JLab CLAS high precision data which finds it to be $-0.106(23)$ [52].

   Such a discrepancy between the global fit of experiments and the lattice calculation of the quark spin from the axial-vector current is unsettling. It was emphasized some time ago that it is essential that a lattice calculation of the flavor-singlet axial-vector current be able to accommodate the triangle anomaly [53, 54]. It was specifically suggested [53] to calculate the triangle anomaly from the VVA vertex and take it as the normalization condition for the axial-vector current in order to determine the normalization factor κ_A on the lattice. To address the discrepancy of the strange quark spin, we used the anomalous Ward identity (AWI) to provide the normalization and renormalization conditions to calculate the strange and charm quark spins. The anomalous Ward identity includes the triangle anomaly in the divergence of the flavor-singlet axial-vector current

   $$\kappa_A \partial^\mu A^0_\mu = 2 \sum_{f=1}^{N_f} m_f \bar{q}_f \gamma_5 q_f + N_f 2q, \tag{1}$$

   where q is the local topological charge operator. We put this identity between the nucleon states and calculate the matrix element on the right-hand side with a momentum transfer $\vec{q}$. Since we use the overlap fermion for the calculation of the pseudoscalar density, its renormalization constant cancels that of the renormalization of the quark mass, i.e. $Z_m\, Z_P = 1$ for the chiral fermion. Also, the topological charge density, when calculated with the overlap operator, i.e. $q(x) = Tr\gamma_5(1 - \frac{1}{2}D_{ov}(x,x))$, does not require multiplicative renormalization. Thus, when the matrix elements on the right-hand side are so calculated, the flavor-singlet axial-vector current is automatically *normalized* non-perturbatively *à la* anomalous Ward identity (AWI). There is a two-loop renormalization for the axial current. However, it is shown that using the overlap fermion and its operator for the topological charge density, the renormalized AWI is the same as that of the unrenormalized but normalized AWI in Eq. (1) [32].

We calculated the strange and charm quark contributions to the nucleon spin from the anomalous Ward identity (AWI), using overlap valence quarks on 2+1-flavor domain-wall fermion gauge configurations on a $24^3 \times 64$ lattice with the light sea mass at $m_\pi = 330$ [32]. For the charm quark, we find that the $2mP$ and the anomaly contributions almost cancel. For the strange quark, the $2mP$ term is somewhat smaller than that of the charm. Instead of taking the $q^2 \to 0$ limit for the $g_P(q^2)$ and $g_G(q^2)$, we calculated the unnormalized $g_A^L(q^2) = g_A(q^2)/\kappa_A$ and the induced pseudoscalar form factor $h_A^L(q^2) = h_A(q^2)/\kappa_{h_A}$ and fit $2m_N g_A(q^2) + q^2 h_A(q^2) = 2m g_P(q^2) + g_G(q^2)$ to obtain $g_A^s(0)$. At $m_\pi = 330$ MeV, κ_A turns out to be l.36(4) and the chiral extrapolation to the physical pion mass gives $\Delta s^+ = \Delta s + \Delta \bar{s} = -0.0403(44)(67)$. This $\kappa_A = 1.36(4)$ is found to be larger than that for the isovector axial-vector current and explains why our result is larger in magnitude than the other lattice results. This implies that it is affected by a large cutoff effect for the point current, presumably due to the triangle anomaly. We now have results for the u or d in the disconnected insertion (DI) on the 24I, 32I, and 32ID lattices and found that it is $\sim -0.102(17)$ with a $\kappa_A \sim 1.7$ from the joint fit of chiral and continuum extrapolations, while $\Delta s^+ = -0.051(14)$. Combined with the CI, this gives the total quark spin to be $\Sigma = 0.312(31)$ which is in agreement with a recent global fit of the DIS experimental which gives $\Sigma = 0.25(10)$ [56]. We note that the lattice calculation at this stage has a much smaller error than that deduced from DIS experiments. We understand the latter will be improved when EIC is online. Also, the isovector $g_A^3 = 1.253(14)$ from our multi-lattice fit is within one sigma of the precise experimental value of 1.2723(23).

There is a way for the overlap fermion to do better than what have been achieved so far for the vector and axial-vector current matrix elements. There is a formulation to obtain the conserved vector and chiral axial-vector currents for the overlap fermion [58]. While there are conserved vector currents for other non-chirally symmetric fermion formulations, there are no chiral axial-vector currents for these fermions. Although it is more complicated to calculate it numerically, as it involves a derivative of an extra $U(1)$ gauge link in the Wilson kernel of the overlap operator, we have managed to implement it in our existing code. We have successfully tested it to show that the proton charge is unity to machine precision. The Ward identity for the axial-vector current is shown to satisfy to sub-percent precision configuration by configuration. We propose to calculate the quark spin with this chiral axial-vector current directly for the forward matrix element for both the connected and disconnected insertions on the 48I, 24IDC, and 32IDC lattices inTable 1 which should have smaller errors than going through the anomalous Ward identity that we have been doing for the non-chiral point current. We will work on the other 3 smaller lattices elsewhere. When all the results are available on these 6 lattices, we will perform a global fit with an aim of making a definite prediction of the quark spin contribution to the proton spin with high precision (i.e. 10% for the combined statistical and systematic errors) at the physical point and with the systematic errors due to continuum and large volume extrapolations under control.

2. Glue spin:

   As we reported in the Progress Report, we have carried out a calculation of the glue spin S_G on the lattice with the same set of $2 + 1$ flavor dynamical domain-wall configurations on the 24I, 32I, 32If, 32ID and 48I lattices which represent ensembles with 4 lattice spacings, 4 volumes, and 4 sea quark masses including one at the physical pion mass [24]. We obtained results for the longitudinal nucleon momenta p_z from 0 to 1.4 GeV. The chiral interpolation of the $p_z = 0$ from 25 valence quarks on these 5 lattices are found to be rather insensitive to the pion mass. The large moment extrapolation is carried out.

   The $O(a^2)$ and chiral extrapolated glue spin result is $\Delta G = 0.251(47)(16)$ at $\mu^2 = 10\text{GeV}^2$ which indicates that half of the proton spin resides in the glue spin. This is consistent with the recent global analysis [4] of COMPASS and RIHC data, albeit it has a large uncertainty for the region $x < 0.05$.

   This work [24] used some glue operators generated on Stampede and it is published in Physical Review Letters as the editor's suggestion of the issue. It is reported in Physics Viewpoint by Steve Bass – Viewpoint: Spinning Gluons in the Proton (https://physics.aps.org/articles/v10/23).

