Implementing noise reduction techniques into the OpenQ*D package

Lucius Bushnaq,[f,*] **Isabel Campos,**[b] **Marco Catillo,**[a] **Alessandro Cotellucci,**[c] **Madeleine Dale,**[d,e,h] **Jens Lücke,**[c,g] **Marina Krstić Marinković,**[a] **Agostino Patella,**[c,g] **Mike Peardon**[f] **and Nazario Tantalo**[d,e]

[a]*Institut für Theoretische Physik, ETH Zürich,*
Wolfgang-Pauli-Str. 27, 8093 Zürich, Switzerland

[b]*Instituto de Física de Cantabria & IFCA-CSIC,*
Avda. de Los Castros s/n, 39005 Santander, Spain

[c]*Humboldt Universität zu Berlin, Institut für Physik & IRIS Adlershof,*
Zum Großen Windkanal 6, 12489 Berlin, Germany

[d]*Università di Roma Tor Vergata, Dip. di Fisica,*
Via della Ricerca Scientifica 1, 00133 Rome, Italy

[e]*INFN, Sezione di Tor Vergata,*
Via della Ricerca Scientifica 1, 00133 Rome, Italy

[f]*School of Mathematics, Trinity College Dublin,*
College Green, Dublin 2, Ireland

[g]*DESY,*
Platanenallee 6, D-15738 Zeuthen, Germany

[h]*Department of Physics, University of Cyprus, 1 Panepistimiou Street, 2109 Aglantzia, Nicosia, Cyprus*
E-mail: bushnaql@tcd.ie, isabel@campos-it.es,
Marco.Catillo@physik.uni-muenchen.de,
alessandro.cotellucci@physik.hu-berlin.de, m.dale@stimulate-ejd.eu,
fritzsch@uni-muenster.de, jens.luecke@hu-berlin.de,
marinama@phys.ethz.ch, agostino.patella@physik.hu-berlin.de,
mjp@maths.tcd.ie, nazario.tantalo@roma2.infn.it

We present the results of testing a new technique for stochastic noise reduction in the calculation of propagators by implementing it in OpenQ*D for two ensembles with O(a) improved Wilson fermion action, with periodic boundary conditions and pion masses of 437 MeV and 331 MeV, for the connected vector and pseudoscalar correlators. We find that the technique yields no speedup compared to traditional methods, owning to the failure of its underlying assumption that the spectra of the spatial Laplacian and Dirac operators are sufficiently similar for the technique's purposes.

The 38th International Symposium on Lattice Field Theory, 2021 Zoom/Gather@Massachusetts Institute of Technology, LATTICE2021 26th-30th July

---

*Speaker

1. Introduction

The distillation operator [1] has seen great success as a smearing tool. This might lead one to suspect that the low-lying spectrum of the spatial Laplacian has substantial overlap with the low-lying spectrum of the Dirac operator. Since the spatial Laplacian is considerably smaller, its spectrum is easier to calculate than that of the full Dirac operator. Thus, if the overlap between the two operator's spectra were good enough, it could make sense to substitute the Distillation spectrum for the Dirac spectrum in many applications.

A potential practical use case and test of this notion would be distillation low-mode averaging. In conventional low-mode averaging, one attempts to reduce the stochastic noise on numeric calculations of propagators by calculating the lowest modes of the Dirac operator explicitly, which are assumed to give the dominant contribution to the observable. The propagator is then broken up into a contribution projected into the space of low-lying modes, which is now available directly, and a space of high-lying modes, which is calculated stochastically [2] [3].

Here, the idea is to do the same, but using the distillation operator instead. The distillation operator is formed out of the eigenmodes of the spatial Laplacian

$$\nabla^2_{ms,ns'}(t) = -6\delta_{ss'}\delta_{mn} + \delta_{ss'} \sum_{j=1}^{3} \left(U_j(m,t)\delta_{m+\hat{j},n} + U_j^\dagger(m-\hat{j},t)\delta_{m-\hat{j},n} \right)$$

$$\square(t) = V(t)V^\dagger(t) = \sum_{k=1}^{N} v^{(k)}(t) v^{(k)\dagger}(t),$$

(1)

where $v^{(k)}(t)$ are the eigenvectors, N is the number of eigenvectors calculated, and $V(t)$ is a matrix with the eigenvectors as column entries. Since the operator does not affect the spin space, the eigenvectors come in degenerate sets of four. We chose the spin-orthogonal basis for these degenerate states.

