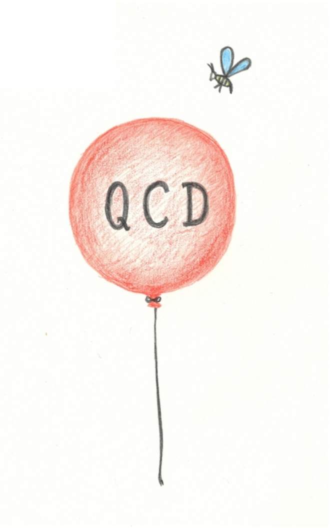
Deflating QCD

Physical Concepts & Algorithms

Martin Lüscher, CERN Physics Department



Motivation and general framework

The basic strategy adopted here is to

- ⋆ Choose a simple lattice formulation of QCD
- ⋆ Simulate large lattices: 64×32^3, 96×48^3, $\ldots$
- ⋆ Try to "teach physics to the algorithms"

Low-mode deflation is an important case where physics insight is used to accelerate the computations

Neff et al. '01, Giusti et al. '03f, DeGrand & Schaefer '04

Bali et al. '05, Foley et al. '05, ML '07

Low quark modes, a source of difficulty

There is a hierarchy of scales

$$m_u, m_d \ll M_\pi \ll 4\pi F_\pi$$

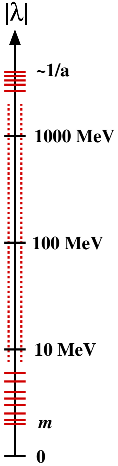
linked to the spontaneous breaking of chiral symmetry

Leutwyler '74

⇒ *the condition number of the lattice Dirac operator D is huge*

⇒ *the numerical solution of $D\psi = \eta$ tends to be very expensive*

⇒ *so are the numerical simulations*



Moreover, the low modes of the Dirac operator are dense

$$\Delta\lambda \sim \frac{1}{\Sigma V}, \qquad \Sigma = \lim_{m\to 0}\lim_{V\to\infty} |\langle \bar{u}u \rangle|, \qquad V\text{: lattice volume}$$

Banks & Casher '80, Leutwyler & Smilga '92

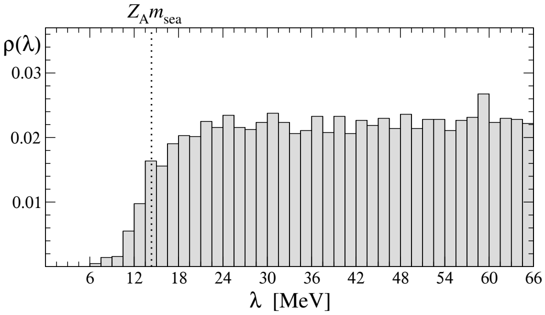
Example:

Spectrum of $(D^\dagger D)^{1/2}$ in two-flavour QCD

48×24^3 lattice

$a = 0.08$ fm

$m_{\text{sea}} \sim \frac{1}{4} m_{\text{s}}$



According to Banks & Casher, the number of low modes is

$$N_{|\lambda|\leq M} = \frac{2}{\pi} M\Sigma V + \mathrm{O}(M^2)$$

For illustration, set $M = 100$ MeV, $\Sigma = (250\ \text{MeV})^3$, $V = 2L^4$

$$\Rightarrow \quad N_{|\lambda|\leq M} = \begin{cases} 21 & L = 2\ \text{fm} \\ 107 & L = 3\ \text{fm} \\ 338 & L = 4\ \text{fm} \end{cases}$$

- ⋆ Actual numbers are ~ 1.4 times larger
- ⋆ Practically continuous spectrum
  $\Rightarrow$ individual modes do not matter so much!

