PROCEEDINGS OF SCIENCE

Variance Reduction in Trace Estimation for Lattice QCD Using Multigrid Multilevel Monte Carlo

Andreas Frommer,[a] **Jose Jimenez-Merchan,**[a,*] **Gustavo Ramirez-Hidalgo**[b] **and Juan Andrés Urrea-Niño**[a]

[a] *Department of Mathematics, University of Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany*

[b] *Institute for Advanced Simulation (IAS), Jülich Supercomputing Centre (JSC) Wilhelm-Johnen-Straße 52428 Jülich*

E-mail: frommer@uni-wuppertal.de, jimenezmerchan@uni-wuppertal.de, g.ramirez.hidalgo@fz-juelich.de, urreanino@uni-wuppertal.de

Trace estimation is a significant challenge in lattice QCD simulations. The Hutchinson method's accuracy scales with the square root of the sample size, resulting in high computational costs for precise estimates. Variance reduction techniques, such as deflating the lowest eigen or singular vectors of the matrix, are employed to alleviate this issue.

This study explores Multigrid Multilevel Monte Carlo (MGMLMC) to reduce the computational cost of constructing the deflation subspace while maintaining efficient application of the deflation projectors and improving variance reduction. In MGMLMC, spectral deflation is accomplished using a projector derived from the multigrid prolongator P used in solving linear systems involving the Wilson-Dirac operator. By utilizing the low-mode spectral information inherent in P, this approach significantly lowers memory requirements while achieving up to a three-fold variance reduction compared to inexact deflation, which relies on a few iterations of the inverse block power method to derive the deflation subspace.

We investigate the efficacy of MGMLMC for computing $\mathrm{tr}(B(t)D^{-1}(t,t))$, where $B(t)$ acts on spin, color, and space indices, for example, a combination of gamma matrices and gauge covariant spatial derivatives.

The 41st International Symposium on Lattice Field Theory (LATTICE2024)
28 July - 3 August 2024
Liverpool, UK

*Speaker

PoS(LATTICE2024)279

1. Introduction

We consider estimating the trace of the inverse of a large sparse matrix $D \in \mathbb{C}^{n\times n}$, possibly after multiplication by an operator B, i.e., computing quantities of the form $\mathrm{tr}(BD^{-1})$. This problem arises in various fields, notably in Lattice Quantum Chromodynamics (QCD), where the disconnected fermion loop contributions to an observable are obtained from the trace of the inverse of the discretized Dirac operator multiplied by an operator Γ, i.e., $\mathrm{tr}(\Gamma D^{-1})$ [1]. The operator Γ fixes the symmetry channel; for example, when Γ is the identity operator, it appears in the two-point function of a scalar meson or in the mass derivative of the action [2]. We also explore quantities of the form $\mathrm{tr}(B(t)D^{-1}(t,t))$, where $B(t)$ is an operator acting on spin, color, and space indices at a fixed time slice t. Here, $D(t,t)$ denotes a time-slice of the Dirac operator, considering only the three spatial coordinates. Examples of $B(t)$ include combinations of gamma matrices and gauge covariant spatial derivatives [2].

Because the $n \times n$ matrix D is exceedingly large, direct computation of its inverse is infeasible. The only practical way to access information about the entries of D^{-1} is through matrix-vector multiplications $D^{-1}x$, that is, by solving linear systems involving the matrix D. This situation necessitates the use of stochastic estimation techniques, such as Hutchinson's method [3], which uses random vectors $x \in \mathbb{C}^n$ whose components x_i follow an isotropic distribution, meaning

$$\mathbb{E}[|x_i|^2] = 1, \quad \mathbb{E}[\bar{x}_i x_j] = 0 \quad \text{for } i, j = 1, \dots, n,\ i \neq j. \tag{1}$$

