Algorithms for lattice QCD IV

Stefan Schaefer

NIC, DESY

December 5th, 2013

Solving the Dirac equation

Classical iterative methods are based on contructing solution of the Dirac equation

$$\boldsymbol{D}\,\psi = \phi$$

in the Krylov space

$$\mathcal{K}_n = \mathbf{span}\{\phi, \boldsymbol{D}\phi, \ldots, \boldsymbol{D}^n\phi\}$$

They tend to converge with an exponential rate on the scale of the inverse condition number.

Need of deflating the system at small quark masses.

Global deflation works, but is prohibitivly expensive on large systems.

Exact deflation with eigenvectors

Elimitate these eigenmodes from the Dirac equation.

$$D\,\psi_i = \lambda_i\,\psi_i$$

Projector on small eigenmodes ψ_i

$$P = \sum_{i=1}^{N_s} \psi_i\,\psi_i^\dagger$$

Using it, we can split the Dirac equation in two

$$D = \begin{pmatrix} P\,D\,P & 0 \\ 0 & (1-P)D(1-P) \end{pmatrix}$$

$$\Rightarrow D^{-1} = \begin{pmatrix} \sum_{i=1}^{N_s} \frac{1}{\lambda_i}\psi_i\,\psi_i^\dagger & 0 \\ 0 & [(1-P)D(1-P)]^{-1} \end{pmatrix}$$

Local deflation

The problem of "classical" deflation is the scaling with the volume.
Need $N_s \propto V$ modes w/ cost/mode at least $\propto V$.

Local coherence

Lüscher'07

Experimental fact:
Locally eigenvectors with $\lambda < 100$ MeV can be constructed from very few components.

Procedure:
Take N_0 lowest eigenmodes.

Decompose the lattice in small blocks Λ_i, e.g., $(0.3\,\mathrm{fm})^4$

Consider space spanned by block projected vectors.

$$\mathcal{R} = \mathrm{span}\{P_{\Lambda_i}\psi_j |\ i = 1,\dots,N_{\mathrm{block}},\ j = 1,\dots,N_0\}$$

Deflation subspace

$$\mathcal{R} = \mathrm{span}\{P_{\Lambda_i}\psi_j |\, i = 1, \ldots, N_{\mathrm{block}},\, j = 1, \ldots, N_0\}$$

Define deficit

$$\epsilon = |P_{\mathcal{R}}\psi_i - \psi_i|$$

with $P_{\mathcal{R}}$ the orthonormal projector to $\mathcal{R}$.

Experimental finding:

The deficit for eigenvectors ψ_i with eigenvalue ≤ 100 MeV is small, $N_0 \sim 12$.

$$\epsilon \approx \text{O(few \%)}$$

Local coherence

This result can be interpreted as consequence of local coherence.

In each point, the IR fields are aligned.

However, the vectors in $\mathcal{R}$ are quite discontinuous.

They can only be decent approximations to the eigenvectors in the centers of the blocks.

Big advantage is that size of deflation space $\propto$ volume.

Eigenvectors do not need to be very exact.
A few inverse iterations suffice.

Implementation in a solver

Decomposition of the Dirac operator

$$D = \begin{pmatrix} (1-P_{\mathcal{R}})\,D\,(1-P_{\mathcal{R}}) & (1-P_{\mathcal{R}})DP_{\mathcal{R}} \\ P_{\mathcal{R}}\,D\,(1-P_{\mathcal{R}}) & P_{\mathcal{R}}\,D\,P_{\mathcal{R}} \end{pmatrix}$$

with the "little Dirac operator"

$$D_{LL} = P_{\mathcal{R}}\,D\,P_{\mathcal{R}}$$

This is a $(N_s N_{\mathrm{block}})^2$ matrix.
Using the usual Schur complement trick

$$\begin{pmatrix} D_{HH} & D_{HL} \\ D_{LH} & D_{LL} \end{pmatrix}^{-1} =$$

$$\begin{pmatrix} 1 & 0 \\ -D_{LL}^{-1}D_{LH} & 1 \end{pmatrix} \begin{pmatrix} (D_{HH} - D_{HL}D_{LL}^{-1}D_{LH})^{-1} & 0 \\ 0 & D_{LL}^{-1} \end{pmatrix} \begin{pmatrix} 1 & -D_{HL}D_{LL}^{-1} \\ 0 & 1 \end{pmatrix}$$

