Affordable low-mode averaging

Roman Gruber[1], Tim Harris[1], Marina Marinkovic[1]
[1]Institut für Theoretische Physik, ETH Zürich, Wolfgang-Pauli-Str. 27, 8093 Zürich, Switzerland

1 Introduction

We present a **new take on low-mode averaging**, where the dimension of the low-mode subspace is increased by **exploiting local coherence** of low modes. The fraction of the quark propagator captured by this subspace can easily be volume averaged or sampled excessively and reaches gauge variance with lower computational cost than using the stochastic estimator. The remainder piece can be sampled stochastically by solving the Dirac equation.

2 Observable of interest

We are interested in the **two-point connected vector correlator** needed to evaluate **HVP contribution to the muon** $g-2$, given by

$$G(t) = \frac{1}{3|\Omega|} \sum_{k=1}^{3} \sum_{y\in\Omega} \sum_{\vec{x}\in\Lambda_3} C_k(y_0+t, \vec{x}|y), \quad (1)$$

$$C_k(x|y) = \mathrm{tr}_{CD}\left[\gamma_k D^{-1}(x|y)\gamma_k D^{-1}(y|x)\right]. \quad (2)$$

→ Usually $\Omega \subset \Lambda$ with $|\Omega| \approx \mathcal{O}(10-100)$

→ Ideally $\Omega = \Lambda$ (full lattice volume average $\implies$ gauge noise reached)

3 Ensembles

We studied the CLS $N_f = 2$ ensembles [2] listed below.

Name	Size $[L^3 \times T]$	Pion mass	a [fm]	# configs
D5d	$24^3 \times 48$	439 MeV	$0.0658(10)$	100
F7	$48^3 \times 96$	268 MeV	$0.0658(10)$	100

4 Quark propagator

We can evaluate the quark propagator in different ways, i.e.

$$D^{-1}(x|y) \approx \frac{1}{N_{st}} \sum_{r=0}^{N_{st}} \psi_{[r]}(x)\eta_{[r]}(y)^\dagger, \quad \text{(stochastic)}$$

$$D^{-1}(x|y) = \sum_{i=1}^{12V} \frac{1}{\lambda_i}\gamma^5 \xi_i(x)\xi_i(y)^\dagger. \quad \text{(spectral)}$$

(stochastic) is an all-to-all propagator by solving $D\psi_{[r]} = \eta_{[r]}$ for every source.

(spectral) is obtained using the eigenmodes of the Hermitian Dirac operator $Q = D\gamma^5$, i.e. $Q\xi_i = \lambda_i \xi_i$.

The goal is to **decompose the propagator** according to the variance of $G(t)$ into two pieces,

$$D^{-1}(x|y) = D_L^{-1}(x|y) + D_H^{-1}(x|y). \quad (4)$$

5 Low-mode averaging

Low-mode averaging (LMA) [6,7] is a technique to decompose the propagator in the following way:

$$D_L^{-1}(x|y) = \sum_{i=1}^{N_s} \frac{1}{\lambda_i}\gamma^5 \xi_i(x)\xi_i(y)^\dagger, \quad (5a)$$

$$D_H^{-1}(x|y) = PD^{-1}(x|y). \quad (5b)$$

and is has found many applications in precise lattice calculations over the years, including the HVP of the muon $g-2$ [1,3,4].

- P is the projector away from the low-mode subspace
- We need to determine N_s exact low-modes of $Q = D\gamma^5$

− $N_s \approx \mathcal{O}(1000)$ on large lattices [1,3,4]

− **Very expensive** to determine $\mathcal{O}(1000)$ low modes

+ Beneficial, since eigenmodes can be recycled for other observables

- Predominantly used in physical pion mass scenarios

− **Huge memory and storage** demands

- Low-mode piece of observables is calculated **exactly**, since the subspace dimension is **only** $\mathcal{O}(1000)$
- Many different ways to calculate remainder terms (X,HH), see eq. (6)

6 Local coherence

The property of local coherence [5,8] says that low modes have *similar* local properties, i.e. **only a few block-projected low modes may represent many low modes very well** (see Figure 1).