To implement this method for the calculation of the connected vector correlator, we begin by splitting the Hilbert space into the low-lying space of the Laplacian modes, and the "rest-space" at both source and sink by inserting $1 - \square(t) + \square(t)$

$$
\begin{aligned}
&\langle \overline{\psi}^A(t)\gamma_\mu \psi^B(t) \overline{\psi}^B(t')\gamma_\nu \psi^A(t') \rangle_F \\
&= \langle \overline{\psi}^A(t)\gamma_\mu \left(1 - \square(t) + \square(t)\right) \psi^B(t) \overline{\psi}^B(t')\gamma_\nu \left(1 - \square(t') + \square(t')\right) \psi^A(t') \rangle_F \\
&= \langle \overline{\psi}^A(t)\gamma_\mu \left(1 - \square(t)\right) \psi^B(t) \overline{\psi}^B(t')\gamma_\nu \left(1 - \square(t')\right) \psi^A(t') \rangle_F \\
&\quad + 2\langle \overline{\psi}^A(t)\gamma_\mu \square(t) \psi^B(t) \overline{\psi}^B(t')\gamma_\nu \left(1 - \square(t')\right) \psi^A(t') \rangle_F \\
&\quad + \langle \overline{\psi}^A(t)\gamma_\mu \square(t) \psi^B(t) \overline{\psi}^B(t')\gamma_\nu \square(t') \psi^A(t') \rangle_F \\
&=: C_{rest}(t) + 2C_{mixed}(t) + C_{dist}(t).
\end{aligned}
$$

(2)

We have used time reversal symmetry to combine the two mixed contributions into one. If we were working with conventional low-mode averaging, and $\square(t)$ consisted of eigenmodes of the Dirac operator, $C_{mixed}(t)$ would vanish since $\square$ and the propagator would commute. But here, we need to calculate it on top of the other two contributions.

Now we perform a one-end insertion to introduce the stochastic sources

$$C_{dist}(t) = \text{tr}[\gamma_5\gamma_\mu V^\dagger(t)D^{-1}(t,t')V(t')\gamma_\nu\gamma_5 V^\dagger(t')D^{-1^\dagger}(t',t)V(t)]$$

$$C_{rest}(t) = \frac{1}{N_r}\text{tr}[\gamma_5\gamma_\mu\left(1-\square(t)\right)D^{-1}(t,t')\eta^{(r)}(t')\eta^{\dagger(r)}(t')\gamma_\nu\gamma_5\left(1-\square(t')\right)D^{-1^\dagger}(t',t)] \qquad (3)$$

$$C_{mixed}(t) = \frac{1}{N_r}\text{tr}[\gamma_5\gamma_\mu\square(t)D^{-1}(t,t')\eta^{(r)}(t')\eta^{\dagger(r)}(t')\gamma_\nu\gamma_5\left(1-\square(t')\right)D^{-1^\dagger}(t',t)] \, .$$

Thus, to calculate $C_{dist}(t)$ we need the vectors $\phi^{(k)} := D^{-1}(t,t')v^{(k)}(t')$ which then form the perambulator [1]. Thus, we need to solve the Dirac equation with the eigenvectors of the spatial Laplacian at the desired source time as sources.

To calculate $C_{rest}(t)$ and $C_{mixed}(t)$, we need

$$\phi^{(r)} := D^{-1}(t,t')\eta^{(r)}(t')$$
$$\overline{\phi}^{(r)} := D^{-1}(t,t')\left(1-\square(t')\right))\eta^{(r)}(t') \, , \qquad (4)$$

making for $N + 2R$ needed solutions to the Dirac equation in total. Any two-point correlation function can then be formed through the contraction of these solutions by calculating the spin structure explicitly, as detailed in [4] adjusted for the distillation-low-mode-averaging method.