   Despite the fact that this work is included in Physics of the American Physical Society as one of the eight 'Highlights of the Year' in 2017 (https://physics.aps.org/articles/v10/137), it is not complete. The one-loop perturbative matching in the Large Momentum Effective Theory from finite momentum to infinite

momentum in the Coulomb gauge turns out to be large and not useable. We propose to consider the temporal gauge instead which does not have one-loop matching correction. We will also apply the variance reduction technique through the cluster decomposition which could improve the statistics by an order of magnitude for the large 48I, 24IDC, and 32IDC lattices to reduce errors.

3. Quark and glue angular momentum fractions in the nucleon:

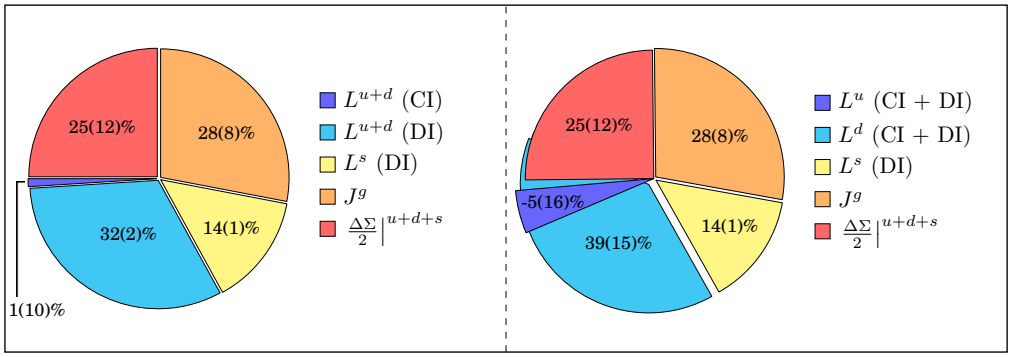

Figure 1: Decomposition of the proton spin in terms of quark spin, glue angular momentum, and quark orbital angular momentum for the u, d and s quarks.

Since we have gone through a complete calculation of the quark and glue momenta and angular momenta in the quenched approximation [33], we have been working on the same calculation with valence overlap fermion on $2+1$-flavor dynamical domain-wall fermion gauge configurations on the smaller 24I, 32I, and 32ID lattices listed in Table 1 in an attempt to obtain more definitive calculations of the proton momentum and angular momentum components at the physical pion mass and in the continuum limit with realistic dynamical quarks in the vacuum with which to compare with experiments and make predictions. In this regard, it is important to evaluate the components on the larger 48I, 24IDC, and 32IDC lattices to control the systematics due to finite volume and chiral interpolation.

Pie charts for the momenta and angular momenta decomposition of the nucleon calculated in the quenched approximation are shown in [33].We would like to produce the corresponding plots with 2+1 flavor dynamical fermion configurations at the physical pion mass and include the systematic errors due to the continuum and infinite volume limits.

We have several subgroups in the collaboration which are specialized in working on different projects as listed above. N. Mathur's group in Mumbai, Y. Chen's group in Beijing, and those of us at University of Kentucky can each carry out independent analysis and fitting of the two-point correlators to study hadron masses and decay constants. We also have several experienced subgroups to tackle the nucleon form factor part of the project. Takumi Doi, Ming Gong, Jain Liang, and Yibo Yang have been working on various nucleon form factors and glue momentum with quark propagators generated on the 2+1 flavor dynamical clover configurations. Frank Lee and Andrei Alexandru at George Washington University have also done calculations with quenched and dynamical fermion gauge configurations. Once the valence quark propagators and loops are generated, they can begin constructing the three-point correlators to carry out the physics calculations.

2 Replies to Reviewers

2.1 Replies to Reviewer # 1

- 'In spite of the excellent quality of the proposed research, the proposers must be much more careful in future proposals to acknowledge XSEDE resources and to clearly delineate in their progress report what they have calculated with XSEDE resources. Of the 3 articles referenced in their progress report, only one of them explicitly acknowledges XSEDE resources. This must be corrected and future proposals will be rejected if this occurs again.'

Thanks for the reminder from the reviewer. We have indeed been sloppy in acknowledging the support of the XSEDE resources in previous years. We have started to be careful about this since last year.

- 'In addition, it is important for large project like this that the proposers clearly delineate what is being done on what resources in what year. There is no mention of computing support fromTitan in the list of "Other Support" and yet eigenvectors on 65 of the congurations were generated on Titan. Clearly delineating what was done on which resource at what time is vital for requests of this size.'

  The eigenvectors on 65 configurations were generated on Titan through an ALCC award in 2014. We thought the "Other Support" meant the current year support from other resources. We have been working on the smaller set of lattices in Table 1 since 2008. It is not clear how far back should we list the various accumulated results on different lattices from different resources. It would be useful to have a guideline on this. We will try to delineate the resources for major production and analysis runs.

- 'In addition to acknowledging XSEDE in the Acknowledgements section, you should also cite theXSEDE paper. Instructions are on the web.'

  Yes, we will adhere to the instructions in citing the XSEDE papers.

2.2 Replies to Reviewer # 2

- 'Some references are made to the used of Oak Ridge computing facility yet this resource is not mentioned in "Other Supercomputing Support". Therefore it is not clear what has been accomplished from past XSEDE support with rather sizeable allocation. This should be clean up in future renewal.'

  The eigenvectors on 65 configurations were generated on Titan through an ALCC award in 2014. We have interpreted (perhaps wrongly) the "Other Support" to mean the current year support from other resources. We have been working on the smaller set of lattices in Table 1 since 2008, It is not clear how far back should we list the various accumulated results on different lattices from different resources. It would be useful to have a guideline on this. We will try to delineate the resources for major production and analysis runs.

- 'Another issue is how they plan to run jobs. The jobs will be bundled to run with 16K for 5 hours a day and they plan to run 3 of this 16K jobs a day to burn their time. There is a big assumption that the 3 16K jobs will get through the queue every day consistently to burn the time. Although theoretically this is possible, I am not so sure if it is practically feasible, but I am willing to let them try it.'

  We have used the maximally allowed 16,234 cores to run last years allocation of 10.4 M hours and have almost exhausted this amount in just 3 and 1/2 months from Apr. 1 to mid July, 2016. Given this record, we fully expect to have been able to use up in a year the 33 M hours that we had requested.

