Frequency-splitting estimators of single-propagator traces

Leonardo Giusti\textsuperscript{1,2}, Tim Harris\textsuperscript{1,2,a}, Alessandro Nada\textsuperscript{3}, Stefan Schaefer\textsuperscript{3}

\textsuperscript{1} Dipartimento di Fisica, Università di Milano-Bicocca, Milan, Italy
\textsuperscript{2} INFN, sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milan, Italy
\textsuperscript{3} John von Neumann Institute for Computing (NIC), DESY, Platanenallee 6, 15738 Zeuthen, Germany

Received: 5 April 2019 / Accepted: 14 June 2019 / Published online: 11 July 2019
© The Author(s) 2019

Abstract Single-propagator traces are the most elementary fermion Wick contractions which occur in numerical lattice QCD, and are usually computed by introducing random-noise estimators to profit from volume averaging. The additional contribution to the variance induced by the random noise is typically orders of magnitude larger than the one due to the gauge field. We propose 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. Depending on the fermion bilinear considered, the cost of computing these diagrams is reduced by one to two orders of magnitude or more with respect to standard random-noise estimators. As two concrete examples of physics applications, we compute the disconnected contributions to correlation functions of two vector currents in the isosinglet \( \omega \) channel and to the hadronic vacuum polarization relevant for the muon anomalous magnetic moment. In both cases, estimators with variances dominated by the gauge noise are computed with a modest numerical effort. Theory suggests large gains for disconnected three and higher point correlation functions as well. The frequency-splitting estimators and their split-even components are directly applicable to the newly proposed multi-level integration in the presence of fermions.

1 Introduction

Disconnected fermion Wick contractions contribute to many physics processes at the forefront of research in particle and nuclear physics: the hadronic contribution to the muon anomalous magnetic moment, \( K \rightarrow \pi \pi \) decays, nucleon form factors, quantum electrodynamics and strong isospin-breaking contributions to hadronic matrix elements, \( \eta' \) propagator to name a few. When computed numerically in lattice Quantum Chromodynamics (QCD) and if the distances between the disconnected pieces are large, their variances are dominated by the vacuum contribution. The latter are well approximated by the product of variances of the connected sub-diagrams the contractions are made of. The recently-proposed multi-level Monte Carlo integration in the presence of fermions \cite{1,2} is particularly appealing for computing disconnected contractions, since the various sub-diagrams can be computed (essentially) independently from each other, thus making the scaling of the statistical error with the cost of the simulation much more favorable with respect to the standard Monte Carlo integration.

The simplest examples of this kind are the disconnected Wick contractions of fermion bilinear two-point correlation functions, where each single-propagator trace is usually computed by introducing random-noise estimators \cite{3–5}. As the action of the auxiliary fields is already factorized, the multi-level integration in the gauge field becomes highly profitable once the variance of each connected sub-diagram is driven by its intrinsic gauge noise. The random-noise contribution, however, is typically orders of magnitude larger than the one due to the gauge field, a fact which calls for more efficient estimators in order to avoid the need of averaging over many random-noise fields with large computational cost.

The aim of this paper is to fill this gap by introducing a new family of stochastic estimators of single-propagator traces which combine the newly introduced split-even estimators with a frequency splitting and a hopping expansion of the quark propagator. We test their efficiency by simulating two-flavour QCD with pions as light as 190 MeV. As a result, depending on the fermion bilinear considered, the cost of computing single-propagator traces is reduced by one to two orders of magnitude or more with respect to the computational needs for standard random-noise estimators. The frequency-splitting estimators can be straightforwardly implemented in any standard Monte Carlo computation of disconnected Wick contractions, as well as directly combined
with the newly proposed multi-level integration in the presence of fermions.

In the next section we summarize basic facts about variances of generic disconnected Wick contractions, while those of single-propagator traces are discussed in Sect. 3. The following section is dedicated to introduce stochastic estimators of single-propagator traces of heavy quarks based on a hopping expansion of the propagator, while in Sect. 5 we introduce the split-even estimators for the difference of two single-propagator traces also relevant for the muon anomalous magnetic moment. The frequency-splitting estimators are introduced in Sect. 6, where also the outcomes of their numerical tests are reported. In Sect. 7 we discuss the impact of these findings on two concrete examples of physics applications: the disconnected contributions to the correlator of single-propagator traces are discussed in Sect. 3. Then followings sections are dedicated to introduce stochastic estimators of generic disconnected Wick contractions, while those of Wick contractions made of two sub-diagrams 1 decreases (essentially) proportionally to the inverse of the square root of the cost of the simulation. The overall statistical error on \( C_{W_1 W_0} \) thus scales with the inverse of the cost rather than with its square root. The above argument can be iterated straightforwardly to multi-disconnected contractions.

Maybe the simplest example of this kind is a disconnected Wick contraction of the correlator of two bilinear operators for which, following Eq. (2.3), the variance is well approximated by the product of variances of two single-propagator trace estimators.

### 3 Single-propagator traces

The single traces we are interested in are

\[
\langle t_{r,s}(x) \rangle = -\frac{\alpha_t}{a^4} \text{tr} \left[ \Gamma D_{m_r}^{-1}(x,x) \right],
\]

where \( D_{m_r} \) is the massive Dirac operator with bare quark mass \( m_r \) (for definiteness we adopt the \( O(a) \)-improved Wilson–Dirac operator, see Appendix A), \( a \) is the lattice spacing, the factor

\[
a_t = \begin{pmatrix} 1 & \Gamma = I, \gamma_5, \gamma_\mu \gamma_5, \sigma_{\mu\nu} \\ -i & \Gamma = \gamma_\mu \end{pmatrix}
\]

is chosen so that \( t_{r,s}(x) \) is real, and \( \sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu] \). We are interested in the zero three-momentum field \( \tilde{t}_{r,s}(x_0) \), whose expectation value is

\[
\langle \tilde{t}_{r,s}(x_0) \rangle = \frac{1}{L^3} \sum_x a^3 \langle t_{r,s}(x) \rangle,
\]

where \( \langle \tilde{t}_{r,s}(x_0) \rangle \) is a quark flavour of mass \( m_r \), and \( L^3 \) is the three-dimensional lattice volume. The variance of \( \tilde{t}_{r,s}(x_0) \),

\[
\sigma^2_{\tilde{t}_{r,s}} = \langle \tilde{t}_{r,s}^2(x_0) \rangle - \langle \tilde{t}_{r,s}(x_0) \rangle^2,
\]

can be written as

\[
\sigma^2_{\tilde{t}_{r,s}} = \frac{a_t^2}{L^3} \sum_x a^3 \langle O_{r,s}(0,x) \rangle \langle O_{r,s}(0) \rangle_c.
\]

\(^1\) Without loss of generality we assume \( W_i(x) \) to be real.
the subscript \( c \) stands as usual for connected, and \( \psi_r \) is a second flavour\(^3\) of mass \( m_r = m_r \). The operator product expansion would predict generically that \( \sigma^2_{\Gamma, \tau} \) diverges as \( a^{-3} \). There are exceptions, however, depending on the symmetries preserved by the regularization and on the operator implemented\(^4\). Moreover \( \sigma^2_{\Gamma, \tau} \) vanishes in the free-theory limit \( g_0 \rightarrow 0 \), and the first non-zero contribution appears at \( O(g^3_0) \) or higher in perturbation theory.

