Domain decomposition, multi-level integration and exponential noise reduction in lattice QCD

Marco Cè

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
and INFN, Sezione di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy
E-mail: marco.ce@sns.it

Leonardo Giusti

Dipartimento di Fisica, Università di Milano–Bicocca,
and INFN, sezione di Milano–Bicocca,
I-20126 Milano, Italy
E-mail: Leonardo.Giusti@mib.infn.it

Stefan Schaefer

John von Neumann Institute for Computing (NIC),
DESY, Platanenallee 6, D-15738 Zeuthen, Germany
E-mail: Stefan.Schaefer@desy.de

Abstract

We explore the possibility of computing fermionic correlators on the lattice by combining a domain decomposition with a multi-level integration scheme. The quark propagator is expanded in series of terms with a well defined hierarchical structure. The higher the order of a term, the (exponentially) smaller its magnitude, the less local is its dependence on the gauge field. Once inserted in a Wick contraction, the gauge-field dependence of the terms in the resulting series can be factorized so that it is suitable for multi-level Monte Carlo integration. We test the strategy in quenched QCD by computing the disconnected correlator of two flavor-diagonal pseudoscalar densities, and a nucleon two-point function. In either cases we observe a significant exponential increase of the signal-to-noise ratio.

1 Introduction

With state of the art techniques, the numerical computation of hadronic correlation functions in lattice Quantum Chromodynamics (QCD) suffers from signal-to-noise ratios which decrease exponentially with the time separation of the sources, notable exceptions being the propagators of non-singlet pseudoscalar mesons. For connected Wick contractions, the problem can be traced back to the fact that, on a typical gauge configuration, the quark propagator decreases approximatively as $\exp\{-M_\pi|y - x|/2\}$ at asymptotically large distances $|y - x|$, while the expectation value of a generic hadron correlator decays much faster [1,2]. For a nucleon two-point function at zero momentum, for instance, the signal-to-noise ratio decreases proportionally to $\exp\{-(M_N - 3M_\pi/2)|y_0 - x_0|\}$, where $|y_0 - x_0|$ is the time-distance of the sources and $(M_N - 3M_\pi/2)$ is as big as 3.7 fm^{-1} at the physical point. The number of configurations needed to reach a given statistical precision thus increases with the square of that exponential factor. For disconnected contractions, the problem is even worse due to vacuum contributions to the variance.

Analogous severe problems afflict the computation of correlators in a large variety of quantum systems, from the harmonic oscillator to Yang–Mills theory. In some cases, multi-level algorithms have been proposed which lead to an impressive acceleration of the simulations [3–8]. They take advantage of the fact that, when the action and the observables depend locally on the integration variables, the signal-to-noise problem can be solved by independent measurements of the local building blocks of the observables. So far, these ideas have been restricted to bosonic theories.

It is not straight forward to formulate multi-level algorithms for systems with fermions. Once they have been analytically integrated out in the path integral, the manifest locality of the action and of the observables is lost. The fermion determinant and propagator are non-local functionals of the background gauge field. In order to make lattice computations with fermions amenable for multi-level algorithms, factorizations of fermionic correlation functions have to be found, where the individual terms depend only on gauge links confined to certain lattice domains. This cannot be achieved with the exact inverse of the Dirac operator, because each of its elements depends on the gauge field over the all lattice. As we will see, however, a series of approximations can be found which exhibit the various degrees of non-locality in the propagator.

In this paper we will pursue two strategies, adapted to different types of correlation functions, both based on domain decomposition techniques [9, 10]. In the first, which we will show to lead to an efficient algorithm for disconnected correlation functions, we define a succession of domains $\Gamma_0 \subset \Gamma_1 \subset \Gamma_2, \ldots$ which are larger and larger and which contain the two end points x and y of the propagator. The latter can then be expanded in series as

$$S(y, x) = S^{(0)}(y, x) + \left[S^{(1)} - S^{(0)}\right](y, x) + \left[S^{(2)} - S^{(1)}\right](y, x) + \ldots, \qquad (1.1)$$

where $S^{(i)}(y, x)$ is the inverse of the Dirac operator restricted to the ith domain, and

depends on the values of the gauge field in Γ_i only. The larger the domain Γ_i, the smaller the corresponding term $[S^{(i+1)} - S^{(i)}]$, the less local is its dependence on the gauge field. By inserting $S^{(0)}$ in the Wick contraction of the disconnected contribution of two pseudoscalar densities located in different domains, it is clear that the gauge-field dependence in the product of the two traces is factorized. The average can then be computed by a two-level Monte Carlo integration scheme.

The second approach leads one step further, and demonstrates that also connected hadron correlation functions can be factorized such that multi-level algorithms can be used. The propagator is approximated by a product of matrices which depend on the gauge field belonging to different domains of the lattice. This in turn leads to factorized correlations and thus to local averaging of them.