In order for the technique to be worth the cost of the increased number of solutions needed per stochastic source, as well as the cost of calculating the distillation modes themselves, $C_{mixed}(t)$ and $C_{rest}(t)$ need to be small relative to $C_{dist}(t)$.

2. Implementation and Simulation setup

The technique was implemented for the connected vector correlator and pseudoscalars in OpenQ*D, using the PRIMME eigensolver [5] to obtain the distillation eigenmodes.

The spatial covariant laplacian is passed to PRIMME as a black-box operator, parallelized with MPI. The solver then returns the specified number of low-lying eigenmodes. This is repeated for the spatial Laplacian on every timeslice of the lattice, for every configuration. No preconditioning for the operator is employed.

We tested the performance of the method on two CLS data sets with two flavours of dynamical O(a) improved Wilson fermions. The computation used periodic boundary conditions and stochastic wall sources. The determinations of the lattice spacing as well as the other configuration parameters were taken from Refs. [6] [7].

Config	V	β	κ	$m_\pi L$	a [fm]	m_π [MeV]
E5	$64x32^3$	5.30	0.13625	4.7	0.0658(7)(7)	437
A5	$64x32^3$	5.20	0.13594	4.0	0.0755(9)(7)	331

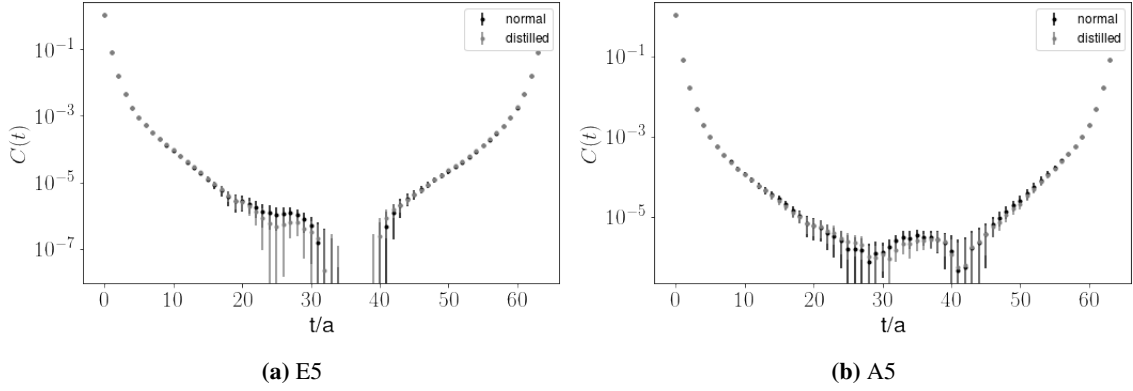**Table 1:** Ensemble parameters

(a) E5

(b) A5

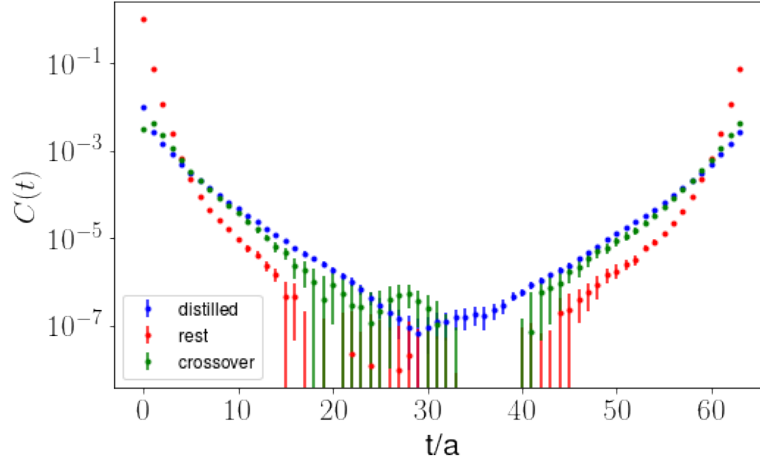**Figure 1:** Vector correlator, 160x4 eigenvectors, 24 wall sources

Figure 2: E5 vector correlator by parts, 160x4 eigenvectors, 24 sources

3. Results

The error analysis for all results was performed with jackknifing. Correlations between the errors at different time slices were not corrected for. Figures 1a and 1b compare the vector correlator calculated with distillation-low-mode-averaging and without, on the E5 and A5 ensembles respectively, with six independent color and even-odd diluted wall sources and 640 Laplacian modes. The results are in good agreement inside the error bars. Figure 2 shows the three individual contributions to the total distilled result in the E5 data set. As one would hope, the C_{rest} contribution quickly falls off in importance relative to C_{dist}. However, C_{mixed} does not fall off to the same extent, and stays close to C_{dist} in size even in the long distance part of the correlator.

