Deflated Multigrid Multilevel Monte Carlo

Gustavo Ramirez-Hidalgo
(Andreas Frommer, Mostafa Nasr Khalil)

Bergische Universität Wuppertal
Department of Applied Mathematics

August 11, 2022

BERGISCHE UNIVERSITÄT WUPPERTAL

STIMULATE
European Joint Doctorates

Table of Contents

1. Short recap
2. Some preliminary results in LQCD
3. Skipping levels in multigrid multilevel Monte Carlo
4. Exact deflation in multigrid multilevel Monte Carlo
5. A (short) note on eigensolving

Short recap

Short recap

Multigrid multilevel Monte Carlo for $\mathrm{tr}(A^{-1})$

The problem at hand is: $\mathrm{tr}(A^{-1})$. The method presented here works for any $f(A)$, in particular any $f(A) = (\Gamma A)^{-1}$: but we focus on $f(A) = A^{-1}$:

We start with the decomposition:

$$\mathrm{tr}(A^{-1}) = \sum_{\ell=1}^{L-1} \mathrm{tr}\left(\hat{P}_\ell A_\ell^{-1} \hat{R}_\ell - \hat{P}_{\ell+1} A_{\ell+1}^{-1} \hat{R}_{\ell+1}\right) + \mathrm{tr}(\hat{P}_L A_L^{-1} \hat{R}_L).$$

Why should this reduce the variance?

$$\mathbb{V}[x^H A^{-1} x] = \|\mathrm{offdiag}(A^{-1})\|_F^2 = \sum_{i=1}^{n} \sigma_i^{-2} - \sum_{i=1}^{n} |(A^{-1})_{ii}|^2$$

BERGISCHE UNIVERSITÄT WUPPERTAL

A bit more specifically:

- two levels: $\mathrm{tr}(A_1^{-1}) = \mathrm{tr}(A_1^{-1}) - \mathrm{tr}(P_1 A_2^{-1} R_1) + \mathrm{tr}(P_1 A_2^{-1} R_1)$
- multilevel ($R_\ell P_\ell = I$):
  $\mathrm{tr}(A^{-1}) = \sum_{\ell=1}^{L-1} \mathrm{tr}\left(A_\ell^{-1} - P_\ell A_{\ell+1}^{-1} R_\ell\right) + \mathrm{tr}(A_L^{-1})$

We can connect this to inexact/approximate deflation:

$$A^{-1} = A^{-1}(I - \Pi) + A^{-1}\Pi$$

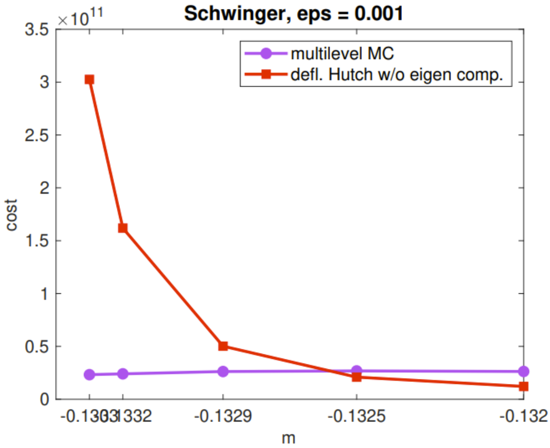
$$\Pi = \hat{U}_k(\hat{V}_k^* A \hat{U}_k)^{-1}\hat{V}_k^* A,\, \hat{U}_k, \hat{V}_k \in \mathbb{C}^{n\times k}$$

which is discussed in *Multigrid deflation for Lattice QCD* (Romero, Stathopoulos and Orginos), and already in Lüscher's local coherence paper.