Typically, the components are independent and identically distributed (i.i.d.) complex numbers z with $\mathbb{E}[z] = 0$ and $\mathbb{E}[|z|^2] = 1$. A notable example is Rademacher vectors, where z is uniformly distributed in $\{-1, 1\}$. For $\mathbb{Z}_4$-vectors, z is instead uniformly distributed in $\{-1, -i, 1, i\}$, where i is the imaginary unit. By averaging $x^\dagger D^{-1}x$ over N independent random vectors x, we obtain an unbiased estimator for the trace:

$$\mathrm{tr}(D^{-1}) \approx \hat{\mathrm{tr}}(D^{-1}) = \frac{1}{N}\sum_{i=1}^{N}(x^{(i)})^\dagger D^{-1} x^{(i)}. \tag{2}$$

The key limitation of this Monte Carlo trace estimation is that its accuracy improves only with the square root of the sample count N, making accurate estimations nearly impractical unless variance can be reduced. In [4] we introduced the multigrid multilevel Monte Carlo (MGMLMC) method, which uses oblique deflation to achieve variance reduction exploiting local coherence embedded in the intergrid transfer operators from the multigrid hierarchy. This has recently been applied in [5] to achieve variance reductions in computations of the isovector vector current correlator.

The primary contribution of this study is the enhancement of MGMLMC through the incorporation of orthogonal projectors. In section 2, we describe deflation as a variance reduction in the Hutchinson estimator. In section 3 we detail MGMLMC as a deflation technique and then introduce the upgraded version of the method with the use of orthogonal projectors. Furthermore, a first approach for the estimation of traces of the form $\mathrm{tr}(B(t)D^{-1}(t,t))$ with MGMLMC is derived. Finally, in section 4 we present the numerical experiments that compare MGMLMC with inexact deflation.

2. Variance Reduction Techniques

For Rademacher vectors, the variance of the Hutchinson estimator for $\mathrm{tr}(D^{-1})$ is given by $\frac{1}{2}\|\,\mathrm{offdiag}(D^{-1}+D^{-\top})\|_F^2$, while for $\mathbb{Z}_4$-vectors, it is $\|\,\mathrm{offdiag}(D^{-1})\|_F^2$ [6]. The heuristics underlying variance reduction techniques typically aim at reducing $\|D^{-1}\|_F^2$.

2.1 Variance Reduction Using Deflation

Given the relation between the Frobenius norm and the singular values σ_i of a matrix [7], $\|D\|_F^2 = \sum_{i=1}^n \sigma_i^2$, deflation techniques aim to remove contributions of the k largest singular values $\sigma_1 > \sigma_2 > \cdots > \sigma_k$ to the variance [8, 9].

From the singular value decomposition $D = U\Sigma V^\dagger$, let $U_k, V_k \in \mathbb{C}^{n\times k}$ hold the k largest right and left singular vectors of D^{-1}, i.e., those belonging to the k smallest singular values of D. Constructing the orthogonal projector $\Pi = V_k V_k^\dagger$, the trace $\mathrm{tr}(D^{-1})$ is split into two terms:

$$\mathrm{tr}(D^{-1}) = \mathrm{tr}\left((I-\Pi)D^{-1}\right) + \mathrm{tr}(\Pi D^{-1}). \tag{3}$$

The first term contains modes corresponding to the remaining $n-k$ smaller singular values of D^{-1}. It has reduced the Frobenius norm and can thus be expected to have a smaller variance.

Using the cyclic property of the trace, the second term in eq. (3) simplifies to $\mathrm{tr}\left(V_k^\dagger D^{-1} V_k\right)$, which, when V_k and U_k are exact, is directly computable as $\Sigma_k^{-1} U_k^\dagger V_k$ with Σ_k the diagonal matrix containing the k deflated singular values. In *inexact* deflation, where V_k only approximates right singular vectors, this term remains computable non-stochastically by solving k linear systems.