### 3.1 Random-noise estimator

We introduce random auxiliary fields (random sources)\([3, 5]\) defined so that all their cumulants are null with the exception of the two-point functions which satisfy

\[
\langle \psi_r(x) \psi^b_r(y) \rangle^s = \delta^{ab} \delta_{yy} \delta_{xx},
\]

where \( a, b \) and \( y, \delta \) are colour and spin indices respectively, and \( x, y \) are lattice coordinates. By using Eq. (3.8), it is straightforward to prove that a random-noise estimator of \( \tau_{\Gamma, \tau} \) is given by

\[
\tau_{\Gamma, \tau}(x) = -\frac{1}{a^4 N_s} \sum_{i=1}^{N_s} \text{Re} \left[ a_r \eta_i(x) \Gamma [D^{-1}_m \eta_i](x) \right],
\]

where \( \eta_i \) are \( N_s \) independent sources (colour and spin indices omitted from now on). The variance of the zero-momentum estimator

\[
\tilde{\sigma}^2_{\Gamma, \tau}(x_0) = \frac{1}{L^3} \sum_{x} a^3 \tau_{\Gamma, \tau}(x),
\]

reads

\[
\sigma^2_{\Gamma, \tau} = \frac{\sigma^2_{\Gamma, \tau}}{2 L^3} \left[ a^2 \sum_{x} a^3 \left\{ O_{\Gamma, \tau}(0, x) O_{\Gamma, \tau}(0) \right\} \right] + \frac{1}{a} \sum_{x} a^4 \langle P_{\Gamma, \tau}(x) P_{\Gamma, \tau}(0) \rangle,
\]

where again \( \psi_r \) and \( \bar{\psi}_r \) are two degenerate flavours of mass \( m_r \), and to simplify the notation we have introduced the usual definition \( P_{\Gamma, \tau} = O_{\Gamma, \tau} \) for the pseudoscalar density (no time-dilution is used since we are interested in the estimator at all times). The random-noise contribution to the variances in Eq. (3.11) diverges proportionally to \( a^{-3} \) like the gauge one. Both integrated correlators on the r.h.s. of Eq. (3.11), however, are colour enhanced with respect to the gauge noise and they are of \( O(1) \) in the free theory, see Appendix B. The \( \Gamma \)-dependent contribution is indeed the flavour-connected counterpart of the disconnected contraction appearing in Eq. (3.6). The \( \Gamma \)-independent term \((PP)\), which is also integrated over the time-coordinate, diverges proportionally to \( m_{\tau}^{-1} \) when \( m_{\tau} \rightarrow 0 \) due to the pion pole, giving large contributions to the stochastic variances of all bilinears indistinctly. It is interesting to notice that if we would not take the real part in Eq. (3.9), the variances would be larger and \( \Gamma \)-independent since the \( \langle O_{\Gamma, \tau}(\tau r, \Gamma r') \rangle \) contributions are dropped, and the prefactor \( 1/(2 N_s) \) goes into \( 1/N_s \).

The random-noise contributions to the variances of the standard stochastic estimators in Eq. (3.9) are thus expected to be much larger than the gauge-noise with large ultraviolet and infrared divergent terms.

### 3.2 Numerical tests

To test the efficiency of the various stochastic trace estimators considered in this paper, we have simulated QCD with two dynamical flavours discretized by the Wilson gluonic action and the \( O(a) \)-improved Wilson–Dirac operator as defined in Appendix A. The details of the ensembles of configurations considered, all generated by the CLS community\([7–9]\), are listed in Table 1. The bare coupling constant is always fixed so that \( \beta = 6/\pi g_0^2 = 5.3 \), corresponding to a spacing of \( a = 0.065 \text{ fm} \). All lattices have a size of \( 2 L \times 3 \), periodic boundary conditions for gluons, (anti-)periodic boundary conditions in (time) space directions for fermions, and spatial dimensions always large enough so that \( M_\pi L \geq 4 \). The pion mass ranges from 190 to 440 MeV. We have always skipped an enough number of molecular dynamics units (MDU) between two consecutive measurements so that gauge-field configurations can be considered as independent in the statistical analysis, see Refs.\([9,10]\) and references therein for more details.

The first primary observables that we have computed are the estimators in Eq. (3.10) with Gaussian random noise. Their variances are shown in Fig. 1 as a function of the number of random-noise sources \( N_s \) for the ensem-
and the subscript \(m\) has been omitted in the block matrices of the even-odd decomposition of the Dirac operator, see Appendix A for further details. The zero three-momentum single-propagator traces in Eq. (3.3) can thus be decomposed as

\[
\tilde{t}_{r,j}(x_0) = \tilde{t}_{r,j}^{M}(x_0) + \tilde{t}_{r,j}^{R}(x_0),
\]

where

\[
\tilde{t}_{r,j}^{M}(x_0) = -\frac{d_{r}}{aL^{3}} \sum_{x} \text{tr}[\Gamma M_{n,m_{r}}(x,x)],
\]

collects the first \(2n\) contributions of the HPE while

\[
\tilde{t}_{r,j}^{R}(x_0) = -\frac{1}{aL^{3}N_{s}} \sum_{x} \sum_{i=1}^{N_{s}} \text{Re} \left\{ a_{x} [\eta \gamma_{i} H^{R}_{m_{r}}(x)] \Gamma \left[ D_{m_{r}}^{-1} H^{n}_{m_{r}}(\eta \gamma_{i})(x) \right] \right\}.
\]

A rough idea of the variance reduction achieved by the HPE can be obtained in the free lattice theory, see Appendix B. For a bare subtracted quark mass of \(am_{q} = 0.3\) and for \(n = 2\), the stochastic variances of the remainder \(\tilde{t}_{r,j}^{R}\) are between one and two orders of magnitude smaller than those of the standard estimators \(\tilde{t}_{r,j}\). For \(n = 4\) a further reduction of approximately 4–8, depending on the bilinear, is obtained. If we had defined the estimator of the remainder by applying \(H^{2n}_{m_{r}}\) to one source only, the variance in the free case would increase approximately by a factor 2 or so. The ultraviolet filtering of \(H^{2n}_{m_{r}}\) on both random sources is thus beneficial with respect to applying \(H^{2n}_{m_{r}}\) to one source only.

4 Hopping expansion of single-propagator traces

To investigate the contribution to trace variances from high-frequency modes of the quark propagator, we first consider single-propagator traces of heavy quarks. In this kinematic regime the hopping expansion (HPE) is known to lead to a significant reduction of the random-noise contribution to trace variances [11–13]. For the \(O(a)\)-improved Wilson-Dirac operator, it is natural to exploit the even-odd decomposition to generalize the hopping parameter expansion [14] to

\[
D_{m}^{-1} = M_{2n,m} + D_{m}^{-1} H^{2n}_{m},
\]

where

\[
M_{2n,m} = \frac{1}{D_{\text{ee}} + D_{\text{oo}}} \sum_{k=0}^{2n-1} H^{k}_{m},
\]

\[
H_{m} = - \left[ D_{\text{ee}} D_{\text{oo}}^{-1} + D_{\text{ee}} D_{\text{ee}}^{-1} \right],
\]

and data for the E5 and the G8 lattices show the same qualitative behaviour. Variances go down linearly in \(1/N_{s}\) until the random-noise contribution becomes negligible, see Eq. (3.11), after which a plateau corresponds to the gauge noise (dashed lines). The first clear message from the data is that the random-noise contribution to the variances is comparable for the various bilinears, as suggested by Eq. (3.11), and it is orders of magnitude larger than the gauge noise. Moreover, the latter can vary by orders of magnitude among the various bilinears, see Sect. 6, with the densities having the largest gauge noise while the currents the smallest one.
A clear picture emerges: the bulk of the random-noise contribution to $\sigma_{\tau_{r,s}}^2$ is due to $M_{2n,m}$, for all bilinears. Once the latter is subtracted from the propagator and its contribution to $\tau_{r,s}$ is computed exactly, the random noise is reduced by approximately one order of magnitude or more. Notice that $\sigma_{\tau_{r,s}}^2$ is from 2 (pseudoscalar) up to 5 (vector) orders of magnitude smaller than $\sigma_{\tau_{r,s}}^2$ for $N_s = 1$.