This is a fundamental problem for the technique. If C_{mixed} does not fall off quickly, the method is not efficient in terms of the computing costs invested. This would seem to indicate that our guess was incorrect, and the spectra of the distillation and Dirac operators are not similar enough to make
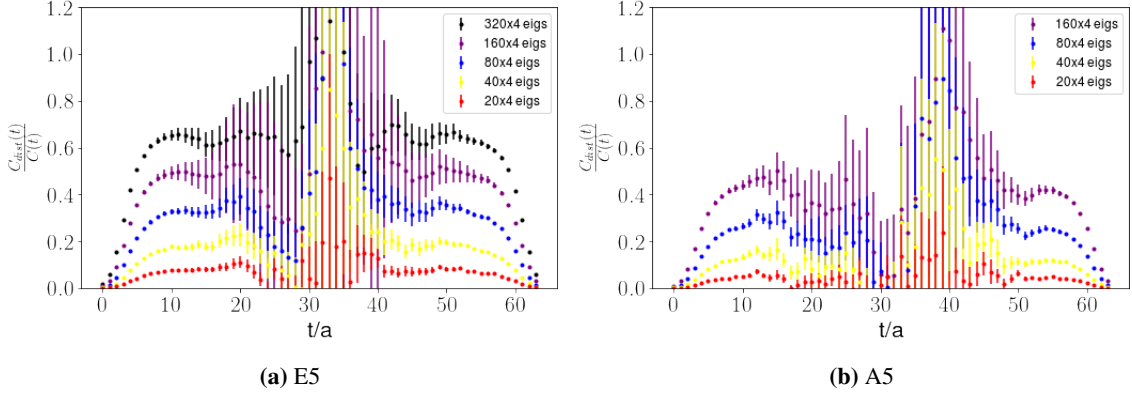
(a) E5   **(b)** A5

Figure 3: Vector correlator, $\frac{C_{dist}}{C}$ by number of distillation modes, 24 sources
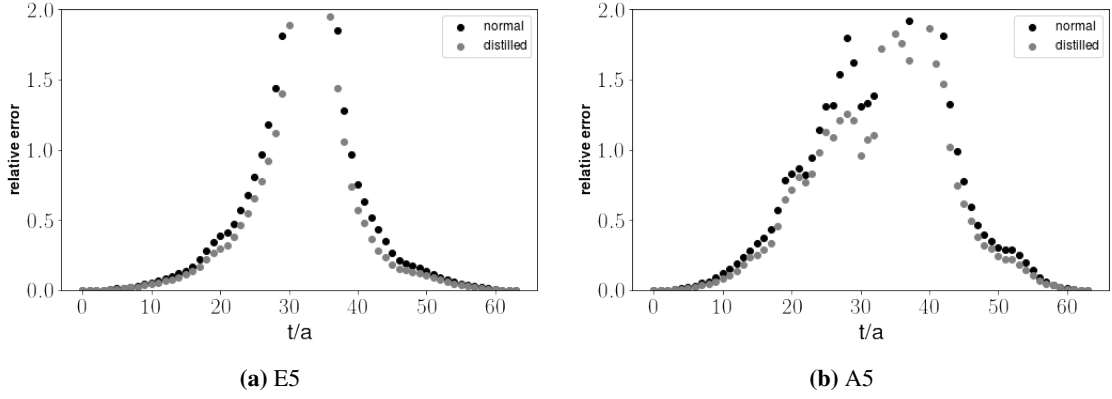

(a) E5   **(b)** A5

Figure 4: Vector correlator, $\frac{\sigma_C}{C}$, 160x4 eigenmodes, 24 sources

this idea less numerically expensive after all. To show this clearly, plots 3a and 3b graph $\frac{C_{dist}}{C}$, the contribution of the distillation sub-space to the whole correlator on the two data sets for different numbers of distillation modes. At low numbers of modes, $\frac{C_{dist}}{C}$ scales linearly with the size of the subspace. But this scaling quickly falls off, and even a computational investment of 1280 distillation modes falls short of the 0.8 mark on both ensembles.

