

Error reduction technique using covariant approximation and application to nucleon form factor

Thomas Blum

Physics Department, University of Connecticut, Storrs, CT 06269-3046, USA

RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, NY 11973, USA

E-mail: tblum@phys.uconn.edu

Taku Izubuchi

Brookhaven National Laboratory, Upton, NY 11973, USA

RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, NY 11973, USA

E-mail: izubuchi@quark.phy.bnl.gov

Eigo Shintani*

RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, NY 11973, USA

E-mail: shintani@riken.jp

We demonstrate the new class of variance reduction techniques for hadron propagator and nucleon isovector form factor in the realistic lattice of $N_f = 2 + 1$ domain-wall fermion. All-mode averaging (AMA) is one of the powerful tool to more effectively reduce the statistical noise of various observables rather than low-mode averaging (LMA). We adopt this technique into the two-point function and three-point function, and comparing with LMA and traditional source-shift method in the same ensembles we see AMA has much advantage of performance for these observables.

*The 30 International Symposium on Lattice Field Theory - Lattice 2012,
June 24-29, 2012
Cairns, Australia*

*Speaker.

1. Introduction

In order to precisely evaluate non-perturbative quantities in lattice calculation, the reduction in noise-to-signal ratio is one of the most important task especially for nucleon electric dipole moment [1, 2], the hadronic contribution to muon anomalous magnetic moment [3], nucleon form factors and structure functions [4], η' meson mass and mixing angle [5] and so on. We consider the new strategy to effectively increase statistics without generating the new gauge configurations using covariance in lattice symmetry.

Traditionally translational symmetry on the lattice has been used as the covariant symmetry of correlation function (correlator) with hadron interpolating operator. Since correlator in different source positions with the same distance between each local operators is exactly invariant in the infinite statistical limit, the average of several source positions can be regarded as several times statistics if there is no correlation between them. In this case, however, the additional computation of conjugate gradient (CG) to obtain quark propagator at each source positions is needed. Considering the reduction of the above computational cost, low-mode-averaging (LMA) has an advantage for low-(eigen)mode dominant observables as pseudoscalar propagator [6, 7, 8, 9, 10, 11]. In order to apply LMA to some kinds of correlator, another computational cost to obtain many low-modes to reach the dominance in each observables may be required. Thus if the *all* mode contribution should be taken into account, LMA might less work than traditional source-shift method. In fact [13, 12, 14] reported that the statistical error reduction of nucleon propagator (or heavy mesons) in LMA is less significant than pseudoscalar correlator.

Here we suggest the new class of variance reduction techniques for nucleon correlator and composite correlator [15]. Our idea is that it is not only able to cover low-mode contribution as well as LMA but also *all* modes which include the excluded modes from LMA (high-mode) are taken into account without statistical bias. All-mode averaging (AMA) will be one of the powerful technique to precisely evaluate the observables including highly composite contribution coming from low-mode and high-mode. In this proceedings we explain this idea and show some numerical results of two-point and three-point correlator in realistic lattice setup.

2. Covariant approximation averaging

The observable \mathcal{O} obtained in lattice calculation used in gauge ensemble $\{U_1, U_2, \dots, U_{N_{\text{conf}}}\}$ is represented as the statistical average;

$$\langle \mathcal{O} \rangle = \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} \mathcal{O}[U_i] + O(N_{\text{conf}}^{-1/2}), \quad (2.1)$$

where the second term denotes the uncertainty due to finite number of available gauge configurations in actual simulation. Under the transformation of $g \in G$, where G is symbolically defined as a group of transformation in lattice symmetry and the element g denotes one of the manipulation in G for the link variable, $U \rightarrow U^g$, the observable should be invariant

$$\langle \mathcal{O} \rangle = \langle \mathcal{O}^g \rangle, \quad (2.2)$$

in the infinite statistics $N_{\text{conf}} \rightarrow \infty$.

