Sea quark contributions to the electric polarizability of hadrons

Walter Freeman, Andrei Alexandru, Mike Lujan, and Frank X. Lee

Phys. Rev. D **90**, 054507 — Published 23 September 2014

DOI: 10.1103/PhysRevD.90.054507

Sea quark contributions to the electric polarizability of hadrons

Walter Freeman, Andrei Alexandru, Mike Lujan, and Frank X. Lee
The George Washington University

We present a lattice QCD calculation of the polarizability of the neutron and other neutral hadrons that includes the effects of the background field on the sea quarks. This is done by perturbatively reweighting the charges of the sea quarks to couple them to the background field. The main challenge in such a calculation is stochastic estimation of the weight factors, and we discuss the difficulties in this estimation. Here we use an extremely aggressive dilution scheme to reduce the stochastic noise to a manageable level. The pion mass in our calculation is 300 MeV and the lattice size is 3 fm. For the neutron, we find that $\alpha_E = 2.70(55) \times 10^{-4}\,\mathrm{fm}^3$, which is the most precise lattice QCD determination of the polarizability to date that includes sea effects.

PACS numbers: 11.15.Ha,12.38.Gc

I. INTRODUCTION

At leading order, the interaction of hadrons with a background electromagnetic field can be parametrized with a variety of electromagnetic polarizabilities which characterize the deformation of the hadron by the field. Of these, the electric polarizability α describes the induced dipole by an external static, uniform electric field. It is defined as the ratio of the electric field and the induced dipole moment: $\boldsymbol{d} = \alpha \boldsymbol{\mathcal{E}}$. Since this deformation is a direct consequence of the composite nature of the hadrons, it is a necessary component of any overall understanding of hadronic structure.

Measuring the electric polarizabilities of hadrons is challenging. Few hadron polarizabilities have been determined, but there are a number of experiments that plan to measure these quantities for various hadrons in the near future. On the theory side, lattice QCD can be used to determine these parameters as predicted by quark-gluon dynamics. These are challenging calculations, and to establish the methodology it is useful to first focus on electrically neutral hadrons, which are not accelerated by the electric field. Since the hadrons are at rest, it is easier to detect the effect of electric polarizability. In this paper we focus on the neutron but we will also present results for neutral kaon and neutral pion.

Neutron electric polarizability is difficult to measure experimentally due to the unavailability of free neutron targets. It has been measured in the laboratory by scattering neutron beams on lead [1] and off of deuterons [2]; the results respectively were $12.0(1.5)(2.0) \times 10^{-4}\,\mathrm{fm}^3$ and $12.5(1.8)\binom{+1.1}{-0.6} \times 10^{-4}\,\mathrm{fm}^3$.

A lattice calculation of the neutron electric polarizability is desirable for at least three reasons. First, the experimental uncertainties in these quantities are still over 10%, and it may be the case that eventually the lattice may prove superior to experiment in attaining a precision measurement of this quantity. Second, if lattice QCD is to be considered a successful approach to simulating the hadronization of quarks and their properties, then the measurement of such a fundamental property of the neutron is something of a basic test. Finally, the flexibility of lattice calculations (the freedom to use nonphysical parameters) may provide some insight into the origins of the neutron polarizability.

The first lattice study of the neutron polarizability was done in 1989 [3], on a $10^3 \times 20$ quenched lattice with $a \simeq 0.11\,\mathrm{fm}$ using unimproved staggered fermions; this study, along with a subsequent early study using both Wilson and clover fermions on a quenched sea [4], showed good agreement with the experimental value. More recently, improved calculations have produced values that are substantially smaller [5–10], suggesting that the early agreement with experiment was coincidental. The early lattice calculations suffered from a number of problems. The most worrying of these was an overall error in the sign of the polarizability coming from the rotation to Euclidean time. Many of these early calculations also used a linearized phase factor, which fails to capture even all perturbatively-relevant interactions (since the polarizability involves a quadratic effect) [6].

It is well understood that neutron polarizability computed from lattice QCD is smaller than the physical value because the quark mass used is heavier than the physical one. Chiral perturbation theory (χPT) predicts that the polarizability of the neutron diverges in the chiral limit. In fact χPT calculations can be used to predict the value of the polarizability for unphysical quark masses [11–13]. The most precise lattice QCD calculation for the neutron polarizability finds a value that is still significantly different from the χPT predictions [14]. The difference is most likely due to a combination of finite volume effects and a systematic correction due to the electric charge of the sea quarks. In this paper we present a method for removing the latter systematic error and use it to compute correct value of the polarizability on one of the ensembles used in our previous study [14].

Since lattice QCD is best able to measure spectroscopic information about hadronic states, we compute the polarizability through the induced interaction energy $\delta E = -\frac{1}{2}\alpha\mathcal{E}^2$. This is achieved using the *background field method*, where the energy shift is computed by comparing the mass of the hadrons in the presence of a static electric field with the one determined when the field is absent. To include the effects of the charged sea, one could generate

two dynamical ensembles, one with a background field and one without, and measure the mass shift. However, for the valence-only calculation these two masses, measured on the same Monte Carlo ensemble and differing only by the effects of a perturbatively-small background field, are highly correlated. Thus, the error on the mass shift is much less than the error on each mass individually. For instance, in the valence-only calculation on the ensemble used here, the mass shift for the neutron is $a\Delta m = 6.28(57) \times 10^{-7}$, while the neutron mass itself is $am = 0.694(4)$, with an error four orders of magnitude larger than the mass shift. Only the strong correlations between the correlators with and without the background field allow us to extract this mass shift with an error 10^5 times smaller than the error on the masses themselves. Generating two separate ensembles would destroy this correlation and greatly inflate the statistical error. What is needed is a way to obtain ensembles generated with different dynamical properties which are correlated; reweighting provides such a technique.

The plan of the paper is the following: in Section II we will review briefly the steps relevant for the valence calculation. In Section III we discuss the perturbative reweighting strategy we use to couple the sea quarks to the background field. In Section IV the stochastic estimators used to compute the derivatives of the reweighting factors are discussed in detail. The results are presented in Section V.

II. VALENCE CALCULATION

A. Simulation parameters

In this study we will use one of the ensembles from a previous study [14], labeled EN1 in that paper. The configurations in this ensemble were generated using $N_f = 2$ flavors of nHYP-smeared Wilson-clover fermions [15]. The ensemble contains 300 lattice configurations of size $24^3 \times 48$. The lattice spacing of 0.1245(16) fm was determined by a fit to the static quark potential to determine the Sommer scale r_0/a [16] using a value of $r_0 = 0.5$ fm. The sea quarks have $\kappa = 0.1282$, corresponding to $m_\pi = 306(1)$ MeV; we use the same κ value for the valence light quarks as well. The valence strange quark for the kaon correlators has $\kappa_s = 0.1266$. The gauge configuration generation was performed with periodic boundary conditions; Dirichlet boundary conditions have been applied for the valence quarks in the direction of the electric field and the time direction. We use an optimized multi-GPU Dslash operator [17] along with an even-odd preconditioned BiCGstab inverter [18] to do the analysis described here.

While we do not attempt a continuum extrapolation, we expect the discretization errors to be substantially smaller than the statistical errors here. In particular, in [19] the authors conduct a continuum extrapolation for the nucleon mass using HEX-smeared clover fermions, quite similar to the action used here. They find a 3% upper bound on the neutron mass shift from $a = 0.125$ fm to the continuum. We may expect the lattice spacing dependence of the polarizability to be similar, and thus do not expect discretization effects to be large in the present calculation.

B. The background field method

Since the ground state energy of the neutron is shifted by an amount $\delta M = -\frac{1}{2}\alpha\mathcal{E}^2$ in an external electric field, spectroscopic measurements on the lattice can provide a direct avenue to access the polarizability. We use the notation δM rather than δE to emphasize that, since we use Dirichlet boundary conditions, we do not measure the actual neutron mass since we have no zero-momentum state. The approach is straightforward: we measure the neutron energy with the background field and without it, then compute δE, which is then converted to δM to compute α.

We introduce the electric field by adding a $U(1)$ phase factor on top of the $SU(3)$ gauge links that corresponds to a uniform background field; this may be done in any convenient choice of gauge. In practice, there are several complications which must be taken into account when applying this method to the lattice. The simplest is the fact that in Euclidean time, applying phase factors of the form $e^{i\theta}$ corresponds to an imaginary electric field; to get a real electric field, one must use an imaginary θ, giving real exponential factors on the links. However, an imaginary electric field presents no real problems; this gives a positive δE as expected, and has little effect on the final result [6].

We also must address the lattice boundary conditions. With periodic boundary conditions, the phase factor corresponding to an arbitrary electric field will have a discontinuity at the lattice edge, giving a non-physical spike in the electric field there. While we can choose values of $\boldsymbol{\mathcal{E}}$ in conjunction with the lattice size and gauge such that the discontinuity is made to vanish, the size of $\boldsymbol{\mathcal{E}}$ required to do this is so large that one is no longer probing only the lowest-order effects proportional to $\mathcal{E}^2$ for which the polarizability is defined. Moreover, even if the discontinuity in the $U(1)$ phase is addressed, the electric scalar potential will not be single-valued, possibly inducing quark lines or charged pions to wind repeatedly around the lattice. It is not clear what the effects of this, or of discontinuities in the $U(1)$ phase, will be.

