Trinity College, Dublin.

Progress in lattice algorithms

Lattice '01, Berlin
August 19$^{\text{th}}$, 2001.

Mike Peardon,
Trinity College, Dublin.

"Great algorithms are the poetry of computation"

Jack Dongarra, Francis Sullivan,
CiSE Jan/Feb 2000.

Overview

Getting the most configurations from a computer
and
Getting the most from the configurations.

1. Current techniques and comparisons.
2. Odd flavour Wilson fermion simulations.
3. HMC algorithm "plug-ins"
4. Exponential error reduction for Yang-Mills
5. All-to-all propagators.

Current Techniques

□ (Pseudofermionic) Hybrid Monte-Carlo (HMC) [Duane et.al.(1987)].

The standard method for most large-scale two-flavour QCD simulations. (SESAM, CP-PACS, UKQCD, ...)

□ Pseudofermion; replace determinant by a gaussian integral;

$$\det M^\dagger M = \int \mathcal{D}\phi\mathcal{D}\phi^* \ \exp\left\{-\phi^* \left[M^\dagger M\right]^{-1} \phi\right\}$$

□ A time coordinate is introduced, in which the system will evolve according to the classical equations of motion derived from the Hamiltonian

$$\mathcal{H} = \frac{1}{2}\pi^2 + S_G(U) + \phi^* \left[M^\dagger M\right]^{-1} \phi$$

π are a new set of momenta, conjugate to the "co-ordinates" U.

□ The Monte Carlo algorithm has two main ingredients;

- A molecular dynamics evolution to propose a new configuration for the ensemble.
- A Metropolis accept-reject test to correct for inexact integration of the Hamiltonian.

Current Techniques

□ Multi-Boson (MB) Algorithm and its derivatives. [M. Lüscher hep-lat/9311007].

- Two-Step Multi-Boson Algorithm (TSMB) [I. Montvay hep-lat/9510042]
- UV filtered Multi-Boson [Ph. de Forcrand hep-lat/9809145].

□ Construct a polynomial approximation to the inverse,

$$\mathcal{P}(x) \approx \frac{1}{x}$$

and apply it to the hermitian +ve definite matrix, Q^2 ($Q = \gamma_5 M$).

□ The polynomial is written in terms of its roots, $\{z_k\}$ as

$$\mathcal{P}(Q^2) = \prod_k^n (Q^2 - z_k)$$

For a suitable choice of polynomial, the roots come in complex conjugate pairs and so the determinant can be rewritten

$$\det \mathcal{P}(Q^2) = \prod_k^n (Q - \mu_k)(Q - \mu_k^*)$$

□ The fermion determinant is replaced with

$$\det M^\dagger M = \det Q^2 = \frac{1}{\det \mathcal{P}(Q^2)}$$

Current Techniques

□ The determinant can be replaced with a gaussian integral representation.

$$\frac{1}{\det \mathcal{P}(Q^2)} = \int \prod_k \mathcal{D}\phi_k \mathcal{D}\phi_k^* \ \exp\left\{ -\sum_k^n \phi_k^*(Q-\mu_k)(Q-\mu_k^*)\phi_k \right\}$$

The method introduces n bosonic fields, but they are coupled to the gauge fields by a local action.

□ The drawback: local updates of the gauge degrees of freedom become very "stiff" as the number of bosonic fields increases.

□ The solutions proposed involved ensuring the algorithm was exact (by introducing a Metropolis test) and devising ways to keep the number of boson fields added to a minimum.

□ Two-step solution: devised for lattice studies of SYM. The basic ingredients are two polynomials $\mathcal{P}_1$ and $\mathcal{P}_2$, built to approximate

$$\mathcal{P}_1(x)\mathcal{P}_2(x) \approx \frac{1}{x^{N_f/2}}$$

$\mathcal{P}_1$ (low-order) is bosonised and $\mathcal{P}_2$ (high-order) is used to perform a noisy Metropolis step.

□ UV-filtered algorithm: try to handle the UV modes of the fermion matrix with a loop expansion and let the polynomial handle the IR modes.

$$\det M = \exp\left\{ -\sum \mathrm{Tr}\ a_i \Delta^i \right\} \times \det\left\{ M e^{+\sum a_i \Delta^i} \right\}$$

Current Techniques

□ Polynomial Hybrid Monte-Carlo (PHMC) [Frezzotti and Jansen hep-lat/9808011]

□ Exact method. Either correct for $\mathcal{P}(x)$ error with an accept-reject step or use a **re-weighting**;

$$\langle \mathcal{O} \rangle_{QCD} = \frac{\langle \mathcal{O} \det X \rangle_{MB}}{\langle \det X \rangle_{MB}}$$

with $X = Q^2 \mathcal{P}(Q^2)$

□ **The stiff dynamics of the** large number of bosons **is avoided by adding** just one field, ϕ **with** non-local action. Use HMC (**which tolerates non-local actions**) to generate configurations.

$$\det Q^2 \approx \frac{1}{\det \mathcal{P}(Q^2)} = \int \mathcal{D}\phi \mathcal{D}\phi^* \;\; \exp\left\{-\phi^* \mathcal{P}(Q^2) \phi\right\}$$

□ **The break-down** of the **polynomial at very** small eigenvalues is seen as an **advantage**.

□ Frezzotti and Jansen advocate the re-weighting method.

- Polynomial weight generates more configurations with low eigenvalues (which are then assigned a lower weight in the ensemble).
- This should lead to a more reliable determination of quantities that are very sensitive to the details of the lowest eigenvalues.

Comparisons

□ Alexandrou et.al. [hep-lat/9906029] tested the UV filtered multiboson scheme against the SESAM HMC code and found a speed-up for simulations at $\beta = 5.6, \kappa = 0.156$ on $16^3 \times 24$ of ≈ 2.

□ Comparison of PHMC vs HMC performed by the ALPHA collaboration [Frezzotti et.al. hep-lat/0009027]. They conclude there is a "slight" advantage to PHMC (for their parameter range). They favour PHMC due to its over-sampling of low eigenvalues.

At the conference...

□ Talk by W. Schroers (Mon 11:50).
Comparison of HMC vs UV-filtered MB method in QCD on a $16^3 \times 32$ lattice.

β	κ	m_π/m_ρ	a (fm)
5.5	0.159	0.8001(42)	0.141
5.5	0.160	0.670(11)	0.117

For $\kappa = 0.159$, the two algorithms are equivalent, but at the lighter fermion mass, the observed gain for UV-MB is ≈ 3.

