

Sea Contributions to Hadron Electric Polarizabilities through Reweighting

Walter Freeman*

The George Washington University

E-mail: wfreeman@gwu.edu

Andrei Alexandru

The George Washington University

E-mail: aalexan@gwu.edu

Frank X. Lee

The George Washington University

E-mail: fxlee@gwu.edu

Michael Lujan

The George Washington University

E-mail: mlujan@gwu.edu

As part of our ongoing lattice study of the electric polarizabilities of hadrons using the background field approach, we use reweighting to examine the effect of the field on the sea quarks. As with other reweighting studies, the chief difficulty lies in the construction of a stochastic estimate of the ratio of the fermion determinants. In contrast to the case of reweighting in the quark mass, these estimators converge extremely slowly, and are resistant to common variance-reduction techniques such as low-mode subtraction. However, it is possible to construct an alternate estimator, taking advantage of the fact that we are interested in only perturbatively small fields; this estimator is susceptible to a variance-reduction technique based on a hopping parameter expansion.

*The 30 International Symposium on Lattice Field Theory - Lattice 2012,
June 24-29, 2012
Cairns, Australia*

*Speaker.

1. Introduction

Recently there has been a great deal of interest in first-principles lattice computations of hadron polarizabilities. The most interesting such quantity is the electric polarizability of the neutron, since it is difficult to access experimentally due to the lack of free neutrons; the best measurements have been obtained by neutron-lead [1] and neutron-deuteron [2] scattering. A basic computation of this quantity is not too hard to do using the background-field method [3]; the difficulties come in the approach to the physical point that can be compared with experiment. Effective field theories predict very strong dependence on m_q near the chiral limit, making the neutron polarizability a good probe of whether or not chiral behavior is accurately captured by lattice simulations; doing so requires lighter quark masses and potentially expensive chiral actions [4]. Similarly, we must address finite-volume effects [5] and finally extrapolate to the continuum.

The most difficult effect to accommodate, however, is the effect of the electric field on the sea quarks. In principle, this could be done from the ground up by including the background field (see Sec. 3) in gauge generation itself. However, these ensembles will necessarily be uncorrelated. To determine the polarizability we examine the mass shift in the neutron (or other hadron) when a small background electric field is applied; since the zero-field and finite-field propagators are strongly correlated, the error on this mass shift can be much smaller than the error on the masses themselves. Comparing two uncorrelated ensembles destroys this correlation. What we would like is to generate two (or more) correlated ensembles with different values of E . This can be achieved *via* reweighting.

The most difficult aspect of a reweighting calculation is stochastic estimation of the weight factors; the present work is chiefly concerned with this problem. For reweighting in the quark masses, a straightforward stochastic estimator coupled with several standard variance reduction techniques is generally successful in obtaining good estimates with a reasonable amount of computer power. These techniques, however, are not useful for reweighting in the background field. We will discuss possible reasons for those failures, which will likely be illuminating for other groups performing reweighting computations. We then present an alternative approach in which we construct a stochastic estimator for the derivatives of the weight factor with respect to the background field which is susceptible to an improvement technique based on a hopping parameter expansion.

The only prior work known to us on sea contributions to polarizability was done with a purely perturbative method [6]. Instead of applying a uniform background field of specified strength throughout the lattice and then computing hadron propagators, the author first expands the path integral in powers of E and then computes the needed diagrams on the lattice using current insertions. In this method the disconnected and connected insertions (corresponding to sea and valence contributions) appear quite naturally. The author applies this method in a mixed-action formulation, computing domain-wall valence propagators on Asqtad dynamical configurations generated by the MILC collaboration [7] and using stochastic estimators to compute the disconnected diagrams.

2. Lattice simulation details

We apply the following methods to a series of gauge ensembles with two flavors of nHYP-smearred clover fermions [8] with $m_\pi \simeq 330$ MeV and a standard Symanzik-improved gauge action

with $\beta = 7.1$, giving $a = 0.1255$ fm (determined by the Sommer scale r_0) [9]. The ensembles have volumes $24^3 \times 48$, $32 \times 24^2 \times 48$, and $48 \times 24^2 \times 48$; these elongated lattices were originally generated for a scattering study [10], but we reuse them here, since the elongation in the x_1 direction is a convenient probe for finite-volume effects associated with Dirichlet boundary conditions. Each ensemble has 300 minimally-autocorrelated configurations.

3. The background field method

The most commonly-used approach for computing hadron polarizabilities is the background field method, allowing the extraction of the polarizability from spectroscopic measurements. The effect of a uniform background field on the mass of a hadron can be parametrized as:

$$M_H = -\vec{\rho} \cdot \vec{E} - \vec{\mu} \cdot \vec{B} - \frac{1}{2}\alpha E^2 - \frac{1}{2}\beta B^2 + \mathcal{O}(E^4) + \dots \quad (3.1)$$

where μ is the magnetic moment, ρ the electric dipole moment, α is the electric polarizability, β is the magnetic polarizability, and the ellipsis includes various higher-order terms as well as spin polarizabilities [11, 12]. While we are mostly interested in the neutron electric polarizability, the reweighting approach we will use to probe the effects of the sea is not specific to any particular type of hadron. The basic approach, then, to extract the polarizability of some hadron is to measure its mass both in the presence and the absence of a perturbatively small electric field (chosen small enough that higher-order effects are small) and examine the mass shift.

