

Error reduction with all-mode-averaging in Wilson fermion

Eigo Shintani*

PRISMA Cluster of Excellence, Institut für Kernphysik and Helmholtz Institute Mainz, Johannes Gutenberg-Universität Mainz, D-55099 Mainz, Germany

E-mail: shintani@kph.uni-mainz.de

We report the recent development of all-mode-averaging technique (AMA) in Wilson-clover fermion. Using local deflation field as a preconditioning for generalized conjugate residual algorithm to compute the approximation of baryon correlation function, it enables us to make speed-up the computation compared to standard source-shift method. In this proceedings, we numerically search the most efficient parameters of AMA in two-flavor configurations generated by CLS group.

*The 32nd International Symposium on Lattice Field Theory
23-28 June, 2014
Columbia University New York, NY*

*Speaker.

1. Introduction

For lattice computation with Monte-Carlo simulation, accumulating enough amount of independent ensemble is necessary to provide prediction as theoretical calculation. Although lattice QCD is one of the most rigorous methods to compute non-perturbative physics corresponding to strong interaction between quark and gluon, the current numerical precision is not so enough for that required from the experiments of hadron physics and high energy physics, for instance a determination of nucleon structure function and search of high precision for the new physics. This is due to relatively large statistical error in Monte-Carlo simulation compared to experimental precision, in addition to systematic errors of unphysical lattice parameters (large pion mass and significant lattice artifact effect).

Pursuing the precision below 1% for observables of nucleon physics is a challenging task of lattice QCD, since there expects the growth of both statistical fluctuation and computational cost to approach to the physical point. Naive expectation of signal-to-noise ratio of nucleon correlation function in lattice QCD with finite N statistics shows the law as

$$S/N \sim \sqrt{N} \exp[-(m_N - 3/2m_\pi)t], \quad (1.1)$$

which indicates that the statistical fluctuation of nucleon correlation function exponentially increases depending on smallness of pion mass and extent of time-separation from source location. We confront a dilemma for the purpose of precise calculation in nucleon physics, since we need the precise value at large distance to distinguish the higher excited state contamination and ground state.

In order to enhance the statistical signal in the computation of correlation function, we recently develop the strategy of all-mode-averaging (AMA) techniques [1, 2, 3]. As a consequence of tuning of parameters in AMA, the total cost in computation of observable which is extracted from correlation function is suppressed especially for the case of large volume and small pion mass. In this proceedings we present the further study from [3] for performance test of AMA in Wilson-Clover fermion configurations. After tuning the parameter of preconditioning and iteration number of iterative solver algorithm, AMA achieves increase in statistics for computation of nucleon correlation function.

2. All-mode-averaging

We define the AMA estimator as

$$\mathcal{O}^{\text{AMA}} = \mathcal{O} - \mathcal{O}^{\text{appx}} + \frac{1}{N_G} \sum_{g \in G} \mathcal{O}^{\text{appx}g}, \quad (2.1)$$

where $\mathcal{O}^{\text{appx}}$ is an approximation of observable \mathcal{O} constructed by the sloppy inversion algorithm of Dirac operator (truncated solver) being around 10^{-3} precision. Using covariant transformation g in symmetry G on the lattice, for instance the transnational symmetry, the second term is regarded as the bias correction for low precision in $\mathcal{O}^{\text{appx}}$ and so that expectation value of \mathcal{O}^{AMA} is consistent with \mathcal{O} itself. AMA advances that, if $\mathcal{O}^{\text{appx}}$ is appropriate observable having strong correlation with original \mathcal{O} , the statistical quality of \mathcal{O}^{AMA} in N statistics is similar to N_G times statistics of

\mathcal{O} , even though the computational cost of $\mathcal{O}^{\text{appx}}$ is much less than \mathcal{O} . We have the formula of the ratio of standard deviation between \mathcal{O} and \mathcal{O}^{AMA} ,

$$\sigma^{\text{ama}}/\sigma \simeq \sqrt{N_G^{-1} + 2\Delta r + R}, \quad (2.2)$$

$$\Delta r = 1 - \frac{\langle \Delta \mathcal{O} \Delta \mathcal{O}^{\text{appx}} \rangle}{\sigma \sigma^{\text{appx}}}, \quad R = \sum_{g, g' \in G} \frac{\langle \Delta \mathcal{O}^{\text{appx} g} \Delta \mathcal{O}^{\text{appx} g'} \rangle}{\sigma^{\text{appx} g} \sigma^{\text{appx} g'}}, \quad (2.3)$$

where we use $\Delta \mathcal{O} = \mathcal{O} - \langle \mathcal{O} \rangle$. The first quantity in Eq. (2.3) represents the magnitude of correlation between \mathcal{O} and $\mathcal{O}^{\text{appx}}$, *i.e.* $\Delta r \simeq 0$ indicates $\mathcal{O}^{\text{appx}}$ is good approximation for the error reduction. The second one is a summation of correlation between $\mathcal{O}^{\text{appx} g}$ with each transformation in N_G , *i.e.* correlation in correlator with different source location.