□ Talk by N. Zverev (Sun 17:00).
Comparison of TSMB vs HMC for QED_4 in the confined and coulomb phases, on $6^3 \times 12$ lattices. Conclude that TSMB is "competitive" in both cases although they suggest their implementation could be tuned further.

Comparisons

□ Most remarkable feature; the race seems so close!

[de Forcrand hep-lat/9903035]
For heavy fermions, we expect

- HMC → Yang-Mills HMC
- MB → local YM update (eg. Cabibbo-Marinari + over-relaxation)

and experience suggests CM+OR-based methods are orders of magnitude more efficient.

□ It would be extremely useful in comparing algorithms if there was a "standard benchmark" dynamical lattice.

□ A very interesting new feature of the algorithms that have "grown" out of polynomial approximations are their ability to simulate an odd number of flavours of (Wilson) quarks.

Odd-flavoured simulations

□ We want to simulate the one-flavour partition function

$$Z_1 = \int \mathcal{D}U \ \det M \ e^{-S_G}$$

□ Standard pseudofermion formulation is unsuited to simulating odd numbers of flavours, since for light fermions, M can have eigenvalues with negative real parts.

□ Polynomial approximations can solve the problem; they can simulate the partition function

$$Z_+ = \int \mathcal{D}U \ |\det M| \ e^{-S_G}$$

if a polynomial approximation to $1/\sqrt{x}$ is constructed.

Non-hermitian polynomials also provide a solution.

□ The sign of the determinant can be included in a *post-hoc* reweighting of observables;

$$\langle \mathcal{O} \rangle_1 = \frac{\langle \mathcal{O} \ \mathrm{sgn}(\det M) \rangle_+}{\langle \mathrm{sgn}(\det M) \rangle_+}$$

□ Sign problem; if a significant fraction of configurations have $\mathrm{sgn}(\det M) = -1$, Monte-Carlo estimators become extremely noisy.

□ For realistic, accessible simulations, the sign problem is expected to be very mild. See the talk of C. Gebert (Sun 15:10, spectrum session)

Odd-flavoured simulations

[Borici & de Forcrand hep-lat/9505021]
[Takaishi & de Forcrand hep-lat/0108012]

□ Polynomial approximation to $1/x$ of even order 2n, applied to the (non-hermitian) fermion matrix, M;

$$M^{-1} \approx \mathcal{P}(M) = \prod_{k=1}^{2n} (M - z_k)$$

□ For a suitably chosen even-order polynomial, roots come in complex conjugate pairs, so

$$\mathcal{P}(M) = \prod_{k=1}^{n} (M - z_k^*)(M - z_k)$$

γ_5 hermiticity means

$$\det(M - z_k^*) = \det(M^\dagger - z_k^*)$$

□ So $|\det M|$ can be replaced by a bosonic partition function,

$$|\det M| \approx \int \mathcal{D}\phi \mathcal{D}\phi^* \ \exp\left\{-\phi^* T_n^\dagger(M) T_n(M) \phi\right\}$$

with

$$T_n(M) = \prod_{k=1}^{n} (M - z_k)$$

Odd-flavoured simulations

□ The scheme can be made exact by a Metropolis test.
1) Propose new configuration, $\{U'\}$ from the approximate action probability distribution.
2) Accept U' with probability

$$P_{\text{acc}} = \min\left(1, \left|\frac{\det M[U']\mathcal{P}(M[U'])}{\det M[U]\mathcal{P}(M[U])}\right|\right)$$

It is expensive to compute the ratio of determinants; use a noisy (Kennedy-Kuti) acceptance test.

□ There are two alternatives;

$$|\det M\mathcal{P}(M)| = \det\sqrt{\mathcal{P}^\dagger(M)M^\dagger M\mathcal{P}(M)}$$

For a good approximation, $\boldsymbol{M\mathcal{P}(M) \equiv W \approx I}$ and so the square root can be computed by Taylor series.

□ [Takaishi and de Forcrand hep-lat/0108012] Estimate the ratio $|C'/C|$ from an unbiased estimator of $(C'/C)^2$ by using a probabilistically terminating series.

□ JLQCD: $O(a)$ improved fermion action.
See the talk by K-I. Ishikawa (Mon 11:30) and M. Okawa (poster).

ILU preconditioned pseudofermions

□ Using the even-odd preconditioned matrix in the pseudofermionic action leads to a higher Metropolis acceptance rate in HMC at a fixed step-size, $d\tau$.

$$\det M^\dagger M = \det M_{ee}^\dagger M_{ee} = \int \mathcal{D}\phi_e \mathcal{D}\phi_e^* \exp\left\{-\phi_e^*[M_{ee}^\dagger M_{ee}]^{-1}\phi_e\right\}$$

□ The even-odd decomposition is one example of an ILU preconditioning scheme.

□ ILU decomposition:

- Assign an index to every site, $s(\underline{x})$.
- Define an ordering on sites:

$$\underline{x} < \underline{y} \text{ if } s(\underline{x}) < s(\underline{y}).$$

- Define the upper (lower) sectors of M according to

$$U_{xy} = \begin{cases} M_{xy} & x < y \\ 0 & \text{otherwise} \end{cases}$$

□ Then $M = I + L + U$ and the matrix

$$\tilde{M} = (I+L)^{-1}(I+L+U)(I+U)^{-1}$$

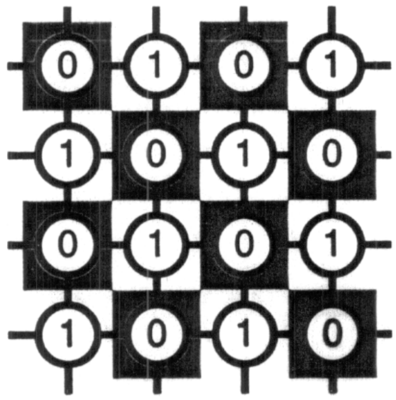
is easier to invert.

□ The "Eisenstat trick": operation with $\tilde{M}$ is no more expensive than operating with M

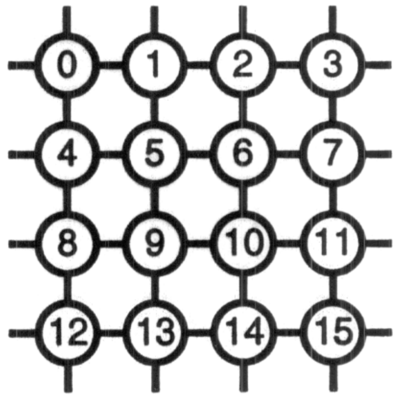
$$\tilde{M} = (I+L)^{-1} + (I+U)^{-1} - (I+L)^{-1}(I+U)^{-1}$$

□ Recent practical example; (SSOR) SESAM collaboration matrix inversion.