To apply a background electromagnetic field to the lattice, one can simply apply a U(1) phase on the gauge links on top of the dynamical SU(3) gauge configurations, making the transformation

$$U_\mu \rightarrow e^{iqaA_\mu} U_\mu. \quad (3.2)$$

We choose here to apply a constant electric field in the x_1 direction. This can be done in any suitable gauge; we choose $A_4 = iEx_1$. The factor of i arises from the Wick rotation to Euclidean time; thus, on the lattice, a real electric field gives a real factor $e^{\eta x_1}$, while an imaginary field gives pure phase factor. Provided that the magnitude of the field is small enough, it should be possible to perform a lattice calculation for an imaginary field and analytically continue the results to the real axis; preliminary studies have confirmed that both methods result in consistent results for the polarizability [12]. It is convenient to use an imaginary field because the links remain unitary and the Dirac operator remains γ_5 -Hermitian. Thus, we apply the electric field by the transformation $U_4 \rightarrow e^{-i\eta x_1}$ where $\eta \equiv a^2 qE$. Note that this value depends on the charge of the quark flavor in question, so to compute the neutron correlator we will need gauge links for two values of η .

While in the limit of infinite statistics there is no order- η shift in the hadron mass, one may appear from statistical fluctuations in a real calculation. To eliminate this source of error, we compute hadron correlators for fields $+\eta \hat{x}_1$ and $-\eta \hat{x}_1$ (which should, in the infinite-statistics ensemble average, be identical) and take their geometric mean to get the two-point function in the presence of the electric field $G(t, \eta) = \sqrt{G_+(t, \eta)G_-(t, \eta)}$. To extract the polarizability, we then fit it along with the zero-field correlator $G(t, 0)$ to the form

$$G(t, \eta) = Ae^{-(M_N + \delta\eta^2)t} \quad (3.3)$$

to get the mass shift δ . The two correlators must be fit jointly since they are drawn from the same gauge configurations and their fluctuations are (strongly) correlated; this correlation greatly reduces the statistical error in δ .

3.1 Boundary conditions

Periodic boundary conditions are preferable for many lattice observables. However, they potentially create substantial problems for measurements using background fields. For the gauge we have chosen and for real electric fields, there will be an unavoidable and large discontinuity in the electric field at the lattice boundary; for imaginary fields, this discontinuity can be avoided only by choosing particular values of E [13]. However, such fields are out of the perturbative regime that we want to examine. For other choices of the gauge, other problems manifest themselves, such as an electric scalar potential that is not single-valued, leading to manifestly nonphysical scenarios involving quark lines winding around the torus in the direction of the electric field. It is not clear how one should address these effects. We can avoid them by applying Dirichlet boundary conditions in the x_1 and x_4 directions. Dirichlet boundary conditions create their own problems, of course; we can now no longer use certain improved nucleon sources, such as Coulomb-wall, and can no longer project completely onto a zero-momentum state. However, these effects can be considered to be finite-size effects which go away in the infinite volume limit, and are manageable in the analysis; we thus use Dirichlet boundary conditions [12].

3.2 Choosing a field strength

We must choose a value of the parameter η to use for the background field. (Since there are two quark flavors, two η 's are required.) Choosing a value which is too large means that we leave the perturbative regime and begin to probe $\mathcal{O}(E^4)$ effects; choosing a value which is too small means that we may encounter issues with numerical precision, either with the accuracy of inverters or (in the extreme case) machine precision. Fig. 1 shows the response of the neutral pion correlator to the background field at different temporal separations as a function of the d quark η for an inverter precision of 10^{-13} ; the onset of nonperturbative behavior is clear. We note that when a similar study was done with an inverter precision of 10^{-9} , the expected quadratic scaling behavior broke down at the smallest η 's. This illustrates the need to choose an η small enough to avoid higher-order effects when probing polarizabilities, and the need to use a sufficient inverter accuracy to avoid numerical artifacts. Within the large flat region in Fig. 1, we are free to choose whatever value of η makes the reweighting process perform best.

4. Reweighting

Reweighting is a technique for extracting physics based on a different action than the one used in Monte Carlo ensemble generation; essentially, it allows for *post hoc* modification of the parameters in the action. In the standard quantum Monte Carlo, we would like to do a path integral of the form

$$\langle \mathcal{O} \rangle = \frac{\int [dU] \mathcal{O} e^{-S_0}}{\int [dU] e^{-S_0}} \quad (4.1)$$

where S_0 is the QCD action.

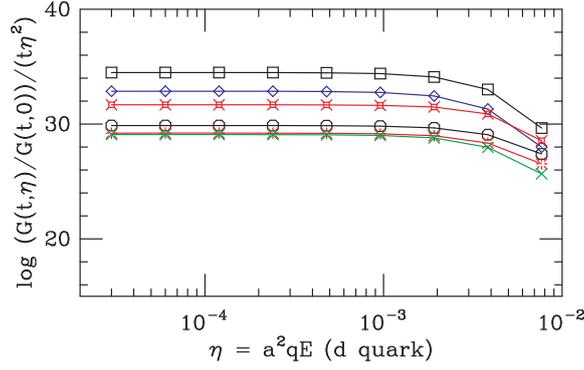


Figure 1: Dependence of $(\log \frac{G(t,\eta)}{G(t,0)})/(t\eta^2)$ for the neutral pion on a $24^3 \times 48$ lattice, as a function of η for different correlator times. This quantity is roughly equivalent to the shift in the effective mass divided by η^2 , and should be constant in the range where η creates a purely quadratic effect. The breakdown of the quadratic behavior is evident at large η .