Thus, we choose to use Dirichlet boundary conditions in time and in the direction of the electric field, which we choose as the $\hat{x}$-direction. While this means that we have no true zero-momentum state, this can be treated as an additional finite-size effect whose effect can be partially compensated for and which will in any case go away in the infinite-volume limit.

We parametrize the electric field with the dimensionless parameter

$$\eta \equiv a^2 q \mathcal{E}\,, \tag{1}$$

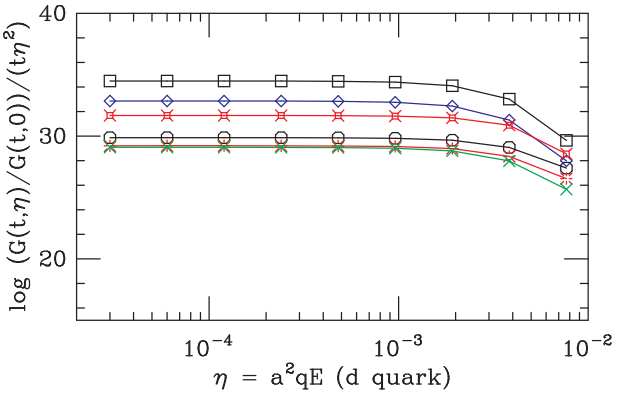

FIG. 1: Dependence of $(\log \frac{G_\pi(t,\eta)}{G_\pi(t,0)})/(t\eta^2)$ for the neutral pion on a $24^3 \times 48$ lattice, as a function of η for different correlator times. This quantity is roughly equivalent to the shift in the effective mass at the chosen correlator time divided by η^2, and should be constant in the range where η creates a purely quadratic effect. As this was only computed on a single configuration there are no error bars; we emphasize that these data are only useful for determining the perturbative regime of η.

noting that η depends both on the quark flavor and $\mathcal{E}$, and choose a gauge for the electric field such that

$$U_4 \to U_4 e^{i\eta x/a}\,. \tag{2}$$

η_d must be chosen small enough that it probes only the lowest-order (quadratic) effects which correspond to the electric polarizability and avoids large $\mathcal{O}(\mathcal{E}^4)$ effects. Since there are two sea quarks, we use a different η for up and down quark propagators, and quote η_d as a measure of the field strength. However, 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, G_π, to the background field as a function of η_d for a few different correlator times on one configuration. The breakdown of quadratic scaling as η becomes large is clear. Note that $G_\pi(t,\eta)$ is symmetrized with respect to η, so that only even powers of η contribute to this correlator. This symmetrization is only valid when the sea quarks are not charged. In this study we used $\eta_d = 10^{-4}$, and the valence correlators on this ensemble were run at this value. Fig. 1 shows that this value is well within the quadratic scaling region, at least for the valence contribution.

C. Extracting the energy shift

To determine the energy shift caused by the external electric field, we compute hadron correlators $G(t,\mathcal{E}_x)$ for positive, zero, and negative values of $\mathcal{E}_x$. This requires the computation of five quark-line propagators: one at $\eta = 0$, used for both up and down quarks in the case of $\mathcal{E}_x = 0$, two at $\eta = \pm\frac{1}{3}a^2\mathcal{E}_x$ for the down quark, and two at $\eta = \mp\frac{2}{3}a^2\mathcal{E}_x$ for the up quark.

The energy shift caused by the external electric field is quite small, smaller than the stochastic error in the hadron energy itself. Thus, in order to resolve it, we must take into account the fact that the correlators measured with and without the electric field are strongly correlated, and only become more strongly correlated as the strength of the electric field is decreased. We cannot simply, then, do independent correlator fits to the three correlators. Just as an ordinary correlator fit must take into account the correlations between $G(t)$ at different t by computing the covariance between them, we must construct a covariance matrix which includes the mixed covariance between zero-field and nonzero-field correlators. This is simply an extension of the standard fitting procedure using the covariance between all pairs of observables.

We then fit all the data at once, using the fit form

$$\langle G(t,\eta)\rangle_\eta = (A + \Delta A)e^{-(E+\Delta E)t}\,, \tag{3}$$

to extract E and the parameter ΔE which is related to α. For details on determining the polarizability from the energy shift, see [14].

For small values of $\mathcal{E}_x$, the covariance matrix is quite poorly conditioned due to the extremely strong correlations. We have observed that in this case both the minimization of χ^2 and the inversion of the covariance matrix must be done in extended precision to get consistent fit results. For $\eta_d = 10^{-4}$, we find that the C `long double` type offers sufficient precision.

III. REWEIGHTING

A. General remarks

As mentioned previously, the simplest way to incorporate the effect of the electric field on the sea quarks would be to include its effects in gauge generation where the sea dynamics are simulated. However, generating a separate Monte Carlo ensemble to compute the correlator in the presence of background field would ruin the correlations which are necessary to achieve a small overall error. Thus, we turn to reweighting as a method of creating two ensembles which have different sea-quark actions yet are correlated. A similar approach has been used before to compute the strangeness of the nucleon using the Feynman-Hellman theorem [20], which requires a measurement of $\frac{\partial M_N}{\partial m_s}$.

Reweighting involves a simple modification of Monte Carlo sampling. Normally, the configurations are sampled using a probability proportional to e^{-S}. Then a Monte Carlo estimate for the expectation value of the correlator $G(t)$

$$\langle G(t)\rangle_0 \equiv \frac{\int \mathcal{D}\mathrm{U}\, G(t) e^{-S_0}}{\int \mathcal{D}\mathrm{U}\, e^{-S_0}} \approx \frac{1}{N_{\mathrm{confs}}} \sum_{i=1}^{N_{\mathrm{confs}}} G(t)_i\,. \tag{4}$$

If we instead want to simulate the physics of a different action S_η (in our case, with the background electric field) but have access to Monte Carlo configurations using the action S_0, we can simply modify the Monte Carlo estimate to correct for the additional portion of the factor e^{-S}:

$$\langle G(t,\eta)\rangle_\eta = \frac{\left\langle G(t,\eta)e^{-(S_\eta - S_0)}\right\rangle_0}{\left\langle e^{-(S_\eta - S_0)}\right\rangle_0} \approx \frac{\sum_i G(t,\eta)_i w_i}{\sum_i w_i}\,, \quad (5)$$

where $\langle\cdot\rangle_0$ indicates the average with respect to e^{-S_0} and $w_i \equiv e^{-(S_\eta - S_0)}$ is the reweighting factor associated with configuration i.

The contribution to the weight factor from the fermion sector, using the standard prescription where the fermions are integrated out, can be written as a ratio of fermion determinants:

$$w_i = \frac{\det M_{\mathcal{E}}(U_i)}{\det M_0(U_i)}\,. \quad (6)$$

We want to include the effect of the electric field on both flavors of sea quarks; this can be done by simply computing weight factors at two values of η (corresponding to the up and down quark charges) and multiplying them.

There are two well-known problems associated with reweighting. The first is that if the overlap between the target and simulated ensemble is poor, the weight factor fluctuates too strongly and the reweighted ensemble will wind up dominated by just a few configurations, leading to a lack of statistical power. The second is that the determinant ratio must be estimated stochastically. The good news is that the since the average over stochastic noises commutes with the gauge average, any unbiased estimator for the weight factor will also produce an unbiased estimate for operators computed on the reweighted ensemble [21], even if it is quite noisy.

When the reweighting factors are close to one, the overlap is good and for most estimators the stochastic noise is also reduced. Since we can get the reweighting factors arbitrarily close to one by decreasing the value of η, none of the issues mentioned above create problems for our calculation. On the other hand, this does not mean that our calculation gets more precise as $\eta \to 0$. This is because the signal we try to measure is encoded in the correlation between the weight factor and the ones in the hadronic correlator. As η is decreased both signal and error decrease in concert, leading to a constant relative error.

B. Perturbative reweighting

As we will see, the most difficult part of performing the reweighting calculation for the electric field is the estimation of the weight factors, as the stochastic estimators for the weight factor in our case are substantially more noisy than in the traditional mass reweighting.

Stochastic estimators for determinant ratios have been used in many studies, more recently as a technique to fine-tune the quark mass in dynamical simulations via reweighting [20–22]. We attempted at first to use a similar method to estimate the weight factors. However, even for large numbers of stochastic noises, we were unable to resolve even the difference of the weight factors from unity on a production-sized lattice [23].