ILU preconditioned pseudofermions



Even-odd



Global lexicographic

□ Since det $I + L = 1$, we can replace the two-flavour fermion determinant with

$$\det M^2 = \det M^\dagger M = \det \tilde{M}^\dagger \tilde{M}.$$

Then use pseudofermions coupled to the gauge field via $\tilde{M}$,

$$\det M^\dagger M = \int \mathcal{D}\phi \mathcal{D}\phi^* \;\; \exp\left\{-\phi^* [\tilde{M}^\dagger \tilde{M}]^{-1} \phi\right\}.$$

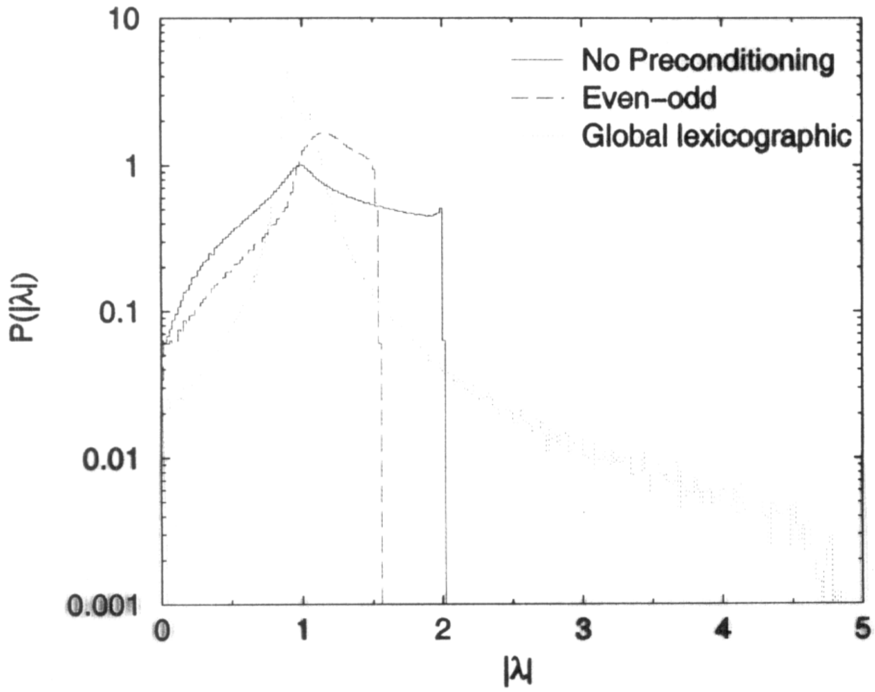
□ Optimal ordering;

- For inversion - global lexicographic.
- For acceptance - even-odd.

□ One possibility; can use different preconditioners in the pseudofermion action and in the matrix inverter.

□ Even-odd matrix is better conditioned, but the eigenvalues of the global lexicographic scheme are sharply peaked around one, with a long tail.

ILU preconditioned pseudofermions



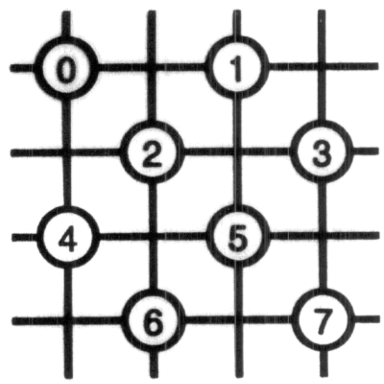
Eigenvalue spectrum of $\gamma_5 M$

2-flavour QED_2.
32×32 lattice,
$\beta = 4.0, \kappa = 0.26$.

[de Forcrand & Takaishi hep-lat/9608093],
[MP hep-lat/0011080]

☐ Even-odd matrix can be ILU preconditioned again. Define an index, $s_e(x_e)$ on the even sites only.

$$U_{xe,ye} = \begin{cases} M_{xe,ye} & x_e < y_e \\ 0 & \text{otherwise} \end{cases}$$



☐ The two-step preconditioned matrix can be used to couple pseudofermions to the gauge fields.

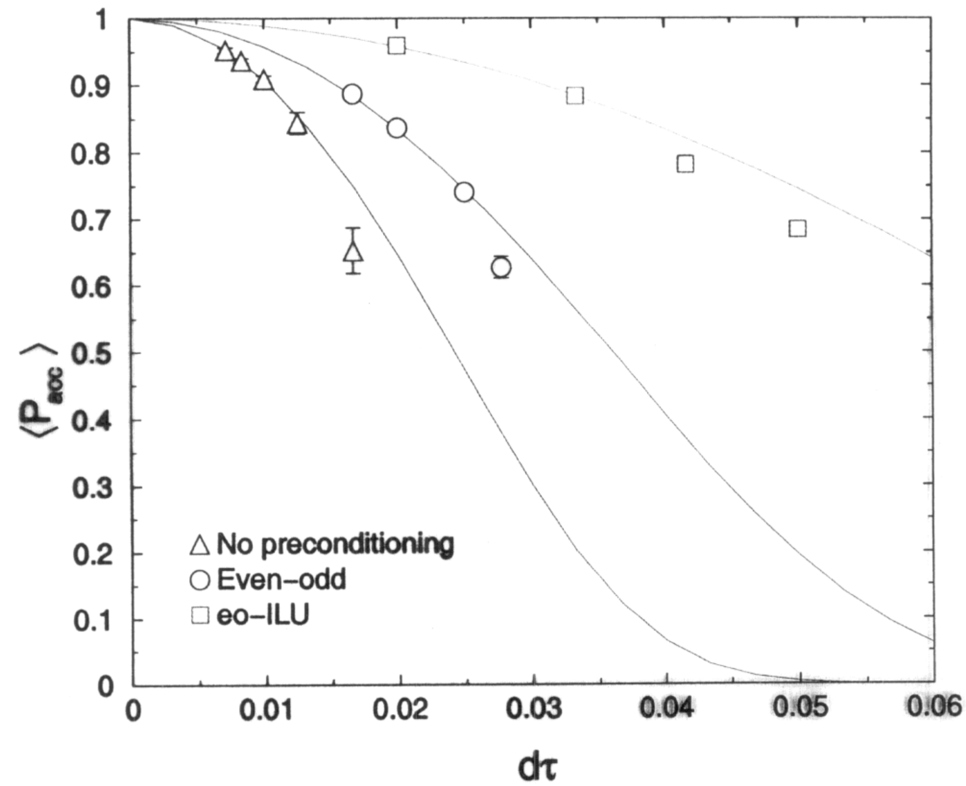
$$\det \tilde{M}_{ee} = \det M_{ee} = \det M,$$