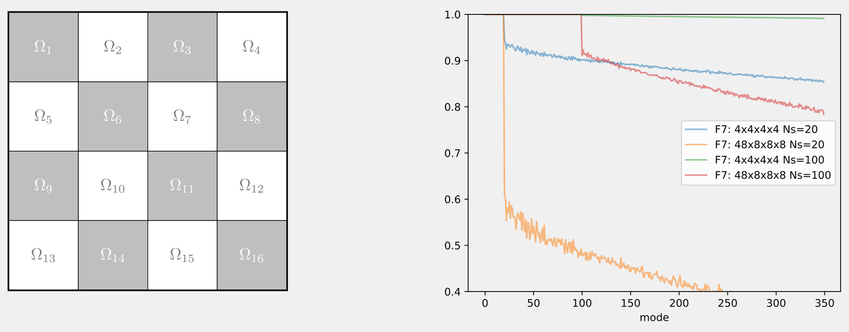

Figure 1: Block decomposition (left) and local coherence of low modes (right). The y-axis shows the percentage of how well the N-th low mode (x-axis) is represented by a basis of only a few N_s block-projected modes.

7 Method

Key idea: *Multiply the LMA-subspace dimension by the number of blocks (motivated by local coherence of low modes).*

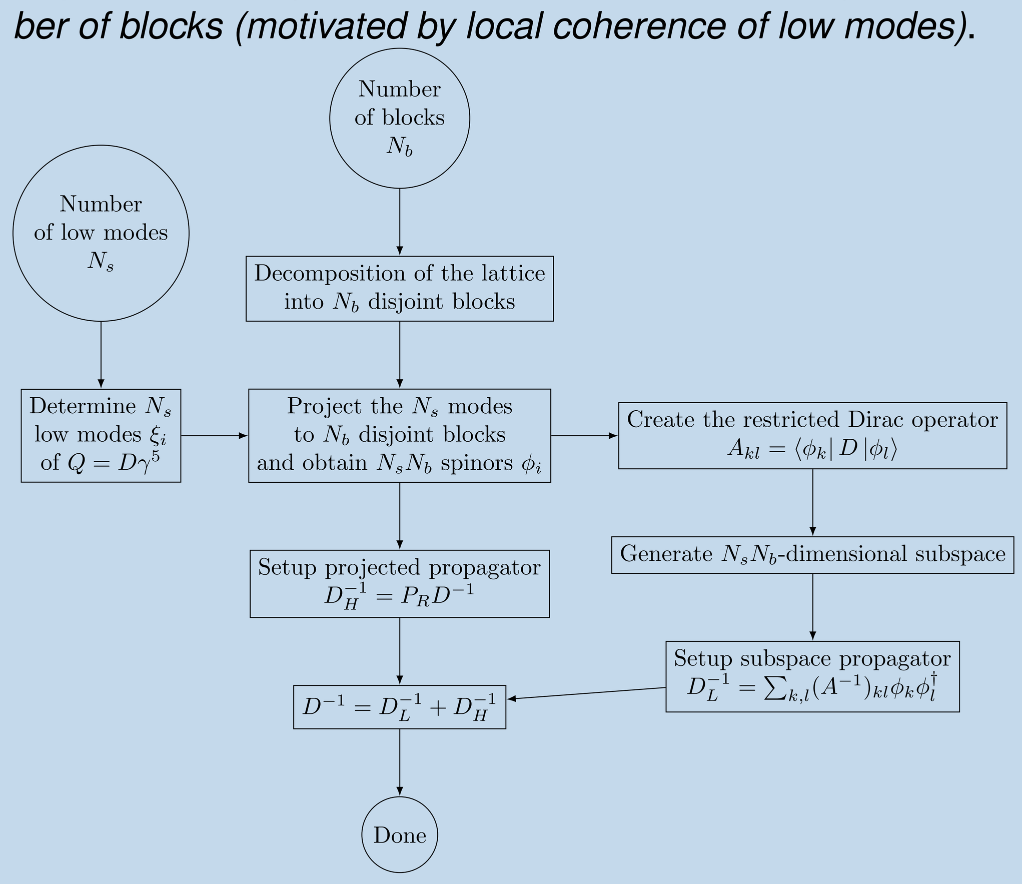

Figure 2: Schematic of the method.