By generating a Monte Carlo ensemble with each configuration weighted by e^{-S_0} , the path integral reduces to the familiar

$$\langle \mathcal{O} \rangle = \frac{\sum \mathcal{O}}{\sum 1} = \frac{\sum \mathcal{O}}{N}. \quad (4.2)$$

Suppose, however, that we wish to use this ensemble (weighted by e^{-S_0}) to learn about the behavior in the presence of the background field, given by the action S_η . Then the Monte Carlo average gives

$$\langle \mathcal{O}(S_\eta) \rangle = \frac{\sum \mathcal{O} e^{-(S_\eta - S_0)}}{\sum e^{-(S_\eta - S_0)}} \quad (4.3)$$

where $e^{-(S_\eta - S_0)}$ is a “weight factor” which may be interpreted as the relative prominence of a particular configuration in the target and source distribution. If there is little overlap between the source and target ensembles, reweighting will fail as the weight factors fluctuate wildly (over many orders of magnitude); in general, reweighting always comes with a decrease in statistical power, as described in [14]. This may not always be immediately evident and it is possible to wind up underestimating statistical errors [15]. Reweighting is used to reweight in the quark mass to approach the chiral limit without incurring the expense of simulating at those light quark masses directly [15], to perform simulations at small but nonzero chemical potential, to couple sea quarks to dynamical photon fields to probe isospin breaking [16], and (most similarly to this project) been used to investigate the intrinsic strangeness of the nucleon [17]. In this last case, the authors sought to evaluate $\frac{\partial M_N}{\partial m_s}$ by measuring M_N at slightly different values of m_s and examining the difference. This is difficult if the errors in M_N are uncorrelated, but by reweighting in the strange quark mass to generate correlated ensembles they were able to measure errors on the difference in M_N more accurately than M_N itself. We intend to do a very similar thing, except with η (which can be thought of as either reweighting in the background field or the quark electric charge) instead of in the strange quark mass.

4.1 Estimating the weight factor

The weight factor w_i is given by $e^{-(S_\eta - S_0)_i} = \det \frac{M_\eta}{M_0} = \det^{-1} M_0 M_\eta^{-1}$, where M_0 and M_η are the fermion matrices corresponding to the actions S_0 and S_η . This determinant is impractical to

compute exactly and must be estimated stochastically. The standard stochastic estimator for the inverse determinant of some matrix Ω is [15]

$$(\det \Omega)^{-1} = \left\langle e^{-\xi^\dagger (\Omega - 1) \xi} \right\rangle. \quad (4.4)$$

While the inverse determinant of $M_0 M_\eta^{-1}$ is real, using this operator with the above estimator will produce complex results. We thus instead apply the stochastic estimator to an operator with the same determinant but which is Hermitian: $\Omega = \sqrt{M_\eta^{\dagger -1} M_0^\dagger M_0 M_\eta^{-1}}$. The square root can be taken by a rational function approximation as done in rational hybrid Monte Carlo, etc. This is somewhat computationally expensive, as it requires inverting a matrix which itself requires the computation of inversions; however, as that matrix is in principle perturbatively close to 1 its inversion should not require too many iterations. As an alternative, the authors of [18] have argued that in such cases one should simply discard the imaginary part of the estimator, saving all the work associated with the rational function approximation. In the analysis that follows, we estimate the inverse determinant of $\Omega = \sqrt{M_\eta^{\dagger -1} M_0^\dagger M_0 M_\eta^{-1}}$ where $w_i = \det^{-1} \Omega$.

While this stochastic estimator may be (and generally is) quite noisy, this noise does not introduce bias because the average over the noise vectors ξ commutes with the gauge average [15]; in principle, if the ensemble is large enough, one need only use a single stochastic estimate per configuration, although with a gauge ensemble of finite size it is often profitable to work harder than this to improve the stochastic estimator to reduce its fluctuations. The simplest way to do this is to average multiple stochastic estimates, but there are other techniques that can result in a greater reduction in the stochastic noise for a given budget of computational power. We note that the ‘‘signal’’ we are trying to extract from these estimates is the true fluctuation of the weight factor from one configuration to the next, and thus we may adopt the rough criterion for reduction of the stochastic noise that $\sigma_{\text{gauge}} \gtrsim \sigma_{\text{noise}}$. This condition is not necessary, but it should be sufficient, for a successful application of reweighting. (Note, however, that if the true values of the weight factor are indeed very similar, it is possible for this test to give an overly-pessimistic description of the quality of the stochastic estimates.)

5. Improving the stochastic estimate

There is a fundamental problem, however: this estimator is tremendously noisy in our case. As one might expect, the distribution of stochastic estimates $e^{-\xi^\dagger (\Omega - 1) \xi}$ is log-normal; when the width is large enough that the skew of the distribution is apparent, sampling the long tail becomes very difficult. For small values of η , the average estimate of the weight factor is very close to unity, while for larger values the long-tail sampling problem becomes tremendously difficult; see Fig. 2.¹ We have confirmed that the estimator does produce the correct result for the determinant on a 4^4 lattice where that determinant can be computed exactly [19], but it required 10^5 - 10^6 noise vectors, something clearly unaffordable for production-size lattices!

¹While it is of course possible that the true value of the weight factor on the configuration shown here is close to unity, tests on other configurations give the same result.

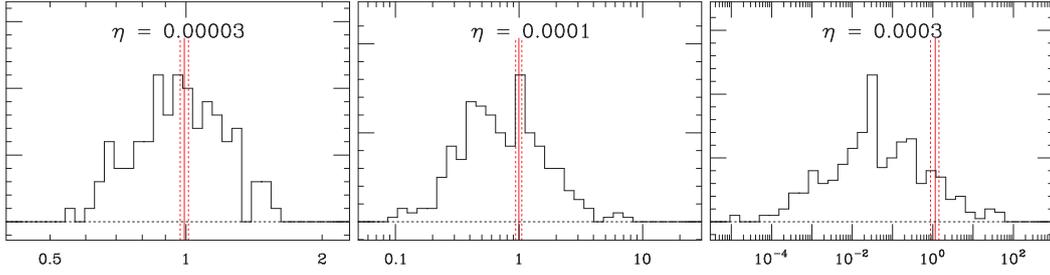


Figure 2: Histograms of the stochastic estimator of the weight factor $e^{-\xi^\dagger(\Omega-1)\xi}$ for three different values of η , along with the mean and its standard error, on a single $24^3 \times 48$ configuration. Even with a large number of noises, the average is indistinguishable from unity; for large values of η , the average is dominated by the few points in the right tail of the distribution, which is roughly log-normal.

