EM sea effects in hadron polarizabilities through reweighting

Walter Freeman, Andrei Alexandru, Michael Lujan, Craig Pelissier

The George Washington University

June 27, 2012

Overview

- Background on nucleon polarizability
  - Status of calculations
  - Background field method
  - Boundary condition considerations
  - Sea contributions
- Background on reweighting
- Simulation details
- The traditional stochastic estimator
  - Failed improvement approaches
- "Perturbative reweighting"
  - Hopping parameter expansion and improvement
- Preliminary results
- Next steps

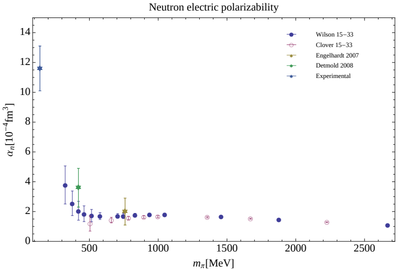
Status of polarizability calculations

This work is part of a long-term project to attack the polarizability of the neutron (or other hadrons).

A "basic" computation of the neutron polarizability is not too hard to do.

Multiple challenging issues required to make contact with experiment:

- Approach to the chiral limit
  - EFT's predict very strong dependence on m_q near chiral limit
  - Lighter quark masses and expensive chiral actions
  - Good probe of chiral behavior on lattice
- Finite volume corrections
- Continuum limit



- Sea effects (this project): simulating interactions between sea quarks and background field is less straightforward

Long-term goal: understand these effects one by one, then put them together

Background on polarizability measurements

Variety of polarizabilities parametrize mass shift in the presence of EM fields:

$$H_{\rm EM} = -\boldsymbol{\mu}\cdot\mathbf{B} - \frac{1}{2}\alpha\mathbf{E}^2 - \frac{1}{2}\beta\mathbf{B}^2 - \frac{1}{2}\gamma_{\rm E1}\boldsymbol{\sigma}\cdot\mathbf{E}\times\dot{\mathbf{E}} - \frac{1}{2}\gamma_{\rm M1}\boldsymbol{\sigma}\cdot\mathbf{B}\times\dot{\mathbf{B}} + \ldots$$

Background field method: measure M_N without and with uniform electric field, examine difference

- Apply electric field by applying appropriate U(1) phase to links in x_4 direction
- Error on mass shift can be much less than error on each mass separately: use extended covariance matrix (error on M_N is 13 MeV; error on ΔM_N is 0.7 MeV)
- Consider fields of both $\vec{E}$ and $-\vec{E}$ to eliminate first-order effects and reduce noise

Apply field by multiplying gauge links by U(1) phases:

Make substitution $U_\mu \to e^{-iqaA_\mu} U_\mu$ with $A = (0, 0, 0, iEx_1)$

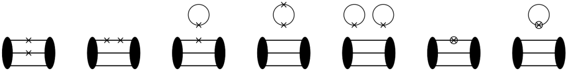
Issues with the background field

- Should really use true exponential rather than linearizing the phase factor ($\exp \alpha = 1 + \alpha$)
- In Euclidian time, U(1) phases represent an *imaginary* electric field!
  - Something of a non-issue: this just flips sign of ΔM
  - Tests on quenched: this works just fine
- Need to address discontinuities at edges of lattice:
- Use Dirichlet boundary conditions in time
- Three options in spatial direction:
  - Just deal with discontinuity at edge of lattice
  - Use larger value of η, well out of perturbative regime
  - Use Dirichlet boundary conditions in x_1, too
    - Avoid multi-valued electric potential, issues with quarks winding around lattice
    - Any effect can be treated as "another finite volume effect": vanishes as $n_x \to \infty$ (we have elongated lattices...)
    - Potentially more important for sea effects
    - Effect of DBC's needs further study: comparison with PBC's as a function of n_x

EM sea contributions to neutron polarizability: status

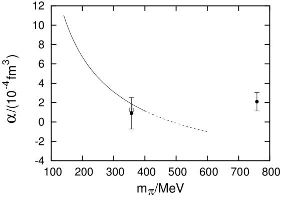
Only aware of one other calculation for the neutron polarizability including sea effects: M. Engelhardt, [0706.3919] and [1011.5044]:

- Uses explicit expansion of the path integral up to order E^2
- Mixed-action calculation: DWF valence on MILC Asqtad 2+1 sea ($a = 0.12$ fm)
- Sea effects appear naturally in list of diagrams that must be evaluated