Deflating the Dirac equation

The Schur complement trick reduces the problem to the solution of

$$D_{LL}\psi_{\parallel} = \phi_{\parallel}$$
$$(D_{HH} - D_{HL}D_{LL}^{-1}D_{LH})\psi_{\perp} = \phi_{\perp}$$

The condition number of the matrix in the second equation is significantly reduced.

Rewrite 2nd eq. in form of preconditioning

$$(1 - DP_{\mathcal{R}}(P_{\mathcal{R}}DP_{\mathcal{R}})^{-1}P_{R})\, D\, \psi_{\perp} = \phi_{\perp}$$

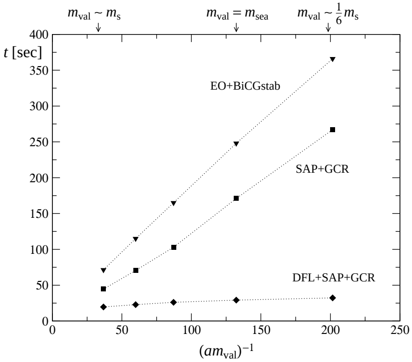
Can be solved with a GCR, but this is still expensive due to solution of the little system.

Still need a good preconditioner to make it feasible
Needs to be effective in the UV
$\rightarrow$ Schwarz alternating procedure

Performance of the deflated GCR

Plot from original paper

M. Lüscher, Local coherence and deflation of the low quark modes in lattice QCD, JHEP0707:081,2007

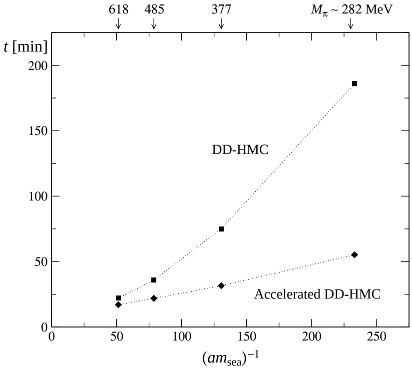


Deflation and the HMC

The construction of the deflation subspace is not cheap.

The more solutions of the Dirac equation, the more it pays.
$\rightarrow$ good for Hasenbusch decomposition

Low-modes evolve slowly in MC time
$\rightarrow$ take subspace in several consecutive time step.



HMC

Momentum and pseudofermion Heatbath

Refresh momenta π
Refresh pseudofermions $\phi \rightarrow$ kept fixed during trajectory
Initialization of deflation subspace

Molecular Dynamics

Solve numerically MD equations for some MC time τ.

Repeated refresh of deflation subspace.

Acceptance Step

Correcting for inaccuracies in integration.

Need to be careful for violation of reversibility.

Commercial

What is presented here is implemented in the publicly available openQCD code.
`http://cern.ch/luscher/openQCD`

Action: fermions

(Un)-improved Wilson fermions

Almost any number of flavors

Twisted mass fermions

limited support (no even-odd preconditioning)

Deflated solver not made for maximal twist.

Action: gauge fields

plaquette and 1×2 rectangles

Commercial (cont.)

Boundary conditions

Periodic b.c. in space

Open or SF b.c. in temporal direction

The code is very flexible:

Action defined in input file

Multiple time-scale integration scheme

Online measurements of gluonic observables.

Easy to extend

Summary

Last decade has seen enormous progress in algorithms.

Starting point were standard, all-purpose techniques.

Physics driven ideas: frequency splitting, local deflation,. . .