5.1 Low-mode subtraction

A commonly-used technique in reweighting in the quark mass is low-mode subtraction. By projecting out the low modes of the operator in question and computing their inverse determinant exactly, we stand to improve the stochastic estimator in two ways. First, we accelerate the inversions required to make the estimates by reducing the conditioning number of the operator being inverted, allowing for the computation of a larger number of estimates with the same amount of computer time. Second, if an appreciable fraction of the fluctuations of the estimator come from the low modes, we may eliminate those fluctuations by treating the low sector exactly (by simply multiplying together the eigenvalues). This is offset, of course, by the additional cost of computing the eigensystem. Can this technique be profitably applied to reweighting in the sea quark charges?

Computation of the eigensystem is more difficult in our case. In the mass-reweighting case, the eigenvectors for all values of κ are the same, so the eigenvectors of Ω are the same as the eigenvectors of the Dirac operator. On the other hand, we must do a separate calculation to compute the eigensystem of $\Omega = \sqrt{M_\eta^\dagger M_0^\dagger M_0 M_\eta^{-1}}$, which is quite computationally expensive. Furthermore, the eigenvectors differ for different values of η , causing problems with the determinant breakup method (see Sec. 5.2).

Since our matrix Ω is in principle close to the identity, we computed eigensystems for both high and low sectors to test the method, generated an ensemble of a few hundred noise vectors, and used those to estimate $\det^{-1} \Omega$ varying the size of the extremal sectors treated exactly. Specifically, if P is a projector onto the space spanned by N extremal modes of Ω with eigenvalues λ_i , and $\bar{P} \equiv 1 - P$, we may write

$$(\det \Omega)^{-1} = \left(\prod_{i=1}^N \frac{1}{\lambda_i} \right) \left\langle e^{-\xi^\dagger \bar{P} (\Omega - 1) \bar{P} \xi} \right\rangle. \quad (5.1)$$

To our surprise, there is essentially no effect whatsoever on the size of the stochastic fluctuations, as shown in Fig. 3. This is strikingly different from the behavior for mass reweighting [15].

In mass reweighting, most of the signal comes from the low modes, as shown empirically in [17]; in fact, almost the entirety of their signal comes from the low modes, and while the high modes are included for correctness their contribution seems to be only (a small amount of) noise. The authors of [20] show this analytically, and furthermore argue that in the case of mass reweight-

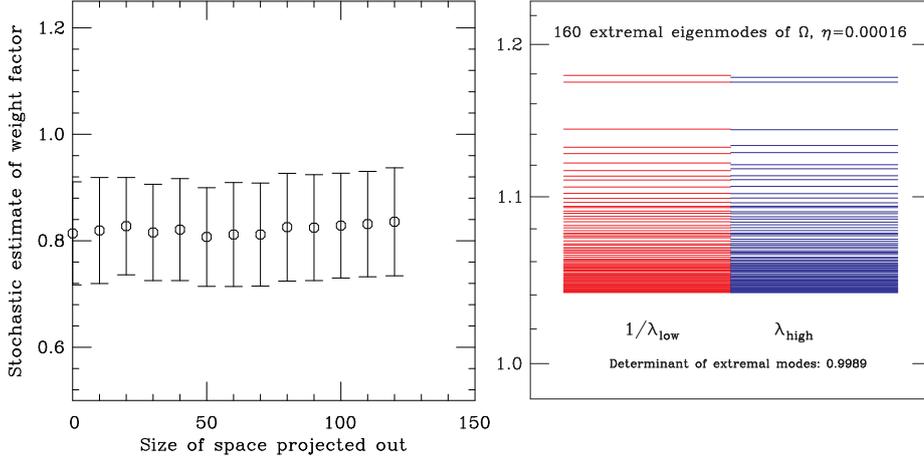


Figure 3: Left: Dependence of the stochastic estimator of the weight factor on the number of extremal modes projected out, showing no meaningful decrease in the variance, on a representative configuration and value of η . The same noise vectors were used for each point. Right: Highest and lowest eigenvalues of $\Omega \equiv \sqrt{M_\eta^\dagger^{-1} M_0^\dagger M_0 M_\eta^{-1}}$ for $\eta = 0.00018$, showing pairing between high and low modes.

ing most of the noise comes from the low modes as well; the estimator is protected from large fluctuations from the high modes. Thus, by removing the low modes from the stochastic estimator, it is possible to also remove most of the noise. This is only guaranteed in the case of reweighting in the quark mass, however; in our case, it fails. Since changing the quark mass does not change the eigenvectors, the low modes of the Dirac operator itself are also eigenvectors of the matrix Ω for quark mass reweighting; this is not true for us, and the extremal modes of Ω may not be related to the low modes of the Dirac operator.

For reweighting in the sea quark charge, the extremal modes of Ω do not contribute much to the determinant. This is because they are very nearly paired; for each high mode with eigenvalue λ , there is a low mode with an eigenvalue close to $1/\lambda$, nearly cancelling their contribution to the determinant; a representative case is shown in Fig. 3. Studies on 4^4 and 6^4 lattices confirm that most of the signal (the difference of the determinant from unity) in our case comes from the bulk “interior” modes of Ω .