- Calculation done relatively far from chiral limit: $m_\pi = 357, 759$ MeV

	α_E, $10^{-4} fm^3$	
M_π	Valence	Valence + sea
759 MeV	1.5(3)	2.1(1.0)
357 MeV	0.5(5)	0.9(1.6)



- Ensembles far enough from chiral limit that result expected to be small
- Effect of sea contributions not yet distinguished from zero

EM sea contributions to neutron polarizability: status

- Related calculation (another plenary at this conference): T. Ishikawa *et al.*[1202.6018], using reweighting to look at EM sea effects on EFT LEC's
  - A bit different, as they are using reweighting to get dynamical QED, rather than a background field
- Conclusions:
  - Reweighting method successful
  - Sea effects not distinguishable from zero
  - Future work can bring down uncertainties
- Both this work and Engelhardt's work unable to resolve effect of charged sea
- It may well be that sea EM effects are not that important, but we'd like to put a smaller error bar on them
  - Zero, after all, is a perfectly reasonable result, if it comes with small errors...

Including sea effects

Can the quark sea effects be considered with our nonperturbative background field approach?

Including sea effects

Can the quark sea effects be considered with our nonperturbative background field approach?

- In principle, yes – just generate two otherwise identical ensembles, one with a background field!
- Problem: requires unaffordably high statistics, since our two mass measurements now no longer have correlated errors

Can we somehow generate two *correlated* ensembles, one with a background field and one without?

Including sea effects

Can the quark sea effects be considered with our nonperturbative background field approach?

- In principle, yes – just generate two otherwise identical ensembles, one with a background field!
- Problem: requires unaffordably high statistics, since our two mass measurements now no longer have correlated errors

Can we somehow generate two *correlated* ensembles, one with a background field and one without?

... yes!

Reweighting, in general

In general, a technique for extracting physics from a different action than the one used in generation: "retroactively change the ensemble parameters"

Standard	Reweighted
$\frac{\int [dU] \mathcal{O} e^{-S_1}}{\int [dU] e^{-S_1}} \to \frac{\sum \mathcal{O}_i}{\sum 1}$	$\frac{\int [dU] \mathcal{O} e^{-S_2}}{\int [dU] e^{-S_2}} = \frac{\int [dU] \mathcal{O} e^{-(S_2 - S_1)} e^{-S_1}}{\int [dU] e^{-(S_2 - S_1)} e^{-S_1}} \to \frac{\sum \mathcal{O}_i w_i}{\sum w_i}$ where $w_i = e^{-(S_2 - S_1)_i}$.

This will only work well if the two ensembles overlap sufficiently.

- If they don't, this manifests as large fluctuations in w_i's (over many orders of magnitude!)
- Large fluctuations decrease the statistical power of the ensemble: can compute the "effective number of configurations": $N_{\rm eff} = N \frac{\overline{w}^2}{\overline{w^2}}$ (Liu et al., [1206.0080])
- Not a problem for reweighting in E, since in principle we are only after perturbatively small changes

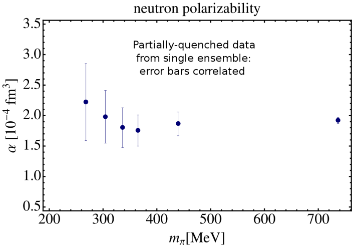
Very useful for us, generate two correlated ensembles with different parameters!

$\to$ Taking advantage of this property of reweighting is nothing new: used by H. Ohki *et al.* to compute $\frac{\partial M_N}{\partial m_s}$ by reweighting in m_s [0910.3271]

Plan of attack

We already know how to do the valence contribution:

- Measure the zero-field neutron correlator as usual
- Compute the link variables with the field: $U_\mu \to e^{-iqaA_\mu} U_\mu$ (different for u and d quarks)
- Use these links to measure the neutron correlator with $\vec{E} = \pm E_0 \hat{x}$
- Do correlated fit to correlators to determine dependence of M_N on E^2



To include the sea contribution:

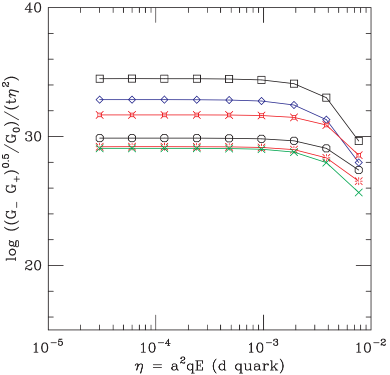
- Determine weight factor $w_i = \frac{\det M_0}{\det M_\eta}$ for each configuration
  - Different for u, d quarks – multiply weight factors together.
- Compute valence propagators as before, using reweighted ensemble

Choosing a value of η

What value of the electric field shall we use?

