PROCEEDINGS OF SCIENCE

Frequency-splitting estimators for single-propagator traces

Leonardo Giusti[a], **Tim Harris***[a], **Alessandro Nada**[b], **Stefan Schaefer**[b]
[a] *Dipartimento di Fisica, Università di Milano-Bicocca, and INFN, Sezione di Milano-Bicocca*
Piazza della Scienza 3, I-20126 Milano, Italy
[b] *John von Neumann Institute for Computing, DESY*
Platanenallee 6, D-15738 Zeuthen, Germany
E-mail: leonardo.giusti@mib.infn.it,
tharris@mib.infn.it,alessandro.nada@desy.de,
stefan.schaefer@desy.de

In these proceedings we address the computation of quark-line disconnected diagrams in lattice QCD. The evaluation of these diagrams is required for many phenomenologically interesting observables, but suffers from large statistical errors due to the vacuum and random-noise contributions to their variances. Motivated by a theoretical analysis of the variances, we introduce a new family of stochastic estimators of single-propagator traces built upon a frequency splitting combined with a hopping expansion of the quark propagator, and test their efficiency in two-flavour QCD with pions as light as 190 MeV. The use of these estimators reduces the cost of the computation by one to two orders of magnitude over standard estimators depending on the fermion bilinear. As a concrete application, we show the impact of these findings on the computation of the hadronic vacuum polarization contribution to the muon anomalous magnetic moment.

DESY 20-010

37th International Symposium on Lattice Field Theory - Lattice2019
16-22 June 2019
Wuhan, China

*Speaker.

PoS(LATTICE2019)157

1. Introduction

Quark-line disconnected Wick contractions are ubiquitous in lattice QCD, and those involving single-propagator traces contribute to many interesting observables, such as Standard Model processes like $K \to \pi\pi$, strong and electromagnetic isospin-breaking corrections and isosinglet spectroscopy. These observables are particularly computationally challenging as they can have large vacuum contributions to their variance, as well as large random noise contributions from the auxiliary fields introduced to evaluate them stochastically [1, 2]. The former can be ameliorated by multi-level integration [3, 4], whereas in the following, we suppress the latter contribution by introducing a new family of reduced-variance stochastic estimators which are constructed using a frequency splitting and a hopping expansion applied at a large quark mass. These estimators can be applied in standard Monte Carlo simulations, and here we report on numerical tests with $N_{\mathrm{f}} = 2$ O(a)-improved Wilson fermions, where we observe between one and two orders of magntitude reduction in the variance (or cost), depending on the fermion bilinear considered. Further details and any unexplained notation can be found in ref. [5].

2. Variances of single-propagator traces

It is often sufficient to study the variance of individual disconnected Wick contractions, e.g. W_1 and W_0 (assumed to be real), whose product comprises a larger correlation function

$$\mathscr{C}_{W_1W_0}(x_1,x_0) = \Big\langle \Big[W_1(x_1) - \langle W_1(x_1)\rangle\Big]\Big[W_0(x_0) - \langle W_0(x_0)\rangle\Big]\Big\rangle \,. \tag{2.1}$$

For large $|x_1 - x_0|$ the variance factorizes into the product of the variances of the individual contractions,

$$\sigma^2_{\mathscr{C}_{W_1W_0}}(x_1,x_0) \approx \sigma^2_{\mathscr{C}_{W_1}}(x_1)\cdot\sigma^2_{\mathscr{C}_{W_0}}(x_0) + \dots\,, \quad \text{where} \quad \sigma^2_{\mathscr{C}_{W_i}}(x_i) = \Big\langle \Big[W_i(x_i) - \langle W_i(x_i)\rangle\Big]^2\Big\rangle, \tag{2.2}$$

and the ellipsis stands for exponentially suppressed terms.