5.2 Determinant breakup

Another commonly-used technique for improving the stochastic estimator is to divide the reweighting up into many steps and compute an independent estimate of each one; this technique is often referred to as “determinant breakup” [15] or “determinant factorization” [14, 21]. For reweighting in the quark charge, for instance, one might imagine estimating the weight factor to reweight from $\eta = 0$ to $\eta = \eta_1$, then the weight factor from $\eta = \eta_1$ to $\eta = \eta_2$, and so forth until the desired value is reached. However, for reweighting in η , this technique also fails to decrease the stochastic noise in the estimator compared to simply using more noise vectors. The authors of [14] find that for reweighting in a substantial shift in m_l , it is more efficient to use more subintervals in the determinant breakup (which they call “steps”) than repetitions of the entire procedure (“hits”) for a given total number of inversions. However, this improvement is limited, as shown in [18], and, as predicted in [14], there is no more to be gained after N_{steps} increases past a certain point. We

suspect that this behavior occurs because determinant breakup obtains its benefits by converting a large reweighting step into a sequence of small ones; once the steps are already sufficiently small, no further benefit is gained by splitting them up further compared to simply increasing N_{hits} . There is no benefit to applying this technique to reweighting in η , however, and this explains why; we are already reweighting by a perturbatively small interval in order to make a valid measurement of the polarizability.

6. The perturbative estimator

Since it is not feasible to perform a direct stochastic estimation of the determinant, and since neither of the most common improvement techniques that are successful for mass reweighting work for reweighting in the quark charges, we turn to an alternative estimator. We are interested only in the weight factors for perturbatively small values of η , so we can expand about $\eta = 0$:

$$w(\eta) = 1 + \eta \left. \frac{\partial w}{\partial \eta} \right|_{\eta=0} + \frac{1}{2} \eta^2 \left. \frac{\partial^2 w}{\partial \eta^2} \right|_{\eta=0} + \dots \quad (6.1)$$

The expansion must be taken to second order in η , since we are interested in effects of order E^2 . The task then becomes to construct estimators for $\left. \frac{\partial w}{\partial \eta} \right|_{\eta=0}$ and $\left. \frac{\partial^2 w}{\partial \eta^2} \right|_{\eta=0}$. Both terms are needed, since the linear term in the weight factor can couple with a linear dependence on η to give a quadratic effect, or the quadratic term can give a second-order effect on its own. Given estimates of these derivatives, we can evaluate the above at any sufficiently-small η to produce a reweighted ensemble on which to apply the valence calculation.

To construct such an estimator for $\left. \frac{\partial w}{\partial \eta} \right|_{\eta=0} = \left. \frac{\partial}{\partial \eta} \frac{\det M_\eta}{\det M_0} \right|_{\eta=0}$, we rewrite $\det M_\eta$ as a Grassmann integral:

$$\begin{aligned} \frac{\partial}{\partial \eta} \det M &= \frac{\partial}{\partial \eta} \int d\psi d\bar{\psi} e^{-\bar{\psi} M \psi} \\ &= \int d\psi d\bar{\psi} \left(-\bar{\psi} \frac{\partial M}{\partial \eta} \psi \right) e^{-\bar{\psi} M \psi} \\ &= \det M \text{Tr} \left(\frac{\partial M}{\partial \eta} M^{-1} \right). \end{aligned} \quad (6.2)$$

This is an expression for $\frac{\partial \det M}{\partial \eta}$, but we want $\left. \frac{1}{\det M_0} \frac{\partial}{\partial \eta} \det M \right|_{\eta=0}$. When the above is evaluated at $\eta = 0$, the determinant in front cancels the $\frac{1}{\det M_0}$ and we get

$$\left. \frac{\partial}{\partial \eta} \frac{\det M_\eta}{\det M_0} \right|_{\eta=0} = \text{Tr} \left(\left. \frac{\partial M}{\partial \eta} \right|_{\eta=0} M_0^{-1} \right) \quad (6.3)$$

The trace involved here must still be evaluated stochastically.

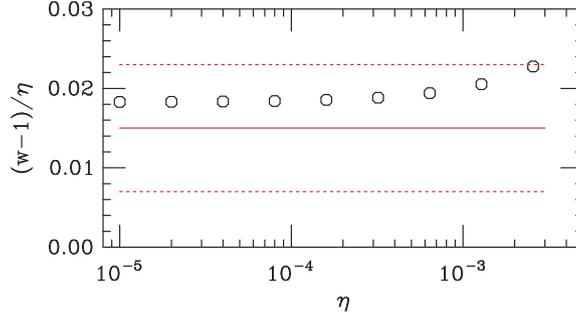


Figure 4: Exact values for $\frac{\det M_\eta / \det M_0 - 1}{\eta}$ on a 4^4 lattice, compared with the value predicted by the stochastic estimator for $\frac{\partial \det M_\eta / \det M_0}{\partial \eta}$ and its error band.

The second derivative term proceeds similarly. Computing $\frac{\partial^2}{\partial \eta^2} \det M$, we get

$$\begin{aligned} \frac{\partial^2}{\partial \eta^2} \det M &= \frac{\partial^2}{\partial \eta^2} \int d\psi d\bar{\psi} e^{-\bar{\psi} M \psi} \\ &= \frac{\partial}{\partial \eta} \int d\psi d\bar{\psi} \left(-\bar{\psi} \frac{\partial M}{\partial \eta} \psi \right) e^{-\bar{\psi} M \psi} \\ &= \int d\psi d\bar{\psi} \left(-\bar{\psi} \frac{\partial^2 M}{\partial \eta^2} \psi \right) e^{-\bar{\psi} M \psi} \end{aligned} \quad (6.4)$$

$$+ \int d\psi d\bar{\psi} \left(\bar{\psi} \frac{\partial M}{\partial \eta} \psi \right)^2 e^{-\bar{\psi} M \psi} \quad (6.5)$$

Evaluating these integrals gives:

$$\frac{\partial^2}{\partial \eta^2} \det M = \det M \left[\text{Tr} \left(\frac{\partial^2 M}{\partial \eta^2} M^{-1} \right) - \left(\text{Tr} \frac{\partial M}{\partial \eta} M^{-1} \right)^2 + \text{Tr} \left(\frac{\partial M}{\partial \eta} M^{-1} \right)^2 \right]. \quad (6.6)$$

As above, this gives

$$\frac{\partial^2}{\partial \eta^2} \frac{\det M}{\det M_0} \Big|_{\eta=0} = \text{Tr}(B M_0^{-1}) + \text{Tr}(A M_0^{-1} A M_0^{-1}) - (\text{Tr}(A M_0^{-1}))^2 \quad (6.7)$$

where $A \equiv \frac{\partial M}{\partial \eta} \Big|_{\eta=0}$ and $B \equiv \frac{\partial^2 M}{\partial \eta^2} \Big|_{\eta=0}$.