In Fig. 3, $\sigma_{\tau_{r,s}}^2$ for $N_s = 1$ and $\sigma_{\tau_{r,s}}^2$ multiplied by 10 both for $n = 2$ are shown as a function of the valence bare subtracted quark mass $am_{q,r}$ for the pseudoscalar density and the spatial component of the vector current. As expected the variance reduction due to the subtraction of $M_{2n,m}$ gets larger and larger at heavier quark masses. In particular at $am_{q,r} = 0.3$ the variance of the remainder is approximately one order of magnitude smaller than at the sea quark mass value of $am_{q,r} = 0.00207$. It is worth noting that even at this light mass, the random-noise contribution to $\sigma_{\tau_{r,s}}^2$ from $M_{2n,m}$ is still significant for all bilinears. The variance reduction due to HPE, however, is only a factor 2 or so.

All in all data suggest that at heavy masses an efficient estimator of $s_{r,s}$ is obtained by computing $\tau_{r,s}$ exactly and the remainder via the stochastic estimator $\tau_{r,s}^{R}$. Which is the optimal order $n$ and how many random sources $N_s$ are required for the remainder depend on the bilinear considered and on the final target observable of interest, see Sect. 6.

5 Differences of single-propagator traces

To analyse the contribution to trace variances from low-frequency modes of the quark propagator, we consider the difference of two single-propagator traces with different
masses. It is worth noting, however, that often the difference itself is a sub-diagram of the correlator of interest, e.g. the disconnected contribution to the hadronic vacuum polarization from the up, down and strange quarks in the exact isospin limit. The estimator of the difference of two single-propagator traces reads

\[ t_{\tau,rs}(x) \equiv t_{\tau,rs}(x) - t_{\tau,rs}(x) = -\frac{a_t}{a^4} \left[ \Gamma(D_{m_2}^{-1}(x,x) - D_{m_3}^{-1}(x,x)) \right] = -\frac{a_t}{a^4}(m_2 - m_3)\left[ \Gamma D_{m_2}^{-1}D_{m_3}^{-1}(x,x) \right], \tag{5.1} \]

where \( m_2 \neq m_3 \). Its expectation value can be written as

\[ s_{\tau,rs} \equiv s_{\tau,rs} = a_t \left\{ \langle O_{\tau,rs}(x) \rangle - \langle O_{\tau,rs}(x) \rangle \right\} = a_t(m_2 - m_3)\sum_y a^4\langle S_{rs}(y) O_{\tau,rs}(x) \rangle, \tag{5.2} \]

where, to simplify the notation, we have introduced the usual notation \( S_{rs} = O_{\tau,rs} \) for the scalar density. If we define the zero three-momentum field as

\[ \tilde{t}_{\tau,rs}(x_0) = \frac{1}{L^3} \sum_x a^3 t_{\tau,rs}(x), \tag{5.3} \]

its variance is given by

\[ \sigma_{\tilde{t}_{\tau,rs}}^2 = \frac{a_t^2}{L^3} (m_2 - m_3)^2 \times \sum y_1 y_2 y_3 a^{11} \langle S_{rs}(y_1) O_{\tau,rs}(0, y_2) S_{\psi'}(y_3) O_{\tau,rs'}(0) \rangle, \tag{5.4} \]

where two extra valence fermions \( \psi_r \) and \( \psi_{r'} \), with masses \( m_{\psi_r} = m_r \) and \( m_{\psi_{r'}} = m_{r'} \), respectively, are introduced if not already present in the theory. This time the operator product expansion generically predicts that \( \sigma_{\tilde{t}_{\tau,rs}}^2 \) diverges as \( a^{-1} \), i.e. two powers less than in Eq. (3.6) thanks to the presence of the squared-mass difference in the prefactor. Analogously to Sect. 3, there are exceptions depending on the symmetries preserved by the regularization and on the discretization chosen for the operator, and \( \sigma_{\tilde{t}_{\tau,rs}}^2 \) vanishes in the free-theory limit with the first non-zero contribution appearing at \( O(g_0^3) \) or higher in perturbation theory.

### 5.1 Standard random-noise estimator

Maybe the simplest random-noise estimator of \( s_{\tau,rs} \) is

\[ \theta_{\tau,rs}(x) = -\frac{(m_2 - m_3)}{a^4 N_s} \sum_i \text{Re} \left\{ a_t \psi^+_i(x) \Gamma(D_{m_r}^{-1}D_{m_r}^{-1} \eta_i)(x) \right\}. \tag{5.5} \]

where the variance of its zero three-momentum counterpart

\[ \theta_{\tau,rs}(x_0) = \frac{1}{L^3} \sum_x a^3 \theta_{\tau,rs}(x) \]

is

\[ \sigma_{\theta_{\tau,rs}}^2 = \sigma_{\tilde{t}_{\tau,rs}}^2 - \frac{(m_2 - m_3)^2}{2L^3 N_s} \times \sum y_1 y_2 y_3 a^{11} \langle S_{rs}(y_1) O_{\tau,rs}(0, y_2) S_{\psi'}(y_3) O_{\tau,rs'}(0) \rangle + \langle P_{\tau,rs}(y_1) O_{\tau,rs'}(0, y_2) P_{\psi'}(y_3) O_{\tau,rs'}(0) \rangle, \tag{5.6} \]

Generically the random-noise contribution on the r.h.s of (5.7) diverges proportionally to \( a^{-1} \) like the gauge variance. The \( \Gamma \)-independent contribution (\( \{S P S P\} \)) is one of the spectral sums introduced in Ref. [15]. It is integrated over one time-coordinate more with respect to the first term, and it gives large contributions to the stochastic variances of all bilinears indistinctly. If we would not take the real part in Eq. (5.5), the variances would be larger and \( \Gamma \)-independent since the \( \langle S O S O \rangle \) contributions are dropped, and the prefactor \( 1/(2N_s) \) goes into \( 1/N_s \).

### 5.2 Split-even random-noise estimator

An alternative random-noise estimator of the difference of two traces is

\[ \tau_{\tau,rs}(x) = -\frac{(m_2 - m_3)}{a^4 N_s} \sum_i \text{Re} \left\{ a_t \psi^+_i(x) \Gamma(D_{m_r}^{-1} \eta_i)(x) \right\}, \tag{5.7} \]

The corresponding zero three-momentum field is

\[ \tilde{\tau}_{\tau,rs}(x_0) = \frac{1}{L^3} \sum_x a^3 \tau_{\tau,rs}(x), \tag{5.8} \]

and its variance reads

\[ \sigma_{\tilde{\tau}_{\tau,rs}}^2 = \sigma_{\tilde{t}_{\tau,rs}}^2 - \frac{(m_2 - m_3)^2}{2L^3 N_s} \times \sum y_1 y_2 y_3 a^{11} \langle \{ S_{rs}(y_1) O_{\tau,rs}(0, y_2) S_{\psi'}(y_3) O_{\tau,rs'}(0) \} + \langle P_{\tau,rs}(y_1) O_{\tau,rs'}(0, y_2) P_{\psi'}(y_3) O_{\tau,rs'}(0) \rangle \rangle, \tag{5.9} \]

where again two extra valence fermion fields \( \psi_r \) and \( \psi_{r'} \) with masses \( m_{\psi_r} = m_r \) and \( m_{\psi_{r'}} = m_{r'} \) are introduced if not already present in the theory. Also this time the operator product expansion predicts generically that \( \sigma_{\tilde{\tau}_{\tau,rs}}^2 \) diverges as \( a^{-1} \), but with respect to the standard random-noise estimator the
(large) $\Gamma$-independent spectral sum $\langle SPM \rangle$ is absent. The first four-point correlation function on the r.h.s. of Eq. (5.10) is the flavour-connected counterpart of the disconnected contraction appearing in Eq. (5.4), the second is analogous but with scalar densities replaced by the pseudoscalar ones and two flavour indices exchanged. Both integrated correlators on the r.h.s. of Eq. (5.10) are colour enhanced with respect to the gauge noise and they are of $O(1)$ in the free theory, see Appendix B.