A simple example is the disconnected contribution to a two-point function of fermion bilinears, whose disconnected components are the single-propagator traces

$$\bar{t}_{\Gamma,r}(x_0) = -\frac{a_\Gamma}{aL^3}\sum_{\boldsymbol{x}} \mathrm{tr}\left[\Gamma D^{-1}_{m_r}(x,x)\right] \,, \tag{2.3}$$

where D_{m_r} is the massive Dirac operator with bare-quark mass m_r, a is the lattice spacing and L^3 is the lattice volume. The factor $a_\Gamma = -\mathrm{i}$ for $\Gamma = \gamma_\mu$ and $a_\Gamma = 1$ for $\Gamma = I, \gamma_5, \gamma_\mu\gamma_5, \sigma_{\mu\nu}$, is chosen so that the Wick contraction is real. The gauge variance can be defined in terms of local operators

$$\sigma^2_{\bar{t}_{\Gamma,r}} = \frac{a^2_\Gamma}{L^3}\sum_{\mathbf{x}} a^3 \langle O_{\Gamma,rr}(0,\mathbf{x})\, O_{\Gamma,r'r'}(0)\rangle_c \,, \tag{2.4}$$

where c stands for connected correlation function, $O_{\Gamma,rs}(x) = \bar{\psi}_r(x)\Gamma\psi_s(x)$, and $m_{r'} = m_r$. It is evident that the gauge variance is itself a disconnected contraction, and so begins only at order g_0^4 or higher in perturbation theory and may be expected to be suppressed. Nevertheless, by the operator product expansion and power-counting the variance has a cubic divergence in the continuum limit.

id	L/a	κ	MDU	N_{cfg}	M_π[MeV]	$M_\pi L$
F7	48	0.13638	9600	100(1200)	268	4.3
G8	64	0.136417	820	25	193	4.1

Table 1: Overview of the ensembles and statistics presented in this study and their simulation and physics parameters.

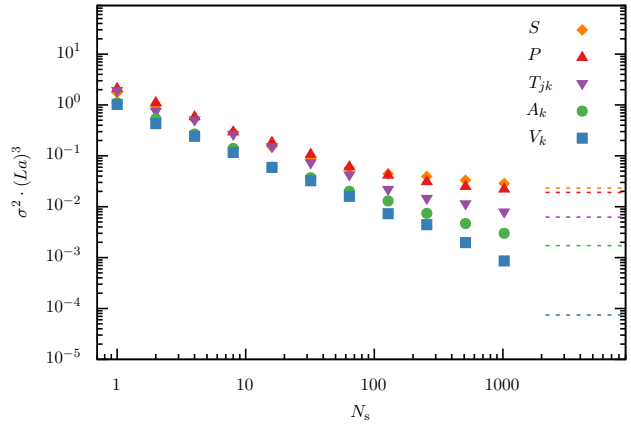

Figure 1: Variances of the standard random noise estimators defined in eq. (2.6) as a function of N_s for the ensemble F7. The symbols S, P, T_{jk}, A_k and V_k stand for $\Gamma = I$, γ_5, σ_{jk}, $\gamma_k\gamma_5$ and γ_k respectively. The dashed lines indicate the gauge noise contributions to the variances computed in sec. 4.

In practice, it is unfeasible to compute the single-propagator trace exactly, to which end we introduce N_s independent auxiliary fields $\eta_i(x)$, whose components must have unit variance and zero mean, which we choose to be drawn from a Gaussian distribution, and a stochastic estimator

$$\bar{\tau}_{\Gamma,r}(x) = -\frac{1}{aL^3N_s}\sum_{i=1}^{N_s}\sum_{\boldsymbol{x}}\text{Re}\left[a_\Gamma\eta_i^\dagger(x)\Gamma\{D_{m_r}^{-1}\eta_i\}(x)\right]. \tag{2.5}$$

The variance of the stochastic estimator receives contributions from the fluctuations of the auxiliary fields

$$\sigma^2_{\bar{\tau}_{\Gamma,r}} = \sigma^2_{\bar{t}_{\Gamma,r}} - \frac{1}{2L^3N_s}\left\{a_\Gamma^2\sum_{\mathbf{x}}a^3\langle O_{\Gamma,rr'}(0,\mathbf{x})O_{\Gamma,r'r}(0)\rangle + \frac{1}{a}\sum_x a^4\langle P_{rr'}(x)P_{r'r}(0)\rangle\right\}, \tag{2.6}$$

where $P_{rs} = O_{\gamma_5,rs}$. The second and third terms in eq. (2.6) are connected diagrams which occur at tree level in perturbation theory, which suggests that they will be larger than the gauge variance unless N_s is large. Note that the third term, which is chirally-enhanced, is independent of Γ.