☐ Matrix inversion is also accelerated. As before,

- Lexicographic schemes optimal for inversion
- Local ordering schemes optimal for acceptance

ILU preconditioned pseudofermions

2-flavour Schwinger model, 64×64 lattice, $\beta = 4.0$



Acceptance probability vs. $d\tau$ for $\kappa = 0.2605$

□ Improvement; Fit $\langle P_{\mathrm{acc}} \rangle = \mathrm{erfc}\ (d\tau/\tau_x)^2$

- Even-odd vs unpreconditioned; $\tau_{eo} \approx 1.5\ \tau_0$
- eo-ILU vs unpreconditioned; $\tau_{eoILU} \approx 3\ \tau_0$

□ Local eo-ILU is parallelisable, although perhaps not efficiently.

Splitting the pseudofermions

□ **Hasenbusch suggests [hep-lat/0107019]** splitting the fermion determinant into two parts;

$$\det M = \det \bar{M} \times \det \bar{M}^{-1} M.$$

Then introduce two auxilliary fields so

$$\det M^\dagger M = \int \mathcal{D}\phi \mathcal{D}\phi^* \mathcal{D}\psi \mathcal{D}\psi^* \exp\{-S_\phi - S_\psi\},$$

with

$$S_\phi = |\bar{M} M^{-1} \phi|^2 \quad \text{and} \quad S_\psi = |\bar{M}^{-1}\psi|^2.$$

□ Choose $\bar{M} = I - \bar{\kappa}\Delta$, then **for a good choice of** $\bar{\kappa}$ ($0 < \bar{\kappa} < \kappa$), the two matrices **are better conditioned**.

Better conditioning $\equiv$ **better acceptance**
(at fixed $d\tau$).

□ At each MD evolution step, **need to compute** $R^{-1}\phi \equiv \bar{M} M^{-1}\phi$.

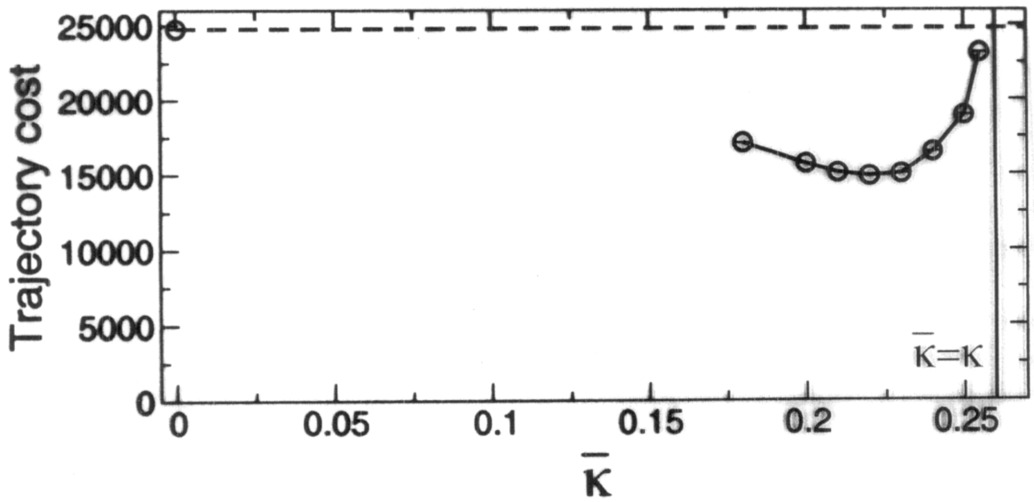
$$\bar{M} = I - \bar{\kappa}\Delta = (1 - \frac{\bar{\kappa}}{\kappa}) + \frac{\bar{\kappa}}{\kappa}(I - \kappa\Delta) = (1 - b) + bM$$

$$\longrightarrow \qquad R^{-1} = (1 - b) M^{-1} + b$$

Splitting the pseudofermions

□ Tests in the Schwinger model, 32×32 lattice, $\beta = 4.0$, $\kappa = 0.26$, even-odd preconditioned pseudofermions.

□ Measure cost (number of $M \times v$ operations for a fixed length trajectory at acceptance $\approx 80\%$) vs $\bar{\kappa}$



□ At minimum, cost $\approx 0.6 \times$ original pseudofermion method.

□ Further runs on bigger lattices with lighter quarks suggest gain increases as $\kappa \to \kappa_{\mathrm{crit}}$. Twice as efficient at lightest quark studied.

□ Martin Hasenbusch & Karl Jansen are studying the algorithm for QCD. See parallel session presentation (Mon 11:10).

Polynomial UV filtered HMC

☐ [MP & Jim Sexton]. Similar splitting to Hasenbusch

$$\det M = \frac{\det M\ \mathcal{P}(M)}{\det\ \mathcal{P}(M)}$$

where $\mathcal{P}(x)$ will be a (crude) polynomial approximation to $1/x$, good close to $x = \lambda_{\text{max}}(M)$.

☐ Polynomial separates determinant into UV and IR components.

☐ Now write a bosonic action;
pseudofermion + PHMC (á la Frezzotti & Jansen).

$$\det M^\dagger M\ e^{-S_G} = \int \mathcal{D}\phi\mathcal{D}\phi^*\mathcal{D}\chi\mathcal{D}\chi^*\ \ \exp\{-S_\phi - S_\chi - S_G\}$$

with gauge action S_G,

$$S_\phi = |\,[M\mathcal{P}(M)]^{-1}\,\phi|^2 \quad \text{and} \quad S_\chi = |\mathcal{P}(M)\chi|^2.$$

☐ The action can be divided into two sectors;