The random-noise contributions to the variances of the split-even estimators in Eq. (5.9) are thus expected to be significantly smaller than for the standard estimators\(^5\) of differences of single-propagator traces. This is not surprising since in this case both sources, $\eta$ and $\eta^\dagger$, are ultraviolet filtered by a quark propagator and the variance has one integral less in the time-coordinate analogously to the case of time-diluted sources.\(^6\) With respect to the gauge variance, however, the random-noise contribution is still expected to be larger.

5.3 Numerical tests

We have computed the two random-noise estimators in Eqs. (5.6) and (5.9) on all ensembles listed in Table 1 and for several pairs of quark masses. For F7 and for the bare valence masses $am_{q,v} = 0.00207$ and $am_{q,s} = 0.0189$, corresponding to the sea and approximately the strange quark masses [9, 16], the variances are shown in Fig. 4 for the pseudoscalar density and for one spatial component of the vector current.

The variance of the standard estimators $\sigma_{\kappa,0}^2$ (red filled symbols) turns out to be essentially $\Gamma$-independent as suggested by Eq. (5.7), and it is dominated by the spectral sum $\langle SPM \rangle$. The split-even estimators $\tau_{\kappa,0}(x_0)$ have much smaller variances.\(^7\) The reduction factor ranges from approximately one order of magnitude for the scalar and pseudoscalar densities up to around two orders of magnitude or more for the axial and vector currents as well as for the tensor bilinear. A similar reduction in the variance using the split-even estimator is observed for ensembles with different sea-quark masses, see Fig. 9 in Appendix D for the analogous figures for the E5 and G8 ensembles. The gauge noise is still smaller than the random noise, but with the split-even estimator the number $N_s$ of random sources needed to approach the gauge noise is moderate. It ranges from a few for the pseudoscalar density up to $O(100)$ for the vector current.

Considering that the estimators have the same cost per noise source, the data show that the split-even random-noise estimator is much more efficient than the standard one for computing differences of single-propagator traces, and it allows one to approach the gauge noise for all bilinears with a moderate number of noisy sources.

6 Frequency-splitting of single-propagator traces

The results of the last two sections suggest to introduce a family of frequency-splitting random-noise estimators of single-propagator zero three-momentum traces defined as

\[^5\] The split-even is an estimator for all times at once, as well as the standard estimator in Eq. (5.6) we compare with. If time-dilution was used in (5.6), the computation of the estimator for all times would have been significantly more expensive.

\[^6\] By the same argument, if a split-line estimator localized in a given region of space is chosen, the sum over $\eta_2$ in Eq. (5.10) is restricted to that region.

\[^7\] The so called one-end trick estimator used in the context of twisted-mass discretization of QCD is a particular case of split-even estimator, for which significant numerical gain has been observed empirically [17, 18]. The analysis of the variances presented here applies straightforwardly to this estimator too, for which a Schwarz inequality between its variance and the one of the standard estimator can also be derived.
Table 2 The total cost of one evaluation of the FS estimator normalized to the cost of one evaluation of the standard estimator, for examples of the FS estimators on all ensembles. The cost of the split-even components, \(\tau_{\text{FS}k}\), is defined as the time for the two required inversions per source, while the cost of the remainder, \(\tau_{\text{FS}m}\), is the single inversion required per source. The standard estimator requires one inversion at the sea-quark mass \(m_{q_{\text{sea}}}\) per source. The relative cost of each component is given with the number of times it is evaluated for each full evaluation of the FS estimator, along with the bare subtracted quark masses which define the estimator. Note that the overhead required to compute the exact hopping terms is not included, which is performed only once per configuration. For \(n = 2\) HPE this quickly becomes negligible when the FS estimator is evaluated multiple times.

| Ensemble | Estimator | \(k\) | \(am_{q_{\text{sea}}}\) | Component | \(N_s\) | Rel. cost/\(N_s\) |
|----------|-----------|-------|-----------------|-----------|-------|----------------|
| E5       | FS1       | 0     | 0.00557         | \(\tau_{\text{FS}0/1}\) | 1     | 1.42           |
|          |           | 1     | 0.08            | \(\tau_{\text{FS}1/2}\) | 3     | 0.66           |
|          |           | 2     | 0.3             | \(\tau_{\text{FS}2}\) | 10    | 0.25           |
|          |           |       |                 | Total rel. cost | 5.81  |                |
| F7       | FS1       | 0     | 0.00207         | \(\tau_{\text{FS}0/1}\) | 1     | 1.26           |
|          |           | 1     | 0.1             | \(\tau_{\text{FS}1}\) | 4     | 0.26           |
|          |           |       |                 | Total rel. cost | 2.30  |                |
| F7       | FS2       | 0     | 0.00207         | \(\tau_{\text{FS}0/1}\) | 1     | 1.49           |
|          |           | 1     | 0.02            | \(\tau_{\text{FS}1/2}\) | 1     | 0.80           |
|          |           | 2     | 0.06            | \(\tau_{\text{FS}2/3}\) | 2     | 0.52           |
|          |           | 3     | 0.15            | \(\tau_{\text{FS}3/4}\) | 3     | 0.36           |
|          |           | 4     | 0.3             | \(\tau_{\text{FS}4}\) | 10    | 0.15           |
|          |           |       |                 | Total rel. cost | 5.92  |                |
| G8       | FS1       | 0     | 0.00108         | \(\tau_{\text{FS}0/1}\) | 1     | 1.33           |
|          |           | 1     | 0.01935         | \(\tau_{\text{FS}1/2}\) | 1     | 0.51           |
|          |           | 2     | 0.1             | \(\tau_{\text{FS}2}\) | 8     | 0.18           |
|          |           |       |                 | Total rel. cost | 3.28  |                |

\[
\tilde{\tau}_{\text{FS}}(\chi_0) = \tilde{\tau}_{\text{FS}M}(\chi_0) + \tilde{\tau}_{\text{FS}R}(\chi_0) + \sum_{k=1}^{m-1} \tilde{\tau}_{\text{FS}_{\text{split}}}(\chi_0)
\]

(6.1)

where \(\tilde{\tau}_{\text{FS}M}\) and \(\tilde{\tau}_{\text{FS}R}\) are defined in Eqs. (4.4), (4.6), and (5.9) respectively. The corresponding variances are given by

\[
\sigma_{\tau_{\text{FS}M}}^2 = \sigma_{\tau_{\text{FS}R}}^2 + \left\{ \sigma_{\tau_{\text{FS}R}}^2 - \sigma_{\tau_{\text{FS}R}}^2 \right\} + \sum_{k=1}^{m-1} \left\{ \sigma_{\tau_{\text{FS}_{\text{split}}}}^2 - \sigma_{\tau_{\text{FS}_{\text{split}}}}^2 \right\}
\]

(6.2)

where the various terms on the r.h.s are defined in Sects. 4 and 5. At high momenta (heavy masses) the contribution from \(\tilde{\tau}_{\text{FS}M}\), responsible for the bulk of the variance of the standard random-noise estimator, is computed exactly with a limited number of probing vectors, and only the remainder \(\tilde{\tau}_{\text{FS}R}\) is estimated by a random-noise estimator. The low-frequency contributions \(\tilde{\tau}_{\text{FS}_{\text{split}}}\) can then be estimated by the very efficient split-even estimator. It is rather clear that splitting the single-propagator traces in several parts whose contributions come from different frequency regions is beneficial. It allows us to design a customized estimator for each contribution which profits from its own peculiarities. An important ingredient in this analysis is the fact that solvers invert the Dirac operator with heavier quark masses at a lower numerical cost.