In this study we use an alternative to the standard stochastic estimator, a perturbative technique for estimation of the weight factor. Since we are interested only in perturbatively small $\mathcal{E}$, we can expand the one-flavor weight factor w_q about $\eta = 0$:

$$w_q(\eta) = 1 + \eta \left.\frac{\partial w_q}{\partial \eta}\right|_{\eta=0} + \frac{1}{2}\eta^2 \left.\frac{\partial^2 w_q}{\partial \eta^2}\right|_{\eta=0} + \mathcal{O}(\eta^3)\,. \quad (7)$$

To obtain the two-flavor weight factor w at some particular value of $\mathcal{E}$ corresponding to η_d for the down quark and $-2\eta_d$ for the up quark, we simply multiply, keeping terms only up to η^2:

$$\begin{aligned} w(\eta_d) &= w_d w_u = w_q(\eta_d) w_q(-2\eta_d) \\ &= 1 - \eta_d \frac{\partial w_q}{\partial \eta} + \eta_d^2 \left[\frac{5}{2}\frac{\partial^2 w_q}{\partial \eta^2} - 2\left(\frac{\partial w_q}{\partial \eta}\right)^2\right]\,. \end{aligned} \quad (8)$$

The derivatives are computed for $\eta_d = 0$. To simplify notation, we will denote the derivatives with respect to η around 0 as w_q' and w_q''. Given estimates of these derivatives, we can evaluate the above at any sufficiently-small η_d to produce a reweighted ensemble on which to apply the valence calculation. This is a *semi-perturbative* calculation, since the sea effects are introduced perturbatively *via* the perturbative estimates of the weight factors, but these weight factors are evaluated at finite η_d and used as inputs to the valence calculation. This differs from the full-perturbative method introduced by Engelhardt [5, 24] in that the hadron correlators are computed non-perturbatively, for a small value of η.

This allows one list of weight factors to be applied to a variety of hadrons, which would not be possible in a fully perturbative calculation. Since determination of the weight factors requires the majority of the computational effort, the numerical effort is greatly reduced when computing the polarizability for a set of hadrons.

Note that we did not include a contribution from the strange quarks in the perturbative expansion. In part this is because the strange sea quarks were not included in the measure used to generate our gauge configurations. Additionally, to include the correction due to the electric charge of the sea strange quarks requires evaluating the derivatives w_s' and w_s'' for a different quark mass, significantly increasing the numerical effort, while their contribution is expected to be extremely small.

While we are looking only for quadratic effects and expect no shift in the neutron mass proportional to η_d (due to reflection symmetry), these can arise in two ways: either by the sole effect of the quadratic term in the weight factor, or by a correlation between the first-order term in the weight factor with a similar linear effect in the

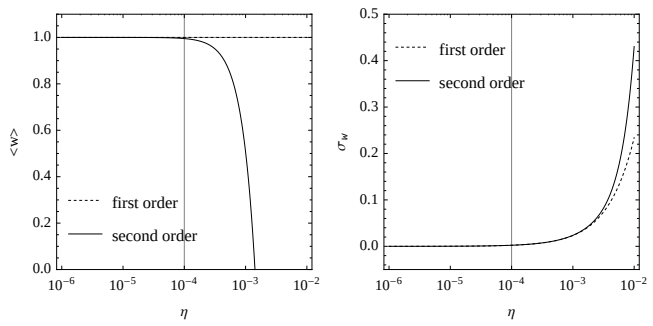

FIG. 2: Dependence of the mean and standard deviation of the reweighting factor as a function of η for the $24^3 \times 48$ ensemble used in this study. The estimator is computed using Eq. 8. The solid lines include the quadratic effects in η, whereas the dotted lines include only the linear term. Vertical lines indicate the η value we used.

neutron correlator. The latter occurs because reflection symmetry is not preserved configuration by configuration, but only in the gauge average. We expect that the gauge average of w'_q is zero, but on individual configurations it will be nonzero.

To evaluate the derivatives we can use Grassman integral techniques and we get

$$w'_q = \frac{\partial}{\partial\eta}\frac{\det M_\eta}{\det M_0} = \mathrm{Tr}\left(M' M_0^{-1}\right) , \tag{9}$$

and

$$\begin{aligned} w''_q =& \frac{\partial^2}{\partial\eta^2}\frac{\det M_\eta}{\det M_0} = \mathrm{Tr}\left(M'' M_0^{-1}\right) \\ &+ \left(\mathrm{Tr}\, M' M_0^{-1}\right)^2 - \mathrm{Tr}\left(M' M_0^{-1}\right)^2 , \end{aligned} \tag{10}$$

where M' and M'' are the derivatives with respect to η at $\eta = 0$ of the one flavor fermionic matrix M.

The chief advantage of perturbative reweighting over conventional reweighting is that computation of the weight factor coefficients requires the evaluation of traces rather than determinants. While these traces must still be evaluated stochastically, and naïve stochastic estimators of them are still quite noisy, there are many improvement techniques available for reducing the stochastic noise in trace estimates; here we make use of dilution. The ability to evaluate Eq. 8 at any desired η after the fact is an auxiliary advantage.

Once the derivatives are computed, we can use them to determine the values of η that are in the small field region. Since we do a perturbative expansion, we want to make sure that the higher order terms are not important. It may seem that given that we only keep the terms of interest, we can set η to any value – the higher order terms are not present. On the other hand, the reweighting is successful only when $\langle w \rangle$ is close to one. If we choose values of η that are too large, the individual reweighting factors could even go negative. In fact, we can choose η such that $\left\langle 1 + w'\eta + w''\eta^2/2 \right\rangle = 0$. It is unclear that the results of the reweighting are meaningful in this case. To set bounds on the η value we used as a guiding principle the requirement that the Taylor expansion of $w(\eta)$ is a good approximation.

For the Taylor expansion to be successful, we expect that the successive terms in the expansion are subdominant. We ask then that η be such that $1 \gg w'\eta \gg w''\eta^2/2$. In Fig. 2 we show both the mean $\langle w \rangle$ and standard deviation σ_w for our ensemble as a function of η when using the first and second order approximations for w. The mean $\langle w \rangle$ when including only the first order term is close to one for all values of η since $\langle w' \rangle \approx 0$, as demanded by symmetries. Note that the mean when including the quadratic term in the approximation deviates quickly from one as we increase η. This is due to the large value of $\langle w'' \rangle$. In fact, a large constant $\langle w'' \rangle$ is not important since it cancels out in the reweighting ratio from Eq. 5. The fluctuations of w about the mean are important and that is why we plot σ_w as a function of η. Note that the standard deviation is dominated by the first order term for values of η much larger than the ones where the mean deviates from one. In any case, the value of $\eta = 10^{-4}$ used in this study is well inside the region where the Taylor expansion is working well.

IV. STOCHASTIC ESTIMATIONS OF THE WEIGHT FACTOR

The traces that appear in the expressions for the determinant derivatives, Eqs. 9 and 10, can be evaluated stochastically in the standard way, that is

$$\mathrm{Tr}\,\mathcal{O} = \left\langle \xi^\dagger \mathcal{O} \xi \right\rangle_\xi , \tag{11}$$

where ξ are Z(4) noise vectors. We note that only three estimators are required—$\mathrm{Tr}\, M' M_0^{-1}$, $\mathrm{Tr}\, M'' M_0^{-1}$, and $\mathrm{Tr}(M' M_0^{-1})^2$—since an estimator for $\left[\mathrm{Tr}(M' M_0^{-1})\right]^2$ can be constructed from two uncorrelated values of the estimator for the first-order term $\mathrm{Tr}(M' M_0^{-1})$. As $\left[\mathrm{Tr}(M' M_0^{-1})\right]^2$ is both computed separately and subdominant, we refer to the combination of the two second-order terms that must be explicitly estimated, $\mathrm{Tr}(M'' M_0^{-1}) - \mathrm{Tr}(M' M_0^{-1} M' M_0^{-1})$, as $\tilde{w}''_q$. Note that there is no bias introduced by using the same stochastic noise vector for the w'_q and $\tilde{w}''_q$, since the ultimate computation of the weight factor involves only linear combinations of these estimates; any correlations in the stochastic fluctuations will not cause the final result to be biased. This reduces the number of inversions required per noise vector to two.

Standard stochastic estimators of these traces are, unfortunately, very noisy. For example, on a 4^4 lattice we need 5×10^6 noise vectors to obtain a signal-to-noise ration greater than one for the first derivative. In Fig. 3 we compare the stochastic result with the exact result computed via direct evaluation [25]. We see an agreement

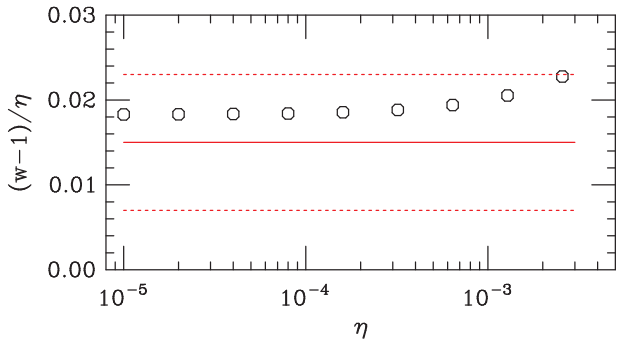

FIG. 3: Exact values for $[w(\eta)-1]/\eta$ on a 4^4 lattice, compared with the value predicted by the stochastic estimator for w' and its error band. The black circles are derived from the exact values of the determinant ratio at different η; the solid red line shows the central value of the stochastic estimator, and the dashed red lines the stochastic error bands.

between $w_q(\eta)$ and the stochastic estimator for w'_q, with the onset of quadratic behavior visible as η is increased.