Although the Dirac operator is a non-Hermitian operator, it exhibits Γ_5-hermiticity, i.e., $(\Gamma_5 D)^\dagger = \Gamma_5 D$. This property ensures the existence of an eigendecomposition $\Gamma_5 D = V\Lambda V^\dagger$, where V is unitary. This yields the SVD of D, $D = \Gamma_5 V\Sigma V^\dagger$. In inexact deflation, we obtain V_k by a few iterations of the inverse block power method on $(\Gamma_5 D)^{-1}$. Then, estimating $\mathrm{tr}(D^{-1})$ using the Hutchinson method (eq. (2)) with deflation and with sample size N gives

$$\widehat{\mathrm{tr}}(D^{-1}) = \frac{1}{N}\sum_{n=1}^{N}\left[(x^n)^\dagger\left(D^{-1} - V_k V_k^\dagger D^{-1}\right)x^n\right] + \mathrm{tr}\left(V_k^\dagger D^{-1} V_k\right). \tag{4}$$

3. Multigrid Multilevel Monte Carlo (MGMLMC)

MGMLMC is a variance reduction technique introduced in [4], which approximates $\mathrm{tr}(D^{-1})$ using a recursive deflation technique based on high-rank projectors obtained from a multigrid hierarchy for D. There, restrictors R_l and prolongators P_l, which are operators tarsnporting variables from level l to level $l+1$ and back, are used to define the coarse-grid (level $l+1$) operators D_{l+1}. In this work, we use the Domain Decomposition aggregation-based αdaptive algebraic multigrid method (DDαAMG), which is a solver for linear systems of equations arising in simulations of lattice QCD involving Wilson or twisted mass fermions [10]. It uses the Galerkin projection where $R_l = P_l^\dagger$, leading to the coarser operators

$$D_{l+1} = P_l^\dagger D_l P_l, \quad l = 1, \ldots, L-1. \tag{5}$$

MGMLMC realizes deflation using the oblique projectors

$$\Pi_l = P_l D_{l+1}^{-1} P_l^\dagger D_l, \tag{6}$$

allowing at each level l to split D_l^{-1} into a deflated and non-deflated term:

$$D_l^{-1} = \left(D_l^{-1} - P_l D_{l+1}^{-1} P_l^\dagger\right) + P_l D_{l+1}^{-1} P_l^\dagger. \tag{7}$$

This leads to the following decomposition of $\mathrm{tr}(D^{-1})$

$$\mathrm{tr}(D^{-1}) = \sum_{l=1}^{L-1} \left[\mathrm{tr}\left(D_l^{-1} - P_l D_{l+1}^{-1} P_l^\dagger\right)\right] + \mathrm{tr}\left(D_L^{-1}\right). \tag{8}$$

Denote $M_l = D_l^{-1} - P_l D_{l+1}^{-1} P_l^\dagger$. MGMLMC stochastically estimates $\mathrm{tr}(M_l)$ at each level l and $\mathrm{tr}(D_L^{-1})$ at the coarsest level using the Hutchinson method (eq. (2)), giving the unbiased estimator

$$\widehat{\mathrm{tr}}(D^{-1}) = \sum_{l=1}^{L-1} \widehat{\mathrm{tr}}(M_l) + \widehat{\mathrm{tr}}(D_L^{-1}). \tag{9}$$

This approach expects the variance for each level difference $D_l^{-1} - P_l D_{l+1}^{-1} P_l^\dagger$ to be small, since the prolongators P_l are built from appproximations to small eigenmodes, and due to local coherence, they span many of the problematic low modes. The term $\mathrm{tr}(D_L^{-1})$ can be cheaply computed either directly or stochastically since the computational cost of the solution of linear systems is reduced by a factor of about 8 every time we go from level l to level $l+1$.

3.1 MGMLMC Through an Orthogonal Projector

An orthogonal projector ($\Pi^\dagger = \Pi$) is guaranteed to reduce the Frobenius norm, whereas an oblique projector does not necessarily do so (see e.g. [7]). The projector in eq. (6) is oblique, motivating us to propose an alternative MGMLMC method using an orthogonal projector.