+ $N_s \approx \mathcal{O}(100)$ on large lattices

+ **Significantly cheaper** than LMA, since we need only a fraction of the low modes

+ **Small memory and storage** requirements

- Low-mode piece of observables is **sampled excessively** using $100-1000$ stochastic sources, since the subspace dimension can go up to $\mathcal{O}(10^6)$ **or larger**

− **Remainder terms (X,HH)** can be estimated exactly/stochastically depending on subspace size

+ The stochastic noise on the remainder is smaller than the stochastic noise on the sum

8 Results

Using the decomposition of the quark propagator, eq. (4), for the two-point connected vector correlator, we find

$$G(t) = G_{LL}(t) + G_X(t) + G_{HH}(t), \quad (6)$$

and we study the contribution of each of the **three terms**, depending on the **number of blocks**, N_b, and the **number of low modes**, N_s.

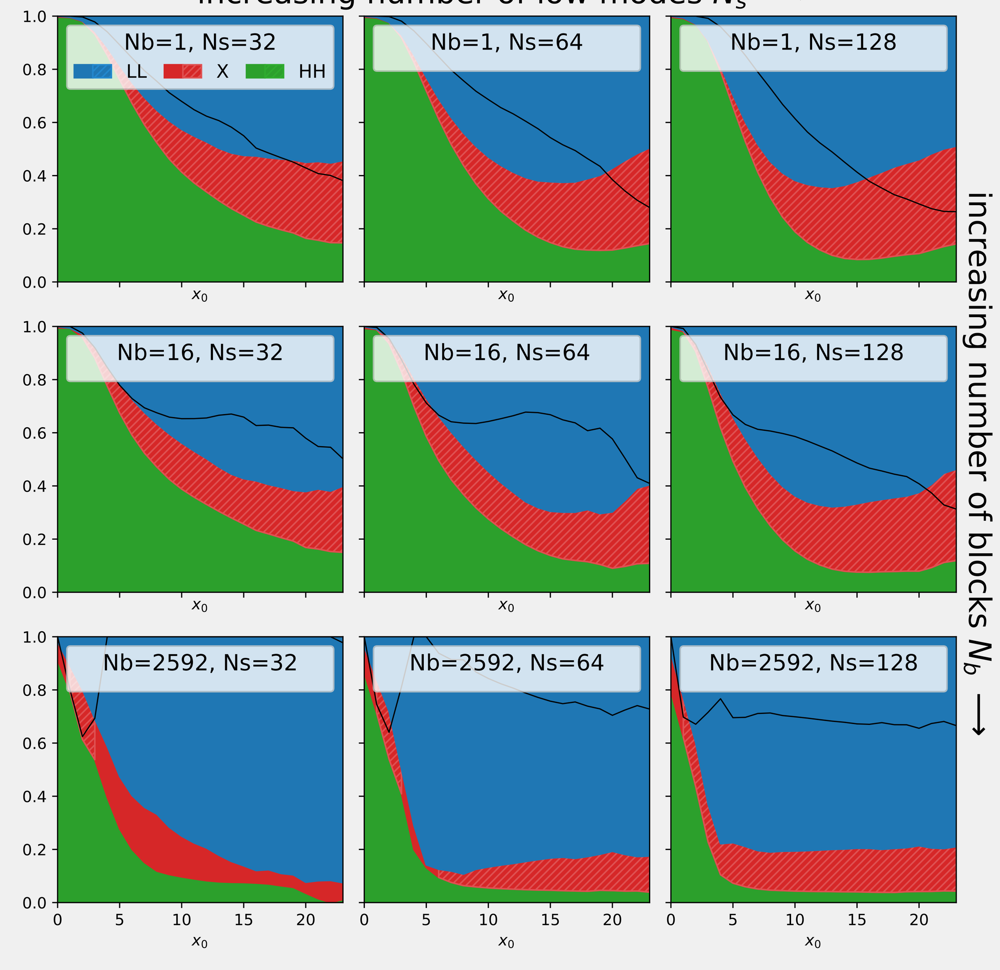

Figure 3: Contributions to the **value** of the correlator $G(t)$, eq. (6), on **D5d** lattice using different number of low modes N_s and different number of blocks N_b. The black solid line is the sum of the three terms.