In order to investigate the relative magnitude of the contributions to the variance, we employed the ensembles listed in table 1, with $N_\text{f} = 2$ flavours of non-perturbatively O(a)-improved Wilson fermions. In fig. 1 we plot the variance as a function of N_s for the scalar and pseudoscalar densities, and the tensor, axial and vector currents for the F7 ensemble. The linear behaviour in N_s^{-1} illustrates that the random noise variance dominates by orders of magnitude for small N_s, and is practically independent of the bilinear, as expected from the arguments outlined above. For large N_s, the gauge variance (dashed lines) is saturated, and is orders of magnitude larger for the densities than the currents. It is therefore highly desirable to examine variance reduction methods in particular for the currents in order to reach the gauge noise.

3. Estimators for differences of single-propagator traces

In this section, we investigate the difference of single-propagator traces $\bar{t}_{\Gamma,rs} = \bar{t}_{\Gamma,r} - \bar{t}_{\Gamma,s}$ with two different bare quark masses, $m_r < m_s$, which contains the infrared contributions to the single-propagator trace. Such differences arise, for example, from the disconnected contraction of the electromagnetic current in the isosymmetric theory, as well as constituting a component of our improved frequency-splitting estimator in sec. 4.

We examine two estimators for the differences of single-propagator traces defined by[1]

$$\bar{\theta}_{\Gamma,rs}(x_0) = -\frac{(m_s - m_r)}{aL^3N_s}\sum_{i=1}^{N_s}\sum_{\boldsymbol{x}}\mathrm{Re}\left[a_\Gamma\eta_i^\dagger(x)\Gamma\{D_{m_r}^{-1}D_{m_s}^{-1}\eta_i\}(x)\right], \tag{3.1}$$

$$\bar{\tau}_{\Gamma,rs}(x_0) = -\frac{(m_s - m_r)}{aL^3N_s}\sum_{i=1}^{N_s}\sum_{\boldsymbol{x}}\mathrm{Re}\left[a_\Gamma\{\eta_i^\dagger D_{m_r}^{-1}\}(x)\,\Gamma\,\{D_{m_s}^{-1}\eta_i\}(x)\right], \tag{3.2}$$

which we denote as the standard [6] and split-even estimators [5] in the following. Note that the cost of the two estimators is the same for a given N_s. Analogously to the previous section, their variances can be defined in terms of local operators

$$\sigma^2_{\bar{\theta}_{\Gamma,rs}} = \sigma^2_{\bar{t}_{\Gamma,rs}} - \frac{(m_s-m_r)^2}{2L^3N_s}\left\{a_\Gamma^2\sum_{y_1,\mathbf{y_2},y_3}a^{11}\langle S_{rs}(y_1)O_{\Gamma,ss'}(0,\mathbf{y_2})S_{s'r'}(y_3)O_{\Gamma,r'r}(0)\rangle + \frac{1}{a}\sum_{y_1,y_2,y_3}a^{12}\langle S_{rs}(y_1)P_{ss'}(y_2)S_{s'r'}(y_3)P_{r'r}(0)\rangle\right\}, \tag{3.3}$$

$$\sigma^2_{\bar{\tau}_{\Gamma,rs}} = \sigma^2_{\bar{t}_{\Gamma,rs}} - \frac{a_\Gamma^2(m_s-m_r)^2}{2L^3N_s}\sum_{y_1,\mathbf{y_2},y_3}a^{11}\Big\{\big\langle S_{rs}(y_1)\,O_{\Gamma,ss'}(0,\mathbf{y_2})\,S_{s'r'}(y_3)\,O_{\Gamma,r'r}(0)\big\rangle + \big\langle P_{rr'}(y_1)\,O_{\Gamma,r's'}(0,\mathbf{y_2})\,P_{s's}(y_3)\,O_{\Gamma,sr}(0)\big\rangle\Big\} \tag{3.4}$$

where $S_{rs} = O_{I,rs}$ and the gauge variance $\sigma^2_{\bar{t}_{\Gamma,rs}}$ is the fully-connected analogue of the first four-point function in eq. (3.3), just as for the single-propagator trace. In this case, the variance has a less severe linear divergence in a. The two estimators differ only between the second four-point functions in eq. (3.3) and eq. (3.4). For the standard estimator, the $\langle SPSP\rangle$ term is independent of Γ and integrated over one more time coordinate compared to the $\langle POPO\rangle$ term in the split-even estimator, which furthermore depends on Γ.