2.3 Replies to Reviewer # 3

- 'I reject this proposal. It is just not feasible to run 16 k jobs for 5 hours every day on Stampede. 33 MSU is twice the size of last year's allocation. I very much appreciate the detailed response to the most recent review. But the response to reviewer #4 makes clear that Stampede is not the right resource for this project. Running 3 15k jobs every day for 5 hours is not feasible. With the queue restrictions the team will not be able to burn through 33 MSU in a year. Also, the response make clear that even large runs are in planning (up to 46 MSU). As pointed out 33 MSU are needed to reach statistical accuracy.The group needs to apply to a different resource where the usage model supports multiple high core-count runs every day.'

  We respectively disagree with the reviewer. We have addressed this issue in the reply to the reviewer in last years proposal, as there was such a question from the review of the previous year, which was acknowledged by this reviewer. We have used the maximally allowed 16,234 cores to run last years allocation of 10.4 M hours and have almost exhausted this amount in just 3 and 1/2 months from Apr. 1 to mid July, 2016. This can be easily checked with the account management. Given this record, we fully expect to be able to use up in a year the 33 M hours that we requested. In fact, had we tried to run two such large jobs on most of the days from two users (we have done so on some days and have not encountered a problem), we could

probably have finished the 10.4 M hours in less than 2 months. We don't understand how the reviewer came up with the conclusion that our proposed mode of job running is not feasible. Our experience on Stampede1 and Skylake this year do not support the reviewer's conception.

3 Computational Strategy and Algorithm

To better explain the computational strategy and the resource request we need to review the properties of the overlap operator. We will discuss first the D_W multiplication routine, the most time consuming part of our codes, then present the relevant details for the overlap operator, and then discuss the eigensolver and inverter we use.

In lattice QCD the space-time is approximated by a four dimensional grid, the quarks are viewed as particles hopping between the grid sites and gluons are represented by parallel transporters that change the internal state of the quarks as they hop along the given link. The quark operator represents a discretization of the covariant derivative $D_\mu = \partial_\mu + igA_\mu$, where A_μ is the color (gluon) field. The quark fields, ψ_n, are represented by 4×3 matrices at each lattice site and the glue link $U_\mu(n) = e^{iagA_\mu(n)}$ by $SU(3)$ matrices. Wilson discretization of $m + \not{D} = m + \gamma_\mu D_\mu$ is given by the following matrix

$$D_W(m;U) = (ma+4)\mathbb{1} - \frac{1}{2}\sum_{\mu=\pm1}^{\pm4} T_\mu(U)\,. \tag{2}$$

where T_μ are the *parallel transporters* for all 8 directions. D_W is a complex $12V \times 12V$ matrix, where V is the number of sites on the grid. The matrix is very sparse since the parallel transporters only connect to nearest-neighbor sites. For numerical simulations, we store only the non-zero elements of this matrix and we implement a `dslash` routine to compute $D_w\psi$ on any given quark field ψ. The storage requirements and the numerical cost for the `dslash` routine are proportional to V. D_W is γ_5-symmetric, i.e. $D_W^\dagger = \gamma_5 D_W \gamma_5$. Thus, $H_W \equiv \gamma_5 D_W$ is hermitian.

Our codes are written in C++ using MPI for interprocess communication. The dslash kernel is implemented using inline SSE instructions. The lattice is divided in equal domains and each MPI process is responsible for one domain: all data associated with the lattice sites in that domain is stored in the memory controlled by the process and all associated calculations are performed by the execution unit it controls. Communication between domains is performed via two-sided MPI calls.

The division strategy we used is designed to minimize the number of sites on the boundary. This is based on the assumption that the communication time will be proportional to the boundary area. This assumption is valid when the communication speed between any two processes is the same. For heterogeneous systems, where the network "distance" between processes varies, this division strategy might not be optimal.

The overlap operator is defined in terms of D_W,

$$D_{\rm ov} = 1 + \gamma_5 {\rm sign}(H_W)\,, \tag{3}$$

where ${\rm sign}(H_W)$ is a matrix sign function of the matrix H_W. We use iterative algorithms that require the ability to compute matrix-vector multiplication, $\psi \to D_{\rm ov}\psi$. To compute the overlap operator, the sign function needs to be evaluated. The sign function cannot be accurately approximated around zero and, to overcome this problem, the strategy is to resolve the spectrum of the H_W operator around zero with eigenvectors $|\lambda_i\rangle$, and separate the source in $\psi = \psi_\parallel + \psi_\perp$. The sign function for the low-lying component $\psi_\parallel = P_\parallel\psi = \sum_i \langle\lambda_i|\psi\rangle|\lambda_i\rangle$ is computed exactly,

$${\rm sign}H_W\,\psi_\parallel = \sum_i {\rm sign}(\lambda_i)\langle\lambda_i|\psi\rangle|\lambda_i\rangle.$$

The sign function for the high-frequency component $\psi_\perp = P_\perp\psi$, on the other hand, is computed using an approximation for the sign function. Two commonly used approximations are the rational approximation and the polynomial approximation. We find that the polynomial approximation

$${\rm sign}H_W\,\psi_\perp = \sum_k c_k H_w^k \psi_\perp\,,$$

is better by a factor of two [59]. The polynomial is determined as an expansion in terms of Chebyshev polynomials using Remez algorithm, taking into account the spectrum of the $P_\perp H_W P_\perp$ operator and the required precision; the matrix multiplication is done in a numerically stable way using Clenshaw recursion [60].

The number of H_W eigenvectors is determined by minimizing the computational cost required to compute $D_{\rm ov}\psi$. Using a large number of vectors speeds up the high-frequency part by lowering the degree of the polynomial required to approximate the sign function. On the other hand, as we increase the number of vectors the cost of calculating the low-lying part increases and the decrease in the polynomial order slows down due to an increase in the density of eigenmodes. For the $48^3 \times 96$ lattice, given the performance of D_W routine on 4096 CPU-cores we find that the optimal number of vectors is 800. The total time for $D_{\rm ov}$ multiplication is 1.7 seconds in this configuration on Stampede.

To compute the quark propagators, we use a conjugate gradient (CG) method, which can be adapted to compute multiple quark masses without extra matrix-vector multiplication [63]. This is possible because the change in mass of the operators amounts to a simple shift and the Krylov spaces are identical. To further increase the efficiency of our codes we use an adaptive CG method [64]: the precision of the $\psi \to D_{\rm ov}\psi$ is decreased as we near convergence. The savings come from using a smaller order polynomial to approximate the sign function and, when the target precision is lower than 10^{-7}, we employ the single precision version of D_W, which speeds the multiplication by a factor of two. For a quark mass corresponding to $m_\pi \approx 200$ MeV the adaptive method is 60% faster than the regular CG.