Here we introduce the approximation $\mathcal{O}^{(\text{appx})}$ which fulfills the following condition:

Appx-1:

$$r = \frac{\langle \Delta \mathcal{O} \Delta \mathcal{O}^{(\text{appx})} \rangle}{\sqrt{\langle (\Delta \mathcal{O})^2 \rangle \langle (\Delta \mathcal{O}^{(\text{appx})})^2 \rangle}} \simeq 1, \quad \langle (\Delta \mathcal{O})^2 \rangle \simeq \langle (\Delta \mathcal{O}^{(\text{appx})})^2 \rangle, \quad (2.3)$$

with $\Delta X = X - \langle X \rangle$. r is regarded as correlation between \mathcal{O} and $\mathcal{O}^{(\text{appx})}$.

Appx-2: The computational cost of $\mathcal{O}^{(\text{appx})}$ is much smaller than original \mathcal{O} .

Appx-3: $\langle \mathcal{O}^{(\text{appx})} \rangle = \langle \mathcal{O}^{(\text{appx})} g \rangle$.

Using $\mathcal{O}^{(\text{appx})}$ we construct improved estimator;

$$\mathcal{O}^{(\text{imp})} = \mathcal{O}^{(\text{rest})} + \mathcal{O}_G^{(\text{appx})}, \quad \mathcal{O}^{(\text{rest})} = \mathcal{O} - \mathcal{O}^{(\text{appx})}, \quad \mathcal{O}_G^{(\text{appx})} = N_G^{-1} \sum_{g \in G} \mathcal{O}^{(\text{appx})} g, \quad (2.4)$$

which is also satisfied with $\langle \mathcal{O}^{(\text{imp})} \rangle = \langle \mathcal{O} \rangle$ (**Appx-3**). For the above estimator the first and the second conditions are to reduce the computational cost at fixed N_{conf} , and the third one is to avoid the statistical bias. When we perform N_G times measurement of $\mathcal{O}^{(\text{appx})}$ after transformation g , for instance, by shifting source locations, the statistical error of improved estimator, $\Delta_{(\text{imp})}$, will be reduced to

$$\Delta_{(\text{imp})} \simeq \Delta \sqrt{2(1-r) + N_G^{-1}}, \quad (2.5)$$

compared with original error Δ ignoring the correlation between different $\mathcal{O}^{(\text{appx})} g$. In the case of $r \simeq 1$ (**Appx-1**), $\Delta_{(\text{imp})}$ becomes nearly $N_G^{-1/2}$ times smaller than Δ . Since the computational cost $\text{Cost}(\mathcal{O}_G^{(\text{appx})})$ is cheaper than $N_G \times \text{Cost}(\mathcal{O})$ from **Appx-2**, total cost of $\mathcal{O}^{(\text{imp})}$ is significantly reduced. The above estimator is defined as *covariant approximation averaging (CAA)*.

LMA is one of the class of CAA; $\mathcal{O}^{(\text{appx})}$ consists of low-mode, and g is a shift of source location under translational symmetry. In LMA $\mathcal{O}^{(\text{appx})}$ is correlator constructed by the inverse of Hermitian Dirac operator $S(x, y) = H^{-1}(x, y)$ (or the even-odd preconditioned counterpart), where we only present formula for the point source case for simplicity,

$$S^{(\text{low})}(x, y) = \sum_{k=1}^{N_\lambda} \lambda_k^{-1} \psi_k(x) \psi_k^\dagger(y), \quad \mathcal{O}_G^{(\text{LMA})} = \frac{1}{N_G} \sum_{g \in G} \mathcal{O}(S^{(\text{low})} g), \quad (2.6)$$

with eigenmode ψ_k and eigenvalue λ_k in $H(x, y) \psi_k(y) = \lambda_k \psi_k(x)$. **Appx-3** is by definition and **Appx-2** is fulfilled when ignoring the I/O time of stored eigenmode in disk storage or memory and time of vector-multiplication to construct $\mathcal{O}(S^{(\text{low})})$. Both conditions are (mostly) independent from observables, however the validity of **Appx-1** depends on magnitude of N_λ (from the practical point of view N_λ usually needs $O(100)$).