6.1 Numerical tests

The best choice of the number of mass differences, the values of the masses, and the order of the HPE for defining the frequency-splitting estimators in Eq. (6.1) depends on many factors: the bilinear of interest, the target mass, the solver chosen for inverting the Dirac operators and its particular implementation, etc. It is not the aim of this paper to optimize with respect to all these factors but, provided a reasonable choice is made, our goal is to give a numerical proof that the frequency-splitting estimators are efficient and allow to reduce significantly the numerical cost for computing single-propagator traces. To this aim we have implemented two such estimators:

- FS1 is the simplest frequency-splitting estimator with one mass difference only. The masses are \(am_q = 0.00207\) and 0.1, \(\tilde{\tau}_{\text{FS}0/1}\) and \(\tilde{\tau}_{\text{FS}1}\) are defined with \(N_s = 1\) and 4 respectively per full evaluation of the estimator. Each

---

8 Computing variances for such an optimization is cheap because it requires a few sources only.
We do not include the preparatory cost for computing $t^{(M)}_{\mu}$, since it becomes quickly negligible after few evaluations of the random-noise components of the estimator.

evaluation of this estimator costs approximately 2.5 times more than evaluating one random source for the standard estimator. See Table 2 for a detailed comparison of the relative cost of the estimators for our particular implementation.

- FS2 is defined by 4 mass splittings corresponding to the masses $am_q = 0.00207, 0.02, 0.06, 0.15, 0.3$, and the corresponding random-noise estimators are defined with $N_s = 1, 1, 2, 3, 10$ random sources respectively. Each application of this estimator costs approximately 6 times with respect to evaluating one random source for the standard estimator; see Table 2.

In both cases the solver used is the generalized conjugate residual (GCR) algorithm preconditioned by a Schwarz alternating procedure (SAP) and local deflation as implemented in openQCD-1.6 [19].

In Fig. 5 we show the variances of FS1, FS2 and of the standard estimator as a function of $N_s$, the number of evaluations of each of them per gauge configuration. A clear message emerges: a large gain is obtained for both frequency-splitting estimators with mild differences in efficiency between them. The FS1 is slightly better for the scalar and pseudoscalar densities, while FS2 is more efficient for the vector, axial-vector and tensor bilinears. In particular, the variance of FS1 is approximately 20 and 15 times smaller than the one of the standard estimators for the scalar and pseudoscalar densities respectively. Taking into account that one application of FS1

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]

\[ N_s \]

\[ \frac{\sigma^2}{\langle \lambda \rangle^3} \]
point functions, which are the simplest correlation functions with a non-trivial time dependence composed only of single-propagator traces. We use the estimators proposed in Sects. 5 and 6 to confirm the expected improvement over the standard propagator traces. We use the estimators proposed in Sects. 5 with a non-trivial time dependence composed only of single-propagator traces as in Eq. (5.6) and (5.9) respectively. A reduction of the variance of up to four orders of magnitude is obtained with the split-even estimator (two orders of magnitude in the cost), which starts to be comparable to the gauge noise for \( N_s \sim 256 \). As expected, the variance is practically constant in \( x_0 \) and well-described by the factorization formula in Eq. (2.3) when the averaging over time and the polarizations of the current are taken into account.

In the right-hand panel of Fig. 6 our best estimate of the correlation function using the split-even estimator is shown using an increased number of gauge configurations, with respect to those used for estimating the variances, of \( N_{cfg} = 1200 \). This in turn corresponds to a relative statistical precision of approximately 10% to the disconnected light-quark part of the muon anomalous magnetic moment coming from contributions to the integral up to time-distances of 1.5 fm. If the integral is computed up to 3.0 fm or so, the relative statistical error grows up to 70%, calling for the multi-level integration to determine the contribution from the long distance part of the integrand. To properly renormalize the correlator each current has to be multiplied by the factor \( Z_V \) which brings a negligible error with respect to the statistical error of the bare correlator.\(^{11}\)

7.2 Frequency-splitting estimator for isoscalar vector currents

In spectroscopic applications, disconnected diagrams arise generically in isoscalar channels. The vector channel, for instance, contains the contribution

\[
C_{VV}^{\gamma_0}(x_0) = -\frac{L^3}{3L_0} \sum_{k=1}^{3} \sum_{y_0} a \left( \bar{t}_{y_k,r} (x_0 + y_0) t_{y_k,s} (y_0) \right),
\]

(7.2)

determines the light disconnected contribution, via the time-momentum representation \([23]\), of the leading-order hadronic vacuum polarization, once each current is renormalized by \( Z_V = 0.74636(70) \) \([24]\) and the correct electric charge factor of 1/9 is included.

In the left-hand panel of Fig. 6, we show the variance of this correlation function for \( x_0/a = 10 \) computed by using the standard (red filled squares) and split-even estimators (blue open squares) in Eqs. (5.6) and (5.9) respectively. A reduction of the variance of up to four orders of magnitude is obtained with the split-even estimator (two orders of magnitude in the cost), which starts to be comparable to the gauge noise after \( N_s \sim 256 \). Right: the disconnected contribution using the split-even estimator from \( N_{cfg} = 1200 \) gauge configurations.

\(^{11}\) Improving the vector current goes beyond the scope of this paper. All formulas, however, can be found in Refs. \([25,26]\).
To evaluate this correlation function, we use the FS2 estimator introduced in Sect. 6 for both single-propagator traces. In the left plot of Fig. 7 we show the variances of the standard estimator (filled symbols) against the number of sources, and the improved FS2 estimator (open symbols) against the number of its evaluations per gauge configuration. The gauge variance is approached with about \( N_s \sim 256 \) evaluations of the FS2 estimator, similarly to the case of the one-point function of Sect. 6. In this case, while the disconnected piece gives only a small contribution to the isoscalar channel at intermediate hadronic distances, its variance quickly dominates the statistical error at large distances. The improved estimator thus allows the full correlation function to be resolved at much larger distances.

8 Conclusions

The numerical computation of disconnected Wick contractions is challenging in lattice QCD because (a) their variances are dominated by the vacuum contribution, which in turn implies that statistical errors remain constant with the distance of the disconnected pieces while the signal typically decreases exponentially, and (b) averaging each disconnected sub-diagram over the volume tends to be numerically expensive because the quark propagators must be re-computed at each lattice point.

A milestone for solving the second problem was the introduction of random-noise estimators [3–5] which allow one to sum over many or all source points stochastically. However for single-propagator traces, the simplest among the disconnected sub-diagrams, such estimators tend to have variances which are typically orders of magnitude larger than the intrinsic gauge noise. An a priori theoretical analysis of the variances is thus mandatory for deciding how to define exactly the stochastic observables.