Now we must construct stochastic estimators for these three traces, as well as the trace in the first derivative. We use the standard stochastic estimator $\text{Tr} \mathcal{O} = \langle \xi^\dagger \mathcal{O} \xi \rangle$; note that two stochastic estimates of the first derivative term can be used to construct an estimate of $(\text{Tr}(A M^{-1}))^2$. Tests on a 4^4 lattice confirm that this estimator for the first derivative is correct; see Fig. 4. However, it is just as noisy as the direct estimator of the determinant; the data in Fig. 4 required 5×10^6 noise vectors to produce! Clearly this new estimator on its own has not gained anything.

6.1 Hopping parameter expansion improvement

However, this estimator for the trace is susceptible to an improvement technique. If other operators \mathcal{O}'_i can be identified such that the stochastic fluctuations in $\xi^\dagger \mathcal{O} \xi$ and $\xi^\dagger \mathcal{O}'_i \xi$ are correlated,

then we can reduce the overall fluctuations by writing

$$\text{Tr } \mathcal{O} = \left\langle \xi^\dagger \left(\mathcal{O} - \sum_i \mathcal{O}'_i \right) \xi \right\rangle + \sum_i \text{Tr } \mathcal{O}'_i. \quad (6.8)$$

Obviously, for this to be useful, the \mathcal{O}'_i 's must themselves have exactly-computable traces.

For the operators needed here, a set of improvement operators can be gotten by performing a hopping parameter expansion of M^{-1} , with each term in the expansion acting as one \mathcal{O}'_i ; this guarantees that the fluctuations in the estimator will be correlated. Specifically, the fluctuations are correlated with the magnitude of the off-diagonal elements of $\frac{\partial M}{\partial \eta} M^{-1}$; the terms in the hopping parameter expansion approximate the largest of these close to the diagonal [22]. For the Wilson-clover fermions considered here, we may write M as $1 - \kappa(\mathcal{D} + C)$, where the hopping term \mathcal{D} is given by

$$\mathcal{D}_{mn} = \sum_\mu [(1 - \gamma_\mu) U_\mu(m) \delta_{m+\hat{\mu},n} + (1 + \gamma_\mu) U_\mu^\dagger(n) \delta_{m,n+\hat{\mu}}] \quad (6.9)$$

and the clover term C is given by

$$C_{mn} = c_{SW} \frac{1}{8} \sum_{\mu\nu} \sigma_{\mu\nu} L_{\mu\nu}(n), \quad (6.10)$$

where $L_{\mu\nu}(n)$ is a sum of the imaginary part of the plaquettes in the $\mu\nu$ plane that include site n .

For the first derivative term, the technique of 6.8 gives

$$\begin{aligned} \text{Tr } \frac{\partial M}{\partial \eta} M^{-1} = & \left\langle \xi^\dagger \left(\frac{\partial M}{\partial \eta} M^{-1} - \frac{\partial M}{\partial \eta} - \kappa \frac{\partial M}{\partial \eta} (\mathcal{D} + C) - \kappa^3 \frac{\partial M}{\partial \eta} (\mathcal{D} + C)^2 - \kappa^2 \frac{\partial M}{\partial \eta} (\mathcal{D} + C)^3 - \dots \right) \xi \right\rangle \\ & + \text{Tr } \frac{\partial M}{\partial \eta} + \text{Tr } \kappa \frac{\partial M}{\partial \eta} (\mathcal{D} + C) + \text{Tr } \kappa^2 \frac{\partial M}{\partial \eta} (\mathcal{D} + C)^2 + \text{Tr } \kappa^3 \frac{\partial M}{\partial \eta} (\mathcal{D} + C)^3 + \dots \end{aligned} \quad (6.11)$$

where $\mathcal{O}'_i = \frac{\partial M}{\partial \eta} M^{-1} \kappa^i (\mathcal{D} + C)^i$. The stochastic component of this can be computed directly as before. The difficult part is to compute $\text{Tr } \kappa^n \frac{\partial M}{\partial \eta} (\mathcal{D} + C)^n$ quickly. In the end we wish to apply this technique to nHYP-smearred Wilson-clover fermions, but it is useful to consider the unimproved Wilson action first, for which $C = 0$.

The trace of $\frac{\partial M}{\partial \eta} \mathcal{D}^n$ can be evaluated most simply by writing \mathcal{D} as the sum of eight hopping terms. $\frac{\partial M}{\partial \eta}$ can be calculated analytically, and consists itself of two hopping terms in the $\pm \hat{x}_4$ direction; the rest are zero. We break $\frac{\partial M}{\partial \eta}$ and \mathcal{D}^n into separate hopping terms and expand the product. All terms which give a nonzero contribution to the trace are products of hops which form closed paths. However, not all such paths contribute; some closed paths (such as those that double back on themselves) have zero Dirac trace and do not contribute. The most efficient way to compute the trace is thus to expand $\frac{\partial M}{\partial \eta} \mathcal{D}^n$ as a sum of products of hopping terms which have definite spatial and Dirac structure, each corresponding to the product of gauge links along a path of a particular shape, and pick out only those paths which are closed. We then compute the Dirac trace of each term. The computationally-intensive step, summing the appropriate products of gauge links over the spatial volume, must only be done for terms with nonzero Dirac trace.