The A5 ensemble also exhibits worse scaling overall than E5. We speculate that this is due to the larger size of the physical volume increasing the number of distillation modes below any given threshold [8]. Figures 4a and 4b plot the relative errors $\frac{\sigma_C}{C}$ of the total correlators with and without the distillation technique applied, for 160 degenerate distillation modes. As might be expected from the analysis above, there is a factor of about $\frac{1}{\sqrt{2}}$ between the two, which is consistent with reducing the stochastic noise on half of the total correlator to approximately zero, since

$$\sigma_C = \sqrt{\sigma_{C_{dist}}^2 + \sigma_{C-C_{dist}}^2} \, . \tag{5}$$

This means that if the contribution of $\sigma_{C-C_{dist}}$ to the total stochastic error budget simply scaled linearly with the size of $C - C_{dist}$, one would expect a $\frac{1}{\sqrt{2}}$ factor for $\frac{C_{dist}}{C} = 0.5$, which is what
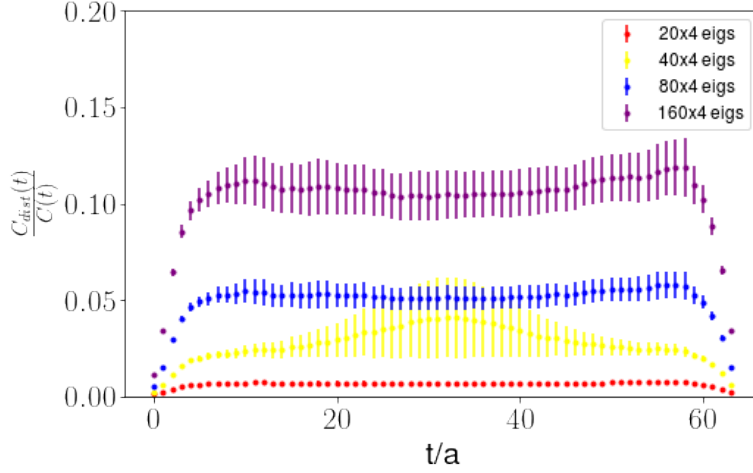
Figure 5: A5 Pseudoscalar correlator, $\frac{C_{dist}}{C}$ by number of distillation modes, 24 sources

we observe. Similar results are observed for the pseudoscalar correlator, except that the scaling of $\frac{C_{dist}}{C}$ with the number of modes is even more unfavourable. This is pictured in figure 5

3.1 Computational costs

To compare computational costs, table 2 shows the time needed to calculate the distillation eigenmodes with Primme, and C_{dist} with OpenQ*D, on a single configuration, for different numbers of eigenmodes. At low numbers of distillation modes, the computational cost is dominated by the

N_{dist}	C_{dist} Dirac solver	Lap. solver	Lap. solver Ortho.	Lap. solver MatVec
20x4	1.18e+02 s	46.3 s	3.5 s	31.1 s
40x4	2.53e+02 s	1.04e+02 s	10.8 s	68.7 s
80x4	4.84e+02 s	2.43e+02 s	41.8 s	1.52e+02 s
160x4	1.12e+03 s	7.77e+02 s	2.51e+02 s	3.46e+02 s
320x4	2.82e+03 s	2.14e+03 s	1.14e+03 s	7.40e+02 s

Table 2: Computational costs

Dirac solver. However, as the number of modes is increased, the time needed by the Primme Laplacian solver for orthogonalisation increases quadratically.

This represents another fundamental obstacle for the technique. Since a very large number of eigenmodes seems needed for C_{dist} to dominate C_{mixed} in the correlator, the time needed to solve for all the eigenmodes of the laplacian will eventually scale quadratically with the number of modes, potentially destroying the supposed advantage over normal low-mode-averaging that individual modes take less time to calculate.