Equation (2.2) suggests that, in order to reduce the statistical error of AMA estimator to being close to maximum reduction $1/\sqrt{N_G}$, $\mathcal{O}^{\text{appx}}$ having the small Δr and less correlation between different transformation g is needed. This relies on the parameter of algorithm in the inversion of Dirac operator with truncated iteration number of conjugate residual (CR) or conjugate gradient (CG) method. In the next section, we show our strategy for the implementation of AMA in Wilson-Clover fermion action.

3. Deflated SAP+GCR in AMA

For Wilson-Clover fermion, it is convenient to employ the combination of local deflation field with approximation generated by Schwartz alternative procedure (SAP) [4, 5]. SAP is able to construct the approximation of Wilson-Dirac operator as 2×2 matrix of which operator is divided by two local domain Λ , Λ^* with local boundary fields $\partial\Lambda$, $\partial\Lambda^*$ as shown in Figure 1, and thus Wilson-Dirac operator is approximated to be

$$D_w \simeq \begin{pmatrix} D_\Lambda & \partial D_\Lambda \\ \partial D_{\Lambda^*} & D_{\Lambda^*} \end{pmatrix}, \quad \begin{cases} D_{\Lambda^{(*)}} : \text{Dirac operator defined in domain } \Lambda^{(*)}, \\ D_{\partial\Lambda^{(*)}} : \text{Dirac operator defined in local boundary } \partial\Lambda^{(*)}, \end{cases}, \quad (3.1)$$

and which is good approximation as long as Wilson-Dirac operator has at least the nearest neighboring hopping term. As a consequence of SAP for approximated Dirac operator in Eq. (3.1), the inversion of Wilson-Dirac operator is represented as the polynomial function of Dirac operator in each domain,

$$D^{-1} \simeq M_{\text{sap}} = K \sum_{v=0}^{n_{\text{cy}}-1} (1 - DK)^v, \quad (3.2)$$

$$K = R_\Lambda^T D_\Lambda^{-1} R_\Lambda + R_{\Lambda^*}^T D_{\Lambda^*}^{-1} R_{\Lambda^*} - R_{\Lambda^*}^T D_{\Lambda^*}^{-1} D_{\partial\Lambda^*} D_\Lambda^{-1} R_\Lambda, \quad R_\Lambda^T = (1, 0), \quad R_{\Lambda^*}^T = (0, 1), \quad (3.3)$$

where n_{cy} is number of cycle of SAP relying on a quality of approximation of inversion. Here we use M_{SAP} as not only preconditioner of generalized conjugate residual (GCR) algorithm but also kernel in smoothing procedure for deflation field.

Deflation field for the preconditioning before starting the inversion of Dirac operator is also efficient to make speed-up the iterative solver. Using the projection with deflation field (see, for

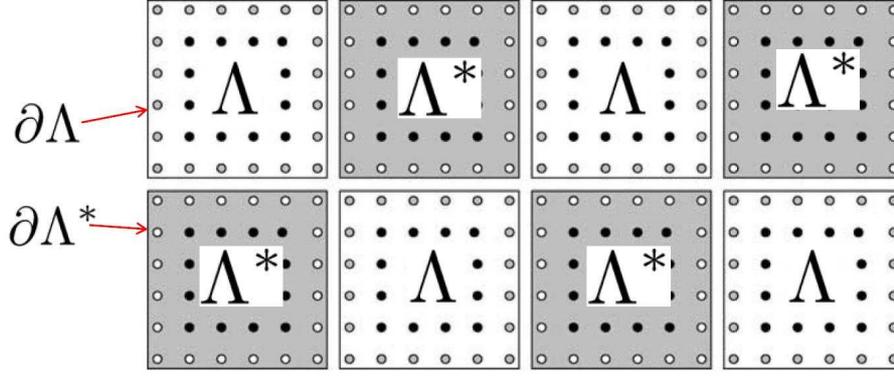


Figure 1: Domain decomposition of lattice site into $\Lambda^{(*)}$ and its boundary $\partial\Lambda^{(*)}$ for use of SAP.