- Use of Dirichlet boundary conditions in x_1 allows use of arbitrarily small fields
  - Choosing $\eta = a^2 qE$ too large will provoke $\mathcal{O}(E^4)$ effects
  - Choosing η too small will run into numerical precision issues

- Plot $\Delta M_N/\eta^2$ against η for a handful of configurations



- Large intermediate area in the middle where correlator response is almost perfectly quadratic
- Some calculations (including ours!) have gotten dangerously close to the edge of the perturbative region
- Have the freedom to choose η at whatever point makes the reweighting easier

Estimating weight factors

Can't compute $w_i = e^{-(S_2 - S_1)_i} = \det \frac{M_2}{M_1}$ exactly

Standard stochastic estimator: $w_i = \left\langle \exp(-\xi^\dagger(\Omega - 1)\xi) \right\rangle_{e^{-\xi^\dagger \xi}}$ where $\det \Omega = \det M_1 M_2^{-1}$ and Ω is explicitly Hermitian

- For reweighting two identical flavors at once, this is easy: $\Omega = M_2^{\dagger -1} M_1^\dagger M_1 M_2^{-1}$
- For us, must take the square root: $\Omega = \sqrt{M_2^{\dagger -1} M_1^\dagger M_1 M_2^{-1}}$
- Use a rational function approximation to the square root
- Somewhat expensive: requires inverting a matrix that itself contains inversions
  - It has been suggested by J. Finkenrath *et al.* that one can just abandon Hermiticity and discard the imaginary part of the estimator
  - This saves all the work associated with the rational function approximation
  - See their talk at this conference for details on more ways to improve estimators
- Stochastic estimator not a problem as far as bias: the average over noises commutes with the gauge average (A. Hasenfratz, [0805.2369])
- How much work should you do for the stochastic estimator, then?
- Rough yardstick: $\sigma_{\mathrm{stoc}} \simeq \sigma_{\mathrm{gauge}}$

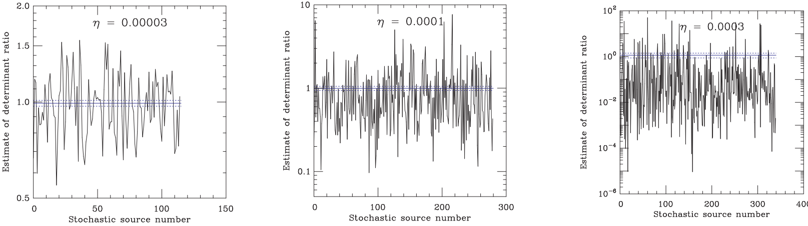
Gauge ensembles and analysis code

- 2 flavors of nHYP-smeared clover fermions, $m_\pi \approx 300$ MeV
- Standard Symanzik-improved gauge action, $\beta = 7.1$
- $24^3 \times 48$, $32 \times 24^2 \times 48$, $48 \times 24^2 \times 48$ volumes; $a = 0.1255$ fm
  - Elongated boxes originally generated for a scattering study; we reuse them here
- $24^3 \times 64$, $m_\pi \approx 230$ MeV in progress
- 300 minimally-autocorrelated configurations per ensemble
- Analysis done with the GWU-QCD general-purpose library
  - Most notable feature: multi-GPU support for many things

Estimating weight factors

Problem: this estimator is tremendously noisy!

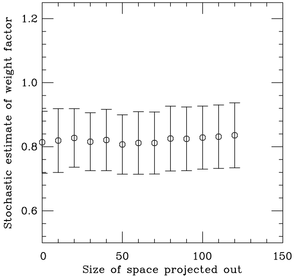
- This is not a problem *per se*: the average over noises commutes with the gauge average (A. Hasenfratz *et al.*, [0805.2369])
- However, too much variance in the stochastic estimator will kill the signal
- Distribution of estimators is log-normal, as you might expect
  - Extreme difficulties in sampling the "long tail" for $\eta = aqE$ large ($\eta \gtrsim 10^{-4}$)
  - Weight factors indistinguishable from unity for smaller η for a "sane" number of noises
  - Can't even estimate how many noises we will need to see a "signal"!