A. Estimator quality

Since the limiting factor for this calculation is the stochastic estimation of the weight factors, it is useful to understand how far we need to reduce the variance in the stochastic estimator. Whether using perturbative or nonperturbative reweighting, it is the variation of the weight factor between gauge configurations that carries the information, and it is this fluctuation that we seek to extract using a stochastic estimator. Thus the gauge variance between configurations in the weight factor amounts to a signal, while the stochastic variance gives the noise in that signal. This immediately suggests a criterion for judging the quality of any given stochastic estimation scheme: the stochastic signal-to-noise ratio

$$w_{SNR} \equiv \frac{\sigma_{\text{gauge}}}{\sigma_{\text{noise}}} \,. \tag{12}$$

Ideally we would like this SNR to be as large as possible. A SNR substantially less than unity means that the stochastic estimator scheme used is insufficient to extract whatever physics differences exist between the original and reweighting ensembles. In our case, this may be because the actual difference is small, or because the estimator is too noisy; the only way to distinguish these cases is to carry the calculation to its conclusion and see how much reweighting increases the overall error bar.

There are two difficulties which make this SNR a guideline, rather than a quantitative measurement:

1. Determining the gauge variance is difficult, since it requires knowledge of the true weight factors, the same quantities whose estimation we are concerned with.
2. When using a highly diluted estimator (which we will choose to use in the end), determination of the stochastic variance requires computation of multiple stochastic estimates. This may involve a substantial amount of computer power.

We will return to these issues later in the discussion of specific estimators in Sec. IV B.

B. Mapping the stochastic noise

It can be shown readily that the variance of the stochastic estimator $\operatorname{Tr}\mathcal{O} = \langle \xi^\dagger \mathcal{O} \xi \rangle$ is

$$\left\langle (\xi^\dagger \mathcal{O} \xi)^2 \right\rangle - \left\langle \xi^\dagger \mathcal{O} \xi \right\rangle^2 = \sum_{i \neq j} |\mathcal{O}_{ij}|^2, \tag{13}$$

the sum of the squares of the off-diagonal elements of $\mathcal{O}$. Understanding which of these elements dominate is useful for designing improvements to the stochastic estimator. As we cannot even afford to compute all of the diagonal elements (to get an exact value for $\operatorname{Tr}\mathcal{O}$), we certainly cannot compute all of the $\mathcal{O}_{ij}$'s. However, we can examine a representative set to see which are dominant. On a single configuration from our $24^3 \times 48$ ensemble, we have computed all $\mathcal{O}_{ij}$ for a set of sources j

$$\mathcal{S} = \{j | j_{x,y,z} \in \{8, 16\}\,, j_t \in \{16, 24, 32\}\} \,. \tag{14}$$

Since we compute all spin-color combinations, the number of sources is $|\mathcal{S}| = 12 \times 24 = 288$. We kept the information only for sinks i such that the vector between i and j has no components larger than 12 (after accounting for periodic boundary conditions in the y and z directions). The number of data points is very large and to produce a more manageable set we bin the points in equivalence classes. For the purpose of this illustration, we assume that all source positions are equivalent, so we average together the squares of all matrix elements corresponding to different source points. We notice no significant effects on matrix elements whose sinks are near the Dirichlet boundary, so we bin together the points where the separation vector between i and j is related by a reflection in any direction or rotation in the (y, z)-plane. We also bin together the points that have the same starting and ending color indices and separately the ones that have different color indices. We treat the directions x and t separately due to the effects of the electric field and collect each of the 16 spinor combinations in a separate bin. All these data points are used to create Fig. 4 and to predict the error of the stochastic estimators for different dilution schemes.

The relative size of the off-diagonal elements as a function of the Euclidean separation between i and j is shown in Fig. 4 for both $M' M_0^{-1}$, the first-order term, and $M'' M_0^{-1} - (M' M_0^{-1})^2$, the second order term. We note that the magnitude of $|\mathcal{O}_{ij}|^2$ decreases as i and j are further apart, as expected. The short-range behavior is the source of our problem. The trace estimator we use works

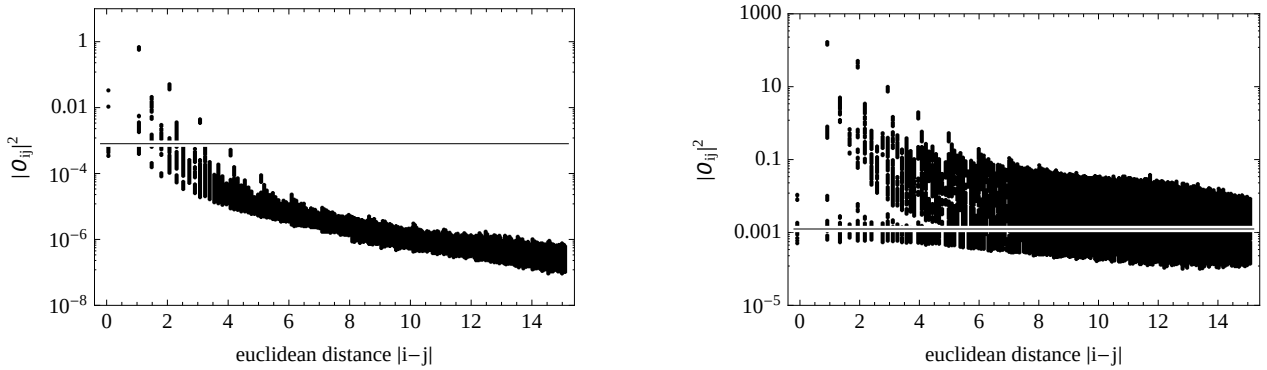

FIG. 4: Mapping of a representative set of the off-diagonal elements of $M'M_0^{-1}$ (left) and $M''M_0^{-1} - (M'M_0^{-1})^2$ (right). The average of the square of the 12 diagonal elements, the ones that define the signal, is shown as a horizontal line.

well for diagonally dominated matrices, where the largest elements of the matrix lie along the diagonal, and decay quickly as we go away from it. Unfortunately, for our matrices the dominant elements are not on the diagonal, as can be easily seen from Fig. 4. Even among the elements at Euclidean separation 0, those off-diagonal elements $\mathcal{O}_{ij}$ where i and j differ in spin and color indices are larger than the diagonal elements. This is a simple depiction of why this stochastic estimator is so difficult: the diagonal elements (the signal) are small, while the near-diagonal elements contributing to noise are much larger. The structure is not unexpected, since M' amounts to a point-split operator in the t direction.

Fig. 4 suggests that the most direct route to reducing the variance is reducing the short-range off-diagonal elements of the operators. There are two somewhat redundant techniques we can use to do this: hopping parameter expansion improvement and dilution. Hopping parameter expansion has the advantage that its numerical cost is relatively modest for small orders, but it only cancels the off-diagonal elements approximately. We explored this technique in a previous study using an expansion up to 7^{th} order, the largest order we could afford [26]. We found that the improvement was insufficient and the signal-to-noise ratio for polarizability was smaller than one. In this work we explore an alternative approach: dilution.

C. Dilution scheme

Dilution is a technique which, with a suitable dilution scheme, can eliminate the noise contribution from near-diagonal elements. It entails partitioning the lattice into N subspaces, estimating the trace over each separately, and adding the estimates; this is done in practice by generating noise vectors with support only on one subspace. This eliminates contributions to the variance from off-diagonal elements $\mathcal{O}_{ij}$ where i and j belong to different subspaces, at the cost of requiring N evaluations of $\mathcal{O}$ to generate a single estimate. Thus there is a fundamental tradeoff involved in dilution. The aim of any stochastic estimation procedure is to minimize the uncertainty in the stochastic estimate for a given computational effort, and that uncertainty, rather than the variance of the estimator itself, should be used as the yardstick for measuring the utility of a dilution technique.

The variance of the diluted estimator should then be compared with the variance of an estimate based on the average of N independent evaluations of the undiluted estimator. The variance of this mean is smaller by a factor of N than the variance of a single evaluation. To be more precise, if we label the partition to which the (spin/color/spatial) index i belongs as $P(i)$, the variance becomes

$$\mathrm{var}(\mathrm{Tr}\,\mathcal{O}) = \sum_{i\neq j} |\mathcal{O}_{ij}|^2 \delta_{P(i)P(j)}, \tag{15}$$

that is, the sum of only those off-diagonal elements that connect indices belonging to the same subspace. If all N subspaces are of equal size (which is generally the case), then this results in a sum with only $1/N$ as many terms. The ratio of uncertainties (the proper figure of merit) between an N-subspace dilution and the mean of N undiluted estimators, is

$$\frac{\sum_{i\neq j} |\mathcal{O}_{ij}|^2 \delta_{P(i)P(j)}}{\frac{1}{N}\sum_{i\neq j} |\mathcal{O}_{ij}|^2}. \tag{16}$$

Thus dilution will only be a success if the average of the off-diagonal elements that survive (belong to the same subspace) is less than the average of all of them. Choosing a dilution strategy, then, must be done with consideration of the form of $\mathcal{O}_{ij}$, as it is entirely possible to partition the lattice in such a way to make the stochastic noise worse.