We propose the new class of CAA as *All-mode averaging (AMA)* [15]. Using the sloppy CG combined with low-mode deflation [10] in which the stopping condition ε of CG is made loose as $\varepsilon_{\text{AMA}} < 10^{-3} - 10^{-4}$ (or restrict a few CG iteration number)¹, the approximation is given by

$$S^{(\text{all})}(x, y) = \sum_{k=1}^{N_\lambda} \lambda_k^{-1} \psi_k(x) \psi_k^\dagger(y) + f_\varepsilon(H(x, y)) \theta(\lambda - \lambda_{N_\lambda}), \quad \mathcal{O}_G^{(\text{AMA})} = \frac{1}{N_G} \sum_{g \in G} \mathcal{O}(S^{(\text{all})} g), \quad (2.7)$$

¹With the fixed the stopping condition of CG in the specification of the approximation, $S^{(\text{all})}$. There could be a small, but finite, probability that a bias will be introduced due to the finite precision (64 bits arithmetic in our case) breaking Appx-3. This bias could be avoided by fixing the iteration number to a constant as pointed out by M.Lüscher and S. Hashimoto independently. In this proceeding, $\mathcal{O}(100)$ configurations were checked that this bias is negligible when $\varepsilon_{\text{AMA}} = 3 \cdot 10^3$.

Table 1: Parameters in LMA/AMA. We represent the maximum and minimum range of CG iteration in each ensembles.

m	N_{conf}	N_G	N_λ	ε	CG iter.	ε_{AMA}	CG iter.(AMA)
0.005	380	32	400	10^{-8}	350–360	3×10^{-3}	70–90
0.01	257	32	180	10^{-8}	600–630	3×10^{-3}	90–130

where f_ε denotes the polynomial function of H created by sequence of CG process. AMA has advantage that $S^{(\text{all})}(x, y)$ takes account of not only low-mode contribution but also (approximately) all-mode contribution which is controlled by the two parameters N_λ and ε_{AMA} . AMA also fulfills the above three conditions (**Appx-1–Appx-3**) for a much wider class of observables than LMA.

3. Numerical results

We use the $N_f = 2 + 1$ domain-wall fermion (DWF) configurations generated by RBC/UKQCD collaboration in $24^3 \times 64$ lattice at $\beta = 2.13$ Iwasaki gauge action [16]. In this configurations 5th dimension size is $L_s = 16$. We carry out CG algorithm with even-odd preconditioning at quark mass $m = 0.01, 0.005$. In the calculation of the exact eigenmode of Hermitian even-odd kernel of DWF operator we implement the implicitly restarted Lanczos algorithm with Chebychev polynomial acceleration [17]. Note that in use of even-odd bases we should take care of shift size of source point to avoid the unexpected bias into LMA/AMA estimator. Since there might not be (trivial) equivalence between even and odd DWF kernel, it is safer to use even (or odd) step of source shift manipulation g in Eq.(2.6) and (2.7) for even (or odd) DWF kernel.

In this proceedings LMA/AMA estimator is obtained in $N_G = 32$ different source locations separated by every 12 for spatial direction and 16 for temporal direction; $(0,0,0,0)$, $(12,0,0,0)$, $(12,12,0,0), \dots, (12,12,12,48)$ in the lattice unit. $(0,0,0,0)$ is an original source location. Stopping condition ε for original observable \mathcal{O} and sloppy CG ε_{AMA} for AMA are defined as $\|Hx - b\|/\|b\| < \varepsilon, \varepsilon_{\text{AMA}}$ with even-site source vector b and even-site solution vector x . Table 1 represents the each parameters. Note that the number of CG iteration is in the case of deflated CG with N_λ low-mode projection. To compare the performance, we set the Gaussian-type smearing source parameter quoted in [4]. Reference [4] has made use of traditional method in which statistics is increased by taking average over \mathcal{O} at 4 different source locations to precisely evaluate the nucleon isovector form factor (and axial-vector form factor, however we do not estimate that), and thus four times of N_{conf} are regarded as their total statistics (In addition [4] has shown the results at $m = 0.005$ with double source method which increases further two times statistics if ignoring the correlation from other source points in temporal direction. Furthermore source used in [4] is non-relativistic one and thus its computational cost reduces to the half.).