In fig. 2, we compare the variances of the standard (filled) and split-even (open) estimators for the pseudoscalar and vector channels for the difference of light- and strange-quark propagators with bare-quark masses $am_{q,r} = 0.00207$ and $am_{q,s} = 0.0189$ on the F7 ensemble. The variance of the split-even estimator is one to two orders of magnitude smaller than the standard one, and reaches the gauge noise with $N_s \sim \mathrm{O}(10)$ for the pseudoscalar channel, and O(100) in the vector channel.

Interestingly, the large gain using the split-even estimator reported here has been confirmed in ref. [7]. The partial cancellation of stochastic noise between the light- and strange-quark traces is already present in the baseline standard estimator, which is not the origin of the significant gain

[1] Here, we have used the identity $D_{m_r}^{-1} - D_{m_s}^{-1} = (m_s - m_r)D_{m_r}^{-1}D_{m_s}^{-1}$ to rewrite the difference as a product.

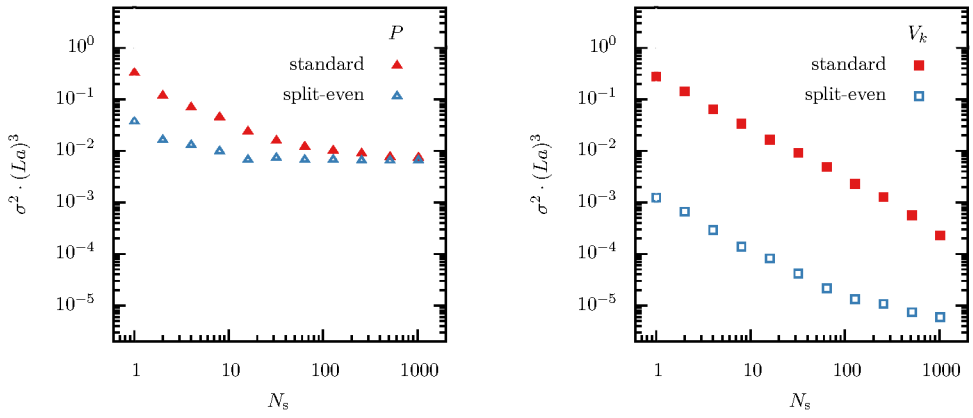

Figure 2: Variances of the standard $\theta_{\Gamma,rs}$ (filled symbols) and the split-even $\tau_{\Gamma,rs}$ (open symbols) estimators for the pseudoscalar density (left) and the vector current (right) for ensemble F7.

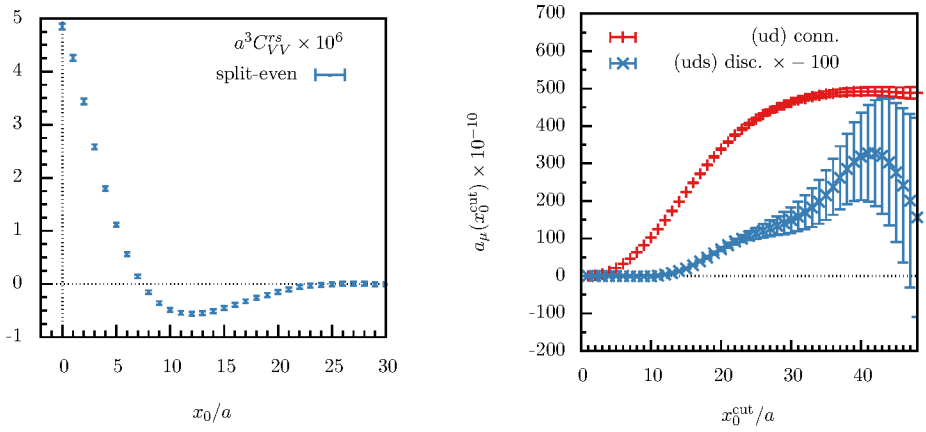

Figure 3: Left: the disconnected contribution to the electromagnetic current correlator using the split-even estimator from $N_{\rm cfg} = 1200$ gauge configurations with $N_s = 512$ noise sources for F7, and (right) the corresponding contribution to a_μ from distances up to $x_0^{\rm cut}$. A striking plateau in $x_0^{\rm cut}$ is visible in the connected case, but no evident for the disconnected contribution and moreover, the error grows quickly due to the vacuum contributions to the gauge noise, which can be tackled with multi-level integration.

as suggested there. The large reduction in the variance is well explained instead by the preceding formulæ for the variances of the two estimators. This analysis also explains the origin of the empirical gains observed for the one-end trick for the pseudoscalar density in twisted-mass QCD [8], in which case it is even possible to show the estimator has a strictly smaller variance than the standard one [5].