- On a 4^4 test lattice we confirmed that the estimator gives the right result... but it takes $10^5 - 10^6$ noises!
- The "signal-to-noise ratio" is essentially independent of η

Variance reduction techniques that don't work, I

- Common technique in mass reweighting: separate out low modes
  - Low modes: compute determinant exactly
  - Remaining high modes: use stochastic estimator
  - Easier for mass reweighting, since eigenvectors of **M** (and thus Ω) don't depend on m_q
  - ... they *do* depend on η, so we must explicitly compute eigensystem of Ω (expensive!)
  - For us Ω close to **1**; might as well extract extremal modes on both ends



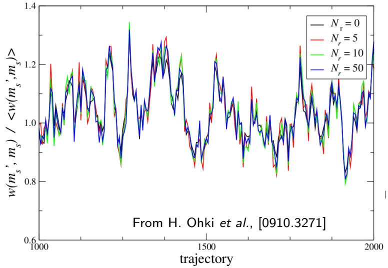
Projecting out extremal modes doesn't reduce the noise at all, unlike for mass reweighting (see A. Hasenfratz *et al.*, [0805.2369] and others)

What is going on?

Variance reduction techniques that don't work, I

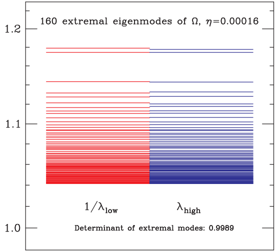
For mass reweighting, most of the signal is carried by the low modes:

- Shown analytically by F. Palombi and M. Lüscher [0810.0946]
- Shown very starkly by H. Ohki *et al.* in their reweighting in m_s – low modes had no effect at all!
- Similar behavior seen by T. Kaneko (talk at this conference), using reweighting to retune m_s



Low and high extremal eigenvalues are very nearly paired
→ Extremal sector contributes very little to $\det \Omega$ for reweighting in the background field!

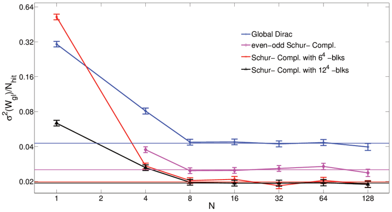
The mechanism that protects the mass reweighting weight factor from high-mode fluctuations does not seem to apply to our problem!



Variance reduction techniques that don't work, II

- Another common technique for improving stochastic estimators is "determinant breakup"
- Instead of reweighting from $\mathcal{S}_0$ to $\mathcal{S}'$ in one step, do it in several

- Used by T. Ishikawa *et al.* [1202.6018] to reweight in α_E
- Investigated in detail by Liu *et al.* [1206.0080] and used by J. Finkenrath *et al.* (Tuesday's talk) for mass reweighting



- Reduces the tendency to underestimate the error associated with highly skew log-normal distributions
- Makes possible "macroscopic" shifts by breaking them into "microscopic" steps
- Not terribly useful for us, since we already want a perturbatively-small change
  - We tried it anyway: using more "steps" is no better than using more "hits"
- Can't be paired efficiently with determinant breakup, since eigenvectors of our Ω change each step

A new pseudo-perturbative approach

- Can we make use of the fact that we only need perturbatively small η?
- Perhaps it is easier to estimate $\left.\frac{\partial w_i}{\partial \eta}\right|_{\eta=0}$ and $\left.\frac{\partial^2 w_i}{\partial \eta^2}\right|_{\eta=0}$ than w_i itself?
- Need to compute form of w_i up to second order in η
  - Linear term in weight factor can combine with linear dependence of $G_N(t)$ on η to give quadratic effect
  - Quadratic term in weight factor by itself can give quadratic effect
- If we can estimate these derivatives instead we can evaluate at any sufficiently-small η we like to get $w_i(\eta)$

Derivation of the estimator

For the first derivative, we want $\frac{\partial}{\partial\eta} \left. \frac{\det M_\eta}{\det M_0} \right|_{\eta=0}$. Rewrite $\det M_\eta$ as a Grassman integral:

$$
\begin{aligned}
\frac{\partial}{\partial\eta} \left. \frac{\det M_\eta}{\det M_0} \right|_{\eta=0} &= \frac{1}{\det M_0} \frac{\partial}{\partial\eta} \int d\psi d\bar{\psi}\, e^{-\bar{\psi} M \psi} \\
&= \frac{1}{\det M_0} \int d\psi d\bar{\psi}\, -\bar{\psi} \frac{\partial M_0}{\partial\eta} e^{-\bar{\psi} M_0 \psi} \\
&= \mathrm{Tr} \left(\frac{\partial M_0}{\partial\eta} M_0^{-1} \right).
\end{aligned}
$$