In the following sections we present the details of the proposed computational strategy for the correlation function of two different flavor-diagonal pseudoscalar densities, and for the nucleon two-point correlation function. We then show numerical evidence of the effectiveness of the strategy in quenched QCD, where only the Wick contractions need to be re-organized in a factorized form. An analogous factorization of the fermion determinant is left for a future publication.

2   Quark propagator and locality

Let us take a lattice Γ with open boundary conditions in the time direction [11], and define the domains

$$\Gamma_0 \subset \Gamma_1 \subset \Gamma_2 \subset \cdots \subset \Gamma \;. \tag{2.1}$$

We adopt here the same block terminology as in Ref. [10]. We choose the Γ_i to be a hyper-cubic domain of lattice points. Its exterior boundary $\partial\Gamma_i$ is defined to be the set of all points that have distance 1 from Γ_i. Each exterior boundary point has a closest "partner" point in the block. The interior boundary $\partial\Gamma_i^*$ of Γ_i consists of all these points. The set of points that are not in the block is denoted by

$$\Gamma_i^* = \Gamma \setminus \Gamma_i \;. \tag{2.2}$$

For a given domain decomposition of the lattice, the Wilson–Dirac operator defined in Appendix A, being a sparse matrix in position space, assumes the block-diagonal form

$$D = \left(\begin{array}{cc} D_{\Gamma_i} & D_{\partial\Gamma_i} \\ D_{\partial\Gamma_i^*} & D_{\Gamma_i^*} \end{array} \right) \;. \tag{2.3}$$

The operator D_{Γ_i} acts on quark fields on Γ_i in the same way as D, except that all terms involving the exterior boundary points $\partial\Gamma_i$ are set to zero (which is equivalent to impose Dirichlet boundary conditions on $\partial\Gamma_i$). By using the decomposition in Eq. (B.11), the exact quark propagator between the points $x, y \in \Gamma_i$ is given by

$$S(y,x) = S^{(i)}(y,x) + \sum_{w_1,w_2 \in \partial\Gamma_i^*} S^{(i)}(y,w_1) \, [D_{\partial\Gamma_i} D^{-1} D_{\partial\Gamma_i^*}](w_1,w_2) \, S^{(i)}(w_2,x) \;, \tag{2.4}$$

where

$$S^{(i)}(y, x) = D_{\Gamma_i}^{-1}(y, x) \tag{2.5}$$

depends on the values of the gauge field in the block Γ_i *only.* It is rather clear at this point that we can generate a succession of approximations $S^{(i)}$ which, by construction, converges to the exact propagator when Γ_i gets larger and larger. For a typical gauge configuration, when the sink and the source of the two $S^{(i)}$ in the sum on the r.h.s. of Eq. (2.4) are at asymptotically large distances, it holds

$$\mathrm{tr}\{S^{(i)}(y, x)\, S^{(i)}(y, x)^\dagger\}^{1/2} \sim \mathrm{tr}\{S(y, x)\, S(y, x)^\dagger\}^{1/2} \sim e^{-\frac{1}{2}M_\pi |y-x|} \tag{2.6}$$

with M_π the mass of the corresponding pseudoscalar meson made of degenerate quarks. A rough estimate of the distance between the exact and the approximated propagator is

$$\mathrm{tr}\{(S(y, x) - S^{(i)}(y, x))\, (S(y, x)^\dagger - S^{(i)}(y, x)^\dagger)\}^{1/2} \sim e^{-M_\pi d_i} , \tag{2.7}$$

with d_i the average of the distances of x and y from the interior boundaries of Γ_i.

3   Multi-level integration of the disconnected pseudoscalar propagator

The decomposition in Eq. (2.4) calls for a multi-level integration of disconnected contributions to correlation functions. We test the idea in SU(3) Yang–Mills theory with open boundary conditions [11] supplemented with a doublet of quenched quarks, u and d, degenerate in mass. Both fermions are discretized with the Wilson–Dirac operator, so that isospin symmetry is exactly preserved. We compute the correlator of two different flavor-diagonal pseudoscalar densities (the generalization to other cases being straightforward)

$$C_{P_d}(y_0, x_0) = \frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \langle \bar{d}(y)\gamma_5 d(y)\, \bar{u}(x)\gamma_5 u(x)\rangle = \frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \langle W_{P_d}(y, x)\rangle , \tag{3.1}$$

where $W_{P_d}(y, x)$ indicates the Wick contraction of the fermion fields, and L is the length of the lattice in the spatial directions[1]. In a standard Monte Carlo simulation, the statistical error of $C_{P_d}(y_0, x_0)$ is constant as a function of $|y_0 - x_0|$ while its expectation value decreases proportionally to $\exp\{-M_\pi |y_0 - x_0|\}$ at large time separations. The number of configurations n_0 required to reach a given relative statistical error thus grows exponentially with the time distance of the densities, i.e. $n_0 \propto \exp\{2M_\pi |y_0 - x_0|\}$.