For each configuration, 48 propagators for 7 different masses (about 8000 GB) need to be written and read from disk for further analysis. This is one of the potential bottlenecks we identified. We profiled the I/O routines, identified the code sections that had to be streamlined and implemented a parallel I/O strategy. We benchmarked the new code and determined that for the $48^3 \times 96$ lattices that we intend to run on 43 skylake nodes, the optimal number of I/O processes is 16, when the Lustre file system is tuned to use 16 stripes of size $48^3 \times 96 \times 8/16$. The disk I/O in the new codes is about 5GB/s. Thus, the I/O time for every propagator is of the order of a few minutes and the overhead associated with it is at the one percent level.

In order to improve the signal to noise ratio (SNR) for the connected insertion, we developed a stochastic sandwich contraction method [65] to remove the need of multiple inversions in the sink-sequential approach and use the current-sequential method for the low modes in the propagator between the current and the sink. This is an extension of the noise grid smeared source with LMS to the three point function. Such a many-to-all correlator with LMS is useful when the low-eigenmode contributions are important in the relevant time windows where the physical quantities are extracted. With this strategy and pointed data buffer techniques, the cost of the connected insertion is reduced by an order of magnitude and we can reach the same precision comparing with the simulation with much cheaper twisted mass action, in the scalar current case [22, 66].

To improve the quark loop calculation, we use low-mode averaging (LMA) for the low-mode contribution by summing over the spatial volume on a time slice, while the high-mode contribution is calculated with Z_4 noise estimator on the $4, 4, 4, 2$ grid with odd-even and time dilution. This strategy has been applied in several cases [22, 67, 68].

4 Request for Computer Time and Storage

We request time to carry out the proposed projects on the 48I, 24IDC, and 32IDC lattices. The 1000 pairs of lowest eigenvectors of the 48I lattice were produced sometime ago which will be used for deflation in matrix inversions and low-mode substitution in correlator construction. The 600 pairs of the lowest eigenvectors of the 24IDC are produced on the GPU cluster at BNL in 2017 and 1000 pairs of them on the 32IDC lattice will be produced on the same cluster during the first half of 2018. We will copy them to TACC to be used for the proposed projects.

For the inversion of a quark propagator of the 48I lattice with 12 columns of color and spin for the grid source, the average time it takes is about 21 node hours on SKX. Since we use low-mode substitution (LMS), the inversion for the high modes can be done with lower precision, i.e. with a 10^{-4} residual. This saves a factor of 3 in time compared to the alternative with LMS not applied in which case higher precision would be needed in inversion. For each configuration, we will need to have 6 noise grid sources on different space-time locations and 2, 3, 4, 5,

and 12 sink sources at 5 different source-sink separations to remove the excited state contamination for a total 26 inversions for the quark propagators from the sink time slices. To construct three-point functions with LMS to calculate quark spin and quark angular momentum will require 2.7 node hours for each quark mass, momentum transfer, noise source and source-sink separation.

The total time needed for calculating the three-point functions in the connected insertions for the 48I lattice includes the following:

1) 6 quark propagators from the source side and 26 from the sink side for the sandwich approach needs 32 inversions which costs $32 \times 21 = 672$ node hours. The construction of the chiral axial-vector current on each time slices cost about one inversion, so that the inversion times is doubled to $1,344$ hours.

2) Constructing the three-point function with 6 sources, 5 source-sink separations, 7 valence quark mass, and 7 momentum transfer will require $6 \times 5 \times 7 \times 7 \times 3.2 = 4{,}704$ node hours.

For the disconnected insertions, we use the low-mode average (LMA) for the exact contribution of the loop from the low modes. For the high modes, we will need an additional 2 Z_4 noise estimator on a 4,4,4,2 grid with odd-even and time dilution. This will take 16 quark inversions on each configuration. The construction of the chiral axial-vector current doubles the inversion to an equivalent of 32 quark inversions. The construction of the loop for 7 valence quark mass and 7 momenta will take 10 node hours. The total time for the disconnect insertion is therefore $32 \times (21 + 10) = 992$ node hours.

The total computer time for 81 configurations of the 48I lattice is $(1,344 + 4,704 + 992) \times 81 = 570,240$ node hours. Using the request for the 48I lattice to scale down to the 24IDC and 32IDC lattices based on the ratio of inversion times will cost 440 and 1,408 node hours per configuration. The total for the 24IDC and 32IDC lattice with 200 configurations is thus $(440 + 1,408) \times 200 = 369,600$ node hours. The total requested time for the 48I, 24IDC and 32IDC is 570,240 + 369,600 = 939,840 node hours.

In the past few months, we have been able to run 25 jobs on SKX almost concurrently (with a queuing time of about an hour) on 43 nodes each for 40 hours. This takes up 1,075 nodes. We presume this is because SKX is not fully occupied. Come next year, we plan to again queue up 25 jobs on 43 nodes each and run close to the limit of 48 hours. Assuming a reduction factor of 0.3 for the wait time in queue and idle time (i.e. only 8 jobs out of 25 are run at the same time), the 939,840 node hours that we quested can be used up in 114 days.

Since last year, we have started to save the quark propagator files to NERSC via globus. This is faster than saving them to tapes at TACC. Thus we will not ask for storage space of the propagators. We only need storage for the eigenvectors. Each eigenvector takes $48^3 \times 96 \times 12 \times 16 = 2.04$ GB. The total storage for 1000 overlap eigenvecrtors and 500 Wilson eigenvectors on 81 configurations takes is $0.002 \times 1500 \times 81 = 243$ TB. Similarly, the storage for 1100 eigenvectors (overlap and Wilson) of the 24IDC lattice for 200 configurations will be 28TB and storage for 1500 eigenvectors of the 32IDC lattice for 200 configurations will need 121 TB. Together, we will need 243 + 28 + 121 = 392 TB of storage

All together, we request 940,000 SUs on Stampede2 at TACC and 392 TB of storage at TACC.

5 Local Computing Environment and Other Supercomputer Support

We are allotted 65 million processor Hopper-equivalent hours on Edison and Cori this year at NERSC to calculate nucleon structure and hadron spectroscopy for the 32ID lattice to be combined with results from other lattices for the continuum, large volume and chiral extrapolations. We are also given 900k GPU hours on the BNL cluster and 12 million hours on JLab cluster from July 1, 2017 to June 30, 2018 from the USQCD Lattice Collaboration. We also have 37 million hours on Titan to calculate the hadronic tensor from the INCITE program this year. In 2014 we were awarded 67 million hours on Titan from the ALCC program which we used to calculate 1000 pairs of eigenvectors for 65 configurations of the 48I lattice. The eigenvectors of the remaining 16 configurations were calculated on Stampede 1 through XRAC in 2014-2015.