If we were to use the computation time invested in distillation for additional stochasic sources instead, we would expect

$$t_{\frac{N}{4},l} + t_{N,D} + 2t_{R_1,D} = t_{R_0,D} \,, \tag{6}$$

where R_1 is the number of stochastic sources with distillation, R_0 the number possible for the same cost without distillation, $t_{N,D}$ the time required to solver the Dirac equation for the given number

of distillation modes N, $t_{R,D}$ the time required to do the same for the stochastic sources and $t_{\frac{N}{4},l}$ the time required to run the Laplacian solver for the given number of distillation modes N. The factor of 2 reflects the fact that introducing distillation forces us to solve the Dirac equation for the stochastic sources twice, with different preconditionings.

Assuming roughly equal time needed to solve the equation for any kind of Dirac source, this implies

$$R_0 > 2R_1 + N\,,\tag{7}$$

since the relative variance of the correlator scales as

$$\frac{\sigma_C{}^2}{C} \propto \frac{1}{R}\,,\tag{8}$$

we would thus need

$$1 - \frac{C_{dist}}{C} \le \frac{R_1}{R_0} \le \frac{R_1}{2R_1 + N} = \frac{1}{\frac{N}{R_1} + 2}\,,\tag{9}$$

even in the limit of assuming the time needed to find the Laplacian modes is negligible, which is clearly not the case.

In our existing tests plotted in figures 3a and 3b, this relation is never fulfilled.

4. Conclusion

We conclude that distillation-low-mode-averaging does not seem computationally efficient for the connected vector and pseudoscalar correlators. Contrary to our hopes, the similarity between the low-lying spectra of the spatial Laplacian and the Dirac operator was not strong enough to allow efficiently substituting the eigenmodes of the former for the later. Since the eigenmodes of the spatial Laplacian become more dense at larger lattice volumes, we would expect the efficiency of the technique to fall off further if the physical lattice volume were increased.

If the underlying idea of finding a more computationally friendly substitute for the Dirac operator in low-mode-averaging were to be pursued further, a possible next step could be testing operators with a non-trivial spin structure.

5. Acknowledgements

References

[1] M. Peardon, J. Bulava, J. Foley, C. Morningstar, J. Dudek, R.G. Edwards et al., *Novel quark-field creation operator construction for hadronic physics in lattice QCD*, *Physical Review D* **80** (2009) .

[2] T. DeGrand and S. Schaefer, *Improving meson two-point functions in lattice QCD*, *Computer Physics Communications* **159** (2004) 185–191.

[3] L. Giusti, P. Hernandez, M. Laine, P. Weisz and H. Wittig, *Low-energy couplings of QCD from current correlators near the chiral limit*, *Journal of High Energy Physics* **2004** (2004) 013–013.

[4] RBC & UKQCD collaboration, P. Boyle, C. Kelly, R. Kenway and A. Jüttner, *Use of stochastic sources for the lattice determination of light quark physics*, *Journal of High Energy Physics* **2008** (2008) 086–086.

[5] L. Wu, E. Romero and A. Stathopoulos, *Primme_svds: A high-performance preconditioned SVD solver for accurate large-scale computations*, *CoRR* **abs/1607.01404** (2016) [1607.01404].

[6] P. Fritzsch, F. Knechtli, B. Leder, M. Marinkovic, S. Schaefer, R. Sommer et al., *The strange quark mass and lambda parameter of two flavor QCD*, *Nuclear Physics B* **865** (2012) 397–429.

[7] M.D. Morte, A. Francis, V. Gülpers, G. Herdoíza, G. von Hippel, H. Horch et al., *The hadronic vacuum polarization contribution to the muon g-2 from lattice QCD*, *Journal of High Energy Physics* **2017** (2017) .

[8] C. Morningstar, J. Bulava, J. Foley, K.J. Juge, D. Lenkner, M. Peardon et al., *Improved stochastic estimation of quark propagation with laplacian heaviside smearing in lattice QCD*, *Physical Review D* **83** (2011) .