Progress matches the development of computer hardware.

Measuring hadronic observables

The goal is to compute hadronic correlation functions on a set of gauge configurations, e.g.

$$\langle P^a(x)\, P^b(y)\rangle$$

with

$$P^a = \overline{\psi}\frac{1}{2}\tau^a\gamma_5\psi \qquad \text{and} \qquad \psi = \begin{pmatrix} u \\ d \end{pmatrix}$$

Use Wick's theorem to eliminate the Grassmann fields

$$\langle P^a(x)\, P^b(y)\rangle = -\frac{1}{2}\delta^{ab}\,\langle \mathrm{tr} S(x,y) S(y,x)^\dagger \rangle$$

where the γ_5 Hermiticity of the Dirac operator has been used

$$S(x,y) = D^{-1}(x,y) = \gamma_5 S(y,x)^\dagger \gamma_5$$

Practical computation

Traditional method:

$$\langle \mathrm{tr}_{d,c} S(x,y) S(y;x)^\dagger \rangle = \sum_{c,d,c',d'} \langle S_{c,d;c',d'}(x,y) S(y;x)^\dagger_{c',d';c,d} \rangle$$

For this one space-time column of the propagator is needed.

$$S_{c,d;c',d'}(x,y) = (S\,\eta^{(y,c',d')})(x)_{c,d}$$

with a point source

$$\eta^{(x_0,c_0,d_0)}_{c,d}(x) = \delta_{x,x_0}\delta_{c,c_0}\delta_{d,d_0}$$

Solve Dirac equation for the 4×3 Dirac–color index combinations

$$D\,\phi = \eta^{(y,c,d)}$$

Get propagator from one point to *all* other points.

Volume average

Pion propagator projected on zero momentum

$$C_{PP}(x_0 - y_0) = -\frac{1}{V} \sum_{\vec{x}} \sum_{\vec{y}} \langle \mathrm{tr}_{d,c} S(x,y) S(y;x)^{\dagger} \rangle$$

Using point sources, the sum over y is difficult to do, would need $12V$ solutions of the Dirac equation.

Translational invariance helps, need sum only at one end.

Still need $\mathrm{O}(L^3)$ inversions in large volume to fully sample information.

Use a stochastic estimate for the traces.

Noise sources

Insert additional complex scalar fields into your partition function.
Here just for one time slice y_0; "one-end trick".

$$Z_\eta = \int [d\eta][d\eta^\dagger]\, e^{-(\eta,\eta)}$$

Each lattice point, Dirac and color index has an independent Gaussian random number

$$\langle \eta_{c,d}(\vec{x})\, \eta^\dagger_{c',d'}(\vec{y}) \rangle_\eta = \delta_{\vec{x},\vec{y}}\, \delta_{d,d'}\, \delta_{c,c'}$$

Insert in correlation function

$$C_{PP}(x_0 - y_0) = -\frac{1}{V} \sum_{\vec{x}} \langle \mathrm{tr}_{d,c} S(x,\cdot)\eta\eta^\dagger S(\cdot;x)^\dagger \rangle$$

here the $\langle\cdot\rangle$ includes average over η fields.

Stochastic estimate

$$C_{PP}(x_0 - y_0) = -\frac{1}{V}\sum_{\vec{x}}\langle \mathbf{tr}_{d,c}\boldsymbol{S}(x,\cdot)\eta\eta^{\dagger}\boldsymbol{S}(\cdot;x)^{\dagger}\rangle$$

As always in Monte Carlo, we replace integrals by a sum over a number of field realizations.