Luckily the random-noise contribution to the variances can be re-expressed in the form of simple integrated correlation functions of local composite operators, a fact which allows us to use the quantum field theory machinery for analyzing the origin of the statistical errors and eventually to reduce them.

As a result, we have introduced new stochastic observables for single-propagator traces: the split-even and the frequency-splitting estimators for difference of two traces and for single traces respectively. The former needs from a few random sources for the pseudoscalar density up to \( O(100) \) for the vector current to approach the gauge noise. The reduction in numerical cost with respect to the standard estimator ranges from one order of magnitude for the scalar and pseudoscalar densities up to around two orders of magnitude or more for the axial and vector currents as well as for the tensor bilinear. Just one or a few evaluations of the frequency-splitting estimators are needed for the variances of the scalar, pseudoscalar and tensor bilinears to be comparable to the gauge noise, while for the axial-vector and vector currents \( O(10) \) and \( O(100) \) evaluations are required to reach the same goal. In this case the reduction of the computational cost with respect to the standard estimator is of one order of magnitude or so depending on the bilinear. In all cases considered the variances of the stochastic estimators reach the level of the intrinsic gauge noise with a moderate number of evaluations per gauge configuration.

The use of these new estimators significantly speeds up the computation of disconnected fermion Wick contractions which contribute to many physics processes at the forefront of
research in particle and nuclear physics: the hadronic contribution to the muon anomalous magnetic moment, $K \to \pi \pi$ decays, nucleon form factors, quantum electrodynamics and strong isospin-breaking contributions to hadronic matrix elements, $\eta'$ propagator, etc. As an example we have shown their potential for computing the disconnected contribution to the light-quark contribution to the muon anomalous magnetic moment and to the correlator of two singlet vector currents. Theory suggests large gains for disconnected three and higher point correlation functions as well. To solve or mitigate the problem (a) alluded to at the beginning of this section, the next step is to combine these estimators with the newly proposed multi-level integration in the presence of fermions [1,2].

**Acknowledgements** Simulations have been performed on the PC clusters Marconi at CINECA (CINECA-INFN and CINECA-Bicocca agreements) and Wilson at Milano-Bicocca. We thank these institutions for the computer resources and the technical support. 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 computing network”.

**Data Availability Statement** This manuscript has no associated data or the data will not be deposited. [Authors’ comment: There are no external data associated with the manuscript.]

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP3.

**Appendix A: $O(a)$-improved Wilson–Dirac operator**

The massive $O(a)$-improved Wilson–Dirac operator is defined as [27,28]

$$D_m = D_w + \delta D_v + m,$$

(A.1)

where $m$ is the bare quark mass, $D_w$ is the massless Wilson–Dirac operator

$$D_w = \frac{1}{2} \left\{ \gamma_\mu (\nabla^+_\mu + \nabla^-_\mu) - \nabla^+_\mu \nabla^-_\mu \right\} ,$$

(A.2)

$\gamma_\mu$ are the Dirac matrices, and the summation over repeated indices is understood. The covariant forward and backward derivatives $\nabla^-_\mu$ and $\nabla^+_\mu$ are defined to be

$$a \nabla^-_\mu \psi(x) = U_{\mu}(x) \psi(x + \hat{\mu}) - \psi(x),$$

$$a \nabla^+_\mu \psi(x) = \psi(x) - U^\dagger_{\mu}(x - \hat{\mu}) \psi(x - \hat{\mu}).$$

(A.3)

where $U_\mu(x)$ are the link fields. The clover term is defined as

$$\delta D_v \psi(x) = a c_{SW} \frac{i}{4} \varepsilon_{\mu\nu} \tilde{F}^\mu_{\nu} \psi(x),$$

(A.4)

where the field strength of the gauge field is

$$\alpha^2 \tilde{F}^\mu_{\nu}(x) = \frac{1}{8} \left\{ Q^\mu_{\nu}(x) - Q_{\nu \mu}^\mu(x) \right\}$$

(A.5)

with

$$Q_{\nu \mu}(x) = U_{\nu}(x) U_\nu(x + \hat{\mu}) U^\dagger_\nu(x + \hat{\nu}) U^\dagger_\nu(x)$$

$$+ U_\nu(x) U^\dagger_\nu(x - \hat{\mu}) U^\dagger_\nu(x - \hat{\nu}) U_\nu(x - \hat{\mu})$$

$$+ U^\dagger_\nu(x - \hat{\mu}) U_\nu(x - \hat{\nu}) U_\nu(x - \hat{\nu}) U^\dagger_\nu(x)$$

(A.6)

### A.1: Hopping expansion

By applying the standard even-odd decomposition of the Wilson–Dirac operator

$$D_m = \begin{pmatrix} D_{ee} & D_{eo} \\ D_{oe} & D_{oo} \end{pmatrix},$$

(A.7)

see Ref. [29] for unexplained notation, it is straightforward to verify that

$$D_m^{-1} = \frac{1}{D_{ee} + D_{oo}} \frac{1}{1 - H_m},$$

(A.8)

where

$$H_m = - \left[ D_{oe} D_{oo}^{-1} + D_{oe} D_{ee}^{-1} \right],$$

(A.9)

and for clarity the subscript $m$ has been omitted in the block matrices of the even/odd decomposition defined in Eq. (A.7). It follows that

$$D_m^{-1} = M_{2n,m} + D_m^{-1} H_m^{2n},$$

(A.10)

where

$$M_{2n,m} = \frac{1}{D_{ee} + D_{oo}} \sum_{k=0}^{2n-1} H_m^{k},$$

(A.11)
Appendix B: Bilinear chains in the free case

The propagator of a free Wilson fermion is

\[ S(x - y) = \langle \psi(x) \bar{\psi}(y) \rangle = \int_{BZ} \frac{d^4 p}{(2\pi)^4} \ K^{-1}(p) \ e^{ip(x - y)}, \]  
\[ \tag{B.1} \]

where

\[ K(p) = i \gamma_\mu \bar{p}_\mu + M(p), \quad M(p) = m + \frac{a}{2} \bar{p}^2, \]  
\[ \tag{B.2} \]

with

\[ \bar{p}_\mu = \frac{1}{a} \sin(p_\mu a), \quad \hat{p}_\mu = \frac{2}{a} \sin \left( \frac{p_\mu a}{2} \right), \]

and as usual \( \bar{p}^2 = \bar{p}_\mu \bar{p}_\mu \) and \( \hat{p}^2 = \hat{p}_\mu \hat{p}_\mu \).