References

[1] DOE/NSF Nuclear Science Advisory Committee, "The 2015 Long Range Plan for Nuclear Science," 2015, https://science.energy.gov/~/media/np/nsac/pdf/2015LRP/2015_LRPNS_091815.pdf

[2] J. Ashman *et al.*, *Phys. Lett.* **B206**, 364 (1988).

[3] M. Stolarski [COMPASS Collaboration], *Nucl. Phys. Proc. Suppl.* **207-208**, 53 (2010); P. Djawotho [STAR Collaboration], *Gluon polarization and jet production at STAR*, *J. Phys. Conf. Ser.* **295**, 012061 (2011).

[4] D. de Florian, R. Sassot, M. Stratmann and W. Vogelsang, Phys. Rev. Lett. **113**, 012001 (2014) [arXiv:1404.4293 [hep-ph]].

[5] P. Djawotho for the STAR Collaboration 2013, [arXiv:1303.0543 [nucl-ex]].

[6] A. Adare et al. (PHENIX Collaboration), Phys.Rev. **D90**, 012007 (2014), [arXiv:1402.6296 [hep-ex]].

[7] D. Kaplan, *Phys. Lett.* **B288**, 342 (1992), [hep-lat/9206013].

[8] Y. Shamir, *Nucl. Phys.* **406**, 90 (1993), [hep-lat/9303005].

[9] R. Narayanan and H. Neuberger, *Nucl. Phys.* **B 443**, 305 (1995).

[10] H. Neuberger, *Phys. Lett.* **B 417**, 141 (1998).

[11] For reviews, see for example, H. Neuberger, *Nucl. Phys.* **B (Proc. Suppl.) 83-84**, 67 (2000); F. Niedermayer, *Nucl. Phys.* **B 73** (Proc. Suppl.), 105 (1999).

[12] P. Hasenfratz, *Nucl. Phys.* **B525**, 401 (1998).

[13] C. Allton et al. (RBC-UKQCD), *Phys. Rev.* **D76**, 014504 (2007), [hep-lat/0701013].

[14] C. Jung (RBC-UKQCD), *PoS(LATTICE 2007)*, 037 (2007), [arXiv:0710.5337].

[15] R. Arthur *et al.* [RBC and UKQCD Collaborations], Phys. Rev. D **87**, no. 9, 094514 (2013), [arXiv:1208.4412 [hep-lat]].

[16] T. Blum *et al.* [RBC and UKQCD Collaborations], Phys. Rev. D **93**, no. 7, 074505 (2016), doi:10.1103/PhysRevD.93.074505, [arXiv:1411.7017 [hep-lat]].

[17] S. J. Dong, F. X. Lee, K. F. Liu and J. B. Zhang, Phys. Rev. Lett. **85**, 5051 (2000), [hep-lat/0006004].

[18] Y. Chen, S. J. Dong, T. Draper, I. Horvath, F. X. Lee, K. F. Liu, N. Mathur and J. B. Zhang, Phys. Rev. D **70**, 034502 (2004), [hep-lat/0304005].

[19] A. Li, A. Alexandru, Y. Chen, T. Doi, S.J. Dong, T. Draper, M. Gong, A. Hasenfratz, I. Horvath, F.X. Lee, K.F. Liu, N. Mathur, T. Streuer, and J.B. Zhang, *Phys. Rev.* **D82**, 114501 (2010), [arXiv:1005.5424].

[20] Y. B. Yang, Y. Chen, A. Alexandru, S. J. Dong, T. Draper, M. Gong, F. X. Lee and A. Li *et al.*, arXiv:1401.1487 [hep-lat].

[21] http://pdg8.lbl.gov/rpp2013v2/pdgLive/Viewer.action

[22] Y. B. Yang, A. Alexandru, T. Draper, J. Liang and K. F. Liu, Phys. Rev. D **94**, no. 5, 054503 (2016), doi:10.1103/PhysRevD.94.054503, arXiv:1511.09089 [hep-lat].

[23] R. S. Sufian, Y. B. Yang, A. Alexandru, T. Draper, K. F. Liu and J. Liang, Phys. Rev. Lett. **118**, no. 4, 042001 (2017), doi:10.1103/PhysRevLett.118.042001, [arXiv:1606.07075 [hep-ph]].

[24] Y. B. Yang, R. S. Sufian, A. Alexandru, T. Draper, M. J. Glatzmaier, K. F. Liu and Y. Zhao, Phys. Rev. Lett. **118**, no. 10, 102001 (2017), doi:10.1103/PhysRevLett.118.102001, [arXiv:1609.05937 [hep-ph]].

[25] Y. Zhao, K. F. Liu and Y. Yang, Phys. Rev. D **93** (2016) no.5, 054006, doi:10.1103/PhysRevD.93.054006, [arXiv:1506.08832 [hep-ph]].

[26] C. Michael *et al.* [UKQCD Collaboration], Nucl. Phys. Proc. Suppl. **106**, 293 (2002), doi:10.1016/S0920-5632(01)01692-9, [hep-lat/0109028].

[27] K. Takeda *et al.* [JLQCD Collaboration], Phys. Rev. D **83**, 114506 (2011) doi:10.1103/PhysRevD.83.114506 [arXiv:1011.1964 [hep-lat]].

[28] G. S. Bali *et al.* [QCDSF Collaboration], Phys. Rev. D **85**, 054502 (2012), doi:10.1103/PhysRevD.85.054502, [arXiv:1111.1600 [hep-lat]]; G. S. Bali *et al.* [RQCD Collaboration], Phys. Rev. D **93**, no. 9, 094504 (2016) doi:10.1103/PhysRevD.93.094504 [arXiv:1603.00827 [hep-lat]].

[29] M. Gong *et al.* [χQCD Collaboration], Phys. Rev. D **88**, no. 1, 014503 (2013), [arXiv:1304.1194 [hep-ph]].

[30] Y. B. Yang, A. Alexandru, T. Draper, M. Gong and K. F. Liu, Phys. Rev. D **93**, no. 3, 034503 (2016), doi:10.1103/PhysRevD.93.034503, [arXiv:1509.04616 [hep-lat]].

[31] G.P. Lepage, "The Analysis of Algorithm for Lattice Field Theory", invited lecture given at TASI 1989 Summer School, Boulder, CO, June 4-20, 1989, Published in Boulder, ASI 1989:97-120 (QCD161: T45:1989).