Textbook deflation

Deflation $\equiv$ separation and special treatment of the low modes

Example:

Let P be the projector to the N lowest modes of $A = D^\dagger D$

$$A = PAP + (1-P)A(1-P) \equiv A_\parallel + A_\perp$$

$$A\psi = \eta \quad \Leftrightarrow \quad \begin{cases} A_\parallel \psi_\parallel = \eta_\parallel & \text{“little system”} \\ A_\perp \psi_\perp = \eta_\perp & \text{“deflated system”} \end{cases}$$

- *The little system can be solved directly*
- *If enough modes are projected away, the deflated system is well conditioned and can be solved rapidly using the CG algorithm*

A recent proposition is to compute the eigenvectors "on the fly",
using the Krylov space implicitly constructed by the solver

Stathopoulos & Orginos '07, Morgan & Wilcox '07

However

- ⋆ Accurate computations of eigenvectors tend to be very expensive
- ⋆ In the large-volume regime of QCD, there are O(V) low modes
  $\Rightarrow$ deflation overhead grows like V^2

Textbook deflation is therefore not a viable option on large lattices

Outline

1. Deflation without eigenvectors
2. Domain-decomposed deflation projectors
3. Applications

Let P be a projector to any subspace of quark fields. Define

PDP : little Dirac operator

$$\underbrace{\left\{1 - DP(PDP)^{-1}P\right\}}_{\text{oblique projector } P_L} D \equiv \hat{D} : \text{ deflated operator}$$

The solution of the Dirac equation $D\psi = \eta$ then splits into two parts

$$\psi(x) = \chi(x) + (PDP)^{-1}P\eta(x)$$

where

$$\hat{D}\chi(x) = P_L\eta(x), \qquad \underbrace{\left\{1 - P(PDP)^{-1}PD\right\}}_{\text{oblique projector } P_R} \chi(x) = \chi(x)$$

The condition number of $\hat{D}$

$$\|\hat{D}\|\,\|\hat{D}^{-1}\| \simeq \|D\|\,\|\hat{D}^{-1}\|$$

$$= \|D\|\,\|(1-P)D^{-1}(1-P)\|$$

$$\leq \|D\|\,\|(1-P)(D^{\dagger}D)^{-1}(1-P)\|^{1/2}$$

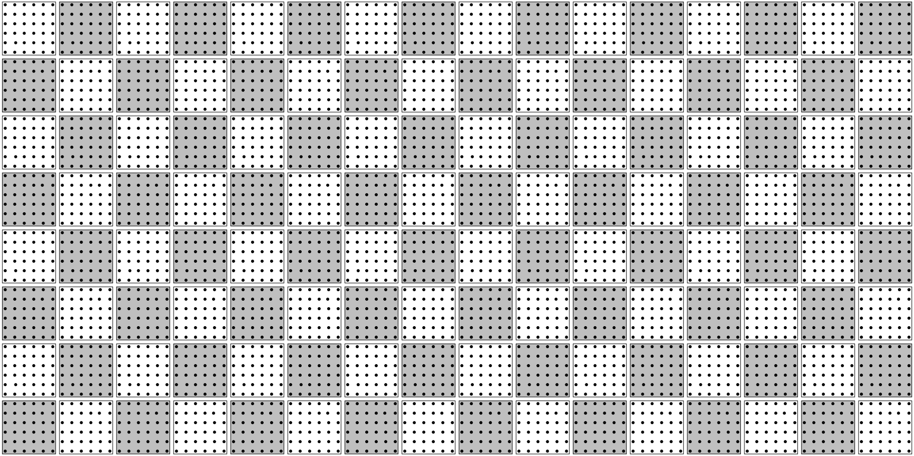
will be much smaller than the one of D if

$$\|(1-P)\psi\|^2 \ll \|\psi\|^2$$

for all low modes ψ of $D^{\dagger}D$

Such projectors P can be constructed w/o knowing the low modes!

Domain-decomposed deflation projectors



96×48^3 lattice, divided into 8192 domains of size 6^4

Try to approximate the low modes in each domain separately, i.e.

$$P = \sum_{\text{blocks } \Lambda} P_\Lambda, \qquad P_\Lambda \psi(x) = \sum_{l=1}^{N_s} \phi_l^\Lambda(x) \left(\phi_l^\Lambda, \psi\right)$$

$\uparrow$ orthonormal, supported on Λ

where N_s does not need be scaled with V

Somewhat surprisingly, this is in fact possible!