The depletion of the signal-to-noise ratio is particularly severe at large time-distances. To take advantage of the locality of the theory, it is therefore natural to divide the lattice Γ in two non-overlapping thick time-slices Γ_0 and Γ_0^*. The first time coordinate of Γ_0^*, x_0^{cut}, is chosen approximatively in the middle between the two densities (see Fig. 1).

---

[1]Throughout the paper dimensionful quantities are always expressed in units of the lattice spacing a unless explicitly specified.
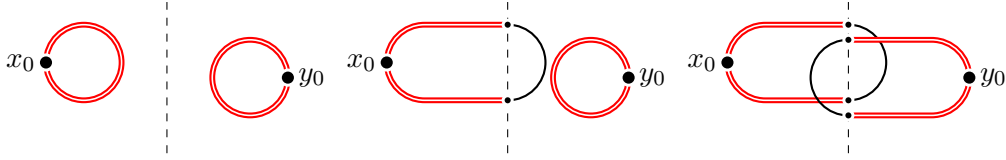
Figure 1: The three type of contributions to the disconnected pseudoscalar propagator in Eqs. (3.3)–(3.6). Black (single) lines are full propagators, red (double) ones are those within a domain.

By using the first and the second diagonal elements in Eqs. (B.11) and Eqs. (B.10) respectively, the Wick contraction can be decomposed as

$$W_{P_d}(y, x) = W_{P_d}^{(\mathrm{f})}(y, x) + W_{P_d}^{(\mathrm{r})}(y, x) \ , \tag{3.2}$$

where

$$W_{P_d}^{(\mathrm{f})}(y, x) = \mathrm{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x, x)\right\} \times \mathrm{tr}\left\{\gamma_5 D_{\Gamma_0^*}^{-1}(y, y)\right\} \ , \tag{3.3}$$

$x \in \Gamma_0$, and $y \in \Gamma_0^*$. The rest of the contraction is given by

$$W_{P_d}^{(\mathrm{r})}(y, x) = \left[W_{P_d}^{(\mathrm{r}_1)}(y, x) + (\Gamma_0, x) \leftrightarrow (\Gamma_0^*, y) \right] + W_{P_d}^{(\mathrm{r}_2)}(y, x) \tag{3.4}$$

with

$$W_{P_d}^{(\mathrm{r}_1)}(y, x) = \mathrm{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x, \cdot)[D_{\partial\Gamma_0} D^{-1} D_{\partial\Gamma_0^*}](\cdot, \cdot) D_{\Gamma_0}^{-1}(\cdot, x)\right\} \times \mathrm{tr}\left\{\gamma_5 D_{\Gamma_0^*}^{-1}(y, y)\right\}, \tag{3.5}$$

$$W_{P_d}^{(\mathrm{r}_2)}(y, x) = \mathrm{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x, \cdot)[D_{\partial\Gamma_0} D^{-1} D_{\partial\Gamma_0^*}](\cdot, \cdot) D_{\Gamma_0}^{-1}(\cdot, x)\right\} \times$$

$$\mathrm{tr}\left\{\gamma_5 D_{\Gamma_0^*}^{-1}(y, \cdot)[D_{\partial\Gamma_0^*} D^{-1} D_{\partial\Gamma_0}](\cdot, \cdot) D_{\Gamma_0^*}^{-1}(\cdot, y)\right\} \ . \tag{3.6}$$

When the spatial gauge links at x_0^{cut} are kept frozen, the dependence of the action and of $W_{P_d}^{(\mathrm{f})}(y, x)$ on the remaining link variables is factorized.

3.1 Two-level integration

When an observable depends only on the link variables in a given sub-lattice and the action of the theory is local, it is useful to define its expectation value restricted to that domain. This is a function of the link variables at the boundary of the sub-lattice only, and do not depend on the gauge field values elsewhere. For the trace of the Wilson–Dirac operator that we are interested in, it reads

$$\left[\mathrm{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x, x)\right\} \right] = \frac{1}{Z_{\Gamma_0}} \int D[U]_{\Gamma_0} e^{-S[U]_{\Gamma_0}} \, \mathrm{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x, x)\right\} \tag{3.7}$$

where $D[U]_{\Gamma_0}$ and $S[U]_{\Gamma_0}$ are the invariant Haar measure and the action restricted to the domain Γ_0, and the sub-lattice partition function is fixed by requiring that $[1] = 1$. By