example reference [5]), the condition number is improved and then total iteration number is significantly reduced. Here incorporating SAP into subspace of deflation field [5], the projection matrix and little Dirac operator is defined as

$$P = \sum_{\Lambda, \Lambda'}^{N_b} \sum_{k, l}^{N_s} R_{\Lambda}^T \phi_k^{\Lambda} (\phi_l^{\Lambda'})^{\dagger} R_{\Lambda'} (A_{kl}^{\Lambda\Lambda'})^{-1}, \quad A_{kl}^{\Lambda\Lambda'} = (\phi_k^{\Lambda}, R_{\Lambda}^T D R_{\Lambda'} \phi_l^{\Lambda'}), \quad (3.4)$$

and then in this method there are totally $N_d = N_s N_b$ deflation fields. In order to make deflation method work well, the so-called smoothing process, which is enhancement of low-mode contribution in deflation subspace, is important. The preconditioner M_{SAP} with n_{cy} is also useful for this process. Therefore, first we generate the N_s deflation field ϕ_k^{Λ} in each domain Λ after smoothing process with a few cycle in M_{SAP} starting from random field, and then use it as projection in iterative solver with Eq. (3.4). The important point is that the performance of preconditioning with deflation field trades off the solving cost of little Dirac operator, *i.e.* large space of deflation field (large number of deflation field and its domain) is required of further cost to compute $(A_{kl}^{\Lambda\Lambda'})^{-1}$ because of increasing the condition number of little Dirac operator.

In AMA, the number of deflation field and domain size is regarded as tuning parameter to control the quality of approximation in addition to the number of iteration N_{iter} in the inversion algorithm. The parameter of N_s and N_b enable us to carefully control the performance of computational cost for $\mathcal{O}^{\text{appx}}$ and magnitude of r in Eq.(2.3), which also depends on not only observable but also target time-separation of correlation function. We notice that, since domain decomposition is assumed that there is Dirichlet boundary condition in each local domain, the translational invariance of inside domain is broken by SAP approximation. This is recovered after several iterations in the inversion algorithm to achieve the enough precision and thus for \mathcal{O} there is no matter. However for $\mathcal{O}^{\text{appx}}$ it is possible to include the bias for violation of translational invariance in M_{SAP} . To avoid this issue, we use the shift of source location between the same domain Λ (or Λ^*) and same local domain site with original one, since there is no violation of translational invariance for M_{SAP} between local site in each Λ (or Λ^*). Our strategy is that we take the constraint on available region of source for $\mathcal{O}^{\text{appx}}$ relying on the size of domain (this is only case of covariant symmetry in AMA, and using in random source location [3] it also does not matter).

4. Numerical test

First we check the covariant symmetry of domain-decomposition for Dirac operator in SAP preconditioning. The consistency with source and gauge shift for the approximation,

$$\delta_c = \mathcal{O}^{\text{appx}g}[U^{\bar{g}}] - \mathcal{O}[U], \quad (4.1)$$

in which a quantity δ_c presents the error of violation of covariance, is convenient to check it configuration-by-configuration. In the numerical test, we use the domain-size of 6^4 and deflation field $N_s = 30$. We use the transformation g as the shift of source location $g = (6, 6, 0, 0)$ and gauge field $\bar{g} = (-6, -6, 0, 0)$, where those shifts correspond to even shift of domain-location. For approximation, we set the threshold of normalized norm of residual vector to 0.01 in GCR method. In Figure 2 we show δ_c for nucleon propagator at each time-separation from sink and source point. One sees that δ_c increases as large time-separation to maximally order of 10^{-5} , which is due to accumulation of round-off error, and also below 25 time-slice in lattice unit, which is appropriate signal region, δ_c is less than 10^{-6} . In the practical point of view, the error of violation of covariant symmetry is negligibly suppressed.