B.1: Two-point correlators

The integrated two-point correlation functions of non-singlet bilinears are

\[ \sum_x a^3 \left\{ O_{r,r'}(0, x) O_{r,r'}(0) \right\} \]
\[ = -12 \int_{BZ} \frac{d^3 \bm{p}}{(2\pi)^3} \left\{ b_r G(p, m_r) G(p, m_{r'}) + c_r p^2 F(p, m_r) F(p, m_{r'}) \right\}, \]  
\[ \tag{B.4} \]

where

\[ F(p, m) = \int_{BZ} \frac{dp_0}{2\pi} \frac{1}{\bar{p}^2 + M^2(p)}, \]
\[ G(p, m) = \int_{BZ} \frac{dp_0}{2\pi} \frac{M(p)}{\bar{p}^2 + M^2(p)}, \]  
\[ \tag{B.5} \]

and

\[ b_r = \begin{cases} 
1 & \Gamma = I, \gamma_5, \gamma_\mu \sigma_{\mu\nu} \\
-1 & \Gamma = \gamma_\mu \gamma_5
\end{cases}, \]

\[ c_r = \begin{cases} 
-1 & \Gamma = I \\
1 & \Gamma = \gamma_5 \\
1 - \frac{2}{3} (1 - \delta_{\mu0}) & \Gamma = \gamma_\mu \gamma_5 \\
-1 + \frac{2}{3} (2 - \delta_{\mu0} - \delta_{\nu0}) & \Gamma = \sigma_{\mu\nu}
\end{cases}. \]  
\[ \tag{B.6} \]

By using

\[ K^\dagger(p) K(p) = K(p) K^\dagger(p) = \bar{p}^2 + M^2(p), \]
\[ K^\dagger(p) = \gamma_5 K(p) \gamma_5, \]  
\[ \tag{B.7} \]

it is straightforward to obtain

\[ \sum_y a^3 \left\{ P_{r,r'}(y) P_{r,r'}(0) \right\} \]
\[ = -12 \int_{BZ} \frac{d^3 \bm{p}}{(2\pi)^3} \left\{ \bar{p}^2 + M_r(p) M_{r'}(p) \right\}, \]  
\[ \tag{B.8} \]

where \( M_r(p) \) is evaluated at the bare mass of \( m_r \).

B.2: Four-point correlators

The integrated four-point correlation functions of non-singlet bilinears we are interested in are

\[ \sum_{x_1, x_2, x_3} a^{11} \left[ P_{r,r'}(x_1) O_{r',r'}(0, x_2) P_{r',r'}(x_3) O_{r',r'}(0) \right] \]
\[ = -12 a^2 \int_{BZ} \frac{d^3 \bm{p}}{(2\pi)^3} \left\{ F(p, m_r) F(p, m_{r'}) \right\}, \]  
\[ \tag{B.9} \]

and

\[ \sum_{x_1, x_2, x_3} a^{11} \left[ S_{r,r'}(x_1) O_{r',r'}(0, x_2) S_{r',r'}(x_3) O_{r',r'}(0) \right] \]
\[ = -12 \int_{BZ} \frac{d^3 \bm{p}}{(2\pi)^3} \left\{ b_r \left\{ K(p, m_r, m_{r'}) \right\}^2 + c_r \bar{p}^2 \left\{ H(p, m_r, m_{r'}) \right\}^2 \right\}, \]  
\[ \tag{B.10} \]

where

\[ H(p, m_r, m_{r'}) = \int_{BZ} \frac{dp_0}{2\pi} \frac{M_r(p) + M_{r'}(p)}{\bar{p}^2 + M_r^2(p)} \]
\[ K(p, m_r, m_{r'}) = \int_{BZ} \frac{dp_0}{2\pi} \frac{M_r(p) M_{r'}(p)}{\bar{p}^2 + M_r^2(p)}, \]  
\[ \tag{B.11} \]

and \( m_r = m_r \) and \( m_{r'} = m_s \). Finally

\[ \sum_{x_1, x_2, x_3} a^{12} \left[ S_{r,r'}(x_1) P_{r',r'}(y_2) S_{r',r'}(y_3) P_{r',r'}(0) \right] \]
\[ = -12 \int_{BZ} \frac{d^3 \bm{p}}{(2\pi)^3} \left\{ \bar{p}^2 + M_r^2(p) \right\}, \]  
\[ \tag{B.12} \]

where again we are interested in the case \( m_{r'} = m_r \) and \( m_{r'} = m_s \).

B.3: Variance of the HPE remainder

In the free theory, the variance of the noisy estimator of the remainder in Eq. (4.6) is

\[ \sigma^2_{\epsilon, r,m} = \frac{6}{L^3 N_s} \int \frac{d^3 \bm{p}}{(2\pi)^3} \left\{ J_{0,n}(\bm{p}, m) J_{1,n}(\bm{p}, m) \right\} \]
\[ + \left( J_{1,2n}(\bm{p}, m) \right)^2 a_r^2 \left( c_1(p) \right)^2 \]  
\[ \tag{B.13} \]

where

\[ J_{0,n}(\bm{p}, m) = \int_{BZ} \frac{dp_0}{2\pi} \left\{ (c_{1,n}(p))^2 \bar{p}^2 + (c_{0,n}(p))^2 \right\}, \]  
\[ \tag{B.14} \]
The non-zero elements of $M_{2n,m}$ are those that connect two lattice sites $x$ and $y$ with $\|x-y\|_1 < na$, while the matrix is dense in the spin and colour indices. For a lattice which can be decomposed in hypercubic blocks of size $(2na)^4$, an obvious scheme to define the set of probing vectors which satisfies the condition (C.1) is

$$v^k(x) = \begin{cases} 1 & k = i_{cs} + 12l_2(n) \\ 0 & \text{otherwise} \end{cases}$$

where $i_{cs} = 1, \ldots, 12$ indicates the spin-colour index and $l_2(n)(x) = (x_0/a) \mod 2n + 2n \cdot [(x_1/a) \mod 2n] + \cdots$ is the lexicographical index labeling the sites in any given block. This scheme, illustrated in Fig. 8 for $n = 2$, requires $K = 192n^4$ probing vectors because one vector is required for each of the spin-colour components for every site in the block. A more efficient scheme, already outlined in Ref. [32], is depicted in the right-hand panel of Fig. 8, where even-probing vectors were introduced in the context of lattice QCD to define stochastic estimators of traces of the full quark propagator.

---

**Fig. 8** Two schemes of probing vectors suitable for a matrix whose generic column has non-zero entries on the sites within the shaded region. Left: the probing vectors have support, e.g. white squares, on one point of each of the shaded blocks. The column of the matrix $M$ corresponding to $(x/a) = (4,4)$ has non-zero entries only in the shaded region where the probing vector is zero. Right: as on the left but for smaller blocks labeled even and odd.
odd blocks of half the linear size of the previous ones are introduced. The probing vectors are defined by

\[ v^k(x) = \begin{cases} 1 & k = i_{cs} + 12 \{ p + 2 l_n(x) \} \\ 0 & \text{otherwise} \end{cases} \]

where as before \( i_{cs} \) indicates the spin-colour index, \( p = 0, 1 \) is the parity of the block, and again \( l_n(x) \) is the lexicographical index labeling the sites in any given block. This scheme requires just \( K = 24 n^4 \) vectors, which is a factor 8 fewer than the first one.

**Appendix D: Numerical tests on E5 and G8 ensembles**

In this appendix we include the results for the split-even estimators and examples of the frequency-splitting estimators for the E5 and G8 ensembles, which have a larger and smaller quark mass than the F7 ensemble described in the text, see Table 1. In addition we provide a more detailed comparison of the relative cost between the frequency-splitting and standard estimators for our particular implementation for each ensemble in Table 2.

D.1: Split-even estimators

Figure 9 depicts the variances of the standard and split-even estimators for the differences of single-propagator traces for the E5 (top) and G8 (bottom) ensembles in the pseudoscalar (left) and vector (right) channels. The quark masses are chosen to be the sea and strange quark masses as before. In the pseudoscalar channel it is again observed that the variance saturates with the gauge variance after just a few random sources. In both channels, the variance increases as the sea quark mass is lowered. However, in each channel, the split-even estimator results in a consistent reduction in the variance which is quite independent of the sea-quark mass.