4

following the standard line of argumentation in multi-level integration technique [3–8], it follows that

$$\left\langle \operatorname{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x,x)\right\} \operatorname{tr}\left\{\gamma_5 D_{\Gamma_0^*}^{-1}(y,y)\right\}\right\rangle = \left\langle \left[\operatorname{tr}\left\{\gamma_5 D_{\Gamma_0}^{-1}(x,x)\right\}\right]\left[\operatorname{tr}\left\{\gamma_5 D_{\Gamma_0^*}^{-1}(y,y)\right\}\right]\right\rangle. \quad (3.8)$$

This suggests that the mean value of $W_{P_d}^{(f)}(y,x)$ can be computed with a two-level algorithm. First, for each level-0 gauge field (spatial components) on the boundary x_0^{cut}, the averages of the two traces are carried out independently on n_1 level-1 configurations generated independently in the two thick time-slices. Then the average over the level-0 configurations of the product of the two means is performed by updating the gauge links over the entire lattice. The crucial question, to be answered numerically in section 6, is whether one can choose n_1 large enough to profit from the level-1 averaging, or if instead the variance of the factorized contribution is dominated by the fluctuations of the spatial components of the gauge field at the boundary. If n_1 can be taken large enough such that the product of the (level-1) mean values is proportional to $\exp\left\{-M_\pi|y_0 - x_0|\right\}$, then a good statistical precision is reached with a number of updates $(n_0 \cdot n_1) \propto \exp\left\{M_\pi|y_0 - x_0|\right\}$. Notice that the factor in the exponent is halved with respect to the standard Monte Carlo.

The contribution from $W_{P_d}^{(r_1)}(y,x)$ is expected to be suppressed, for a typical configuration, by a factor $\exp\{-M_\pi|x_0 - x_0^{\text{cut}}|\}$ at large time separations. Measuring it over the $n_0 \cdot n_1$ configurations generated in the two-level update, by blocking the results and averaging over the n_0 of them, may be enough to reduce the error at the same level of the one of $W_{P_d}^{(f)}(y,x)$ (up to a pre-factor that have to be quantified numerically). The last contribution, $W_{P_d}^{(r_2)}(y,x)$, is expected to be already proportional to $\exp\{-M_\pi|y_0 - x_0|\}$. This is of the same order of the expected signal, and therefore the standard level-0 average is adequate.

4 Factorization of the approximated quark propagator

The decomposition discussed in section 2 can be generalized for approximating the propagator between two points with a large temporal separation. A simple domain decomposition, where this can be done in practice, is the one where the lattice is divided in thick time-slices Λ_i all[2] of thickness Δ, with $i = 0, \ldots, n-1$, $n = T/\Delta$ and T being the time-extension of the lattice (see Fig. 2). The block structure of the Wilson–Dirac

---

[2]We choose this set up for simplicity. The lattice can, of course, be divided in domains of different sizes if required by a specific problem.

matrix[3] is then given by

$$
D = \begin{pmatrix}
\cdots & \cdots & D_{\Lambda_{i-1,i}} & 0 & 0 & 0 \\
0 & D_{\Lambda_{i,i-1}} & D_{\Lambda_{i,i}} & D_{\Lambda_{i,i+1}} & 0 & 0 \\
0 & 0 & D_{\Lambda_{i+1,i}} & D_{\Lambda_{i+1,i+1}} & D_{\Lambda_{i+1,i+2}} & 0 \\
0 & 0 & 0 & D_{\Lambda_{i+2,i+1}} & \cdots & \cdots
\end{pmatrix} , \tag{4.1}
$$

where $D_{\Lambda_{i,i}}$ acts on quark fields in Λ_i in the same way as D, except that all terms involving the exterior boundary points are set to zero. The off-diagonal terms on the r.h.s. of Eq. (4.1) are given by

$$
D_{\Lambda_{i,i-1}}(\boldsymbol{x}, \boldsymbol{y}) = P_+ U_0^\dagger(x_0^i - 1, \boldsymbol{x}) \, \delta_{\boldsymbol{x}, \boldsymbol{y}} , \tag{4.2}
$$

$$
D_{\Lambda_{i,i+1}}(\boldsymbol{x}, \boldsymbol{y}) = P_- U_0(x_0^{i+1} - 1, \boldsymbol{x}) \, \delta_{\boldsymbol{x}, \boldsymbol{y}} ,
$$

where $U_0(x)$ are the temporal links, $P_\pm = (1 \pm \gamma_0)/2$, and $x_0^j = j \cdot \Delta$ is the first time slice of the block Λ_i. By using Eq. (A.4), it is easy to show that

$$
D_{\Lambda_{i,i}} = \gamma_5 \, D_{\Lambda_{i,i}}^\dagger \, \gamma_5 , \qquad D_{\Lambda_{i,i-1}} = \gamma_5 \, D_{\Lambda_{i-1,i}}^\dagger \, \gamma_5 . \tag{4.3}
$$

The blocking in Eq. (4.1) and the decomposition in Eq. (B.10), or equivalently in Eq. (B.11), are the basic ingredients for constructing an approximated propagator between two points whose distance is much larger than Δ. This can be achieved as described in the following three steps.