The most common sort of dilution is spin/color dilution, where each noise vector has support for a single spin and

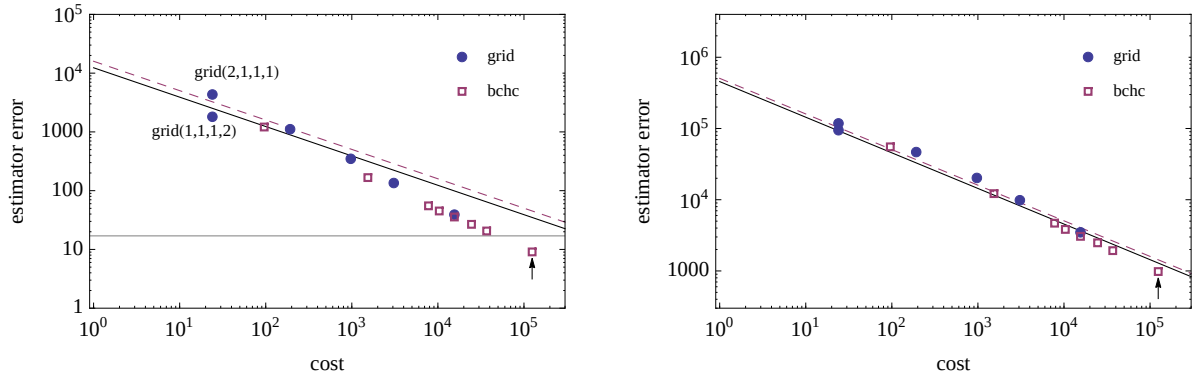

FIG. 5: Uncertainty estimates for different dilution schemes for the first-order term (left) and second-order term (right) as a function of the the estimator cost, given by the number of partitions in each dilution scheme. The solid (dashed) line indicates the expected uncertainty for the estimates based on repeated use of the undiluted (spin-color diluted) estimator. The horizontal line in the left panel corresponds to the standard deviation of the gauge fluctuations, as estimated in the next section. The point in the bottom-right labelled by a black arrow is the BCHC-6^4 dilution scheme actually used in the computation.

color over the entire lattice. As we can see from Fig. 5 this dilution scheme alone does not help us; it must be used alongside other dilution schemes in which the subspace structure also involves spatial separation.

To construct the spatial structure for a dilution scheme for an operator whose off-diagonal elements are expected to decrease with increasing Euclidean distance, we want to allocate sites among the N subspaces so as to maximize the minimum Euclidean distance separating two sites belong to the same subspace. We investigate two schemes: regular grid and body-centered hypercubic (BCHC) scheme.

For a regular grid two points belong to the same partition if

$$p_1 - p_2 = 0 \pmod{\Delta}. \tag{17}$$

The four-dimensional vector Δ defines the steps of the grid in the four spatial directions. The number of partitions, which is proportional to the cost of the dilute estimator, is controlled by the volume of one grid cell $N = \prod_i \Delta_i$. When used in conduction with the spin-color dilution, we have $N = 12 \prod_i \Delta_i$. The minimum Euclidean distance between two points on the same grid is the smallest grid step $\min_i \Delta_i$.

In the BCHC scheme, two points belong to the same partition if

$$p_1 - p_2 \in \{0, \Delta\} \pmod{2\Delta}. \tag{18}$$

This can be thought of as two regular grids of steps 2Δ displaced by vector Δ, so that the origin of the second grid lies in the middle of the unit cell 2Δ, creating a body-centered hypercubic pattern with unit cell 2Δ. The number of partitions for this scheme is $N = 8 \prod_i \Delta_i$, or $N = 96 \prod_i \Delta_i$ when spin-color dilution is also used; it is half that of a grid dilution scheme with the same nearest-neighbor distance. The minimal distance between two points from the same partition depends on the relative magnitude of the components of Δ; when all components are the same $\Delta_i = b$, the minimal distance is $\|\Delta\| = 2b$. Note that for a regular grid we would need N twice as large to achieve this minimal distance. We note that the "hierarchical probing" method [27] generates, as part of the hierarchy, both regular and BCHC grids with $\Delta_i = 2^n$.

A disadvantage of large-N dilution strategies is the need for a large numerical effort to compute even a single estimate of the trace, even if that single estimate has greatly reduced stochastic error. This makes it difficult to empirically determine the variance of a large-N dilution scheme by repeated application, because the cost of repeating the estimator enough times to achieve a sufficiently low error on the variance becomes prohibitive. However, we can use the off-diagonal element mapping data to estimate the variance for any dilution scheme. Under the assumptions outlined above, we estimate

$$\mathrm{var}(\mathrm{Tr}\,\mathcal{O}) \approx \frac{12 \times V_{\mathrm{lat}}}{N} \sum_{i \in V}^{N} \sum_{j \neq i} |\mathcal{O}_{ij}|^2 \delta_{P(i)P(j)}\,, \tag{19}$$

where V_{lat} is the lattice volume and the positions and spin/color of the N sources i are chosen randomly. The sum over the sinks j extends over the entire lattice, rather than a limited hypercube as in Fig. 4, eliminating any effects from small points beyond the horizon on the variance. Additionally, scattering the points over the entire lattice, rather than confining them to a central region away from the Dirichlet walls, correctly incorporates the finite-size effects from the Dirichlet boundary conditions into the estimator variance. For N large enough, the result should quickly converge to the true variance of $\mathcal{O}$. Using 300 randomly chosen lattice points and evaluating all 12 color-spin indices at these points, we determine the standard deviation for our estimators with percent-level error on a few configurations. We find that the standard

deviation varies very little from configuration to configuration. The mean value over the configurations is used for the data in Fig. 5.

This is a useful tool to use in planning a dilution scheme. In Fig. 5 we compare the predicted uncertainty for our stochastic estimator using various dilution schemes. Except for the solid line, all estimators use spin-color dilution. As noted before, the spin-color dilution by itself (indicated by the dashed line) is inferior to the undiluted estimator. At first order, moderately-aggressive dilution schemes essentially keep pace with the decline in the estimator variance caused by simple repetition. Dilution begins to win out once the minimum Euclidean distance between adjacent points in the same subspace, reflected in the increasing cost, increases. At second order, only an extremely small improvement is seen; this is due to the substantially slower falloff of the offdiagonal elements seen in Fig. 4. Either a more aggressive dilution scheme or an operator-improvement technique used in tandem with dilution is needed to see much improvement over simple repetition of the naïve estimator.

The BCHC dilution schemes should show at best a reduction in the cost by a factor of 2 compared to grid schemes, since they achieve the same minimum distance with half as many partitions. The actual gain is less than this, because a grid source has only eight nearest neighbors, while the BCHC source has sixteen. Nonetheless, for both the first and second order estimators, the BCHC dilution outperforms grid dilution by a small amount.

To reduce the stochastic variance to a level comparable with the gauge variance we need a large grid spacing. In the left panel of Fig. 5 we see that this happens for the first-order derivative only when $\Delta = \{6, 6, 6, 6\}$. This is the dilution scheme used in the subsequent calculation. In this scheme, the minimal Euclidean distance between two points in the same partition is 12 and the number of partitions is $N = 96 \prod_i \Delta_i = 124,416$. Evaluating the needed traces requires two inversions per partition per configuration; the evaluation of them all on this ensemble required approximately 150,000 GPU-hours.

D. Gauge variance

Off-diagonal element data allows us to determine the expected variance for our estimators. However, it provides no indication as to the level of gauge variance, which we also need to know to determine whether a dilution scheme noise is smaller that the expected signal, as discussed in Section IV A. To estimate gauge variance we did two tests: an extrapolation from small lattices where we can compute the operators exactly and a more computational intensive study where we evaluated our expensive high quality estimator (BCHC with $\Delta = \{6, 6, 6, 6\}$) on a couple of lattices from our ensemble.

We discuss first direct evaluation of our estimator on $24^3 \times 48$ lattices. For the first two lattice configurations in our ensemble we run several evaluations of our estimator.

config	N_{est}	w'_q		$\tilde{w}''_q$	
		mean	std-dev	mean	std-dev
2	6	-2.8(2.7)	6.5(2.1)	-196,362(468)	1147(371)
3	4	-19.9(4.7)	9.4(4.0)	-197,399(324)	648(274)

TABLE I: Repeated trials of the BCHC diluted estimator for two configurations. The standard deviation field indicates the stochastic error, which we determined in Section IV C to be 9.5(6) for the first-order estimator and 1038(76) for the second-order one.

The results of this test are shown in Table I. We first note that the standard deviation for the stochastic estimators is consistent with the estimate from the previous section. For the first-order term $w'_q = \operatorname{Tr} M' M_0^{-1}$ the gauge fluctuations are 16(8). This estimate takes into account the fact the gauge average is zero, by reflection symmetry, for the first-order term. A correction factor is used to account for the bias in the standard deviation estimator. The stochastic fluctuations are smaller than the gauge fluctuations. This suggests that this estimator is precise enough to follow the gauge fluctuations.