The clover term can be dealt with by considering a generalization of the procedure used for the pure Wilson action. Another way to describe that procedure is as a separation of the \mathcal{D} operator into eight pieces (hops in $\pm x, y, z, t$), each with a definite Dirac and spatial structure. The need to separate out terms with a definite spatial structure is readily apparent since in taking the trace we are only concerned with closed paths. In doing so we got a definite Dirac structure that does not depend on the gauge links and can be factored out of the sum over sites along the way. The same principle applies to the clover term even though the spatial structure is trivial: separate C into six pieces each with definite Dirac structure. These are the individual terms $\sigma_{\mu\nu}L_{\mu\nu}$ in Eq. 6.10, allowing us to factorize the trace into Dirac and SU(3) pieces.

We must also compute $\frac{\partial L_{\mu\nu}}{\partial \eta}$, as it appears in $\frac{\partial M}{\partial \eta}$. This can be done analytically for “plain” clover fermions, but for nHYP-smearred clover fermions, the smearing process complicates computation of $\frac{\partial \mathcal{D}}{\partial \eta}$ and $\frac{\partial C}{\partial \eta}$. While this can in principle still be done analytically, it is simpler and not expensive to do it numerically. Our procedure for evaluating $\text{Tr} \kappa^n \frac{\partial M}{\partial \eta} (\mathcal{D} + C)^n$ is thus as follows:

1. Write \mathcal{D} and $\frac{\partial \mathcal{D}}{\partial \eta}$ (the latter calculated numerically) as the sum of eight separate hopping terms. Similarly, break C and $\frac{\partial C}{\partial \eta}$ into six terms, each with the Dirac structure of $\sigma_{\mu\nu}$.
2. Expand $\frac{\partial (\mathcal{D} + C)}{\partial \eta} (\mathcal{D} + C)^n$, giving 14^{n+1} terms, each with a definite Dirac and spatial structure.
3. Compute spatial part: The majority of these terms correspond to paths that are not closed, and can be discarded right away.
4. Each term now can be factorized into a SU(3) part (which depends on the gauge links, and must be summed over sites), and a Dirac part (which is the same for each lattice site and can be factored out).
5. Compute Dirac part: the majority of terms will have zero Dirac trace and can be discarded.
6. Compute SU(3) part: For the remaining terms (roughly one in 500, regardless of order, after the first few), do the hard work of computing products of links and $L_{\mu\nu}$'s over all sites.

The second-derivative traces can be evaluated in a very similar way. For the $\text{Tr}(BM^{-1})$ term, we must additionally compute $B \equiv \frac{\partial^2 (\mathcal{D} + C)}{\partial \eta^2}$, but this is not difficult to do numerically.

This expansion can be in principle carried out to arbitrarily high order. The evaluation of the stochastic estimates of additional \mathcal{O}_i^l 's is not difficult, as the cost is dominated by the inversion which only must be performed once for any number of orders. The limiting factor is the evaluation of $\text{Tr} \kappa^i \frac{\partial M}{\partial \eta} \mathcal{D}^i$, as the cost increases exponentially with i . However, it is possible to examine the reduction in stochastic noise without computing the exact traces. We find that the degree to which this improvement procedure reduces the stochastic noise, especially after the first few orders, is strongly dependent on the value of m_π ; this is not surprising in light of the use of a hopping parameter expansion. This behavior is shown for several different values of m_π in Fig. 5. Because the improvement slows down after the first few orders and the cost increases substantially, we compute the exact traces only up to seventh order.

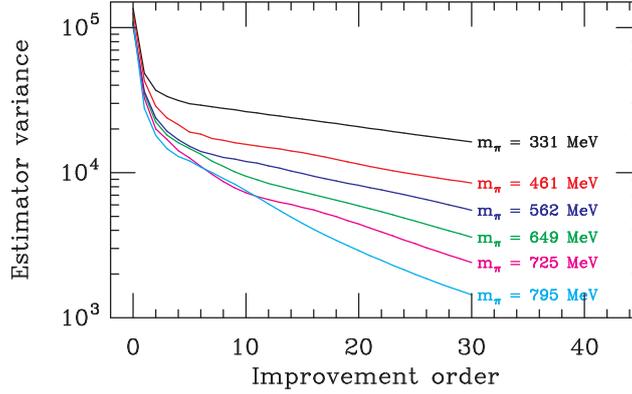


Figure 5: Improvement in the variance of the stochastic estimator of $\frac{\partial w}{\partial \eta}$ as a function of the order of the hopping parameter expansion for different values of m_π .

7. Applicability to the polarizability and future plans

While the improvement technique carried out to seventh order yields a substantial reduction in the variance of the stochastic estimator, this estimator is still quite expensive to run; we aim for around 10^3 stochastic sources per configuration. These estimates are almost complete on the $24^3 \times 48$ ensemble. Once they are complete, we will use the estimates of $\frac{\partial w}{\partial \eta}$ and $\frac{\partial^2 w}{\partial \eta^2}$ to evaluate $w(\eta) = \frac{\det M_\eta}{\det M_0}$ at particular values of η (corresponding to the charges of the up and down sea quarks), chosen small enough that they lie within the perturbative regime of Fig. 1 but large enough to avoid any subsequent numerical precision issues, and then use the resulting weight factors to generate a reweighted ensemble in the presence of the background field. The remainder of the polarizability computation proceeds as in Ref. [3], but using the reweighted ensemble to compute the nonzero-field neutron correlator. It is likely that the influence of the charged sea is greater at lighter quark masses. We have such an ensemble available (with $m_\pi \approx 200$ MeV) and intend to repeat the calculation on it, although the variance reduction from the hopping-parameter expansion will not be as strong due to the lower m_π . Finally, once the calculation detailed above is complete, if the variance in the stochastic estimator of the weight factors leads to a large increase in the overall error, we are considering using deflation to accelerate the thousands of inversions required to further improve the estimator.