D.2: Frequency-splitting estimators

The analogous figures to Fig. 5 for examples of frequency-splitting estimators are shown in Fig. 10 for the E5 (top) and G8 (bottom) ensembles for the pseudoscalar (left) and vector (right) channels. The cost for one full evaluation of the estimator relative to the cost of the evaluation of the standard estimator is shown in Table 2 along with the values of the quark masses and choices of \( N_s \) for each component, which
define the estimators. For the lighter sea-quark mass ensemble, the G8 ensemble, accounting for the increased relative cost of about a factor 3.3, the gain in the computational cost for the FS1 estimator is approximately a factor 3 in the pseudoscalar channel, and about a factor 30 in the vector channel, one of the best cases. This demonstrates the effectiveness of the strategy as the physical point is approached even without any particular fine-tuning of the parameters.

References

1. M. Cè, L. Giusti, S. Schaefer, Domain decomposition, multi-level integration and exponential noise reduction in lattice QCD. Phys. Rev. D 93(9), 094507 (2016). arXiv:1601.04587
2. M. Cè, L. Giusti, S. Schaefer, A local factorization of the fermion determinant in lattice QCD. Phys. Rev. D 95(3), 034503 (2017). arXiv:1609.02419
3. K. Bitar, A.D. Kennedy, R. Horsley, S. Meyer, P. Rossi, The QCD finite temperature transition and hybrid Monte Carlo. Nucl. Phys. B 313, 348–376 (1989)
4. S.-J. Dong, K.-F. Liu, Stochastic estimation with Z(2) noise. Phys. Lett. B 328, 130–136 (1994). arXiv:hep-lat/9308015
5. UKQCD Collaboration, C. Michael, J. Peisa, Maximal variance reduction for stochastic propagators with applications to the static quark spectrum. Phys. Rev. D 58, 034506 (1998). arXiv:hep-lat/9802015
6. M. Lüscher, Topological effects in QCD and the problem of short distance singularities. Phys. Lett. B 593, 296–301 (2004). arXiv:hep-th/0404034
7. L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, N. Tantalo, QCD with light Wilson quarks on fine lattices (I): first experiences and physics results. JHEP 02, 056 (2007). arXiv:hep-lat/0610059
8. L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, N. Tantalo, QCD with light Wilson quarks on fine lattices. II. DD-HMC simulations and data analysis. JHEP 02, 082 (2007). arXiv:hep-lat/0701009
9. P. Fritzsch, F. Knechtli, B. Leder, M. Marinković, S. Schaefer, R. Sommer, F. Virotta, The strange quark mass and Lambda parameter of two flavor QCD. Nucl. Phys. B 865, 397–429 (2012). arXiv:1205.5380
10. G.P. Engel, L. Giusti, S. Lottini, R. Sommer, Spectral density of the Dirac operator in two-flavor QCD. Phys. Rev. D 91(5), 054505 (2015). arXiv:1411.6386
11. C. Thron, S.J. Dong, K.F. Liu, H.P. Ying, Pade—Z(2) estimator of determinants. Phys. Rev. D 57, 1642–1653 (1998). arXiv:hep-lat/9707001
12. UKQCD Collaboration, C. McNeile, C. Michael, Mixing of scalar glueballs and flavor singlet scalar mesons. Phys. Rev. D 63, 114503 (2001). arXiv:hep-lat/0010019
13. G.S. Bali, S. Collins, A. Schäfer, Effective noise reduction techniques for disconnected loops in lattice QCD. Comput. Phys. Commun. 181, 1570–1583 (2010). arXiv:0910.3970
14. V. Gülpers, G. von Hippel, H. Wittig, Scalar pion form factor in two-flavor lattice QCD. Phys. Rev. D 89(9), 094503 (2014). arXiv:1309.2104

15. L. Giusti, M. Lüscher, Chiral symmetry breaking and the Banks–Casher relation in lattice QCD with Wilson quarks. JHEP 03, 013 (2009). arXiv:0812.3638

16. M. Della Morte, A. Francis, V. Gülpers, G. Herdoíza, G. von Hippel, H. Horch, B. Jäger, H.B. Meyer, A. Nyffeler, H. Wittig, The hadronic vacuum polarization contribution to the muon $g - 2$ from lattice QCD. JHEP 10, 020 (2017). arXiv:1705.01775

17. E.T.M. Collaboration, P. Boucaud et al., Dynamical twisted mass fermions with light quarks: simulation and analysis details. Comput. Phys. Commun. 179, 695–715 (2008). arXiv:0803.0224

18. E.T.M. Collaboration, S. Dinter, V. Drach, R. Frezzotti, G. Herdoíza, K. Jansen, G. Rossi, Sigma terms and strangeness content of the nucleon with $N_f = 2 + 1 + 1$ twisted mass fermions. JHEP 08, 037 (2012). arXiv:1202.1480

19. The openQCD package. http://luscher.web.cern.ch/luscher/openQCD/

20. L. Giusti, P. Hernández, M. Laine, P. Weisz, H. Wittig, Low-energy couplings of QCD from current correlators near the chiral limit. JHEP 04, 013 (2004). arXiv:hep-lat/0402002

21. T.A. DeGrand, S. Schaefer, Improving meson two point functions in lattice QCD. Comput. Phys. Commun. 159, 185–191 (2004). arXiv:hep-lat/0401011

22. T. Blum, T. Izubuchi, E. Shintani, New class of variance-reduction techniques using lattice symmetries. Phys. Rev. D 88(9), 094503 (2013). arXiv:1208.4349

23. D. Bernecker, H.B. Meyer, Vector correlators in lattice QCD: methods and applications. Eur. Phys. J. A 47, 148 (2011). arXiv:1107.4388

24. M. Dalla Brida, T. Korzec, S. Sint, P. Vilaseca, High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks. Eur. Phys. J. C 79(1), 23 (2019). arXiv:1808.09236

25. A. Gérardin, T. Harris, H.B. Meyer, Nonperturbative renormalization and $O(a)$-improvement of the nonsinglet vector current with $N_f = 2 + 1$ Wilson fermions and tree-level Symanzik improved gauge action. Phys. Rev. D 99(1), 014519 (2019). arXiv:1811.08209

26. T. Bhattacharya, R. Gupta, W. Lee, S.R. Sharpe, J.M.S. Wu, Improved bilinears in lattice QCD with non-degenerate quarks. Phys. Rev. D 73, 034504 (2006). arXiv:hep-lat/0511014

27. B. Sheikholeslami, R. Wohlert, Improved continuum limit lattice action for QCD with Wilson fermions. Nucl. Phys. B 259, 572 (1985)

28. M. Lüscher, S. Sint, R. Sommer, P. Weisz, Chiral symmetry and $O(a)$ improvement in lattice QCD. Nucl. Phys. B 478, 365–400 (1996). arXiv:hep-lat/9605038

29. M. Lüscher, Computational Strategies in Lattice QCD, in Modern Perspectives in Lattice QCD: Quantum Field Theory and High Performance Computing, Proceedings, International School, 93rd Session, Les Houches, France, August 3–28, 2009, pp. 331–399 (2010). arXiv:1002.4232

30. J.M. Tang, Y. Saad, A probing method for computing the diagonal of a matrix inverse. Numer. Linear Algebra Appl. 19(3), 485–501 (2012). https://onlinelibrary.wiley.com/doi/pdf/10.1002/nla.779

31. A. Stathopoulos, J. Laeuchli, K. Orginos, Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices. arXiv:1302.4018

32. L. Giusti, T. Harris, A. Nada, S. Schaefer, Multi-level integration for meson propagators, in 36th International Symposium on Lattice Field Theory (Lattice 2018) East Lansing, MI, United States, July 22–28, 2018 (2018). arXiv:1812.01875