- This trace still must be evaluated stochastically; perhaps it can be made more tractable?
- It certainly seems cheaper (no square root)

The second derivative proceeds similarly:

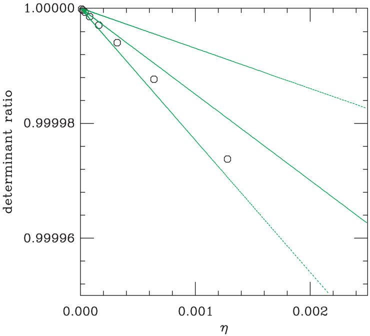
$$
\frac{\partial^2}{\partial\eta^2} \left. \frac{\det M_\eta}{\det M_0} \right|_{\eta=0} = -\mathrm{Tr} \frac{\partial^2 M}{\partial\eta^2} M_0^{-1} + \left(\mathrm{Tr} \frac{\partial M}{\partial\eta} M_0^{-1} \right)^2 - \mathrm{Tr} \left(\frac{\partial M}{\partial\eta} M_0^{-1} \right)^2
$$

Fluctuations in the new stochastic estimator

Estimate the traces in the conventional way: $\mathrm{Tr}\mathcal{O} = \left\langle \xi^\dagger \mathcal{O} \xi \right\rangle$ with either Gaussian or Z(4) noises

- The bad news: the new estimator is still awfully noisy!
- As before, use a coarse 4^4 lattice as a test
  - Can actually afford the statistics needed to get a "signal"
  - Can also do the determinant exactly to test the estimator

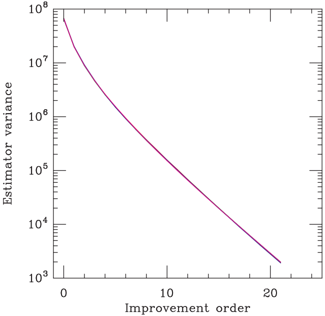
- Estimate for $\left.\frac{\partial w_i}{\partial \eta}\right|_{\eta=0}$: $-0.015(8)$
- ... using 5×10^6 Gaussian noises!
- Good agreement with exact calculation
- ... so what have we gained?



(Possible) salvation: estimator improvement

- Key idea of improvement technique: traceless parts of $\frac{\partial M}{\partial \eta} M^{-1}$ contribute only noise
- Identify and subtract them
- Do this by making hopping parameter expansion of M^{-1}:
  - $AM^{-1} = A + \kappa A\Gamma + \kappa^2 A\Gamma^2 + \kappa^3 A\Gamma^3 + \ldots$, where $A = \frac{\partial M}{\partial \eta}$ and Γ incorporates both the hopping and (if included) clover terms
  - Subtract these terms from the stochastic estimator (cheap)
  - Many have zero trace already and can be just ignored
  - Compute the trace of the rest exactly and add it back in (maybe not cheap)

- Can examine improvement in estimator without doing exact traces
- Estimator is much (much!) faster: no square root
- Examine improvement on several $24^3 \times 48$ lattices
- Improvement of roughly a factor of ten in statistics per two orders!
- How many orders can we compute exactly?



Computing exact traces: Wilson quarks

- Need to compute traces exactly of the $A\Gamma^n$ terms subtracted from estimator
- Problem much simpler without clover term: then A and Γ only contain hopping terms
- Even powers ($\mathrm{Tr} A$, $\mathrm{Tr} A\Gamma^2$) are trivially zero
- $\mathrm{Tr} A\Gamma^2 = 0$ as well due to Dirac structure
- Lowest nonzero contribution is $\mathrm{Tr} A\Gamma^3$ – works out to be the sum of imaginary parts of (1,4) plaquettes
- In general, trace over loops of a given shape proportional to the product of links times electric flux through them