[32] M. Gong *et al.* [?QCD Collaboration], Phys. Rev. D **95**, no. 11, 114509 (2017) doi:10.1103/PhysRevD.95.114509 [arXiv:1511.03671 [hep-ph]].

[33] M. Deka, T. Doi, Y. B. Yang, B. Chakraborty, S. J. Dong, T. Draper, M. Glatzmaier and M. Gong, H.W. Lin, K.F. Liu *et al.*, Phys. Rev. D **91**, no. 1, 014505 (2015), doi:10.1103/PhysRevD.91.014505, [arXiv:1312.4816 [hep-lat]].

[34] H. Ying, S.J. Dong, and K.F. Liu, *Nucl. Phys. (Proc. Suppl.)* **53**, 993 (1997).

[35] S.J. Dong, F.X. Lee, K.F. Liu, J.B. Zhang, *Phys. Rev. Lett.* **85**, 5051-5054 (2000).

[36] A. Abdel-Rehim, C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou and A. V. Aviles-Casco, Phys. Rev. Lett. **116**, no. 25, 252001 (2016), doi:10.1103/PhysRevLett.116.252001, [arXiv:1601.01624 [hep-lat]].

[37] B. Yoon *et al.*, Phys. Rev. D **93**, no. 11, 114506 (2016), doi:10.1103/PhysRevD.93.114506, [arXiv:1602.07737 [hep-lat]].

[38] S. Syritsyn, J. Phys. Conf. Ser. **640**, no. 1, 012054 (2015). doi:10.1088/1742-6596/640/1/012054.

[39] K. F. Liu, J. Liang and Y. B. Yang, arXiv:1702.04384 [hep-lat].

[40] A. Accardi *et al.*, Eur. Phys. J. A **52**, no. 9, 268 (2016) doi:10.1140/epja/i2016-16268-9 [arXiv:1212.1701 [nucl-ex]].

[41] D. de Florian, R. Sassot, M. Stratmann and W. Vogelsang, Phys. Rev. D **80**, 034030 (2009), [arXiv:0904.3821 [hep-ph]].

[42] S. J. Dong, J.-F. Lagae and K. F. Liu, Phys. Rev. Lett. **75**, 2096 (1995), doi:10.1103/PhysRevLett.75.2096, [hep-ph/9502334].

[43] M. Fukugita, Y. Kuramashi, M. Okawa and A. Ukawa, Phys. Rev. Lett. **75**, 2092 (1995), [hep-lat/9501010].

[44] S. Gusken *et al.* [TXL Collaboration], Phys. Rev. D **59**, 114502 (1999).

[45] G. S. Bali *et al.* [QCDSF Collaboration], Phys. Rev. Lett. **108**, 222001 (2012), [arXiv:1112.3354 [hep-lat]].

[46] R. Babich, R. C. Brower, M. A. Clark, G. T. Fleming, J. C. Osborn, C. Rebbi and D. Schaich, Phys. Rev. D **85**, 054510 (2012), [arXiv:1012.0562 [hep-lat]].

[47] M. Engelhardt, Phys. Rev. D **86**, 114510 (2012), [arXiv:1210.0025 [hep-lat]].

[48] A. Abdel-Rehim, C. Alexandrou, M. Constantinou, V. Drach, K. Hadjiyiannakou, K. Jansen, G. Koutsou and A. Vaquero, Phys. Rev. D **89**, no. 3, 034501 (2014), [arXiv:1310.6339 [hep-lat]].

[49] A. J. Chambers *et al.*, arXiv:1508.06856 [hep-lat].

[50] C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou, A. Vaquero Avils-Casco and C. Wiese, Phys. Rev. Lett. **119** (2017) no.14, 142002, doi:10.1103/PhysRevLett.119.142002, [arXiv:1706.02973 [hep-lat]].

[51] E. Leader, A. V. Sidorov and D. B. Stamenov, Phys. Rev. D **91**, no. 5, 054017 (2015), [arXiv:1410.1657 [hep-ph]].

[52] D. Stamenov and E. Leader, private communication.

[53] L. H. Karsten and J. Smit, Nucl. Phys. B **183**, 103 (1981).

[54] J. F. Lagae and K. F. Liu, Phys. Rev. D **52**, 4042 (1995) [hep-lat/9501007].

[55] D. Espriu and R. Tarrach, Z. Phys. C **16**, 77 (1982).

[56] E. R. Nocera *et al.* [NNPDF Collaboration], Nucl. Phys. B **887** (2014) 276, doi:10.1016/j.nuclphysb.2014.08.008, [arXiv:1406.5539 [hep-ph]].

[57] A. J. Chambers *et al.* [QCDSF Collaboration], Phys. Lett. B **740**, 30 (2015), doi:10.1016/j.physletb.2014.11.033, [arXiv:1410.3078 [hep-lat]].

[58] P. Hasenfratz, V. Laliena and F. Niedermayer, Phys. Lett. B **427**, 125 (1998) doi:10.1016/S0370-2693(98)00315-3 [hep-lat/9801021].

[59] A. Alexandru, M. Lujan, C. Pelissier, B. Gamari, and F. X. Lee, *Efficient implementation of the overlap operator on multi-GPUs*, in *Application Accelerators in High-Performance Computing (SAAHPC), 2011 Symposium on*, pp. 123 –130, July, 2011, [arXiv:1106.4964].

[60] L. Giusti, C. Hoelbling, M. Luscher, and H. Wittig, *Numerical techniques for lattice QCD in the epsilon regime*, *Comput.Phys.Commun.* **153** (2003) 31–51 [hep-lat/0212012].

[61] D. C. Sorensen, *Implicit application of polynomial filters in a k-step Arnoldi method*, *SIAM J. on Matrix Analysis and Applications* **13** (1992), no. 1 357–385.

[62] H. Neff, *Efficient computation of low lying eigenmodes of nonHermitian Wilson-Dirac type matrices*, *Nucl.Phys.Proc.Suppl.* **106** (2002) 1055–1057 [hep-lat/0110076].

[63] B. Jegerlehner, *Krylov space solvers for shifted linear systems*, *unpublished* (1996) [hep-lat/9612014].

[64] N. Cundy, J. van den Eshof, A. Frommer, S. Krieg, T. Lippert, *et. al.*, *Numerical methods for the QCD overlap operator. 3. Nested iterations*, *Comput.Phys.Commun.* **165** (2005) 221–242 [hep-lat/0405003].

[65] Y. B. Yang, A. Alexandru, T. Draper, M. Gong and K. F. Liu, arXiv:1509.04616 [hep-lat].