Step 1: If $x \in \Lambda_m$ and $y \in \Lambda_l$, with $l > m$ for instance, choose

$$
\Gamma_0 = \Lambda_{m-1} \cup \Lambda_m \cup \cdots \cup \Lambda_l \cup \Lambda_{l+1} . \tag{4.4}
$$

Thanks to the results in section 2

$$
S(y, x) = D_{\Gamma_0}^{-1}(y, x) + \dots . \tag{4.5}
$$

Step 2: Decompose Γ_0 in overlapping domains

$$
\Omega_i = \Lambda_{m-1} \cup \Lambda_m \cup \cdots \cup \Lambda_{i-1} , \qquad i = m+2, \dots, l+1 , \tag{4.6}
$$

and[4]

$$
\Omega_i^* = \Lambda_i \cup \Lambda_{i+1} , \tag{4.7}
$$

with the boundary operators of Ω_i given by

$$
D_{\partial\Omega_i} = D_{\Lambda_{m-1,m-2}} + D_{\Lambda_{i-1,i}} , \qquad D_{\partial\Omega_i^*} = D_{\Lambda_{m-2,m-1}} + D_{\Lambda_{i,i-1}} . \tag{4.8}
$$

---

[3]The same decomposition applies to the $O(a)$-improved Wilson-Dirac operator as well.
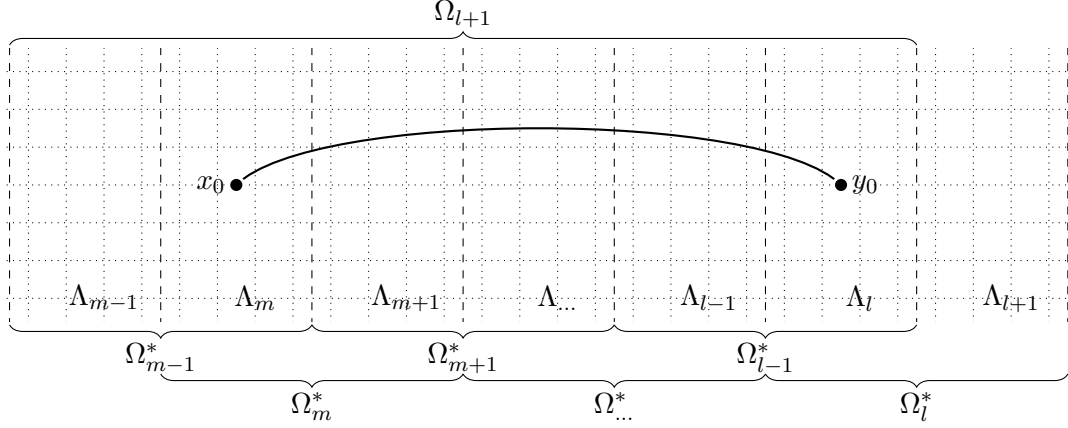[4]Notice that $\Omega_l \cup \Omega_l^* = \Gamma_0$ is valid for $i = l$ only.

Figure 2: Domain decomposition of the lattice in thick time-slices, with the sink and the source of the quark propagator belonging to blocks distant in time.

By taking the bottom-left off-diagonal element in Eq. (B.10), one arrives at

$$D_{\Gamma_0}^{-1}(y, x) = - \sum_{\substack{w_1 \in \partial\Omega_l \\ w_2 \in \partial\Omega_l^*}} D_{\Omega_l^*}^{-1}(y, w_1) \, D_{\partial\Omega_l^*}(w_1, w_2) \, D_{\Gamma_0}^{-1}(w_2, x) \,. \tag{4.9}$$

Step 3: Since w_2 and x in Eq. (4.9) are both at least at a distance Δ from the exterior boundary of Ω_{l+1}, one can replace

$$D_{\Gamma_0}^{-1}(w_2, x) = D_{\Omega_{l+1}}^{-1}(w_2, x) + \ldots \tag{4.10}$$

and arrive to

$$D_{\Gamma_0}^{-1}(y, x) = - \sum_{\substack{w_1 \in \partial\Omega_l \\ w_2 \in \partial\Omega_l^*}} D_{\Omega_l^*}^{-1}(y, w_1) \, D_{\Lambda_{l,l-1}}(w_1, w_2) \, D_{\Omega_{l+1}}^{-1}(w_2, x) + \ldots \,. \tag{4.11}$$

The boundary operator $D_{\partial\Omega_l^*}$ has been replaced by $D_{\Lambda_{l,l-1}}$ since this is the only component acting on fields in Ω_l^*.