Note that the application of P requires only $N_s V$ operations
$\Rightarrow$ *V^2-problem will be solved in this way!*

In the free quark theory, choosing constant block modes works well

Lattice size L, block size b

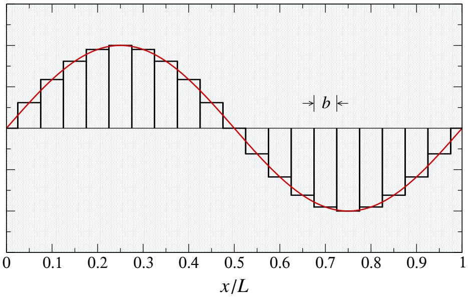
Low modes

$$\psi_p(x) = u_p \, \mathrm{e}^{ipx}$$

Deflation deficits

$$\|(1-P)\psi_p\|^2 = \epsilon_p \|\psi_p\|^2$$

$$\epsilon_p = \tfrac{1}{12} p^2 \left(b^2 - a^2\right) + \mathrm{O}(p^4 b^4)$$

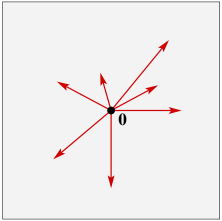


- ⋆ Deficits are independent of V
- ⋆ Deflation modes are far from being eigenvectors of D

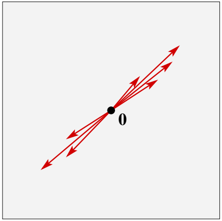
Application to (classical) PDE solvers: Frank & Vuik '01, Nabben & Vuik '06

The block projection can only work out if the low modes of $D^\dagger D$ are locally coherent

That is, on each block Λ, they must align to a low-dimensional subspace of fields



incoherent



coherent

Numerical experiments show this to be so, for every individual gauge-field configuration!

Example

64×32^3 lattice, $a = 0.08$ fm, $m_{\rm sea} \simeq \frac{1}{4} m_{\rm s}$

Block size 4^4, $N_s = 12$ modes per block

$\phi_l^{\Lambda}(x) =$ lowest 12 eigenmodes of $D^\dagger D$ projected to Λ

$\Rightarrow$ The next 36 eigenmodes have deflation deficits $\epsilon = 3 - 6\%$

In fact, any quark field $\psi(x)$ that satisfies

$$\|D\psi\| \leq M\|\psi\|, \qquad M \sim 100 \text{ MeV}$$

is locally coherent with the low modes

⇒ **The projectors P_Λ can be constructed using N_s such modes instead of exact eigenmodes!**

Can use approximate inverse iteration to generate them rapidly

Start: $\psi =$ random

Iterate: $\psi \rightarrow$ "D^{-1}" ψ
$\uparrow$
MR, SAP, . . .

See JHEP 0707 (2007) 081 [arXiv:0706.2298] and arXiv:0710.5417 for details

$\Rightarrow$ Recipe for the generation of domain-decomposed deflation projectors

- *Choose* N_s *random quark fields* $\psi_1, \ldots, \psi_{N_s}$
- *Apply* ~ 10 *approximate inverse iteration steps to them*
- *On each block* Λ, *set* $\phi_l^{\Lambda}(x) = \psi_l(x)|_{x \in \Lambda}$ *and orthonormalize*

The projector is then given by

$$P = \sum_{\text{blocks } \Lambda} \left\{ \sum_{l=1}^{N_s} \phi_l^{\Lambda} \otimes (\phi_l^{\Lambda})^{\dagger} \right\}$$

Applications

√ Quark propagator computations

√ Acceleration of the HMC algorithm

? Low-mode averaging ("all-to-all propagators")

Giusti et al. 04, DeGrand & Schaefer '04, Bali et al. 05, Foley et al. 05

Quark propagator computations

The Dirac equation $D\psi = \eta$ must typically be solved very many times in this case

Ideal for deflation, because the same projector P can be used for all source fields η