As an application, we consider the disconnected contribution to the electromagnetic current correlator for an isodoublet of light quarks and a valence strange quark, $C^{rs}_{VV}(x_0) \sim \langle \bar{t}_{\gamma_k,rs}(x_0 + y_0) \bar{t}_{\gamma_k,rs}(y_0) \rangle$. The disconnected contribution is shown in fig. 3 (left) using $N_{\rm cfg} = 1200$, and a good signal is observed up to 1.0 fm. The current correlator determines the hadronic vacuum polarization contribution to a_μ, and in fig. 3 (right) we show the resulting disconnected contribution (times

a factor 100, for visibility) computed using the split-even estimator, along with the light-quark connected contribution, from distances up to x_0^{cut} using the time-momentum representation [9]. In particular, this variance-reduction technique represents an important computational advance for precision computations of hadronic contributions to muon $g-2$, and as suggested in ref. [5], the use of this estimator can significantly speed up many other computations such as of hadronic matrix elements and their strong and electromagnetic isospin-breaking corrections, see ref. [10] for a first application to isospin-breaking corrections.

4. Frequency-splitting estimators

In order to construct a reduced-variance estimator for the single-propagator trace, we combine the split-even estimator and for the large quark-mass contribution use the order-$2n$ hopping expansion of the Dirac operator [11, 12], $D_{m_r}^{-1} = M_{2n,m_r} + D_{m_r}^{-1} H_{m_r}^{2n}$, where H_{m_r} denotes the hopping part of the Dirac operator and M_{2n,m_r} the first $2n-1$ terms in the hopping expansion. We denote the corresponding decomposition of the trace $\bar{t}_{\Gamma,r} = \bar{t}^M_{\Gamma,r} + \bar{t}^R_{\Gamma,r}$. The first term can be computed exactly with $24n^4$ applications of M_{2n,m_r} onto probing vectors, while an efficient stochastic estimator for the remainder of the hopping expansion is given by

$$\bar{\tau}^R_{\Gamma,r}(x_0) = -\frac{1}{aL^3N_s}\sum_{i=1}^{N_s}\sum_{\mathbf{x}} \text{Re}\left\{ a_\Gamma \left[\eta_i^\dagger H^n_{m_r}\right](x)\,\Gamma\left[D^{-1}_{m_r} H^n_{m_r}\eta_i\right](x)\right\} . \tag{4.1}$$

We therefore define the frequency-splitting estimator for the target quark mass m_1 by using the split-even estimator for $K-1$ differences with $m_{r_k} < m_{r_{k+1}}$ and applying the hopping decompositon at the largest quark mass m_{r_K} which controls the ultraviolet fluctuations,

$$\bar{\tau}^{\text{fs}}_{\Gamma,r_1}(x_0) = \sum_{k=1}^{K-1} \bar{\tau}_{\Gamma,r_k r_{k+1}}(x_0) + \bar{t}^M_{\Gamma,r_K}(x_0) + \bar{\tau}^R_{\Gamma,r_K}(x_0)\,, \tag{4.2}$$

In fig. 4, we investigate two frequency-splitting estimators for single-propagator traces at the sea-quark mass for the G8 (left) and F7 (right) ensemble respectively. For G8 we use $K=3$ and apply the hopping expansion at $am_{r_K} = 0.1$, while for F7 we take $K=5$ and $am_{r_K} = 0.3$, and $n=2$ in both cases[2]. In both cases, we see around two orders of magnitude reduction in the random-noise contributions to the variance, but due to the estimated increase of about 3.3 and 6 in the cost for G8 and F7, the cost reduction in the vector channel is in the region of $10-30$ depending on the mass.