By iterating $(m - l)$ times steps 2 and 3, it is easy to show that one can define an approximated propagator

$$S^{(\mathrm{f})}(y, x) = (-1)^{m-l} \Big[\prod_{i=l}^{m+1} D_{\Omega_i^*}^{-1} D_{\Lambda_{i,i-1}} \Big](y, \cdot) \, D_{\Omega_{m+2}}^{-1}(\cdot, x) \tag{4.12}$$

which satisfies γ_5-hermiticity. Since in each step the (inverse) matrix factors have been approximated so that the source and the sink coordinates are at least at a distance Δ from the Dirichlet boundary conditions, we expect

$$\mathrm{tr}\{(S(y, x) - S^{(\mathrm{f})}(y, x)) \, (S(y, x)^\dagger - S^{(\mathrm{f})}(y, x)^\dagger)\}^{1/2} \sim e^{-M_\pi \Delta} \,. \tag{4.13}$$

7

The crucial property of the r.h.s. of Eq. (4.12) is that the dependence on the gauge field is factorized. The various propagators $D_{\Omega_i^*}^{-1}$ depend on the values taken by the gauge field in two thick slices only, while the last one $D_{\Omega_{m+2}}^{-1}$ on three. Remarkably the formula (4.12) is a systematic approximation of the *exact* formula in Eq. (B.7) derived from the LU decomposition of the Dirac operator, see Appendix B. A succession of approximations of the type in Eq. (1.1) can finally be constructed by taking larger and larger values of Δ.

4.1 Factorization and baryon symmetry

To insert baryon projectors in the partition function at the boundaries of the blocks, we can introduce in the action of the theory a time-dependent $U(1)_V$ field $\theta(x_0)$ constant in space

$$
\theta(x_0) = \begin{cases} \theta_i & x_0 = \Delta\,(i+1) - 1 \\[2mm] 0 & x_0 \neq \Delta\,(i+1) - 1 \end{cases}
\tag{4.14}
$$

by replacing

$$
D_{\Lambda_{i,i+1}} \longrightarrow e^{-i\theta_i} D_{\Lambda_{i,i+1}} \;, \quad D_{\Lambda_{i+1,i}} \longrightarrow e^{i\theta_i} D_{\Lambda_{i+1,i}} \;.
\tag{4.15}
$$

Since the A_i in Eq. (B.2) are θ-independent, the dependence on θ of the exact propagator can be easily deduced from the Eq. (B.12). It is given by

$$
S(y,x) \longrightarrow \exp\left\{ i \sum_{i=m}^{l-1} \theta_i \right\} S(y,x) \;, \quad x \in \Lambda_m \;, \quad y \in \Lambda_l \;, \quad l > m \;.
\tag{4.16}
$$

The $D_{\Lambda_{i,i}}$ are also θ-independent, and therefore the approximated quark propagator $S^{(\mathrm{f})}(y,x)$ in Eq. (4.12) inherits the very same θ-dependence of the exact one in Eq. (4.16), i.e. the approximation preserves the baryon symmetry.

5 Multi-level integration of pion and baryon propagators

There is no unique way to design a multi-level integration algorithm by starting from Eqs. (4.9)–(4.12). A simple possibility to start with is to divide the full lattice Γ in only two overlapping thick time-slices: $\bar{\Gamma}_0$ which includes the time slices $[0, x_0^{\mathrm{cut}} + \Delta - 1]$ (it includes one more thick time-slice of thickness Δ with respect to the Γ_0 defined in section 3), where $x_0^{\mathrm{cut}} = i_c\,\Delta$, and Γ_0^* which include those in $[x_0^{\mathrm{cut}}, T-1]$. Open boundary conditions in time for the full lattice are again assumed for simplicity. By taking the bottom-left off-diagonal element of the decomposition (B.10), and by replacing the full propagator with $D_{\bar{\Gamma}_0}^{-1}$, the approximated factorized propagator can be written as