In the computation of nucleon form factor, for instance axial charge g_A , with AMA, we tune the parameter of solver and deflated SAP method as to close to the optimal one to make the maximum reduction of cost without large correlation Δr in Eq.(2.3). Table 1 shows the parameters for approximation in the computation of nucleon form factor (more details of lattice parameter can refer to [6]). We use a fixed number of iteration in GCR method for two-point function and quark propagator with sequential source as shown in column of ‘‘GCR iter.’’ in order to reduce the cost of computation. In three-point function, it is acceptable to make more relaxed the approximation of propagator between sink and operator point than propagator between sink and source without increase of Δr . In Figure 3, we present the comparison of performance with exact and approximation in the same ensemble. Choosing those parameters, the rate of consuming time for the inversion of Dirac matrix in the approximation is reduced by factor 5 and more, and as a consequence the total wall-time per one configuration is 34% of the exact calculation. We also numerically check the value of Δr in our parameter. In the right figure of Figure 3, Δr is below $1/N_G$ in $N_G = 64$, and also it is comparable with the value in $N_G = 128$, This result indicates that AMA is able to work for error reduction even in use of $N_G = 128$.

Figure 4 plots the scaling test of error reduction in change of the number of configuration and N_G . From $N_G = 8$ to 64 on F7 ensemble, the correlator has an effect increasing correlation between the propagators with different source locations because of decreasing the distance between them, and so that the computational cost may increase. In our test case, $N_G = 8$ case has about 50% smaller cost than other N_G , and from $N_G = 16$ there is no significant difference. On the other

Table 1: Lattice and AMA parameters for approximation.

Label	Lattice size	a fm	m_π GeV	Domain size	N_s	GCR iter. (2pt, 3pt)	N_G	at_s
F7	96×48^3	0.063	0.28	6^4	30	(3,4)	64	13
N6	96×48^3	0.05	0.33	6^4	30	(3,4)	32	22

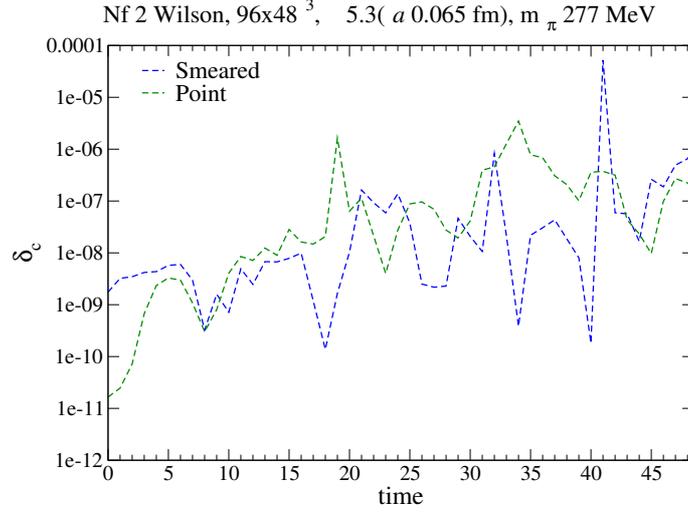


Figure 2: Error of covariance for the nucleon propagator as a function of time-separation at one configuration in 96×48^3 lattice with $m_\pi = 0.277$ GeV. Different colored lines denote the result in smeared source-sink (blue) and point source-sink (green).

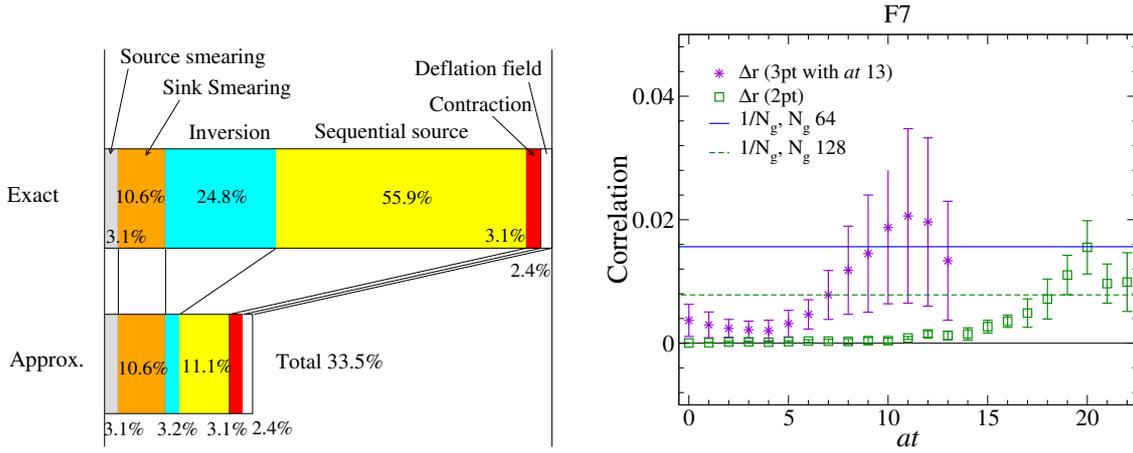


Figure 3: (Left) Comparison of computational cost with the exact calculation and approximation. (Right) Δr in two-point and three-point function of nucleon for the computation of g_A , in which the axial current is inserted. The straight lines denote the reference value for the order of error reduction in AMA.

hand, on N6 ensemble the difference between N_G is not so clear, although the error scaling is clear. Compared to the conventional method, we have roughly 1.5–3 times gain by use of AMA.