5. Conclusions

In these proceedings, we have investigated improved split-even estimators for differences of single-propagator traces, which can be applied to efficiently evaluate the disconnected contributions arising from the electromagnetic current, so that the leading contribution to the variances arises from fluctuations of the gauge field. The cost is reduced by between one to two orders of magnitude depending on the bilinear. Furthermore, these can be used to construct frequency-splitting estimators for single-propagator traces by combining them with the hopping expansion

[2] See ref. [5] for the relative N_s used in each component.

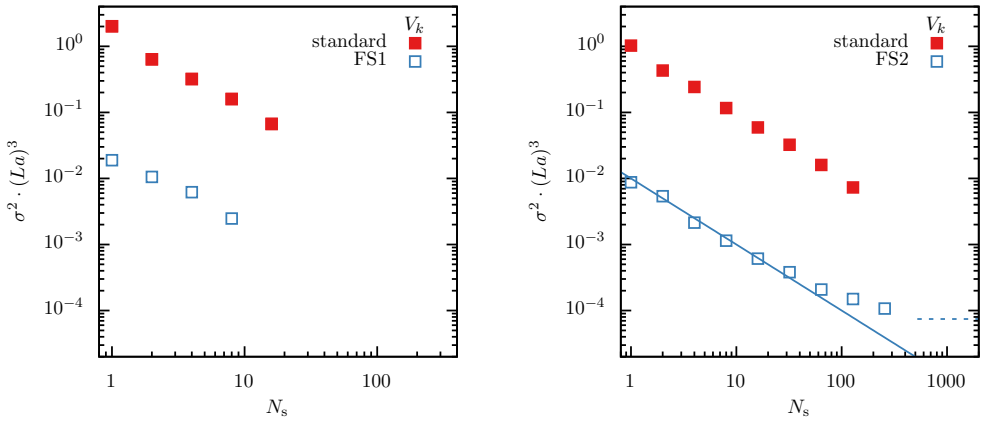

Figure 4: Variances of the frequency-splitting estimators (open symbols) for the vector current, compared with the standard random-noise estimators (filled symbols) on the G8 ensemble (left) and F7 ensemble (right). The cost of one iteration of the FS estimator is about 3.3 and 6 times the standard one on F7 and G8 respectively.

for large quark masses, which reduces the cost by more than an order of magnitude for the vector current close to the physical point. These techniques are compatible with other variance-reduction techniques, such as low-mode averaging [13, 14] and dilution [15].

Acknowledgments Simulations have been performed on the PC clusters Marconi at CINECA (CINECA-INFN and CINECA-Bicocca agreements) and Wilson at Milano-Bicocca. We are grateful to our colleagues within the CLS initiative for sharing the ensembles of gauge configurations with two dynamical flavours. L.G. and T. H. acknowledge partial support by the INFN project "High performance data network".

References

[1] K. Bitar et al. *Nucl. Phys.* B313 (1989), pp. 348–376.

[2] C. Michael and J. Peisa. *Phys. Rev.* D58 (1998), p. 034506.

[3] M. Cè et al. *Phys. Rev.* D93.9 (2016), p. 094507.

[4] M. Cè et al. *Phys. Rev.* D95.3 (2017), p. 034503.

[5] L. Giusti et al. *Eur. Phys. J.* C79.7 (2019), p. 586.

[6] V. Gülpers et al. *PoS* LATTICE2014 (2014), p. 128.

[7] H. Wittig et al. *FCCP2019 Anacapri, Capri Island, Italy, August 29-31, 2019*. 2019.

[8] P. Boucaud et al. *Comput. Phys. Commun.* 179 (2008), pp. 695–715.

[9] D. Bernecker and H. B. Meyer. *Eur. Phys. J.* A47 (2011), p. 148.

[10] A. Risch and H. Wittig. *Lattice 2019 Wuhan, Hubei, China, June 16-22, 2019*. 2019.

[11] G. S. Bali et al. *Comput. Phys. Commun.* 181 (2010), pp. 1570–1583.

[12] V. Gülpers et al. *Phys. Rev.* D89.9 (2014), p. 094503.

[13] L. Giusti et al. *JHEP* 04 (2004), p. 013.

[14] T. A. DeGrand and S. Schaefer. *Comput. Phys. Commun.* 159 (2004), pp. 185–191.

[15] J. Foley et al. *Comput. Phys. Commun.* 172 (2005), pp. 145–162.