Taking the prolongator operators P_l from the multigrid construction in DDαAMG, which are unitary, we build the orthogonal projector $\Pi_l = P_l P_l^\dagger$. Then, we split D_l^{-1} into a deflated and non-deflated term:

$$D_l^{-1} = (I - P_l P_l^\dagger) D_l^{-1} + P_l P_l^\dagger D_l^{-1}, \quad l = 1, \ldots, L, \tag{10}$$

leading to a new MLMC construction for the trace computation:

$$\mathrm{tr}(D_l^{-1}) = \mathrm{tr}\left((I - P_l P_l^\dagger) D_l^{-1}\right) + \mathrm{tr}\left(P_l^\dagger D_l^{-1} P_l - D_{l+1}^{-1}\right) + \mathrm{tr}\left(D_{l+1}^{-1}\right). \tag{11}$$

The first term is a deflation of D_l^{-1} from the left, but using $(I - P_l P_l^\dagger)^2 = I - P_l P_l^\dagger$ and the cyclic property of the trace one can simultaneously deflate from the right:

$$\mathrm{tr}(D_l^{-1}) = \mathrm{tr}\left((I - P_l P_l^\dagger) D_l^{-1} (I - P_l P_l^\dagger)\right) + \mathrm{tr}\left(P_l^\dagger D_l^{-1} P_l - D_{l+1}^{-1}\right) + \mathrm{tr}\left(D_{l+1}^{-1}\right). \tag{12}$$

Denoting F_l the "full-rank difference operator" $P_l^\dagger D_l^{-1} P_l - D_{l+1}^{-1}$ and O_l the orthogonally deflated operator $(I - P_l P_l^\dagger) D_l^{-1} (I - P_l P_l^\dagger)$ at each level l, and computing each term stochastically, we obtain the unbiased stochastic estimator

$$\widehat{\mathrm{tr}}(D^{-1}) = \sum_{l=1}^{L-1} \widehat{\mathrm{tr}}(O_l) + \sum_{l=1}^{L-1} \widehat{\mathrm{tr}}(F_l) + \widehat{\mathrm{tr}}(D_L^{-1}). \tag{13}$$

PoS(LATTICE2024)279

3.2 Estimation of traces for the measurement of observables

In Lattice QCD, the measurement of observables involves traces of the form $\text{tr}(BD^{-1}(t,t))$ [2]. We can write

$$\text{tr}(BD^{-1}(t,t)) = \text{tr}(P_{3D}BD^{-1}P_{3D}^{\dagger}), \tag{14}$$

where P_{3D} orthogonally projects onto the time-slice t. We here present a first proof of concept for a multilevel approach by applying the two-level MGMLMC method from eq. (8) directly on D^{-1}:

$$\begin{aligned}\text{tr}(D_1^{-1}(t,t)) = \text{tr}\left(P_{3D}B\left(D_1^{-1} - P_1D_2^{-1}P_1^{\dagger}\right)P_{3D}^{\dagger}\right)\\ + \text{tr}\left(P_{3D}BP_1D_2^{-1}P_1^{\dagger}P_{3D}^{\dagger}\right),\end{aligned} \tag{15}$$

which can be extended to three levels:

$$\begin{aligned}\text{tr}(D_1^{-1}(t,t)) = \text{tr}\left(P_{3D}B\left(D_1^{-1} - P_1D_2^{-1}P_1^{\dagger}\right)P_{3D}^{\dagger}\right)\\ + \text{tr}\left(P_{3D}P_1B\left(D_2^{-1} - P_2D_3^{-1}P_2^{\dagger}\right)P_1^{\dagger}P_{3D}^{\dagger}\right)\\ + \text{tr}\left(P_{3D}BP_1P_2D_3^{-1}P_2^{\dagger}P_1^{\dagger}P_{3D}^{\dagger}\right).\end{aligned} \tag{16}$$

4. Numerical Experiments

In this section, we present the results of using MGMLMC for two types of problems: the estimation of $\text{tr}(D^{-1})$ and $\text{tr}\left(B(t)D^{-1}(t,t)\right)$. For the first problem, we demonstrate the benefits of our oblique and orthogonal MGMLMC methods over inexact deflation on very large configurations from the CLS collaboration [11] (see Table 1). We implemented our code[1] as en extension of DDαAMG, and ran the computations using up to to 54 nodes, each consisting of 2 sockets with AMD EPYC 7452 32-Core processors.