8. Acknowledgements

We would like to acknowledge Craig Pelissier for his assistance with the GWU-QCD code, and Jacob Finkenrath and Michael Engelhardt for helpful discussions. We are grateful to GWU IMPACT (Institute for Massively Parallel Applications and Computing Technologies) for CPU time and to the Columbian College for logistical support of our GPU cluster. This work was supported in part by DoE grant DE-FG02-95ER40907 and NSF CAREER grant PHY-1151648.

References

- [1] J. Schmiedmayer, P. Riehs, J. A. Harvey, and N. W. Hill, *Measurement of the electric polarizability of the neutron*, *Phys.Rev.Lett.* **66** (1991) 1015–1018.
- [2] K. Kossert, M. Camen, F. Wissmann, J. Ahrens, J. Annand, *et. al.*, *Quasifree Compton scattering and the polarizabilities of the neutron*, *Eur.Phys.J.* **A16** (2003) 259–273, [[nucl-ex/0210020](#)].
- [3] A. Alexandru and F. X. Lee, *Neutron electric polarizability*, *PoS LAT2009* (2009) 144, [[arXiv:0911.2520](#)].
- [4] M. Lujan, A. Alexandru, and F. Lee, *Electric polarizability of hadrons with overlap fermions on multi-GPUs*, *PoS LATTICE2011* (2011) 165, [[arXiv:1111.6288](#)].
- [5] A. Alexandru and F. Lee, *Hadron electric polarizability – finite volume corrections*, *PoS LATTICE2010* (2010) 131, [[arXiv:1011.6309](#)].
- [6] M. Engelhardt, *Progress toward the chiral regime in lattice QCD calculations of the neutron electric polarizability*, *PoS LAT2009* (2009) 128, [[arXiv:1001.5044](#)].
- [7] A. Bazavov, D. Toussaint, C. Bernard, J. Laiho, C. DeTar, *et. al.*, *Nonperturbative QCD simulations with 2+1 flavors of improved staggered quarks*, *Rev.Mod.Phys.* **82** (2010) 1349–1417, [[arXiv:0903.3598](#)].
- [8] A. Hasenfratz, R. Hoffmann, and S. Schaefer, *Hypercubic smeared links for dynamical fermions*, *JHEP* **0705** (2007) 029, [[hep-lat/0702028](#)].
- [9] R. Sommer, *A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory*, *Nucl.Phys.* **B411** (1994) 839–854, [[hep-lat/9310022](#)].
- [10] C. Pelissier and A. Alexandru, *Resonance parameters of the rho-meson from asymmetrical lattices*, [arXiv:1211.0092](#).
- [11] W. Detmold, B. C. Tiburzi, and A. Walker-Loud, *Extracting Electric Polarizabilities from Lattice QCD*, *Phys.Rev.* **D79** (2009) 094505, [[arXiv:0904.1586](#)].
- [12] A. Alexandru and F. X. Lee, *The Background field method on the lattice*, *PoS LATTICE2008* (2008) 145, [[arXiv:0810.2833](#)].
- [13] H. Rubinstein, S. Solomon, and T. Wittlich, *Dependence of lattice hadron masses on external magnetic fields*, *Nucl.Phys.* **B457** (1995) 577–593, [[hep-lat/9501001](#)].
- [14] Q. Liu, N. H. Christ, and C. Jung, *Light Quark Mass Reweighting*, [arXiv:1206.0080](#).
- [15] A. Hasenfratz, R. Hoffmann, and S. Schaefer, *Reweighting towards the chiral limit*, *Phys.Rev.* **D78** (2008) 014515, [[arXiv:0805.2369](#)].
- [16] T. Ishikawa, T. Blum, M. Hayakawa, T. Izubuchi, C. Jung, *et. al.*, *Full QED+QCD low-energy constants through reweighting*, *Phys.Rev.Lett.* **109** (2012) 072002, [[arXiv:1202.6018](#)].
- [17] H. Ohki, S. Aoki, H. Fukaya, S. Hashimoto, T. Kaneko, *et. al.*, *Nucleon sigma term and strange quark content in 2+1-flavor QCD with dynamical overlap fermions*, *PoS LAT2009* (2009) 124, [[arXiv:0910.3271](#)].
- [18] J. Finkenrath, F. Knechtli, and B. Leder, *Application of Domain Decomposition to the Evaluation of Fermion Determinant Ratios*, [arXiv:1211.1214](#).
- [19] A. Alexandru and U. Wenger, *QCD at non-zero density and canonical partition functions with Wilson fermions*, *Phys.Rev.* **D83** (2011) 034502, [[arXiv:1009.2197](#)].
- [20] M. Luscher and F. Palombi, *Fluctuations and reweighting of the quark determinant on large lattices*, *PoS LATTICE2008* (2008) 049, [[arXiv:0810.0946](#)].
- [21] A. Alexandru and A. Hasenfratz, *Partial global, stochastic metropolis update for dynamical smeared link fermions*, *Phys.Rev.* **D66** (2002) 094502, [[hep-lat/0207014](#)].
- [22] C. Thron, S. Dong, K. Liu, and H. Ying, *Pade - Z(2) estimator of determinants*, *Phys.Rev.* **D57** (1998) 1642–1653, [[hep-lat/9707001](#)].