[66] A. Abdel-Rehim, C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou and A. V. Aviles-Casco, arXiv:1601.01624 [hep-lat].

[67] M. Gong *et al.* [XQCD Collaboration], Phys. Rev. D **88**, 014503 (2013) doi:10.1103/PhysRevD.88.014503 [arXiv:1304.1194 [hep-ph]].

[68] M. Sun, Y. Yang, K. F. Liu and M. Gong, PoS LATTICE **2014**, 142 (2015) [arXiv:1502.05482 [hep-lat]].

[69] B. Joo, D. Kalamkar, K. Vaidyanathan, M. Smelyanskiy, K. Pamnany, V. Lee, P. Dubey, W. Watson, *Lattice QCD on Intel® Xeon PhiTM Coprocessors*, *Supercomputing*, (2013) 40–54.

1 Progress Report

We should first note that our work on the glue spin, which is described below in more detail, has been selected as one of the eight 'Highlights of the Year' by Physics of the American Physical Society in 2017 (https://physics.aps.org/articles/v10/137).

1.1 Glue Spin

It is shown [1] that the glue spin in the nucleon can be calculated with the operator

$$\vec{S}_G = \int d^3x \, Tr(\vec{E}_c \times \vec{A}_c) \tag{1}$$

where $\vec{E}_c$ is the electric field in the Coulomb gauge. Although it is gauge invariant, it is frame dependent and thus depends on the proton momentum. It is realized [1] that its infinite momentum frame (IMF) value corresponds to ΔG which is measurable experimentally from STAR, PHENIX, and COMPASS proton-proton scattering. The important outcome of the derivation is that glue spin content is amenable to lattice QCD calculation. To the extent that it can be calculated at large enough momentum frame of the proton with enough precision, it can be compared to the experimental glue helicity ΔG.

We have carried out a calculation of S_G on a set of $2+1$ flavor dynamical domain-wall configurations on the 24I, 32I, 32If, 32ID and 48I lattices (Details of these lattices are described in Table 1 in the main proposal.) They represent ensembles with 4 lattice spacings, 4 volumes, and 4 sea quark masses including one at the physical pion mass [2]. The electric field $\vec{E}$ is constructed from the overlap Dirac operator. The gauge potential $\vec{A}$ is obtained from the smeared gauge link. We obtained results for the longitudinal nucleon momenta p_z from 0 to 1.4 GeV. The chiral interpolation of the $p_z = 0$ from 25 valence quarks on these 5 lattices are presented in the left panel of Fig. 1. They are rather insensitive to the pion mass. The large-moment extrapolation is shown in the right panel of Fig. 1.

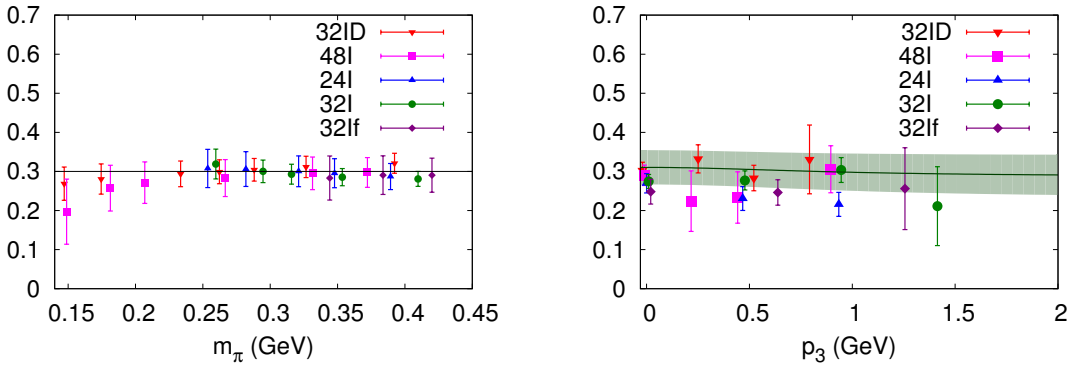

Figure 1: (Left panel): The pion mass dependence of the glue spin in the proton for the case where $p_z = 0$. These dependencies are fairly mild and can be well described with a linear fit. (Right panel): The momentum dependence of the glue spin at the physical pion point.

We see that the results [2] in Fig. 1 have good signals. This is largely due to the highly improved nucleon propagator with sources on multiple time slices and that we used the overlap operator definition for the electric field which is less noisy than that from the link variables. The $O(a^2)$ and chiral extrapolated glue spin is $\Delta G = 0.251(47)(16)$ at $\mu^2 = 10\text{GeV}^2$ which indicates that half of the proton spin resides in the glue spin. This is consistent with the recent global analysis [3] of COMPASS and RIHC data, albeit it has a large uncertainty for the region $x < 0.05$.

This work [2] used glue operators generated on Stampede and it is published in Physical Review Letters as the editor's suggestion of the issue. It is reported in Physics Viewpoint by Steve Bass – Viewpoint: Spinning Gluons in the Proton (https://physics.aps.org/articles/v10/23). It is reported by Emily Conover in Science-News (https://www.sciencenews.org/article/theres-still-lot-we-dont-know-about-proton) and also by Ethan Siegel

in Forbes (https://www.forbes.com/sites/startswithabang/2017/04/19/why-does-the-proton-spin-physics-holds-a-surprising-answer/#71defe9d2c3a). It is also included in Physics of the American Physical Society as one of the eight ' 'Highlights of the Year' in 2017 (https://physics.aps.org/articles/v10/137).

1.2 Strange Quark Magnetic Moment of the Nucleon at Physical Point

We have made a lattice QCD calculation of the strange quark contribution to the proton's magnetic moment and charge radius. This analysis presents the first direct determination of strange electromagnetic form factors including the physical pion mass with chiral fermions.

The determination of the strange (s) quark contribution to nucleon electromagnetic (EM) form factors is of immense importance since this is a pure sea quark effect. A nonzero value of Sachs strange electric form factor G_E^s at any $Q^2 \neq 0$ would mean that the spatial distribution of s and $\bar{s}$ quarks are not the same in the nucleon. The world data constrains that the Sachs strange magnetic form factor $G_M^s(0)$ contributes less than 6% and $\langle r_s^2 \rangle_E$ contributes less than 5% to the magnetic moment and the mean-square charge radius of the proton, respectively [4]. However, the experimental results are limited by rather sizable error bars. The two recent global analyses give $G_M^s(Q^2 = 0.1\,(\mathrm{GeV/c})^2) = 0.29 \pm 0.21$ [5] and -0.26 ± 0.26 [6] which are consistent with zero and differ in sign in their central values.