3.1 Two-point function

In Figure 1 we compare the three different time-slices for nucleon (N), pseudoscalar (P) and vector (V) meson correlator between original and LMA/AMA analysis. As mentioned before LMA takes account of statistical fluctuation of lowmode distribution of Dirac matrix in Eq.(2.6), and thus

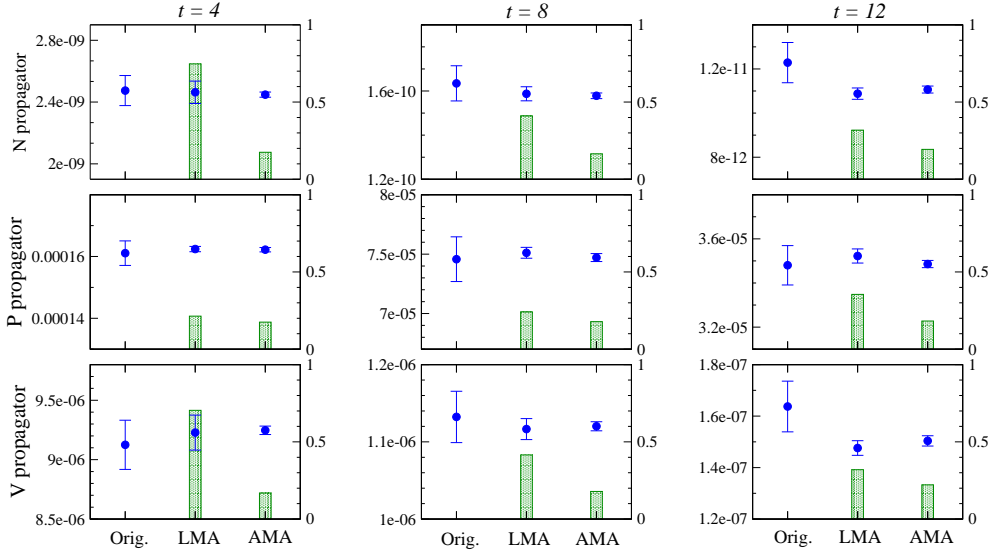


Figure 1: The comparison between LMA/AMA and original analysis for nucleon (N), pseudoscalar (P) and vector (V) meson propagator at different time-slices $t = 4, 8, 12$. The colored bar indicates the ratio of relative error between original and LMA/AMA. This is the case at $m = 0.005$.

Table 2: The nucleon mass obtained by global fitting of N correlator with Gaussian smeared sink. We use GeV unit. Computational costs of LMA/AMA and previous study [4] are estimated in the unit of original computation of CG iteration (Org) with deflation. Note that for Cost(TY) we naively scale assuming that we carry out the same implementation as [4] with deflation in 3728(1424) measurement in $m = 0.005(0.01)$. Bracket in Cost(TY) is an estimate with double source method. Last column is a gain for AMA which is ratio of scaled Cost(TY) to achieve AMA accuracy with Cost(AMA).

m_π	$m_N(\text{Orig})$	$m_N(\text{LMA})$	$m_N(\text{AMA})$	Cost(AMA)	$m_N(\text{TY})$	Cost(TY)	Gain
0.33	1.1242(223)	1.1451(87)	1.1390(38)	8.2	1.1481(100)	9.8[4.9]	8.3
0.42	1.2207(171)	1.2192(111)	1.2334(42)	6.7	1.2169(93)	5.5	4.0

it turns out that the error reduction of LMA may be significant for long distance. On the other hand, since AMA approximately includes the all mode contribution by using the sloppy CG, AMA will work well for both short and long distance effects. The above statement is clearly seen in Figure 1; From $t = 4$ to $t = 12$, the (relative) error reduction is drastically changed in LMA, however for AMA such reduction keeps close to ideal reduction rate, $1/N_G^{1/2} \simeq 0.18$, for every channels. Note that P channel seems to show a similar error reduction between LMA and AMA because this channel mostly dominates the lowmode contribution as expected in [6, 7, 8].

