Deflation and polynomial preconditioning in the application of the overlap operator at nonzero chemical potential

Lattice Conference, July 2024, Liverpool

Gustavo Ramirez-Hidalgo
(Bergische Universität Wuppertal)
&
Jacob Finkenrath
(CERN)

FOR 5269

BERGISCHE UNIVERSITÄT WUPPERTAL

Table of Contents

The overlap discretization

- keeps chiral symmetry when moving to the lattice

The overlap discretization

- keeps chiral symmetry when moving to the lattice
- take D_w to be the usual Wilson-Dirac operator

The overlap discretization

- keeps chiral symmetry when moving to the lattice
- take D_w to be the usual Wilson-Dirac operator
- then, the overlap operator is: $D_N = \rho I + \Gamma_5 \text{sign}\left(Q(m_0^{ker})\right)$, with $\text{sign}(A) = A(A^2)^{-1/2}$ and $Q(m_0^{ker}) = \Gamma_5 D_w(m_0^{ker})$

The overlap discretization

- keep chiral symmetry when moving to the lattice
- take D_w to be the usual Wilson-Dirac operator
- then, the overlap operator is: $D_N = \rho I + \Gamma_5 \mathrm{sign}\left(Q(m_0^{ker})\right)$, with $\mathrm{sign}(A) = A(A^2)^{-1/2}$ and $Q(m_0^{ker}) = \Gamma_5 D_w(m_0^{ker})$
- we further equip here the Dirac operator with a chemical potential i.e. $D_w(\mu, m_0^{ker})$, and then $Q_\mu^H = Q_{-\mu}$

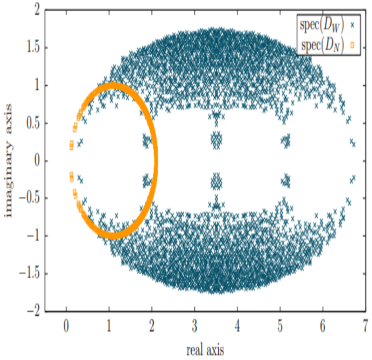
The overlap discretization

- keeps chiral symmetry when moving to the lattice
- take D_w to be the usual Wilson-Dirac operator
- then, the overlap operator is: $D_N = \rho I + \Gamma_5 \mathrm{sign}\left(Q(m_0^{ker})\right)$, with $\mathrm{sign}(A) = A(A^2)^{-1/2}$ and $Q(m_0^{ker}) = \Gamma_5 D_w(m_0^{ker})$
- we further equip here the Dirac operator with a chemical potential i.e. $D_w(\mu, m_0^{ker})$, and then $Q_\mu^H = Q_{-\mu}$
- **task**: solve linear systems with D_N

The overlap discretization

- keeps chiral symmetry when moving to the lattice
- take D_w to be the usual Wilson-Dirac operator
- then, the overlap operator is: $D_N = \rho I + \Gamma_5 \text{sign}\left(Q(m_0^{ker})\right)$, with $\text{sign}(A) = A(A^2)^{-1/2}$ and $Q(m_0^{ker}) = \Gamma_5 D_w(m_0^{ker})$
- we further equip here the Dirac operator with a chemical potential i.e. $D_w(\mu, m_0^{ker})$, and then $Q_\mu^H = Q_{-\mu}$
- **task**: solve linear systems with D_N

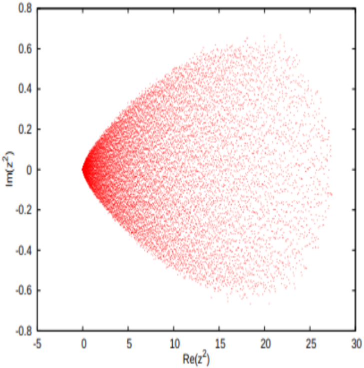
- on a 4^4 lattice (no chemical potential):



The overlap discretization

- keeps chiral symmetry when moving to the lattice
- take D_w to be the usual Wilson-Dirac operator
- then, the overlap operator is: $D_N = \rho I + \Gamma_5 \mathrm{sign}\left(Q(m_0^{ker})\right)$, with $\mathrm{sign}(A) = A(A^2)^{-1/2}$ and $Q(m_0^{ker}) = \Gamma_5 D_w(m_0^{ker})$
- we further equip here the Dirac operator with a chemical potential i.e. $D_w(\mu, m_0^{ker})$, and then $Q_\mu^H = Q_{-\mu}$
- **task**: solve linear systems with D_N

- on a 6^4 lattice ($\mu = 0.3, \beta_g = 5.1$), spectrum of $Q_\mu^2(m_0^{ker})$:



Preconditioning at the solver level

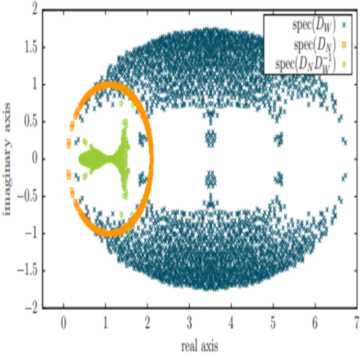
- a first approach at boosting these computations is preconditioning at the level of D_N, i.e. solving linear systems with $D_N \cdot \mathrm{AMG}(D_w(m_0^{prec}), \epsilon)$

Preconditioning at the solver level

- a first approach at boosting these computations is preconditioning at the level of D_N, i.e. solving linear systems with $D_N \cdot \mathrm{AMG}(D_w(m_0^{prec}), \epsilon)$
- motivation for this: if we assume that $D_w(m_0^{prec})$ is normal, then: $\mathrm{spec}(D_N(D_w(m_0^{prec}))^{-1}) = \left\{ \frac{\rho + \mathrm{csign}(\lambda + m_0^{ker})}{\lambda + m_0^{prec}},\ \lambda \in \mathrm{spec}(D_w(0)) \right\}$, and there is an analytic expression for m_0^{prec} which is quite close to optimal

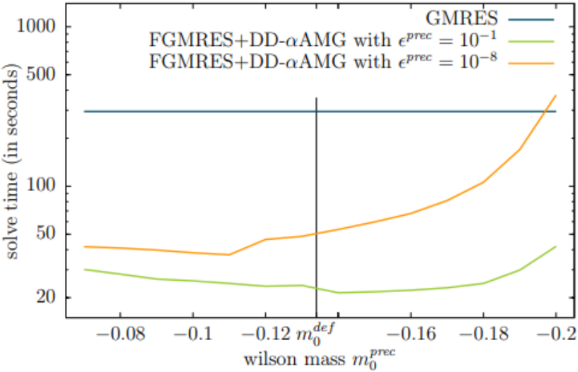
Preconditioning at the solver level

- a first approach at boosting these computations is preconditioning at the level of D_N, i.e. solving linear systems with $D_N \cdot \mathrm{AMG}(D_w(m_0^{prec}), \epsilon)$
- motivation for this: if we assume that $D_w(m_0^{prec})$ is normal, then: $\mathrm{spec}(D_N(D_w(m_0^{prec}))^{-1}) = \left\{ \frac{\rho + \mathrm{csign}(\lambda + m_0^{ker})}{\lambda + m_0^{prec}},\ \lambda \in \mathrm{spec}(D_w(0)) \right\}$, and there is an analytic expression for m_0^{prec} which is quite close to optimal
- on a 4^4 lattice (no chemical potential):



Preconditioning at the solver level

Impact of preconditioning, at the level of D_N, via $\mathrm{AMG}(D_w(m_0^{prec}), \epsilon)$ (no chemical potential, 32×32^3, smeared, work by Brannick, Frommer, Kahl, Leder, Rottmann, Strebel - 2014):



Preconditioning at the sign-function level

But, the sign function still has to be applied and that is the most expensive part.

Preconditioning at the sign-function level

But, the sign function still has to be applied and that is the most expensive part.

In recent work (Frommer, R-H, Schweitzer, Tsolakis - 2024), we have introduced polynomial preconditioning for the inverse square root:

Preconditioning at the sign-function level

But, the sign function still has to be applied and that is the most expensive part.

In recent work (Frommer, R-H, Schweitzer, Tsolakis - 2024), we have introduced polynomial preconditioning for the inverse square root:

$$A^{-1/2}b = q(A)(A(q(A))^2)^{-1/2}b \text{ with } q(A) \approx A^{-1/2}.$$

Preconditioning at the sign-function level

But, the sign function still has to be applied and that is the most expensive part.

In recent work (Frommer, R-H, Schweitzer, Tsolakis - 2024), we have introduced polynomial preconditioning for the inverse square root:

$$A^{-1/2}b = q(A)(A(q(A))^2)^{-1/2}b \text{ with } q(A) \approx A^{-1/2}.$$

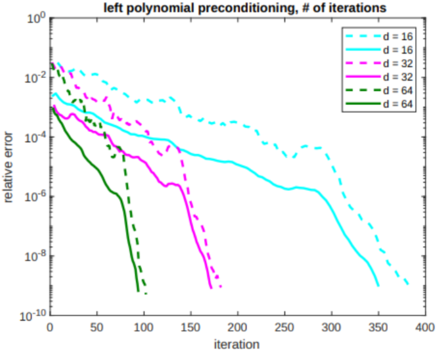
In the LQCD context, this means:

$$(Q_\mu^2)^{-1/2}b = q(Q_\mu^2)(Q_\mu^2(q(Q_\mu^2))^2)^{-1/2}b \text{ with } q(Q_\mu^2) \approx (Q_\mu^2)^{-1/2}$$

(note how Arnoldi is done with Q_μ^2).