For the second-order term, the gauge average value is $-196,881(491)$. The standard deviation of the gauge fluctuations is $\sigma_{\text{gauge}} = 919(694)$, of similar order with the stochastic uncertainty. It is not clear whether the signal-to-noise ratio is good enough for this estimator, especially since our determination is also compatible with small values for the gauge fluctuations. We will see that the extrapolation from small volumes predicts that σ_{gauge} is on the small side of the estimate. This suggests that the second order estimator is noisy. Note that the cost of this seemingly-simple study is 2.5 million inversions, about 3% of the cost of the entire calculation.

We turn now to the extrapolation from small volumes. We generated a set of small lattice of different geometries and computed the first and second order derivatives exactly using the compression method for Wilson fermions [25, 28]. More precisely, we computed the fermionic determinant on these lattices exactly for 7 different values of the electric field parameter η and then evaluated the derivatives numerically using a $\mathcal{O}(\eta^6)$ finite difference scheme

$$f' \approx \frac{1}{\eta} \sum_{k=-3}^{3} c'_k f(k\eta)\,, \quad f'' \approx \frac{1}{\eta^2} \sum_{k=-3}^{3} c''_k f(k\eta)\,, \qquad (20)$$

where $f(\eta) = \log\det M(\eta)$. It is straightforward to relate these derivatives to the derivatives of the reweighting factor: $w'_q = f'$ and $\tilde{w}''_q = f''$. The coefficients for these approximations are given in Table II. We use a value of $\eta = 0.01$ which is sufficiently precise.

For each lattice geometry we generated 10 configurations. We used Wilson pure gauge action with $\beta = 6.0$. The lattice spacing is $a/r_0 = 0.186$ [29], which is similar to the lattice spacing for our large configurations. For the

k	-3	-2	-1	0	1	2	3
c'_k	-1/60	3/20	-3/4	0	3/4	-3/20	1/60
c''_k	1/90	-3/20	3/2	-49/18	3/2	-3/20	1/90

TABLE II: Coefficients for the finite difference derivatives.

fermionic matrix, we use nHYP fermions with $\kappa = 0.1267$. The parameter κ was adjusted to produce a pion mass around 300 MeV to match the sea quark mass on the large configurations.

To make sure that we are not in the deconfined phase, we have to keep $r_0/L < r_0 T_c = 0.7498(50)$ [30]. This means that all of our lattice dimensions $n_i = L_i/a$ should satisfy $n_i \geq 8$. Since this is already at the upper range of lattice volumes where we can compute the determinant exactly, to investigate a wider range of volumes we have to use geometries that do not satisfy this constraint. For these lattices, we take advantage of the Dirichlet boundary conditions in the x and t directions and cut out these lattices from larger ones, with $n_x = n_t = 12$, that are in the confined phase. The only delicate step in this process is that we have to smear the links on the larger lattice and then cut it, so that the boundary do not introduce discontinuities. We use 72 different lattice geometries: $n_y, n_z \in \{8, 10, 12\}$, $(n_x|n_t) \in \{4|4, 4|6, 6|4, 4|8, 8|4, 4|10, 10|4, 6|6\}$.

For each ensemble we determine the gauge standard deviation for both derivatives and mean for the second derivative. We analyzed the dependence of each of these three quantities as we varied the dimension of the lattice in each direction. In most cases we found that these quantities vary linearly with the dimension (either relatively constant or raising linearly). The only exception is the mean of the second order derivative which requires quadratic terms to describe its dependence on n_x, the extent of the lattice in the direction of the external field. Based on these observations and taking into account the rotational symmetry in the (y, z)-plane, the fit functions we use in our extrapolations are

$$\begin{aligned} \sigma_{w'_q} &= \alpha(n_x + \beta_x)(1 + \gamma n_y)(1 + \gamma n_z)(n_t + \beta_t)\,, \\ \left\langle \tilde{w}''_q \right\rangle &= \alpha(n_x + \beta_x + \gamma n_x^2) n_y n_z (n_t + \beta_t)\,, \\ \sigma_{\tilde{w}''_q} &= \alpha(n_x + \beta_x)(1 + \gamma n_y)(1 + \gamma n_z)(n_t + \beta_t)\,. \end{aligned} \tag{21}$$

The results of the fits are presented in Table III. Using these coefficients and their cross-correlations, we estimate that for a $24^3 \times 48$ lattice the gauge averages and standard deviations should be

$$\begin{aligned} \sigma_{w'_q} &= 17(4)\,, \\ \left\langle \tilde{w}''_q \right\rangle &= -212(2) \times 10^3\,, \\ \sigma_{\tilde{w}''_q} &= 164(62)\,. \end{aligned} \tag{22}$$

We note that all these results are compatible with the values determined via repeated evaluation of the stochastic estimator on two full-size configurations. As we mentioned earlier, the gauge standard deviation for $\tilde{w}''_q$ is lower than the stochastic uncertainty, indicating a noisy estimator.

V. RESULTS

A. Reweighting factors

Before we turn to the main results in this paper, hadron polarizabilities, we present the results for the reweighting factors, as evaluated on the full ensemble using the estimators described in the previous section.

The resulting estimates for w'_q, $\tilde{w}''_q$, and $(\mathrm{Tr}\, M' M_0^{-1})^2$ are given in Fig. 6. We discuss here briefly the estimator for $w'^2_q = (\mathrm{Tr}\, M' M_0^{-1})^2$. When more than one estimate per configuration of the first-order term w'_q is available, such as in the previous study using hoping parameter expansion improvement where we used thousands of cheap estimators per configuration [26], we may construct one estimate for w'^2_q out of two independent estimates of w'_q. However, in this study we used an expensive BCHC-diluted estimator and there is no second estimate of w'_q available. Constructing a second one in the same manner as the first, using the $N = 124, 416$ dilution scheme, would require a large extra effort. However, we observed from the previous study that the stochastic fluctuations of this term compared to the fluctuations of the rest of the traces involved in $\tilde{w}''_q$ are small. Thus it is acceptable to use a less labor-intensive method to estimate it. Since we have the estimates of w'_q from the prior run saved to disk, we use them in combination with the new diluted estimates of w'_q to produce an estimate of w'^2_q on each configuration.

For the first order term we find that the standard deviation is $\sigma_{w'_q} = 23(1)$. This includes both the stochastic noise and the gauge fluctuations. The determination is compatible with our estimates described in the previous section. For the second order term, $\tilde{w}''_q$, the mean value is $\left\langle \tilde{w}''_q \right\rangle = -197,549(83)$ and the standard deviation is $\sigma_{\tilde{w}''_q} = 1429(58)$, again in agreement with the values estimated in the previous section. We note that the combined standard deviation is larger than the gauge one estimated from the extrapolation from small volumes, indicating that the stochastic noise is dominant for this estimator.

For the w'^2_q estimator we find that the standard devi-

	Q	α	β_x	β_t	γ
$\sigma_{w'_q}$	0.15	0.0017(10)	-1.6(3)	1.9(1.2)	0.08(4)
$\left\langle \tilde{w}''_q \right\rangle$	0.85	-0.09(2)	-3.95(3)	-1.61(7)	0.118(5)
$\sigma_{\tilde{w}''_q}$	0.27	0.09(3)	-3.1(1)	3(2)	0.014(16)

TABLE III: Fit parameters for extrapolation from small volumes. Q is the conventional confidence level of the fit: one minus the cumulative χ^2 distribution evaluated at the obtained value of χ^2, taking into account the corrections due to the use of the same data to both estimate the covariance matrix and perform the fit [31].

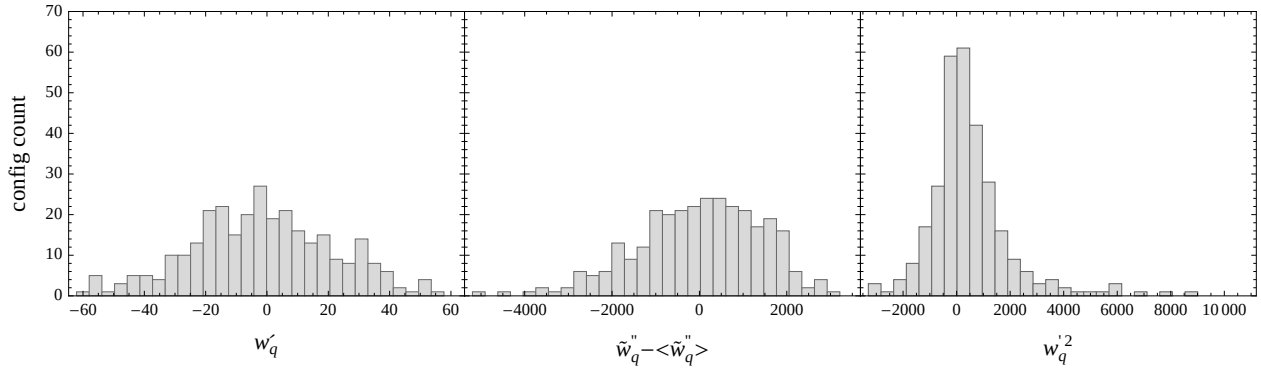

FIG. 6: Distribution of the values of the stochastic estimators on the full 300-configuration ensemble. The two leftmost panels are the stochastic estimators for w'_q and $\tilde{w}''_q$ using BCHC dilution. The right panel is the estimator for w'^2_q using a combination of BCHC-dilution and hopping parameter expansion estimators, as described in the text. Note that for the center panel we shifted the distribution by subtracting its mean.

ation is $\sigma_{w'^2_q} = 1655(68)$. This is comparable with the standard deviation for $\tilde{w}''_q$ and it would seem that this term will add significant variance to the final result. To see why this term is subdominant we have to expand Eq. 8 in terms of traces:

$$w(\eta_d) = 1 - \eta_d w'_q + \frac{1}{2}\eta_d^2 \left(5\tilde{w}''_q + w'^2_q\right) . \qquad (23)$$

We see that in the final result, the quadratic term $w'' = 5\tilde{w}''_q + w'^2_q$ is dominated by $\tilde{w}''_q$. Indeed the total standard deviation for the quadratic term is $\sigma_{w''} = 7231(295)$, compared to the contribution coming from $\tilde{w}''_q$ alone, $5 \times \sigma_{\tilde{w}''_q} = 7147(292)$.