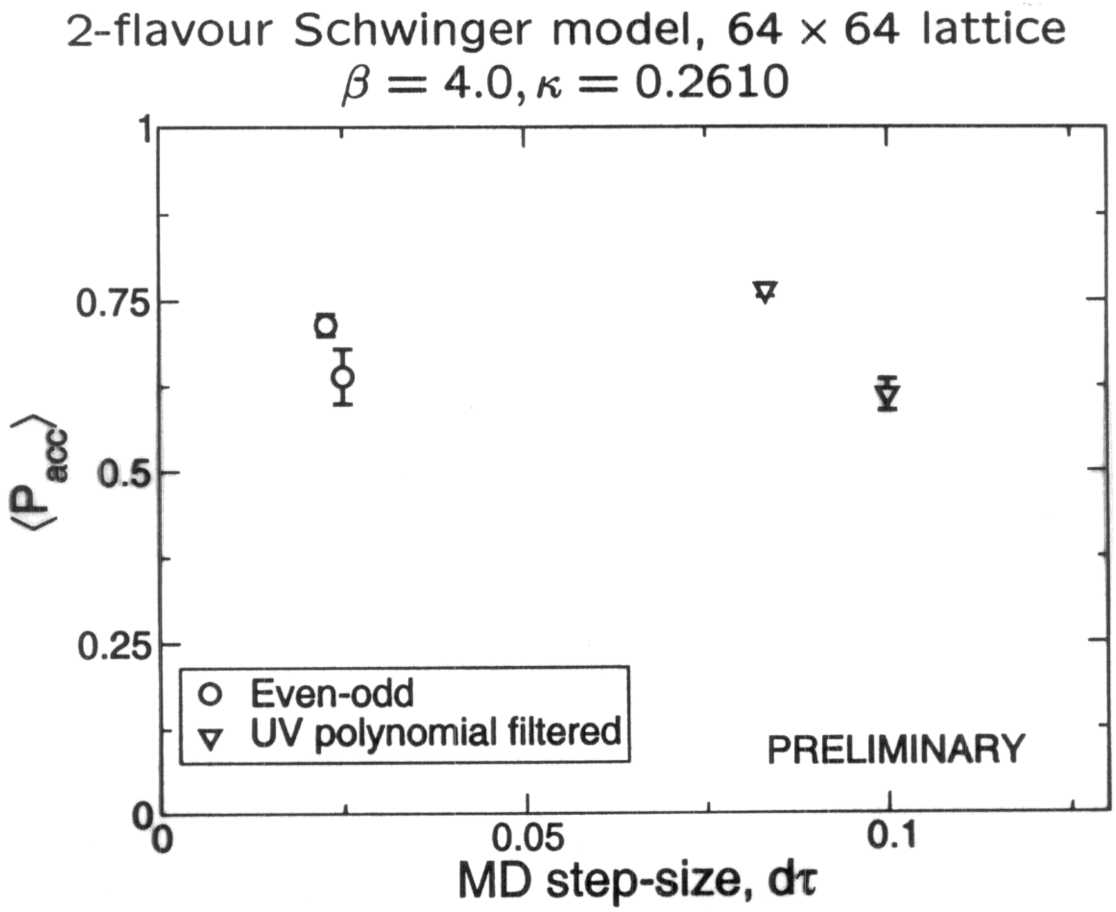
$$S = \underbrace{S_\phi}_{\substack{\text{“IR”} \\ \partial S \text{ costly}}} + \underbrace{S_\chi + S_G}_{\substack{\text{“UV”} \\ \partial S \text{ cheap}}}$$

☐ Sexton-Weingarten multi-scale integrator;
integrate “IR” modes with larger step-size $d\tau$, and “UV” modes with finer step-size $d\tau/m$ $(m \in \mathcal{Z})$.

Trinity College, Dublin.

Polynomial UV filtered HMC

- A first look at the algorithm...



- Chebyshev polynomial; $n = 8, \epsilon = 0.3$
- The UV modes are updated with a step-size 3 times smaller than IR.
- For an acceptance of 70%, can use a step size 4 times bigger.
- Still exploring optimisation.
- First data from (two nodes of) the Dublin PC cluster!

Trinity College, Dublin.

Exponential error reduction for YM theories

[M.Lüscher & P.Weisz. hep-lat/0108014]

□ New scheme for non-abelian Yang-Mills theory; enhance Monte Carlo estimators of large Wilson loops, Polyakov loops, loop correlators etc.

□ Wilson loop; Signal-to-noise falls exponentially with area of loop, $\propto e^{-\sigma A}$

□ Extend the idea of "multihit" technique [Parisi *et.al* (1983)]. Essential tool for measuring potential between static colour sources.
Basic idea of "multihit"; replace temporal links in Wilson loops with their average computed in a fixed background of all other links.

□ Extensions

- Build average of the pairs of links that propagate both the static sources.
- Construct a heirarchy of averaging levels.
- The heirarchy builds averages over successively larger regions.

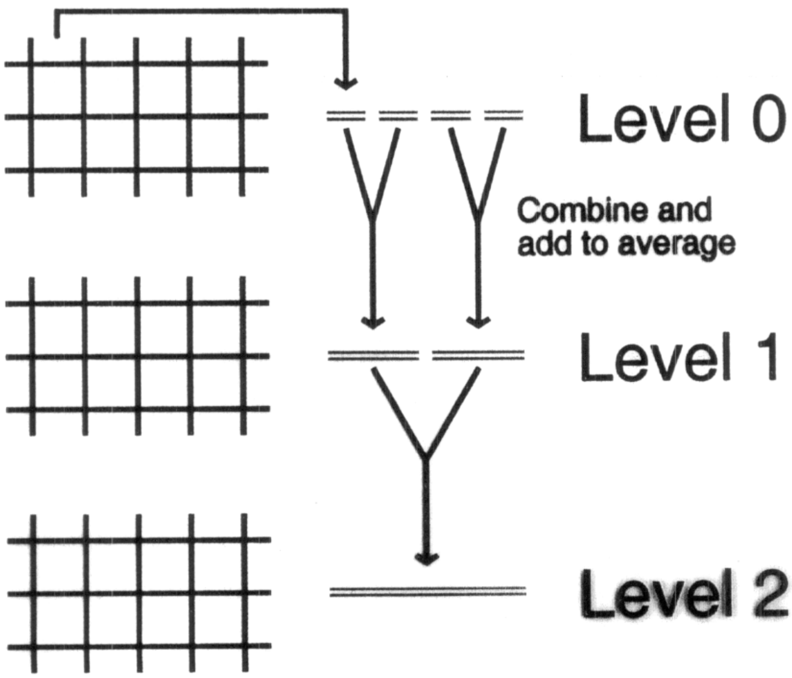
□ For a Polyakov line pair (separated by r), the building block is the two-link transporter,

$$\mathcal{T}(x;t)_{\alpha\beta\gamma\delta} = U_t(x;t)^*_{\alpha\beta}\, U_t(x+r,t)_{\gamma\delta}$$