Computing exact traces: clover quarks

- Problem much more difficult once clover term is added
  - Now $\Gamma = \not{D} + \mathcal{C}$, $A = \frac{\partial \not{D}}{\partial \eta} + \frac{\partial \mathcal{C}}{\partial \eta}$, where $\mathcal{C}$ is the clover term
  - Can no longer discard half of the orders outright, and only look at loop shapes
- Problem: How can we compute this efficiently?
- Solution: separate Γ into hopping (in each of eight directions) and clover term
- Further separate clover term into its six "leaves":
  $\mathcal{C} = c_{sw} \left(\sigma_{12} C_{12} + \sigma_{13} C_{13} + \sigma_{14} C_{14} + \sigma_{23} C_{23} + \sigma_{24} C_{34} + \sigma_{12} C_{34} \right)$
- Expand $A\Gamma^n$, giving 3×14^n terms – each has a simple spatial and Dirac structure
  - Only need to evaluate SU(3) products for terms that are both closed and have nonzero Dirac trace
  - Example:
    $\mathrm{Tr} A\Gamma^2 = \sum \mathrm{Tr}_{SU(3)} \left(4i \frac{\partial C_{14}}{\partial \eta} C_{12} C_{24} + 4i \frac{\partial C_{14}}{\partial \eta} C_{13} C_{34} - 4i \frac{\partial C_{14}}{\partial \eta} C_{24} C_{12} - 4i \frac{\partial C_{14}}{\partial \eta} C_{34} C_{13} \right)$
  - For $A\Gamma^7$: 316M terms, 7M of them closed, 0.5M nonzero Dirac trace
  - Evaluation of loop products: 40 CPU-hours/config (to this order)

→ This variance reduction method can be applied to other estimators on other actions:

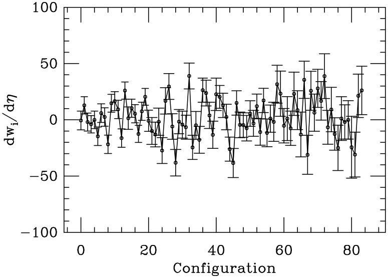
- Estimators of ex. charm quark condensate, which are very noisy
- May be simpler on "pure hopping" actions, like staggered

Finally, a signal

Applying this technique allows extraction of a signal with "modest" amounts of work.

Preliminary data:

- ~ 5000 noises per configuration (10 GPU-hours on "slower" GPU cluster)
- 7th-order improvement (40 CPU-hours)
- Can easily do better than this; this is just a proof-of-concept



- Estimates for $\frac{\partial w_i}{\partial \eta}$ significantly different than zero: success!
- For values of η we are interested in, w_i close to unity $\rightarrow$ ensembles will be highly correlated. This suggests reweighting won't increase error on ΔM_N too much.

Completing the calculation to get α_E

- Need to do the second derivative term too
  - Exact computations of weight factors on 4^4 lattice suggest it doesn't matter that much
  - Same sort of estimator and improvement technique can be used; cost should be comparable
  - Same technology can be used to compute the exact traces
  - Modify stochastic estimator for $(\mathrm{Tr}\mathcal{O})^2$
- Evaluate form for $w_i(\eta)$ for both $\eta = \eta_0$ and $\eta = -2\eta_0$ to get overall sea weight factor
- Use ensemble of weight factors to compute hadron correlators for $\eta \neq 0$
- Construct jackknife estimator for extended covariance matrix and fit to determine ΔM_N as before

Next steps

- Further refine the technique:
  - Port exact trace code to GPU's (take advantage of GWU-QCD code and GPU resources)
  - Exact trace code can be further optimized (at the cost of memory)
- Focusing on α_E:
  - May still be unable to resolve sea effect on polarizability from zero if it is small, but that's okay
  - Fold technique into broader GWU-QCD program for polarizability: approach to chiral limit, higher statistics, bigger volumes, finer lattices
  - New ensemble currently being generated: sea effects predicted to be larger for lower m_π
- Investigate other applications of improvement technique
  - Can be applied to many things involving a trace of an operator containing M^{-1}
  - Charm quark estimators come to mind – estimator noise a major factor in large error in nucleon charm
  - May be difficult to apply to more complicated actions: DWF, overlap?

Conclusions

Reweighting becoming an increasingly-popular technique. Two main uses:

- Make small shift in a parameter (m_q, $\vec{E}$...) to study effect (this work, Ohki's strangeness study) or finetune parameters
- Make large shift as a way of *e.g.* getting closer to chiral limit

Determinant breakup only useful to turn large shifts into small ones (J. Finkenrath's talk)

Low-mode subtraction tremendously powerful for mass reweighting (H. Ohki *et al.*'s work on strangeness, T. Kaneko's talk at Latt2012, useless for us

Estimator for background field reweighting tremendously noisy, can't be tamed by common techniques

New pseudo-perturbative method shows promise:

- Compute $\frac{\partial w_i}{\partial \eta}$ and $\frac{\partial^2 w_i}{\partial \eta^2}$ instead
- Hopping parameter expansion of M^{-1} and exact trace calculations improves estimator drastically
- Preliminary studies show this allows affordable calculation of weight factors
- Potentially broadly applicable