$$
S^{(\mathrm{f})}(y,x) = -D_{\Gamma_0^*}^{-1}(y,\cdot) D_{\partial\Gamma_0^*}(\cdot,\cdot) D_{\bar{\Gamma}_0}^{-1}(\cdot,x)
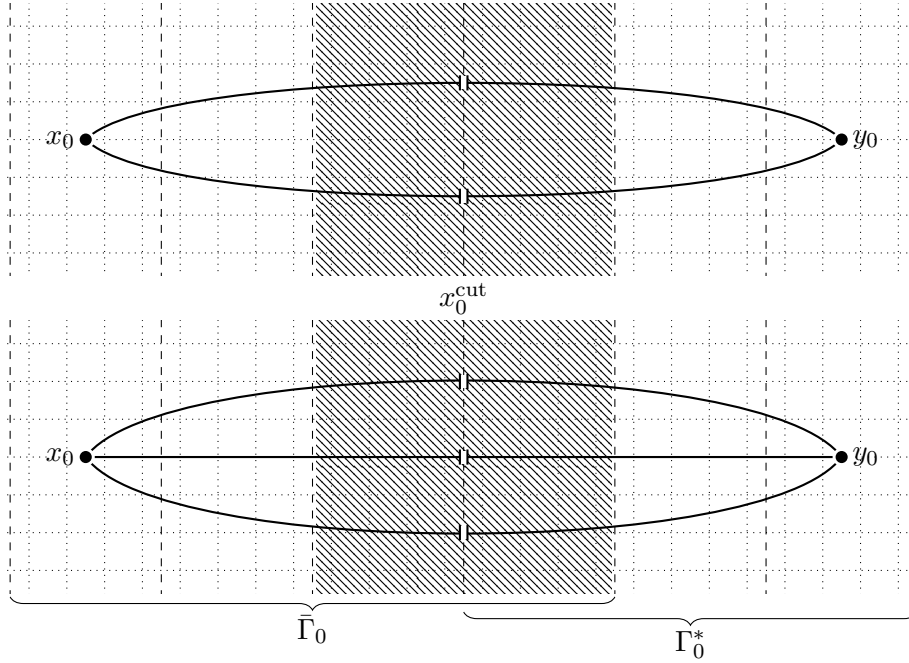\tag{5.1}
$$

Figure 3: Sketch of the factorized approximation of the pion and baryon propagators.

where $x \in \bar{\Gamma}_0$ and $y \in \Gamma_0^*$, see Fig. 3. In this setup Δ *is simply the thickness of* $\bar{\Gamma}_0 \cap \Gamma_0^*$. In order to cut the fermion lines in the Wick contractions, so to transform the matrix products into ordinary ones, we introduce the projector

$$P_L(y, x) = \sum_{i=1}^{N_m} \phi_i(x) \, \phi_i^\dagger(y) \tag{5.2}$$

where ϕ_i are N_m orthonormal vectors. The projector is then used to define a further approximated propagator

$$\tilde{S}^{(\mathrm{f})}(y, x) = -\sum_i [D_{\Gamma_0^*}^{-1} D_{\partial \Gamma_0^*} \phi_i](y) \, [\phi_i^\dagger D_{\bar{\Gamma}_0}^{-1}](x) \ . \tag{5.3}$$

In the following sections we will use two different set of vectors ϕ_i: those which span the deflation subspace as defined in Ref. [12], and N_m orthonormal vectors constructed by applying 10 inverse iterations of the Wilson-Dirac operator defined in the domain $\Omega_{i_c-1}^* = \Lambda_{i_c-1} \cup \Lambda_{i_c}$ with Dirichlet boundary conditions on its exterior boundaries.

5.1 Connected pseudoscalar propagator

The correlation function of two flavor non-diagonal pseudoscalar densities is defined as

$$C_{P_c}(y_0, x_0) = \frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \langle \bar{d}(y)\gamma_5 u(y) \, \bar{u}(x)\gamma_5 d(x) \rangle = -\frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \langle W_{P_c}(y, x) \rangle \ , \tag{5.4}$$

9

where $W_{P_c}(y, x)$ is the corresponding Wick contraction. By using the factorized propagator in Eq. (5.1), the approximated Wick contraction can be written as

$$W_{P_c}^{(\mathrm{f})}(y, x) = \mathrm{tr}\left\{ S^{(\mathrm{f})}(y, x) S^{(\mathrm{f})\,\dagger}(y, x) \right\}. \tag{5.5}$$

When $\bar{\Gamma}_0$, or equivalently Δ, gets larger and larger, the Eq. (5.5) generates a succession of approximations whose rest

$$W_{P_c}^{(\mathrm{r})}(y, x) = W_{P_c}(y, x) - W_{P_c}^{(\mathrm{f})}(y, x) \tag{5.6}$$

converges exponentially fast to zero. Even if the gauge field appears in a factorized form, it is difficult to implement a multi-level integration scheme by using Eq. (5.5). To this aim we can use the approximated propagator in Eq. (5.3), and define

$$\tilde{W}_{P_c}^{(\mathrm{f})}(y, x) = \sum_{i,j} \tilde{B}_{ij}(y) B_{ji}(x) \tag{5.7}$$

where

$$B_{ij}(x) = [\phi_i^\dagger D_{\bar{\Gamma}_0}^{-1} \gamma_5](x)\, [D_{\bar{\Gamma}_0}^{-1} \gamma_5 \phi_j](x)\ ,$$

$$\tilde{B}_{ij}(y) = [\phi_i^\dagger \gamma_5 D_{\partial \Gamma_0^*} D_{\Gamma_0^*}^{-1} \gamma_5](y)\, [D_{\Gamma_0^*}^{-1} D_{\partial \Gamma_0^*} \phi_j](y)\ , \tag{5.8}$$

and the vectors ϕ_i depend only on the gauge links belonging to the thick time-slice $\Omega_{i_c-1}^*$. When the gauge links in $\Omega_{i_c-1}^*$ are kept frozen, the dependence of the action and of $\tilde{W}_{P_c}^{(\mathrm{f})}(y, x)$ on the remaining link variables in $\bar{\Gamma}_0$ and Γ_0^* is factorized. A multi-level similar to the one designed in section 3 can thus be easily implemented.