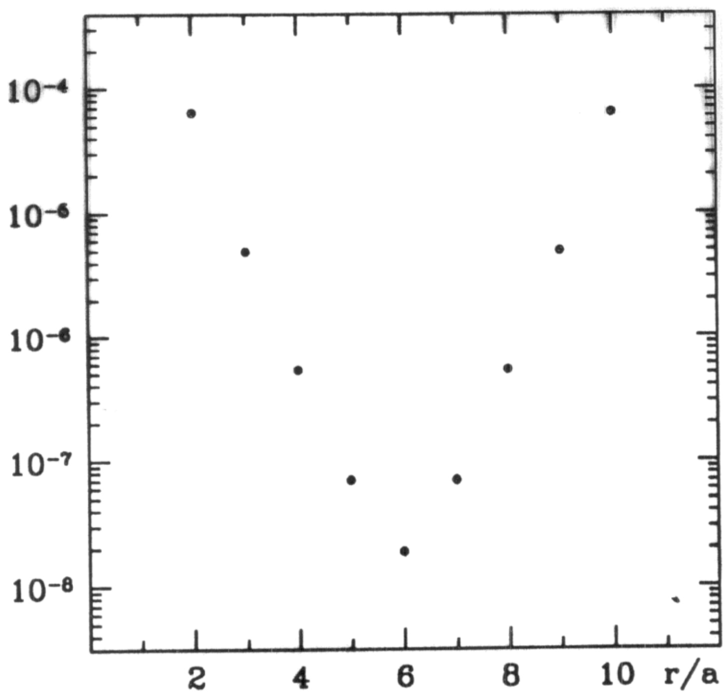
This operator is a 9×9 matrix, acting on colour tensors in $3^* \otimes 3$. The two-hop operator is

$$\{\mathcal{T}(x;t)\mathcal{T}(x;t+a)\}_{\alpha\beta\gamma\delta} = \mathcal{T}(x;t)_{\alpha\lambda\gamma\epsilon}\mathcal{T}(x;t)_{\lambda\beta\epsilon\delta}$$

Exponential error reduction for YM theories



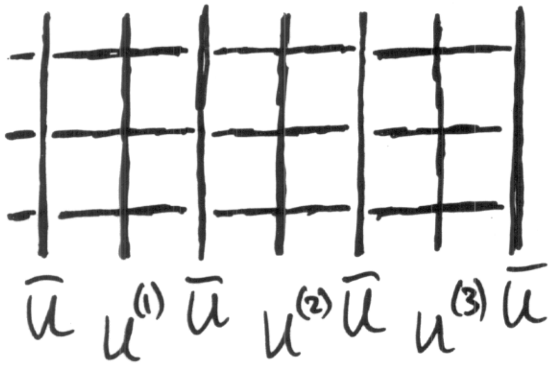
□ Polyakov loop correlation for $\beta = 5.7$, 12^4 lattice. Error bars are smaller than the plot symbols.



WHAT ABOUT DYNAMICAL FERMIONS?

- THE LÜSCHER-WEISZ SCHEME WORKS BY DIVIDING UP THE PARTITION FUNCTION...

$$Z_{YM} = \int \mathcal{D}\bar{U} \int \mathcal{D}U^{(1)} e^{-S^{(1)}[U^{(1)},\bar{U}]} \int \mathcal{D}U^{(2)} e^{-S^{(2)}} \cdots e^{-\bar{S}[\bar{U}]}$$



- THE FERMION DETERMINANT INDUCES NON-LOCAL COUPLINGS BETWEEN $U^{(a)}$ AND $U^{(b)}$.
- BUT LÜSCHER'S MULTI-BOSON METHOD TELLS US HOW TO RE-WRITE THE QCD PARTITION FUNCTION AS A LATTICE THEORY WITH LOCAL COUPLINGS

$$S_{MB} = S_g[U] + \sum_k \phi_k^* (Q-\mu_k^*)(Q-\mu_k)\phi_k$$

- THE LÜSCHER-WEISZ PARTITIONING WORKS AGAIN!

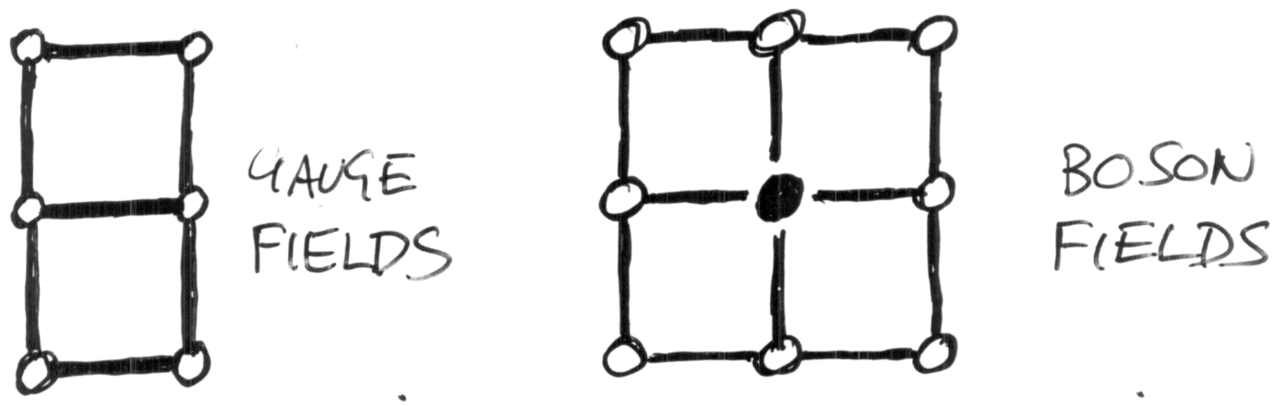
$$Z_{QCD} = \int \mathcal{D}\bar{U}\, \mathcal{D}\bar{\phi}_k\, \mathcal{D}\bar{\phi}^*_k \int \mathcal{D}U^{(1)}\, \mathcal{D}\phi^{(1)}\, \mathcal{D}\phi^{*(1)}\, e^{-S^{(1)}[U^{(1)}, \phi^{(1)}, \phi^{*(1)}; \ldots \bar{U}, \bar{\phi}, \bar{\phi}^*]} \ldots e^{-\bar{S}[\bar{U}, \bar{\phi}, \bar{\phi}^*]}$$

- WHAT ABOUT THE BOUNDARIES?
  WE HAVE AN ACTION WITH A $\phi^* Q^2 \phi$ (TWO-HOP) TERM.