BERGISCHE UNIVERSITÄT WUPPERTAL

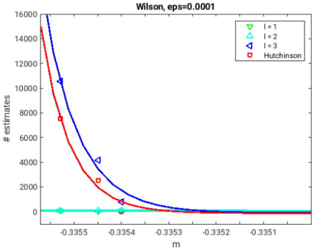
Short recap

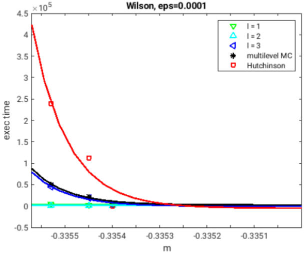
Multigrid multilevel Monte Carlo for $\mathrm{tr}(A^{-1})$



Some preliminary results in LQCD

Wilson: 64×32^3 (work in collaboration with Jose Jimenez-Merchan).
Twisted mass: 96×48^3.





	measurement	Mult. MC	Hutchinson
$\ell = 1$	# estimates	63	1164
	time (seconds)	17,010.0	221,160.0
$\ell = 2$	# estimates	172	-
	time (seconds)	15,480.0	-
$\ell = 3$	# estimates	903	-
	time (seconds)	9,030.0	-
$\ell = 4$	# estimates	7715	-
	time (seconds)	1,543.0	-
total	# estimates	-	1164
total	time (seconds)	43,063.0	221,160.0

BERGISCHE UNIVERSITÄT WUPPERTAL

Skipping levels in multigrid multilevel Monte Carlo

Let's fully write a four-level expansion:

$$\operatorname{tr}(A^{-1}) = \operatorname{tr}\left(A_1^{-1} - P_1 A_2^{-1} R_1\right) + \operatorname{tr}\left(P_1 A_2^{-1} R_1 - P_1 P_2 A_3^{-1} R_2 R_1\right) + \operatorname{tr}\left(P_1 P_2 A_3^{-1} R_2 R_1 - P_1 P_2 P_3 A_4^{-1} R_3 R_2 R_1\right) + \operatorname{tr}\left(P_1 P_2 P_3 A_4^{-1} R_3 R_2 R_1\right)$$

Skip second level-difference:

$$\operatorname{tr}(A^{-1}) = \operatorname{tr}\left(A_1^{-1} - P_1 P_2 A_3^{-1} R_2 R_1\right) + \operatorname{tr}\left(P_1 P_2 A_3^{-1} R_2 R_1 - P_1 P_2 P_3 A_4^{-1} R_3 R_2 R_1\right) + \operatorname{tr}\left(P_1 P_2 P_3 A_4^{-1} R_3 R_2 R_1\right)$$

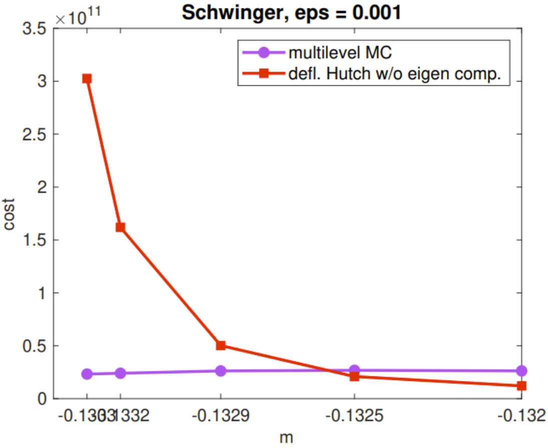
Then, reduce further:

$$\operatorname{tr}(A^{-1}) = \operatorname{tr}\left(A_1^{-1} - P_1 P_2 A_3^{-1} R_2 R_1\right) + \operatorname{tr}\left((A_3^{-1} - P_3 A_4^{-1} R_3)(R_2 R_1 P_1 P_2)\right) + \operatorname{tr}\left(A_4^{-1}(R_3 R_2 R_1 P_1 P_2 P_3)\right)$$

BERGISCHE UNIVERSITÄT WUPPERTAL

Exact deflation in multigrid multilevel Monte Carlo

Mostafa (talk on Th. @ 10:00) already talked about inexact deflation on top of multigrid multilevel Monte Carlo. Let's talk about exact deflation.