In Figure 2 we observe that the effective mass of nucleon in AMA becomes more stable for both point and Gaussian smeared sink rather than LMA. In Table 2 we see that the precision of nucleon mass in AMA is higher than previous study [4] while the computational cost is roughly less by 1/4 times. The detailed comparison between them including computational time of lowmode is discussed in [15].

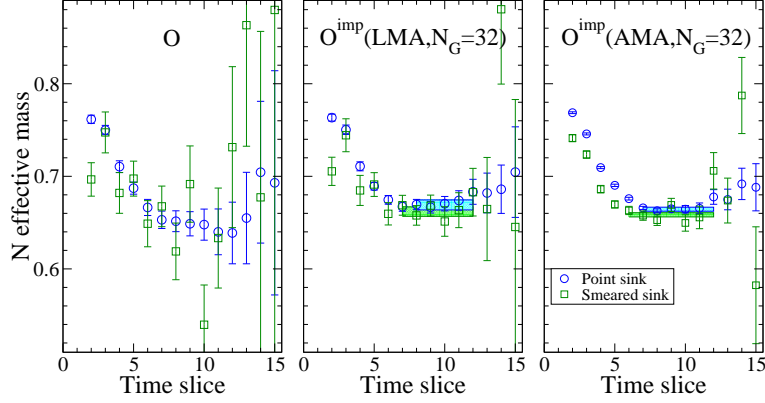


Figure 2: Effective mass plot of nucleon correlator in original, LMA/AMA analysis with point sink and Gaussian smeared sink. The colored bound shows the statistical error when globally fitting the propagator.

3.2 Nucleon isovector form factor

In order to check LMA/AMA in more complicated observables, we try to compare the three-point function in AMA and traditional method. Nucleon isovector form factor is appropriate to investigate the working of AMA because such signal is very clear and it is given by a complicated ratio of quark propagator;

$$R_\mu(t_1, t, t_0 | p_1, p_0) = K \frac{C_{J_\mu}^N(\vec{q}, t)}{C_G^N(t_1 - t_0, 0)} \left[\frac{C_L^N(t_1 - t, \vec{q}) C_G^N(t - t_0, 0) C_L^N(t_1 - t_0, 0)}{C_L^N(t_1 - t, 0) C_G^N(t - t_0, \vec{q}) C_L^N(t_1 - t_0, \vec{q})} \right]^{1/2} \quad (3.1)$$

with two-point function of nucleon of local (L) or Gaussian (G) sink, $C_{L,G}^N(t, \vec{q})$ at three dimensional momentum \vec{q} , three-point function $C_{J_\mu}^N(\vec{q}, t)$ with current J_μ , and $K = \sqrt{2(E_N + m_N)/E_N}$. Following [4] we evaluate the isovector form factor extracted from Eq.(3.1) as shown in Figure 3. Precision of $F_{1,2}(q^2)$ in AMA are more accurate than the previous results [4] at each transfer momenta q^2 . This consistency indicates that AMA will be effective in reducing errors for many lattice observables on the lattice.

4. Summary

In this proceedings we show several results of the new class of error reduction techniques in covariant approximation averaging (CAA). We suggest all-mode-averaging (AMA) in which all mode contribution is taken into account by using sloppy CG with deflation as the improved estimator instead of lowmode in low-mode averaging (LMA). AMA is applicable to broad observables including nucleon spectrum, three-point function and other composite correlator rather than LMA. We compare the nucleon mass and isovector form factor with realistic lattice size (2.5 fm^3) and light quark mass ($m_\pi \simeq 0.3\text{--}0.4 \text{ GeV}$) in $N_f = 2 + 1$ DWF configurations, and show the significant reduction of computational cost to obtain similar precision with traditional one. Calculations of the nucleon electric dipole moment and the hadronic contribution to the muon g-2 are underway.