B. Hadron polarizabilities

The power series expansion given in Eq. 8 can be used to determine the weight factor at any desired η_d on each configuration; these weight factors can then be combined with the valence correlators computed previously to complete the calculation. We note that one set of weight factor estimates may be used without modification for all hadrons; this is a strength of the reweighting approach. Full details of the valence correlators are given in [14]; we repeat only the essential elements here. We use point interpolators for both source and sink. To improve the signal-to-noise ratio, we use 28 sources per configuration; in any case the expense of the many sources is dwarfed by the cost of the weight factor estimates. These sources are spread evenly in the (y, z)-plane but are along the centerline $x = 12$ to avoid the Dirichlet walls. We determined suitable correlator fitting ranges by varying the minimum fit distance and looking for a plateau in the extracted mass; this indicated a fitting range of 9-21 for the neutron and 14-30 for the pion and kaon. Detailed information about the choice of these fitting ranges and the effect of varying them can be found in [14]. We note that similar plateaux also appear in the polarizability itself as a function of $t_{\rm min}$, although those data are noisier.

It is informative to turn on the reweighting one order at a time; we additionally add the extra second-order term, w'^2_q, separate from the others that comprise $\tilde{w}''_q$. Using these approximations for the reweighting factors we compute the hadron propagators using Eq. 5 and do a correlated fit for zero field and non-zero field propagators using the model in Eq. 3. The fit ranges for these fits are the same as in our previous valence study [14]. The results for these fits are presented in Table IV. Focusing on the energy shift, $a\Delta E$, note that the uncertainty remains relatively constant when including only the first order terms, indicating that our estimator adds very little noise. The second order term, in particular $\tilde{w}''_q$, introduces significantly more uncertainty, doubling or tripling the size of the error bars. In principle this could be due to either the gauge fluctuations of the second-order term causing a large fluctuation in the weight factor or to the stochastic noise in the estimator. However, in our case the estimated stochastic error for $\tilde{w}''_q$ is fairly large compared to the overall variation of the estimator, so we suspect that the largest share of the fluctuations in our estimates are due to stochastic noise, despite the substantial effort involved in the estimator. As discussed previously, the addition of the w'^2_q estimate has very little effect both on the value of the energy shift and its error.

Fig. 7 shows the effective mass of the neutron as a function of correlator separation, as well as the shift in the effective mass induced by the electric field with full reweighting. We caution that the shift in the effective mass played no part in our analysis, which was done with the full correlated fit described above, and is presented only as an illustration of trends in the data.

To convert the energy shift to a value for the polariz-

	Valence only			1st order			$\tilde{w}_q''$ only			2nd order		
	aE	$a\Delta E_{\times 10^8}$	Q	aE	$a\Delta E_{\times 10^8}$	Q	aE	$a\Delta E_{\times 10^8}$	Q	aE	$a\Delta E_{\times 10^8}$	Q
π^0	0.245(1)	-5.4(3.4)	0.17	0.245(1)	-6.0(3.4)	0.18	0.245(1)	5.4(5.6)	0.15	0.245(1)	5.6(5.7)	0.15
K^0	0.352(1)	4.2(0.8)	0.12	0.352(1)	3.7(1.0)	0.07	0.352(1)	10.5(3.4)	0.03	0.352(1)	11.1(3.4)	0.02
Neutron	0.694(4)	62.8(5.7)	0.65	0.694(4)	63.9(6.5)	0.57	0.695(4)	72.5(16.4)	0.53	0.695(4)	67.0(16.3)	0.43

TABLE IV: Results for the energy and energy shift for the pion, neutron, and kaon with differing orders of reweighting: none (the valence-only calculation), first-order in η_d, second-order including only the dominant contribution in $\tilde{w}_q''$, and the full calculation to second order. For the energy shifts, the values are in units of 10^{-8}. Q is the confidence level for the fits corrected to account for the sample size [31].

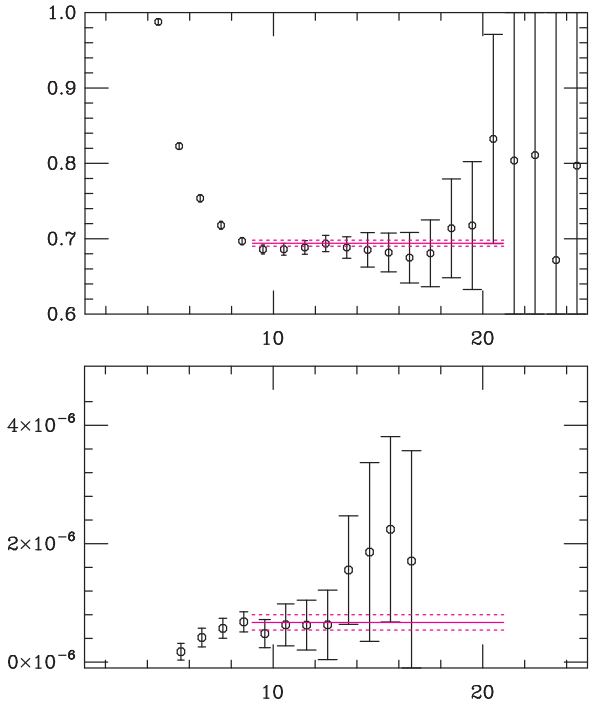

FIG. 7: Effective mass (above) and effective mass shift (below) for the neutron. For the effective mass, the mass shift is sufficiently small that the plot symbols with and without the electric field are indistinguishable. In both cases, the value of the mass or mass shift extracted from the full correlated fit is shown as a colored error band; its horizontal extent indicates the chosen fitting range of 9-21 for the neutron. For the effective mass shift, the data at larger values of t lie outside the plot range and are not shown.

ability we use the relation:

$$\alpha = \frac{2a^3e^2}{9\eta_d^2}(a\Delta m) = \frac{2a^3e^2}{9\eta_d^2}\frac{aE}{aM}(a\Delta E)\,, \qquad (24)$$

where M is the mass of the hadron of interest, computed using periodic boundary conditions. These masses were computed for this ensemble in a previous study [14]. For the neutron, a correction due to the magnetic moment is required, $\alpha_c = \alpha + \mu^2/(2m)$ [10, 14, 32]. The polarizability values are given in Table V. We will discuss now each hadron separately.

We remind the reader that the neutral pion correlator used in this study does not include the disconnected diagrams that are required due to the isospin breaking introduced by the electric field. This is a common limitation for lattice calculations, since the inclusion of these terms is computationally expensive. For the neutral pion, chiral perturbation theory predicts a polarizability around $\alpha_{\pi^0} \approx -0.5\times10^{-4}\,\mathrm{fm}^3$. In the absence of the disconnected contributions, the prediction is that the polarizability would be positive and an order of magnitude smaller in absolute value [9]. Lattice calculations of this quantity indicate that the connected neutral pion polarizability turns negative as the pion mass is lowered below 400 MeV, contradicting these expectations [9, 14, 33]. It was suggested that this discrepancy is due to finite volume corrections [9], but this does not seem to be the case [33]. The correction associated with charging the sea quarks might also be responsible for this discrepancy. As we can see from Table V, the polarizability for the neutral pion seems to change signs as we charge the sea quarks. However, the current errors are too large, relative to the size of polarizability, so no definitive conclusions can be drawn. We also measured the change in the energy shift induced by the reweighting taking into account the correlations between the original and reweighted measurements; the result was consistent with zero. We note that the valence-only result is very close to zero; the ensemble in question happens to lie very near the value of m_π where the polarizability of the neutral pion changes sign [14, 33].