For the second problem, we perform experiments on a lattice of size 16^4 using a single process with a MATLAB implementation[2] of our approach, comparing it against inexact deflation.

4.1 Estimation of $\text{tr}(D)^{-1}$

ensemble	temporal b.c.	Size	m_π [MeV]	m_K [MeV]	a [fm]
J501	Open	192×64^3	336	448	0.39
E250	Periodic	192×96^3	131	493	0.63

Table 1: Comparison of lattice QCD parameters for ensembles J501 and E250.

We computed the estimated variance for the trace estimates of the operators in all the introduced methods. In the oblique MGMLMC method (eq. (9)), the variance was measured for the difference operators M_l and the coarsest-level inverse operator D_L^{-1}. In the orthogonal MGMLMC method (eq. (12)), we evaluated the variance for the full-rank operators F_l and the orthogonal operators

[1] Available at https://github.com/jomjimenezme/MLMC_ci.
[2] Available at https://github.com/Gustavroot/MGMLMC_for_mat_with_displ.

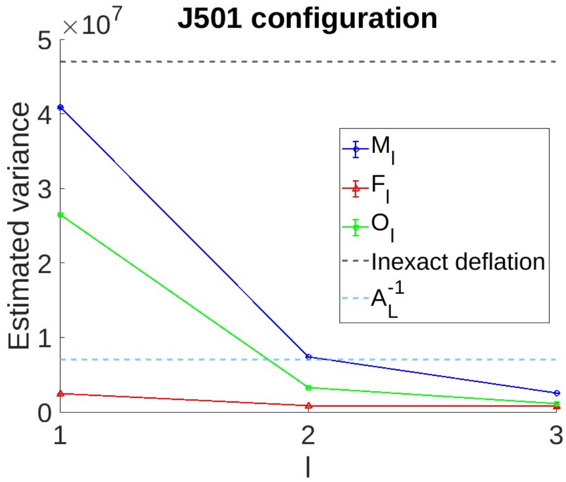

Figure 1: Configuration J501. Estimated variances for the operators: oblique O_O, full rank O_F, orthogonal O_O, and coarsest A_L^{-1} at different levels l. Compared against inexact deflation. The used sample size is $N = 500$.

O_l. For Hutchinson's inexactly deflated method eq. (3), we similarly computed the variance for the trace estimates.

A sample size of $N = 500$ was used across all variance measurements for all trace estimations of the operators at levels $l = 1, 2, 3$, and the coarsest level $L = 4$. The variance was estimated using the unbiased estimator for the variance of the sample mean of a complex random variable X with estimated mean $\hat{\mu}$:

$$\widehat{V} = \frac{1}{N-1} \sum_{i=1}^{N} |X_i - \hat{\mu}|^2. \tag{17}$$

For each operator sample, we needed to solve linear systems involving the matrices A_l or A_{l+1}. The $DD\alpha$AMG multigrid hierarchy established for the matrix A served as the framework for these solves. The consistent use of the prolongator operators P_l across all levels allowed for a seamless multigrid solver at each level without additional cost or overhead. At the coarsest level L, flexible GMRES was employed. All solves were performed with a relative residual tolerance of 1×10^{-8}.

For Hutchinson's inexactly deflated method, we used $k = 128$ deflation vectors, appearing as the columns of V_k, and those were obtained using five iterations of the inverse block power method on $(\Gamma_5 D)$, with each inversion aiming for an accuracy of 1×10^{-6}. The second term in eq. (3) was computed using the necessary k solves and k vector-vector products.