Numerical calculations were performed using the RICC at RIKEN and the Ds cluster at FNAL. This work was supported by the Japanese Ministry of Education Grant-in-Aid, Nos. 22540301

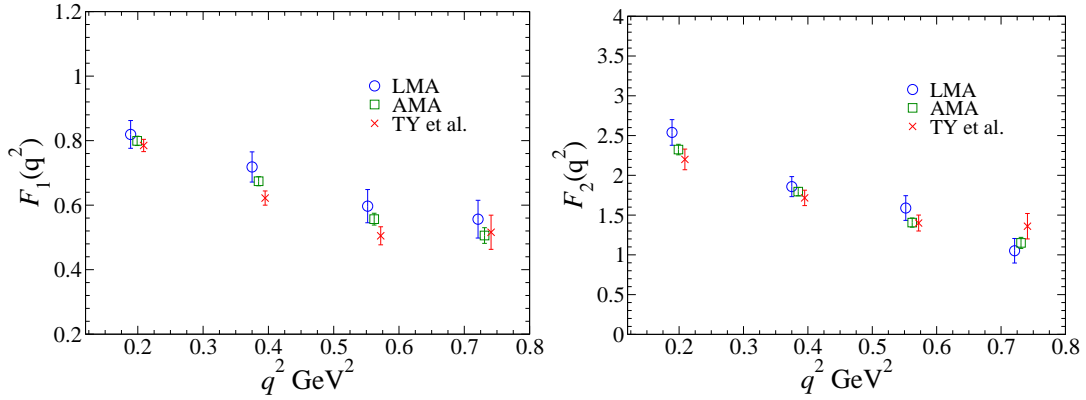


Figure 3: Isovector form factor $F_1(q^2)$ and $F_2(q^2)$ obtained in LMA/AMA and presented in TY et al.[4] at $m = 0.005$.

(TI), 23105714 (ES), 23105715 (TI) and U.S. DOE grants DE-AC02-98CH10886 (TI) and DE-FG02-92ER40716 (TB). We also thank BNL, the RIKEN BNL Research Center, and USQCD for providing resources necessary for completion of this work.

References

- [1] E. Shintani, *et al.*, Phys. Rev. D **72**, 014504 (2005); E. Shintani, *et al.*, Phys. Rev. D **75**, 034507 (2007); E. Shintani, S. Aoki and Y. Kuramashi, Phys. Rev. D **78**, 014503 (2008).
- [2] F. Berruto, T. Blum, K. Orginos and A. Soni, Phys. Rev. D **73**, 054509 (2006).
- [3] C. Aubin and T. Blum, Phys. Rev. D **75**, 114502 (2007); C. Aubin, T. Blum, M. Golterman and S. Peris, arXiv:1205.3695 [hep-lat].
- [4] T. Yamazaki, *et al.*, Phys. Rev. D **79**, 114505 (2009);
- [5] N. H. Christ, *et al.*, Phys. Rev. Lett. **105**, 241601 (2010).
- [6] L. Giusti, C. Hoelbling, M. Luscher and H. Wittig, Comput. Phys. Commun. **153**, 31 (2003).
- [7] L. Giusti, P. Hernandez, M. Laine, P. Weisz and H. Wittig, JHEP **0404**, 013 (2004).
- [8] T. A. DeGrand and S. Schaefer, Comput. Phys. Commun. **159**, 185 (2004).
- [9] T. A. DeGrand and S. Schaefer, Phys. Rev. D **72**, 054503 (2005).
- [10] M. Luscher, JHEP **0707**, 081 (2007).
- [11] H. Fukaya *et al.* [JLQCD Collaboration], Phys. Rev. Lett. **98**, 172001 (2007).
- [12] A. Li *et al.* [χ QCD Collaboration], Phys. Rev. D **82**, 114501 (2010).
- [13] L. Giusti and S. Necco, PoS LAT **2005**, 132 (2006) [hep-lat/0510011].
- [14] G. Bali, L. Castagnini and S. Collins, PoS LATTICE **2010**, 096 (2010) [arXiv:1011.1353 [hep-lat]].
- [15] T. Blum, T. Izubuchi and E. Shintani, arXiv:1208.4349 [hep-lat].
- [16] Y. Aoki *et al.* [RBC and UKQCD Collaborations], Phys. Rev. D **83**, 074508 (2011).
- [17] H. Neff, N. Eicker, T. Lippert, J. W. Negele and K. Schilling, Phys. Rev. D **64**, 114509 (2001).