Neutral kaon polarizability is not shifted by the first order reweighting. When the second order is included, both the central value and its uncertainty increase. In this

	Valence only	1st order	$\tilde{w}_q''$ only	2nd order
π^0	-0.21(14)	-0.24(14)	0.21(22)	0.22(23)
K^0	0.14(3)	0.13(3)	0.36(12)	0.38(12)
Neutron	2.56(19)	2.60(22)	2.89(55)	2.70(55)

TABLE V: Electric polarizability for the pion, neutron, and kaon with differing orders of reweighting: none (the valence-only calculation), first-order in η_d, second-order including only the dominant contribution in $\tilde{w}_q''$, and the full calculation to second order. Values are in units of $10^{-4}\,\mathrm{fm}^3$.

case the shift in polarizability is statistically significant. It is interesting to note that this behavior is consistent with the features we observed in our previous study: kaon polarizability was insensitive to the change in mass of the valence light quarks, but it shifted significantly when the mass of the light sea quarks was changed [14]. This was in contrast with the pion polarizability which seems to depend strongly on the valence quark mass, but it was fairly insensitive to the sea. It is then not surprising that the kaon polarizability should be sensitive to charging the light sea quarks. In any case, the chiral extrapolation performed in our previous study for kaon polarizability needs to be revisited, given the significant shift induced by charging the sea.

The neutron, the benchmark hadron for this type of calculations, shows no statistically significant change when the coupling to the sea is turned on via reweighting. This is a bit puzzling since the chiral perturbation theory expectation is that the neutron polarizability increases by $1\text{–}2 \times 10^{-4}\,\mathrm{fm}^3$ when the sea quarks are charged [34] and our errorbars, even after including the second order correction, are small enough to resolve this difference. It is still possible that this increase shows up after removing the finite volume effects, that are expected to be significant for this quantity. We note that our calculation of neutron polarizability, including sea effects, improves upon the precision of the only other such calculation known to us [5, 24] in both precision and pion mass.

While the effects of charging the sea quarks are not statistically significant here, with the exception of the kaon, we expect them to be enhanced both by enlarging the lattice volume and by approaching the chiral limit. Considering the chiral limit: when the pion mass is reduced it is easier to create virtual pion loops which increases the size of the pion cloud and its contribution to polarizability. Similarly, increasing the size of the box reduces the momentum of the lowest pion state (recall that we use Dirichlet boundary conditions), reducing the cost of exciting pions, with similar consequences. We thus expect the effect of charging the sea to be substantially larger at lower pion mass and on larger boxes.

VI. CONCLUSION

While the result for the neutron here is physically significant, as it improves on the previously-attained precision, we treat it more as a proof of concept for the perturbative reweighting method which will soon be applied to ensembles with larger volumes and smaller pion masses, where we expect the effect to be larger. We caution that this result is evaluated on only one ensemble without the necessary continuum, chiral, and infinite volume extrapolations required to produce a physical result. We expect the effects of the finite volume to be significant and are currently conducting a similar calculation on a larger ensemble. Experience with a similar action in [19] indicates that the effect of the continuum extrapolation is likely to be small compared to the statistical uncertainty. The perturbative estimate for the weight factor correctly predicts the slope of the exact determinant ratio on small lattices where it can be computed exactly, but like the conventional reweighting estimator it is quite noisy. However, dilution can be used to reduce its variance. Strong dilution with the body-centered hypercubic pattern outperforms hopping parameter expansion and it is certainly simpler to formulate and more flexible.

Our results suggest that while these estimates of the first-order term w'_q are sufficient, a reduction in the stochastic noise from the second-order term would be welcome, given that the other ensembles in the study will be inherently more expensive. Dilution completely eliminates the near-diagonal contributions, at the cost of indirectly increasing the contributions away from the diagonal since we no longer average together many estimates. The long-distance behavior of the off-diagonal elements is exponential and its slope is governed by m_π. We are exploring the use of low-mode subtraction to eliminate the lowest lying modes of the Dirac operator from the operators in question and thus increase the exponent of the falloff; preliminary studies of this technique look promising.

Acknowledgments

We would like to thank Craig Pelissier for generating the gauge ensemble used in this work. The computations were done in part on the IMPACT GPU cluster and Colonial One at GWU, the GPU cluster at Fermilab. This work is supported in part by the NSF CAREER grant PHY-1151648 and the U.S. Department of Energy grant DE-FG02-95ER40907.

[1] J. Schmiedmayer, P. Riehs, J. A. Harvey, and N. W. Hill, *Phys.Rev.Lett.* **66** (1991) 1015–1018.
[2] K. Kossert, M. Camen, F. Wissmann, J. Ahrens, J. Annand, *et. al.*, *Eur.Phys.J.* **A16** (2003) 259–273, [{nucl-ex/0210020}].
[3] H. Fiebig, W. Wilcox, and R. Woloshyn, *Nucl.Phys.* **B324** (1989) 47.
[4] J. C. Christensen, W. Wilcox, F. X. Lee, and L. Zhou, *Phys.Rev.* **D72** (2005) 034503, [{hep-lat/0408024}].
[5] **LHPC Collaboration** Collaboration, M. Engelhardt, *Phys.Rev.* **D76** (2007) 114502, [{arXiv:0706.3919}].
[6] A. Alexandru and F. X. Lee, *PoS* **LATTICE2008** (2008) 145, [{arXiv:0810.2833}].
[7] A. Alexandru and F. X. Lee, *PoS* **LAT2009** (2009) 144, [{arXiv:0911.2520}].
[8] W. Detmold, B. C. Tiburzi, and A. Walker-Loud, *PoS*

LATTICE2008 (2008) 147, [{arXiv:0809.0721}].
[9] W. Detmold, B. C. Tiburzi, and A. Walker-Loud, *Phys.Rev.* **D79** (2009) 094505, [{arXiv:0904.1586}].
[10] W. Detmold, B. C. Tiburzi, and A. Walker-Loud, *Phys.Rev.* **D81** (2010) 054502, [{arXiv:1001.1131}].
[11] R. P. Hildebrandt, H. W. Grießhammer, T. R. Hemmert, and B. Pasquini, *Eur.Phys.J.* **A20** (2004) 293–315, [{nucl-th/0307070}].
[12] J. McGovern, D. Phillips, and H. Grießhammer, *Eur.Phys.J.* **A49** (2013) 12, [{arXiv:1210.4104}].
[13] V. Lensky and V. Pascalutsa, *Eur.Phys.J.* **C65** (2010) 195–209, [{arXiv:0907.0451}].
[14] M. Lujan, A. Alexandru, W. Freeman, and F.X. Lee, *Phys.Rev.* **D89** (2014) 074506, [{arXiv:1402.3025}].
[15] A. Hasenfratz, R. Hoffmann, and S. Schaefer, *JHEP* **0705** (2007) 029, [{hep-lat/0702028}].
[16] R. Sommer, *Nucl.Phys.* **B411** (1994) 839–854, [{hep-lat/9310022}].
[17] A. Alexandru, M. Lujan, C. Pelissier, and F. X. Lee, in *Application Accelerators in High-Performance Computing (SAAHPC), 2011 Symposium on*, pp. 123-130, [{arXiv:1106.4964}]
[18] A. Alexandru, C. Pelissier, B. Gamari, and F. Lee, *J.Comput.Phys.* **231** (2012) 1866-1878 [{arXiv:1103.5103}]
[19] S. Durr *et* al., JHEP **1108** (2011) 148, [{arXiv:1011.2711}].
[20] H. Ohki, S. Aoki, H. Fukaya, S. Hashimoto, T. Kaneko, *et. al.*, *PoS* **LAT2009** (2009) 124, [{arXiv:0910.3271}].
[21] A. Hasenfratz, R. Hoffmann, and S. Schaefer, *Phys.Rev.* **D78** (2008) 014515, [{arXiv:0805.2369}].
[22] Q. Liu, N. H. Christ, and C. Jung, {arXiv:1206.0080}.
[23] W. Freeman, A. Alexandru, F. Lee, and M. Lujan, *PoS* **LATTICE2012** (2012) 015, [{arXiv:1211.5570}].
[24] M. Engelhardt, *PoS* **LAT2009** (2009) 128, [{arXiv:1001.5044}].
[25] A. Alexandru and U. Wenger, *Phys.Rev.* **D83** (2011) 034502, [{arXiv:1009.2197}].
[26] W. Freeman, A. Alexandru, M. Lujan, and F. X. Lee, *PoS* **LATTICE2013** (2013) 288, [{arXiv:1310.4426}].
[27] A. Stathopoulos, J. Laeuchli, and K. Orginos, SIAM J. Sci. Comput., **35(5)**, S299S322, [{arXiv:1302.4018}].
[28] K. Nagata and A. Nakamura, *Phys.Rev.* **D82** (2010) 094027, [{arXiv:1009.2149}].
[29] S. Necco and R. Sommer, *Nucl. Phys.* **B622** (2002) 328–346, [{hep-lat/0108008}].
[30] S. Necco, *Nucl.Phys.* **B683** (2004) 137–167, [{hep-lat/0309017}].
[31] D. Toussaint and W. Freeman, {arXiv:0808.2211}.
[32] A. L'vov, *Int.J.Mod.Phys.* **A8** (1993) 5267–5303.
[33] A. Alexandru and F. Lee, *PoS* **LATTICE2010** (2010) 131, [{arXiv:1011.6309}].
[34] W. Detmold, B.C. Tiburzi, and A. Walker-Loud, *Phys.Rev.* **D73** (2006) 114505, [{hep-lat/0603026}].