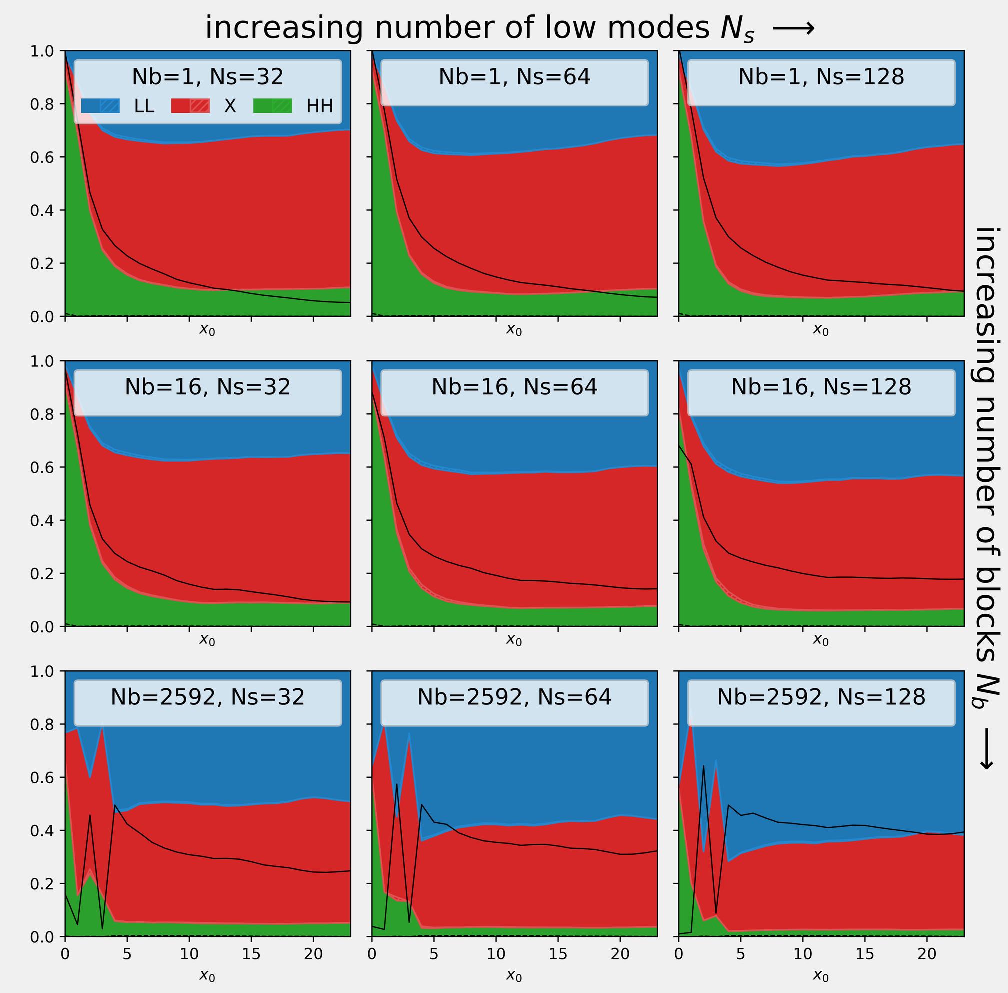

Figure 4: The relative size of the **variances** on each component (LL, X, HH) compared with their sum when estimated stochastically. Note the **large cancellation in the variance on the sum** (solid black line) due to large covariances. The block sizes $(L^3 \times T)$ are; 1st row: no blocking (LMA), 2nd row: $12 \times 12 \times 12 \times 24$, 3rd row: $4 \times 4 \times 4 \times 4$.