Exact deflation in MGMLMC

Deflating non-deflated modes

Exact deflation in "plain" Hutchinson:

$$\mathrm{tr}(A^{-1}) = A^{-1}(I - U_1 U_1^H) + \mathrm{tr}(U_1^H \Sigma_1^{-1} V_1)$$

In LQCD, to compute the singular vectors we can use:

$$A = U\Sigma V \Rightarrow Q = \Gamma_5 A$$

$$Q = X\Lambda X^H$$
$$\Rightarrow A = (\Gamma_5 X \mathrm{sign}(\Lambda))|\Lambda| V$$

Exact deflation in multigrid multilevel Monte Carlo:

instead of A, we have a more complicated operator to deflate from in multigrid multilevel Monte Carlo:

$$M_\ell = A_\ell^{-1} - P_\ell A_{\ell+1}^{-1} P_\ell^H$$

BERGISCHE UNIVERSITÄT WUPPERTAL

Can we construct a symmetrized difference-level operator?

To do this easily, the multigrid hierarchy needs to be built via a Γ_5-compatible aggregation:

$$\Gamma_5^{\ell} P_{\ell} = P_{\ell} \Gamma_5^{\ell+1}, \quad P_{\ell}^H \Gamma_5^{\ell} = \Gamma_5^{\ell+1} P_{\ell}^H$$

and then:

$$\begin{aligned} J_{\ell} &= M_{\ell} \Gamma_5^{\ell} \\ &= A_{\ell}^{-1} \Gamma_5^{\ell} - P_{\ell} A_{\ell+1}^{-1} P_{\ell}^H \Gamma_5^{\ell} = Q_{\ell}^{-1} - P_{\ell} A_{\ell+1}^{-1} \Gamma_5^{\ell+1} P_{\ell}^H \\ &= Q_{\ell}^{-1} - P_{\ell} Q_{\ell+1}^{-1} P_{\ell}^H \end{aligned}$$

BERGISCHE UNIVERSITÄT WUPPERTAL

Exact deflation in MGMLMC

Deflating non-deflated modes

$m_0 = -0.1320$, Schwinger. Single core, Python, laptop.

Method	exec. time	eigsolv. time
Hutch. + exact defl. (384)	1663.74	119.92
MGMLMC	2701.05	0.0
MGMLMC + exact defl. (32,32,32)	1067.77	1544.12
MGMLMC + exact defl. (32,32) + skip	530.26	948.35

BERGISCHE UNIVERSITÄT WUPPERTAL

A (short) note on eigensolving

We need to compute large modes of the operator:

$$J_\ell = Q_\ell^{-1} - P_\ell Q_{\ell+1}^{-1} P_\ell^H$$

Done naively (as we have in our first tests with exact deflation), this is extremely expensive. It's probably better to transform the eigenproblem into a generalized Hermitian eigenproblem:

$$(Q_\ell^{-1} - P_\ell Q_{\ell+1}^{-1} P_\ell^H)v = \lambda v$$
$$\Rightarrow (I - Q_\ell P_\ell Q_{\ell+1}^{-1} P_\ell^H)v = \lambda Q_\ell v$$

BERGISCHE UNIVERSITÄT WUPPERTAL

A (short) note on eigensolving

Using Jacobi-Davidson

$$(Q_\ell^{-1} - P_\ell Q_{\ell+1}^{-1} P_\ell^H) v = \lambda v$$

$$\Rightarrow (I - Q_\ell P_\ell Q_{\ell+1}^{-1} P_\ell^H) v = \lambda Q_\ell v$$

With the GHEP at hand, when using Jacobi-Davidson (which can be realized efficiently and scalably for Hermitian systems), we will need to apply the following operators at every iteration:

$$I - Q_\ell P_\ell Q_{\ell+1}^{-1} P_\ell^H$$

$$Q_\ell$$

and (very approximately e.g. 10^{-1}) solve a linear system with the operator:

$$(I - Q_\ell P_\ell Q_{\ell+1}^{-1} P_\ell^H) - \theta Q_\ell$$

BERGISCHE UNIVERSITÄT WUPPERTAL

Thank you!

BERGISCHE UNIVERSITÄT WUPPERTAL