5. Summary

We report the numerical test of all-mode-averaging technique in Wilson-clover fermion using deflated SAP preconditioning. After tuning the parameter of domain size in SAP and number of deflation field in addition to fixed iteration number of solver, we have certain gain of reduction of

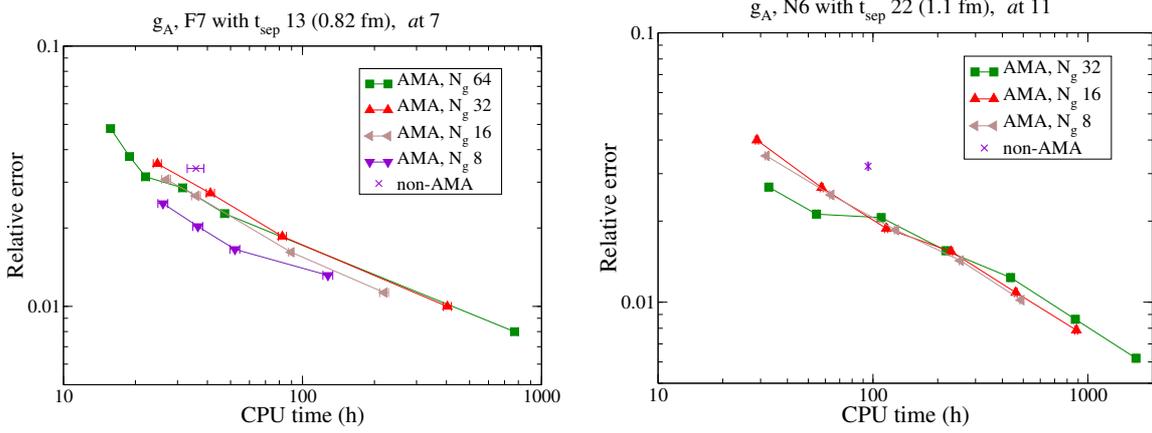


Figure 4: The relative error of axial charge at time-slice $ta = 7$ in the lattice unit versus CPU time, on which we measure in the cluster machine in Mainz University called as “Wilson”. Different symbols denote the error scaling with different number of N_G when changing the number of configurations. Left is in F7 and right is in N6 ensemble.

statistical error for the nucleon form factor calculation. Since those parameters rely on the lattice parameters, for instance lattice spacing, lattice size and quark mass, it is necessary to demonstrate it for other ensemble. Furthermore, especially for the computation of nucleon form factor, it should carefully monitor the correlation Δr in formula Eq.(2.3), since it tends to increase as large time-separation between sink, source and operator location.

Acknowledgments

Our calculations were performed on the “Wilson” and “Clover” HPC Clusters at the Institute of Nuclear Physics, University of Mainz. We thank Dalibor Djukanovic and Christian Seiwerth for technical support. This work was granted access to the HPC resources of the Gauss Center for Supercomputing at Forschungszentrum Jülich, Germany, made available within the Distributed European Computing Initiative by the PRACE-2IP, receiving funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement RI-283493 (project PRA039).

References

- [1] T. Blum, T. Izubuchi and E. Shintani, Phys. Rev. D **88**, no. 9, 094503 (2013).
- [2] T. Blum, T. Izubuchi and E. Shintani, PoS LATTICE **2012**, 262 (2012).
- [3] E. Shintani, R. Arthur, T. Blum, T. Izubuchi, C. Jung and C. Lehner, arXiv:1402.0244 [hep-lat].
- [4] M. Luscher, Comput. Phys. Commun. **156**, 209 (2004).
- [5] M. Luscher, JHEP **0707**, 081 (2007).
- [6] S. Capitani, M. Della Morte, G. von Hippel, B. Jager, A. Juttner, B. Knippschild, H. B. Meyer and H. Wittig, Phys. Rev. D **86**, 074502 (2012).