Figure 1 shows that both of our approaches (oblique and orthogonal MGMLMC) achieve a significant variance reduction compared to inexact deflation at the finest (most expensive) level. This directly translates to a smaller required sample size (i.e., fewer solves) at the finest level for a given accuracy. On the coarser levels, we also observe reduced variance in the deflated operators and in the coarsest-level operator A_L^{-1}. As previously noted, the cost of solving decreases significantly as we move to coarser levels, with each transition reducing the computational cost by approximately a factor of 8. Furthermore, the plot shows that the operators from the orthogonal MGMLMC method, O_l and F_l, indeed present a lower variance than the operators M_l of the oblique one.

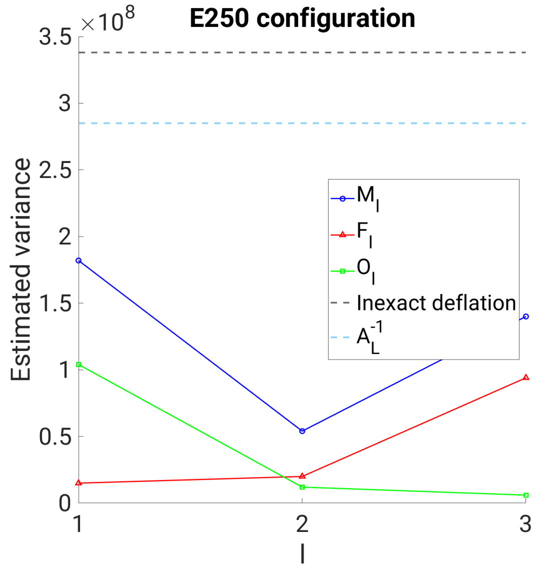

Figure 2: E250 configuration. Estimated variances for the operators: oblique O_O, full rank O_F, orthogonal O_O, and coarsest A_L^{-1} at different levels l. Compared against inexact deflation. The used sample size is $N = 500$.

In contrast, inexact deflation not only exhibits higher variance but also incurs additional overhead from precomputing the deflation space at the finest level. Specifically, this involves multiple solves during the inverse block power iteration, where 128 deflation vectors are used with 6 inverse iterations, totaling around 700 inexact solves. Consequently, the overall cost of the MGMLMC methods is lower, as they avoid this precomputation overhead while keeping the extra computational effort on coarser levels minimal.

Figure 2 demonstrates that for the lighter mass E250 configuration, the previously discussed reasoning remains applicable. However, at levels 3 and 4, we observe an increase in the variance for the deflated operators M_l and F_l. On these levels solves are approximately 100 to 1000 faster than at the finest level, which significantly offsets the effect of increased variance. Consequently, the overall computational effort required at these levels remains minimal, ensuring that the efficiency and scalability of the MGMLMC methods are preserved.

4.2 Estimation of $\mathrm{tr}(B(t)D(t,t))$

We analyze the variance reduction achieved by the oblique MGMLMC compared against inexact deflation in estimating traces relevant to the measurement of observables $\mathrm{tr}(B(t)D(t,t))$ [2]. In this first approach we consider two cases, first $B = I$ and then $B = \sum_i \Gamma_i \nabla_i$.

For the case $B = I$, the measurements are presented in in Table 2, the non-deflated Hutchinson method is shown for reference. We first compare the variance at the finest level, which is the most computationally expensive level. To achieve the same variance reduction as MGMLMC, inexact deflation needs to build a deflation space made up by $k = 64$ vectors. MGMLMC proves to be more cost efficient since it avoids the cost of constructing the deflation space. In MGLMC, the variance shifts to the second level, where solvers have less computational cost, and at the third level, the

k	Inexact-def	MGMLMC
0	75278	1^{st}_{lvl}: 1862, 2^{nd}_{lvl}: 9905, 3^{rd}_{lvl}: 0
8	2107	-
16	1994	-
32	1939	-
64	1863	-

Table 2: Variance results for the estimation of $\mathrm{tr}(D(t,t))$ using inexact deflation with k deflation vectors compared against oblique MGMLMC.

variance is zero because the last term in eq. (16) is computed exactly (non-stochastically).