Preconditioning at the sign-function level

With a chemical potential $\mu = 0.3$, 64×32^3, non-smeared, d is the degree of the polynomial, the **dashed** lines are a cheap approximation of the error. Tolerances are 10^{-5} for the table and 10^{-9} for the figure.



d	iterations	mvms	inner products	time 64 nodes (in s)	time 256 nodes (in s)
1	1 600	3 200	1 279 200	127.8	105.8
8	296	8 910	42 510	25.7	9.8
16	140	8 742	9 991	12.3	7.8
32	72	9 198	3 125	11.6	7.4
64	33	8 636	2 578	10.6	5.5

Preconditioning at the sign-function level

A short note on LR deflation for the sign function (Bloch, Frommer, Lang, Wettig - 2007):

- let: $AR_m = R_m \Lambda_m$

Preconditioning at the sign-function level

A short note on LR deflation for the sign function (Bloch, Frommer, Lang, Wettig - 2007):

- let: $AR_m = R_m \Lambda_m$
- let: $L_m^H A = \Lambda_m L_m^H$

Preconditioning at the sign-function level

A short note on LR deflation for the sign function (Bloch, Frommer, Lang, Wettig - 2007):

- let: $AR_m = R_m \Lambda_m$
- let: $L_m^H A = \Lambda_m L_m^H$
- then: $f(A)b = f(A)(I - R_m L_m^H)b + f(A)R_m L_m^H b$ (note: $f(A)R_m L_m^H b = R_m f(\Lambda_m) L_m^H b$)

Preconditioning at the sign-function level

A short note on LR deflation for the sign function (Bloch, Frommer, Lang, Wettig - 2007):

- let: $AR_m = R_m \Lambda_m$
- let: $L_m^H A = \Lambda_m L_m^H$
- then: $f(A)b = f(A)(I - R_m L_m^H)b + f(A) R_m L_m^H b$ (note: $f(A) R_m L_m^H b = R_m f(\Lambda_m) L_m^H b$)

- main problem with this approach: eigensolving can be extremelly expensive

Preconditioning at the sign-function level

A short note on LR deflation for the sign function (Bloch, Frommer, Lang, Wettig - 2007):

- let: $AR_m = R_m\Lambda_m$
- let: $L_m^H A = \Lambda_m L_m^H$
- then: $f(A)b = f(A)(I - R_m L_m^H)b + f(A)R_m L_m^H b$ (note: $f(A)R_m L_m^H b = R_m f(\Lambda_m) L_m^H b$)

- main problem with this approach: eigensolving can be extremelly expensive
- our numerical experiments show that expensive eigensolving is not a critical issue anymore, if we **combine polynomial preconditioning with LR deflation**

Preconditioning at the sign-function level

A short note on LR deflation for the sign function (Bloch, Frommer, Lang, Wettig - 2007):

- let: $AR_m = R_m \Lambda_m$
- let: $L_m^H A = \Lambda_m L_m^H$
- then: $f(A)b = f(A)(I - R_m L_m^H)b + f(A)R_m L_m^H b$ (note: $f(A)R_m L_m^H b = R_m f(\Lambda_m) L_m^H b$)

- main problem with this approach: eigensolving can be extremelly expensive
- our numerical experiments show that expensive eigensolving is not a critical issue anymore, if we **combine polynomial preconditioning with LR deflation**

Analyzing the interplay of polynomial preconditioning with deflation, in the context of the sign function, is **ongoing work**.

“Modern” overlap solver

while not converged **do**
    for $i = 1 : m_{out}$ **do**
        …
        for $j = 1 : m_{in}$ **do**
            …
            $v_{in}^{(1)} \leftarrow \mathrm{AMG}\left(D_w(m_0^{prec}, \epsilon)\right) w_{in}^{(1)}$ (hp)
            …
            $v_{in}^{(2)} \leftarrow \left(Q_\mu^2 \left(q(Q_\mu^2)\right)^2\right)^{-1/2} \left(w_{in}^{(2)} - R_m L_m^H w_{in}^{(2)}\right)$ (sp/hp)
            …
        end for
        …
        $v_{out} \leftarrow \left(Q_\mu^2 \left(q(Q_\mu^2)\right)^2\right)^{-1/2} \left(w_{out} - R_m L_m^H w_{out}\right)$ (dp)
        …
    end for
end while

Thank you!