- FOR WILSON PARAMETER, $r=1$. THE $\phi^* Q^2 \phi$ STRAIGHT LINE TWO-HOP HAS SPIN STRUCTURE

$$\gamma_5 (1+\gamma_\mu) \gamma_5 (1+\gamma_\mu) = 0$$

SO THE INTERACTION STENCIL FOR THE THEORY IS



AND THE SAME ONE-TIMESLICE THICK WALL WORKS AS BEFORE

- FIRST PIECE OF BAD NEWS:
  REWEIGHTING AND METROPOLIS $\equiv$ NON-LOCAL

All-to-all propagators

□ Generating dynamical configurations is (despite our best efforts!) still expensive. Clearly, we want to extract as much information as possible from our ensembles.

□ A wide variety of interesting physical observables are impractical to compute using point sources. We need the propagator from all (or many) points to all points on the lattice → stochastic estimators.

□ Examples to date come from studies of disconnected diagrams (nucleon form-factors), η', $G \leftrightarrow q\bar{q}$, string breaking.

□ Error in a stochastically estimated observable on N configurations, using M samples per configuration is

$$\sigma = \sqrt{\frac{\nu_M^2}{NM} + \frac{\nu_G^2}{N}}.$$

Diminishing returns once the two error sources are about equal.

□ Not all estimators are equal; we want to find the scheme for which the balance point is reached for the lowest cost.

□ Similarly, not all observables are equal. The optimisation must be done for each quantity we look at. We need as diverse a set of tools as we can build.

Gaussian representation

□ A gaussian representation of the fermion propagator from any point to any other can be written as

$$Q_{ij} = \int \mathcal{D}\phi\mathcal{D}\phi^* \;\; \phi_i(\phi^* Q)_j \;\; \exp\left\{-\phi^* Q^2 \phi\right\}$$

□ Advantage: Cheap local update sweeps are possible, but may suffer from critical slowing down. For many observables, measurements decorrelate quickly and variance is quickly dominated by gauge fluctuations.

See talk by A.Duncan (Mon 9:20)

□ Alternative: global heatbath step. Needs a matrix inversion so can be expensive. de Forcrand [cond-mat/9811025] suggests a trick to accelerate the process; perform inexact inversion then use Metropolis to correct.

The scheme is discussed, along with a proposal to mix gaussian and Z_N noise in the talk of W. Wilcox (Mon 9:40)

□ C. Michael & J. Peisa [hep-lat/9802015] suggest reducing the noise in the estimator by integrating out as many of the gaussian degrees of freedom as possible in disconnected sub-lattices (reminiscent of the Lüscher Weiss idea).

One restriction; the fermion source and sink can not be in the same sector so *eg.* closed fermion loops are inaccessible.

Z_N estimators

[S-J Dong & K-F Liu hep-lat/9308015]

□ A "less noisy" noise. (Usually) generates stochastic estimators of Tr $Q^{-1}\Gamma$ with lower variance than gaussian sources.

□ $Z_2 = \{-1, +1\}$. Pick elements with equal probability to fill a vector then,

$$\langle \eta_i \rangle = 0 \quad \text{and} \quad \langle \eta_i \eta_j \rangle = \delta_{ij}$$

□ For complex matrices, Z_4 is often used.

$$\langle \eta_i^* \eta_j \rangle = \delta_{ij}$$

□ After generating the source, solve

$$Q\chi = \eta$$

Then

$$\langle \eta^* \cdot \Gamma\chi \rangle = \text{Tr } Q^{-1}\Gamma$$

□ Gaussian has the advantage of the accept/reject scheme of de Forcrand. See talk by W. Wilcox (Mon 9:40)

□ Variance of these estimators can be reduced by breaking the vector space into sub-spaces, spanned by a sub-set of the basis $\{e^{(i)}\}$ (with $e_j^{(i)} = \delta_{ij}$) [SESAM hep-lat/9710050].

Thin estimators

□ The vector space can be decomposed and the trace within each sector can be estimated separately.

$$\text{Tr } Q^{-1} \approx \sum_{(i)} \langle \eta^{(i)*} Q^{-1} \eta^{(i)} \rangle$$

Illustrate with 4×4 matrix. Unthinned;

$$\eta = \left\{ \begin{pmatrix} +1 \\ +1 \\ -1 \\ +1 \end{pmatrix}, \begin{pmatrix} -1 \\ +1 \\ +1 \\ -1 \end{pmatrix}, \begin{pmatrix} +1 \\ -1 \\ +1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \\ +1 \end{pmatrix}, \dots \right\}$$

Decompose into even and odd components ...

$$\eta^{(0)} = \left\{ \begin{pmatrix} -1 \\ 0 \\ +1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \dots \right\} \eta^{(1)} = \left\{ \begin{pmatrix} 0 \\ +1 \\ 0 \\ +1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ +1 \end{pmatrix}, \dots \right\}$$

... decompose again ...

$$\eta^{(0)} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \eta^{(1)} = \begin{pmatrix} 0 \\ +1 \\ 0 \\ 0 \end{pmatrix}, \eta^{(2)} = \begin{pmatrix} 0 \\ 0 \\ +1 \\ 0 \end{pmatrix}, \eta^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ +1 \end{pmatrix}$$

... and the exact result is recovered.

□ Not surprising! The trace of an $n \times n$ matrix can be computed exactly in n matrix × vector operations.

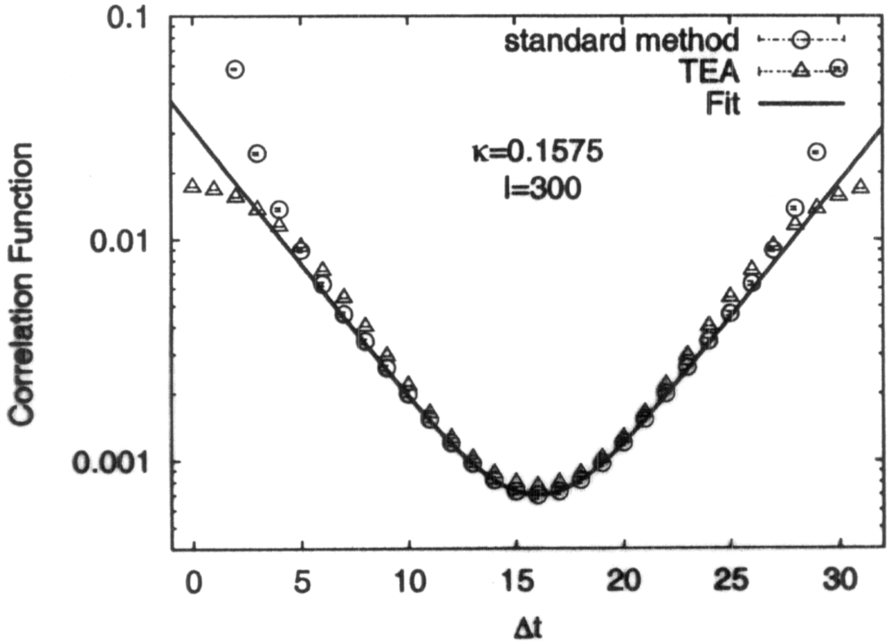
□ Unthinned; error falls off as $1/\sqrt{n}$ for n operations, so eventually, thinning MUST become more efficient.