When considering the operator $B = \sum_i \Gamma_i \nabla_i$, Table 3 illustrates the difficulty of the problem since the trace has a small value and the variance of the Hutchinson estimator is large variance. While both deflation methods present only a modest reduction of the variance, inexact deflation reduces the variance by 3%, whereas MGMLMC achieves a variance reduction of 12% at the first level.

Method	Trace	Variance
Hutchinson	12.1	8906
Inexact Deflation ($k = 64$)	11.8	8660
MGMLMC	11.5	Level 1: 7896, Level 2: 497

Table 3: Variance results for the estimation of $\mathrm{tr}(\sum_i \Gamma_i \nabla_i D(t,t)$ using inexact deflation with $k = 64$ deflation vectors compared against oblique MGMLMC.

5. Conclusion

We developed a new MGMLMC method which uses additional orthogonal projectors and demonstrated its ability to further reduce the variance when compared with its oblique counterpart for the estimation of $\mathrm{tr}(D^{-1})$ in very large configurations.

Furthermore, we presented a first approach for the computation of traces of operators of the form $B(t)D^{-1}(t,t)$ using oblique MGMLMC and showed how it compares to inexact deflation. Future work could explore further improvements of MGMLMC for computing $\mathrm{tr}(B(t)D^{-1}(t,t))$, such as also deflating B or using probing techniques that exploit the structure of B.

Acknowledgments

The authors of this work were supported by Deutsche Forschungsgemeinschaft (DFG) through Research Unit FOR 5269. Computations were carried out on the PLEIADES cluster at University of Wuppertal, funded by DFG grant No. INST 218/78-1 FUGG and the Bundesministerium für Bildung und Forschung (BMBF).

References

[1] C. Gattringer and C.B. Lang, *Quantum Chromodynamics on the Lattice*, vol. 788, Springer, Berlin (2010).

[2] J.J. Dudek, R.G. Edwards, N. Mathur and D.G. Richards, *Charmonium excited state spectrum in lattice qcd*, *Phys. Rev. D* **77** (2008) 034501 [0707.4162].

[3] M.F. Hutchinson, *A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines*, *Commun. Stat. Simul.* **19** (1990) 433.

[4] A. Frommer, M.N. Khalil and G. Ramirez-Hidalgo, *A multilevel approach to variance reduction in the stochastic estimation of the trace of a matrix*, *SIAM J. Sci. Comput.* **44** (2022) A2536 [2108.11281].

[5] R. Gruber, T. Harris and M.K. Marinkovic, *Multigrid low-mode averaging*, 2412.06347.

[6] S. Bernardson, P. McCarty and C. Thron, *Monte Carlo methods for estimating linear combinations of inverse matrix entries in lattice QCD*, *Comput. Phys. Commun.* **78** (1994) .

[7] L.N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA (1997), 10.1137/1.9780898719574.

[8] A.S. Gambhir, A. Stathopoulos and K. Orginos, *Deflation as a method of variance reduction for estimating the trace of a matrix inverse*, *SIAM J. Sci. Comput.* **39** (2017) A532 [1603.05988].

[9] R.A. Meyer, C. Musco, C. Musco and D.P. Woodruff, *Hutch++: Optimal stochastic trace estimation*, in *Proceedings of the 2021 Symposium on Simplicity in Algorithms (SOSA)*, pp. 142–155, SIAM, 2021 [2010.09649].

[10] A. Frommer, K. Kahl, S. Krieg, B. Leder and M. Rottmann, *Adaptive aggregation-based domain decomposition multigrid for the lattice wilson–dirac operator*, *SIAM J. Sci. Comput.* **36** (2014) A1581 [2406.1303.1377].

[11] D. Mohler, S. Schaefer and J. Simeth, *CLS 2+1 flavor simulations at physical light- and strange-quark masses*, *EPJ Web Conf.* **175** (2018) 02010 [1712.04884].