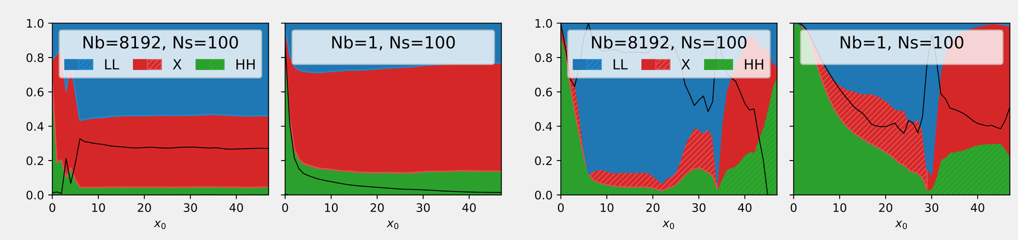

Figure 5: Contributions to the **variance (left)** and **value (right)** of the two-point vector correlator on **F7**. The block sizes $(L^3 \times T)$ are; left: $6 \times 6 \times 6 \times 6$, right: no blocking (LMA).

In Figure 6, we see the variance of the correlator, equation (1), with LMA and using the new method. For the evaluation of the X- and HH-term, we use in all cases the **same statistics (i.e. same cost, but cost of determination of low-modes not counted!)**.

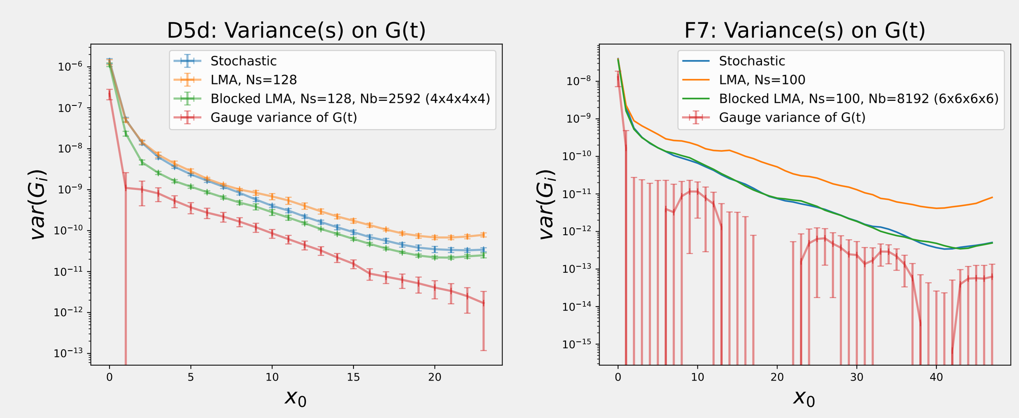

Figure 6: Comparison of variances of the pure stochastic, LMA and blocked LMA estimators. **D5d (left)** and **F7 (right)** with $N_s = 128, 100$ low modes and $N_{st} = 32, 96$ stochastic spin-diluted wall-sources, respectively.

9 Conclusion

- G_{LL} is much cheaper to sample; needs only **inversions of the little Dirac operator A** $\implies$ sample excessively
- G_{HH} has a small variance, when subspace is blocked (more blocks → smaller variance) $\implies$ sampled with low statistics
- G_X is expensive to sample and has large variance, **dominates the total variance** $\implies$ **still under investigation**, but if the LL-subspace is driven very large, we observe a reduction of variance for this term (bottom rightmost plot in Figure 4)
- On D5d: we see **small improvement** over the stochastic and the traditional LMA-estimator with same cost
- On F7: we see **no improvement**

References

[1] Aubin et al. *Physical Review D*, 101(1):014503, 2020.
[2] Blossier et al. *arXiv:2212.11206*, 2022.
[3] Blum et al. *arXiv:2301.08696*, 2023.
[4] Borsanyi et al. *Nature*, 593(7857):51–55, apr 2021.
[5] Clark et al. *EPJ Web Conf.*, 175:14023, 2018.
[6] DeGrand et al. *CPC*, 159(3):185–191, 2004.
[7] Giusti et al. *JHEP*, 2004(04):013–013, apr 2004.
[8] Lüscher. *JHEP*, 2007(07):081–081, jul 2007.