□ Choice of how to decompose the vector space - colour, spin, even-odd, time-slice... Optimal scheme depends on observable of interest.

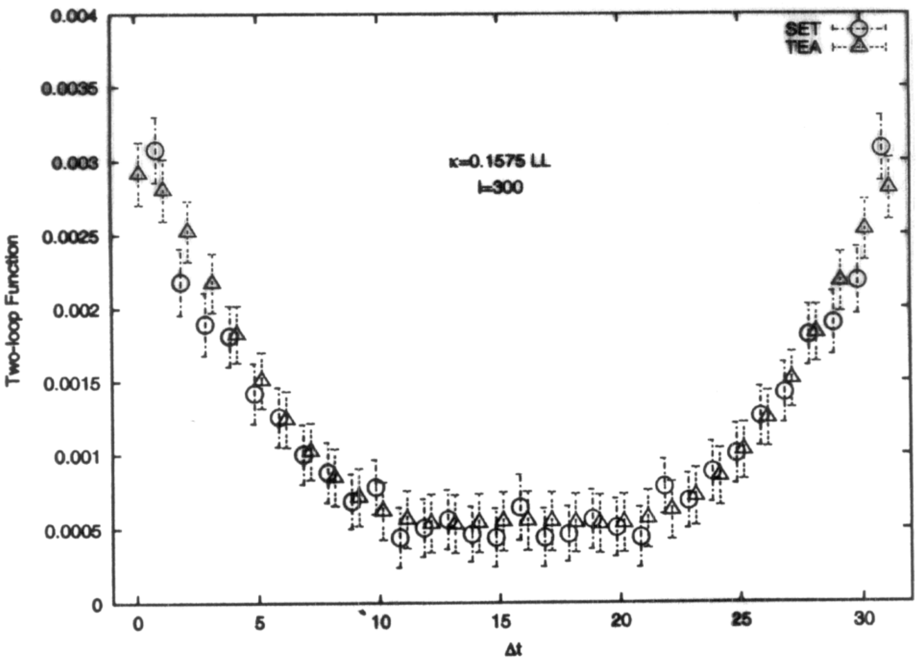
Eigenvector decomposition

[Neff et.al. (SESAM & MIT) hep-lat/0106016]

The π correlation function from 300 eigenvectors.



The disconnected correlation function.



□ Talk by K. Schilling (Sun 17:00 spectrum session).

Trinity College, Dublin.

Eigenvector decomposition

[Neff et.al. hep-lat/0106016]
See talk by H. Neff (Sun 17:20)

□ The trace of a matrix inverse can be computed if the eigenvalues, $\{\lambda_i\}$ (and eigenvectors, $\{v_i\}$) of Q are known.

$$\text{Tr } Q^{-1}\Gamma = \sum_i \frac{1}{\lambda_i} \; v_i^* \, \Gamma v_i$$

□ Sum is dominated by lowest eigenvalues/vectors. If just the lowest m eigenvectors are known, then an approximation to the trace can be constructed.

$$\text{Tr } Q^{-1}\Gamma \approx \sum_{i=1}^{m\ll n} \frac{1}{\lambda_i} \; v_i^* \, \Gamma v_i$$

□ Noisy estimators can correct the error made by truncation. Write the matrix as

$$Q = \underbrace{\sum_{i=1}^{m\ll n} \lambda_i \; v_i \, v_i^*}_{Q_{(0)}} + \underbrace{\sum_{i=m+1}^{n} \lambda_i \; v_i \, v_i^*}_{Q_{(1)}}$$

And define the (sub-space) inverses and projectors,

$$\bar{Q}_{(0)} = \sum_{i=1}^{m\ll n} \frac{1}{\lambda_i} \; v_i \, v_i^* \quad \text{and} \quad \mathcal{P}_{(0)} = \sum_{i=1}^{m\ll n} v_i \, v_i^* .$$

Then the trace we want to compute is

$$\text{Tr } Q^{-1}\Gamma = \text{Tr } \bar{Q}_{(0)}\Gamma + \text{Tr } \bar{Q}_{(1)}\Gamma$$

Eigenvector decomposition

□ Tr $\bar{Q}_{(0)}\Gamma$ is our previous low-eigenvalue approximate. The other piece can be estimated stochastically;

$$\begin{aligned} \text{Tr } \bar{Q}_{(1)}\Gamma &= \langle \eta^* \Gamma \bar{Q}_{(1)} \eta \rangle \\ &= \langle \eta^* \Gamma (Q^{-1} - \bar{Q}_{(0)}) \eta \rangle \\ &= \langle \eta^* \Gamma Q^{-1} (I - \mathcal{P}_{(0)}) \eta \rangle \end{aligned}$$

□ The projection is just an orthogonalisation of η with respect to all the m known eigenvectors.

□ Matrix inversion is accelerated, since the low-lying eigenvectors now do not appear in $(I - \mathcal{P}_{(0)})\eta$.

□ This is an efficient example of a subtraction scheme.

□ Thinning of the stochastic estimator, previously often slow to converge due to low-lying modes may well be accelerated.

□ Multiple RHS solvers could be useful for stochastic estimators. See the poster by W. Wilcox on deflated GMRES, which combines multiple RHS with eigenvector calculations.

Summary

1. Current techniques and comparisons.
    - Multi-Boson methods have advanced significantly.
    - TSMB and UV-filtered MB are competing with HMC.
2. Odd flavour Wilson fermion simulations.
    - Polynomial approximations lead to useful techniques for one-flavour simulations.
3. HMC algorithm "plug-ins"
    - Many ways to improve the standard algorithm.
    - Compatible "plug-ins" could lead to an order-of-magnitude improvement.
4. Exponential error reduction for Yang-Mills
    - New scheme for computing large Wilson, Polyakov loops with orders-of-magnitude higher precision.
5. All-to-all propagators.
    - A wide variety of competing and complementary methods.
    - Thin noise should be explored more fully.
    - Eigenvalue decompositions can be combined with stochastic estimators.