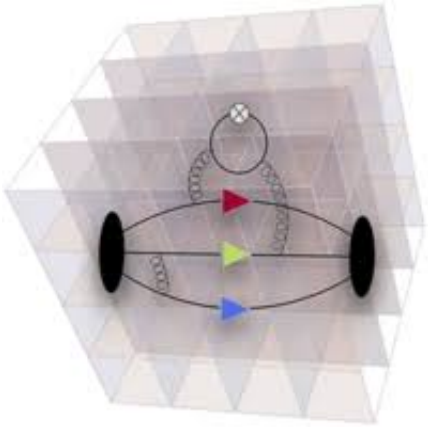

Figure 2: Cartoon depiction of the stangeness magnetic moment in the lattice calculation as a disconnected insertion of the strange quark loop of the vector current correlated with the nucleon propagator.

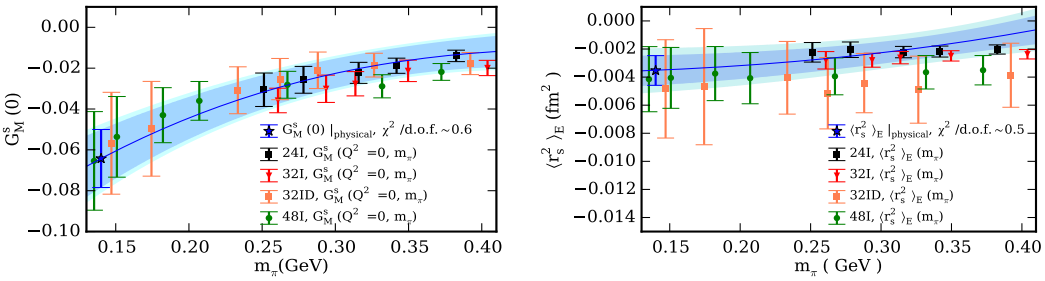

Figure 3: Strange magnetic moment (left panel) and strange charge radius (right panel) at 24 quark masses on 24I, 32I, 32ID, and 48I ensembles as a function of the pion mass. The curved bands on the figures show the behavior in the infinite volume and continuum limits.

We have performed a robust first-principles lattice QCD calculation using four different $2+1$ flavor dynamical fermion lattice ensembles including, for the first time, the physical pion mass with chiral fermion to explore the quark mass dependence and with finite lattice spacing and volume corrections to determine the strange quark matrix elements in the vector channel. This entails a disconnected insertion calculation correlating the strange quark loop with the nucleon propagator as depicted in Fig. 2.

A model-independent extraction of the strange magnetic moment and the strange charge radius from the electromagnetic form factors is carried out in the momentum transfer range of $0.051\,\mathrm{GeV}^2 \lesssim Q^2 \lesssim 1.31\,\mathrm{GeV}^2$. The finite lattice spacing and finite volume corrections are included in a global fit with 24 valence quark masses on four lattices with different lattice spacings, different volumes, and four sea quark masses including one at the physical pion mass. They are plotted in Fig. 3. We obtain the strange magnetic moment $G_M^s(0) = -0.064(14)(09)\,\mu_N$. The 4-sigma precision in statistics is achieved partly due to the low-mode averaging of the quark loop, the use of multiple smeared grid sources, and low-mode substitution of the nucleon propagator to improve the statistics of the nucleon propagator. We also obtain the strange charge radius $\langle r_s^2 \rangle_E = -0.0043(16)(14)\ \mathrm{fm}^2$.

Fig. 4 compares our result of $G_M^s(0)$ and $G_M^s(Q^2 = 0.1\,\mathrm{GeV}^2) = -0.047(11)(06)$ with some other measurements of $G_M^s(0)$ and a global fit of G_M^s at $Q^2 = 0.1\,\mathrm{GeV}^2$. We should note that the error of our *ab initio* calculation is an order of magnitude smaller than those of the global analyses (red circles) from experiments.

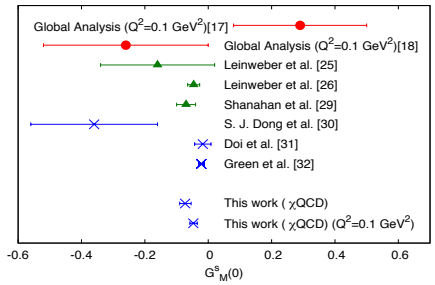

Figure 4: Comparison of some of the many determinations of strange magnetic moment. Results in red are from global analysis of world data, results in green are from indirect calculations, and results in blue are from lattice QCD calculations.

The analysis of this work was carried out at NERSC using quark propagators generated and stored there.

This work [7] is published in Physical Review Letters as the editor's suggestion of the issue. It is reported in Nature by Ross Young – Strangeness in the proton (https://www.nature.com/articles/nature21909).

References

[1] X. Ji, J. H. Zhang and Y. Zhao, Phys. Rev. Lett. **111**, 112002 (2013) [arXiv:1304.6708 [hep-ph]].

[2] Y. B. Yang, R. S. Sufian, A. Alexandru, T. Draper, M. J. Glatzmaier, K. F. Liu and Y. Zhao, Phys. Rev. Lett. **118**, no. 10, 102001 (2017) doi:10.1103/PhysRevLett.118.102001 [arXiv:1609.05937 [hep-ph]].

[3] D. de Florian, R. Sassot, M. Stratmann and W. Vogelsang, Phys. Rev. Lett. **113**, no. 1, 012001 (2014) doi:10.1103/PhysRevLett.113.012001 [arXiv:1404.4293 [hep-ph]].

[4] D. S. Armstrong, R. D. McKeown, "Parity-Violating Electron Scattering and the Electric and Magnetic Strange Form Factors of the Nucleon," Ann. Rev. Nucl. Part. Sci. **62** (2012) 337-359.

[5] Jianglai Liu, Robert D. McKeown, M. Musolf, "Global Analysis of Nucleon Strange Form Factors at Low Q^2," Phys. Rev. **C76** (2007) 025202.

[6] R. Gonzlez-Jimnez, J. Caballero, and T.W. Donnelly, "Global analysis of parity-violating asymmetry data for elastic electron scattering," Phys. Rev. **D90** (2014) 033002, [arXiv:1403.5119 [nucl-th].

[7] R. S. Sufian, Y. B. Yang, A. Alexandru, T. Draper, J. Liang and K. F. Liu, Phys. Rev. Lett. **118**, no. 4, 042001 (2017) doi:10.1103/PhysRevLett.118.042001 [arXiv:1606.07075 [hep-ph]].