However, it is advisable to combine deflation with a preconditioner

$$\psi(x) = \chi(x) + (PDP)^{-1} P\eta(x) \quad \text{(as before)}$$

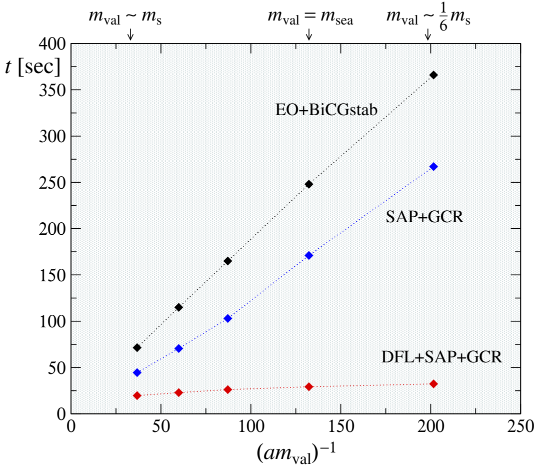
$$P_L D M \phi(x) = P_L \eta(x), \quad \chi(x) = P_R M \phi(x) \quad (M\text{: preconditioner})$$

Required time/solution on a 64×32^3 lattice, $a \simeq 0.08$ fm, $m_{\rm sea} \sim \frac{1}{4} m_{\rm s}$

EO: even-odd preconditioning

SAP: multiplicative Schwarz preconditioner

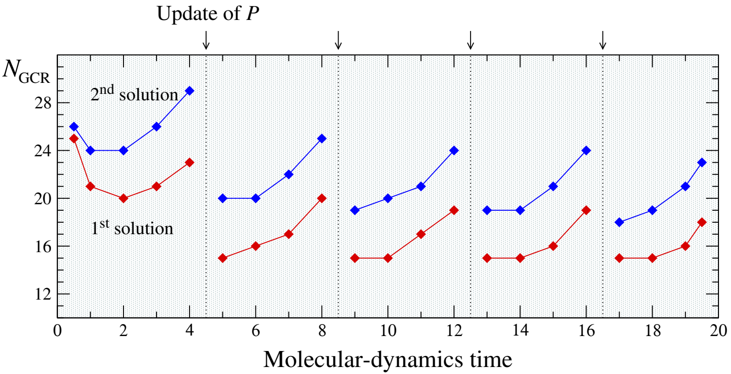
Timings taken on a PC cluster with 64 CPUs



Acceleration of the HMC algorithm

The deflation efficiency of P can be maintained along the MD trajectories by applying an inverse iteration step from time to time

Typical history of the deflated solver iteration numbers along a trajectory

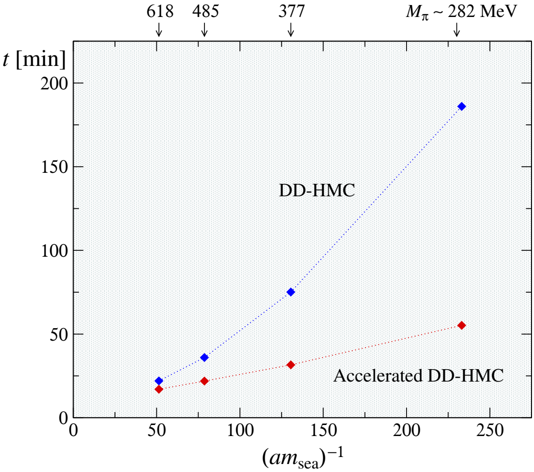


Required time/trajectory on a 64×32^3 lattice, $a \simeq 0.08$ fm

Two-flavour QCD,
n.p. O(a)-improved

DD-HMC:
Domain Decomposition
HMC algorithm

Timings taken on a PC
cluster with 64 CPUs



Conclusions

- ⋆ Deflation works very well in lattice QCD
  $\Rightarrow$ flat quark mass dependence

- ⋆ Domain-decomposed projectors provide a solution to the V^2-problem

- ⋆ Variance-reduction methods using these remain to be explored