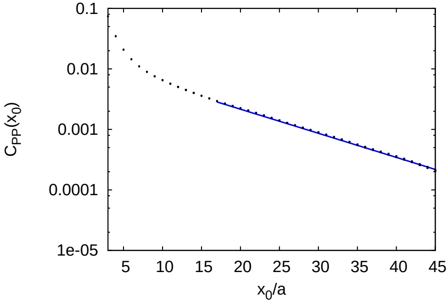
$$-\frac{1}{V}\frac{1}{N_s}\sum_{i=1}^{N_s}\sum_{\vec{x}}\langle \mathbf{tr}_{d,c}\boldsymbol{S}(x,\cdot)\eta_i\eta_i^{\dagger}\boldsymbol{S}(\cdot;x)^{\dagger}\rangle$$

Unbiased estimator, no need to take N_s large.
Also $N_s = 1$ is correct, but take new source on each configuration. Integrals commute.

Need to solve one Dirac equation per source.
For pions $\mathrm{O}(10)$ give a very good signal.

For mesons, no need to use more sources as $V \to \infty$.

Pion propagator



Exponential fall-off for $x_0 \to \infty$

$$C_{PP}(x_0) = \sum_n A_n e^{-m_n x_0} \to A_0 e^{-m_\pi x_0}$$

Source couples to all states with given quantum numbers.
Excited states clearly visible at small x_0.

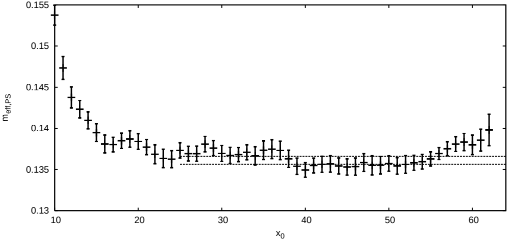
Effective mass

Since exponential fall-off is difficult to judge, one typically looks at quantities that show a plateau.

$$\frac{C(x_0)}{C(x_0+a)} = \frac{A\,e^{-mx_0}}{A\,e^{-m(x_0+a)}} = e^{am}$$

Effective mass

$$am_{\text{eff}} = \log \frac{C(x_0)}{C(x_0+a)}$$



Statistical error

Reminder:
The square of the error of a measuremnt is proportional to the variance of the observable

$$\sigma^2(A) = \langle A^2 \rangle - \langle A \rangle^2$$

Parisi'83

The variance is a physical observable, the exponential fall-off can be predicted.

$$\underbrace{\langle A \rangle = \langle P^a(x) P^b(y) \rangle}_{\propto e^{-E_\pi |x-y|}} \quad \rightarrow \quad \sigma^2(A) \quad \underbrace{= \langle P^a(x) P^a(x) P^b(y) P^b(y) \rangle - \langle A \rangle^2}_{\propto e^{-E_{2\pi} |x-y|}}$$

In large volume, $E_{2\pi} = 2m_\pi = 2E_\pi$
Constant signal-to-noise ratio

$$\frac{\langle A \rangle}{\sigma(A)} \propto \frac{e^{-m_\pi |x-y|}}{\sqrt{e^{-2m_\pi |x-y|}}} = \text{const.}$$

Signal-to-noise problem

For the nucleon, one considers

$$\langle A \rangle = \Gamma_{\alpha\beta} \langle N_\alpha(x) \bar{N}_\beta(y) \rangle \propto e^{-E_N |x-y|}$$

Variance

$$\langle A^2 \rangle - \langle A \rangle^2 \text{``} = \text{``} \langle N(x) \bar{N}(x)\, N(y) \bar{N}(y) \rangle - (\langle N(x)\, \bar{N}(y) \rangle)^2 \rangle \propto e^{-E_{3\pi} |x-y|}$$

Matches quantum numbers of three pions and therefore the signal-to-noise ratio is

$$\frac{\langle A \rangle}{\sigma(A)} = e^{-(m_N - \frac{3}{2} m_\pi)|x-y|}$$

Exponential reduction once $m_N > \frac{3}{2} m_\pi$.

Makes calculations of proton properties exceedingly difficult.

Summary fermions

Most effort goes into fermions.

Deflation of Dirac equation brought great progress.
Is there even more possible?

Computation of PS meson two-point functions well-established.

Significant challenges in baryon sector.