5.2 Nucleon propagator

A possible choice for baryon interpolating operators is

$$N = \left[u^{aT} C \gamma_5 d^b \right] d^c \epsilon^{abc}\ , \qquad \bar{N} = \bar{d}^e \left[\bar{d}^f C \gamma_5 \bar{u}^{gT} \right] \epsilon^{feg}\ . \tag{5.9}$$

The corresponding two-point function reads

$$C_N(y_0, x_0) = \frac{1}{L^3} \sum_{\mathbf{x}, \mathbf{y}} \mathrm{tr}\left[\langle N(y) \bar{N}(x) \rangle P_- \right] = \frac{1}{L^3} \sum_{\mathbf{x}, \mathbf{y}} \left\{ \langle W_{N1}(y, x) \rangle - \langle W_{N2}(y, x) \rangle \right\}, \tag{5.10}$$

where $C = i\gamma_0 \gamma_2$, and the trace is over the nucleon spinor indices. The $W_{N1}(y, x)$ and $W_{N2}(y, x)$ are the two Wick contractions

$$W_{N1}(y, x) = \mathrm{tr}[S_u^{ag\,T}(y, x) C \gamma_5 S_d^{bf}(y, x) C \gamma_5]\, \mathrm{tr}[S_d^{ce}(y, x) P_-] \epsilon^{abc} \epsilon^{feg}\ ,$$

$$W_{N2}(x, y) = \mathrm{tr}[S_u^{ag\,T}(y, x) C \gamma_5 S_d^{be}(y, x) P_- S_d^{cf}(y, x) C \gamma_5] \epsilon^{abc} \epsilon^{feg}\ , \tag{5.11}$$

10

where again the trace is on the spinor index only. The factorized approximation of the nucleon Wick contractions $W_{N1}^{(f)}$ and $W_{N2}^{(f)}$ are defined analogously to W_{N1} and W_{N2} by replacing on the r.h.s. of Eqs. (5.11) each quark propagator by its approximated factorized expression in Eq. (5.1). Finally the corresponding $\tilde{W}_{N1}^{(f)}$ and $\tilde{W}_{N2}^{(f)}$ are defined as

$$\tilde{W}_{N1}^{(f)}(y, x) = -\sum_{ijk} \mathbb{B}[\xi_j, \xi_i, \xi_k; x]^T \gamma_5 P_- \mathbb{B}[\eta_i, \eta_j, \eta_k; y]$$

$$\tilde{W}_{N2}^{(f)}(y, x) = -\sum_{ijk} \mathbb{B}[\xi_k, \xi_i, \xi_j; x]^T \gamma_5 P_- \mathbb{B}[\eta_i, \eta_j, \eta_k; y] \tag{5.12}$$

where

$$\xi_i^T = \phi_i^\dagger D_{\bar{\Gamma}_0}^{-1} , \qquad \eta_i = D_{\Gamma_0^*}^{-1} D_{\partial \Gamma_0^*} \phi_i , \tag{5.13}$$

and the colorless spinors

$$\mathbb{B}[s_i, s_j, s_k; x]_\alpha = \epsilon^{abc} \Big\{ [s_i^T]^a(x) \, C \gamma_5 \, [s_j]^b(x) \Big\} [s_k]_\alpha^c(x) \tag{5.14}$$

have been introduced. By choosing again the vectors ϕ_i to depend only on the gauge links belonging to the thick time-slice $\Omega_{i_c-1}^*$, the two colorless spinors on the r.h.s. of Eqs. (5.12) depend only on the gauge links belonging to $\bar{\Gamma}_0$ and $\Gamma_0^* \cup \Omega_{i_c-1}^*$ respectively. When the links belonging to $\Omega_{i_c-1}^*$ are kept frozen, the dependence of the action and of the two approximated Wick contractions is factorized. Also in this case their mean value can then be computed with a two-level algorithm analogous to what has been described in section 3 for the disconnected pseudoscalar propagator. Contrary to the pion propagator, the signal-to-noise ratio in the baryon correlator is exponentially suppressed with the distance of the sources. If it turns out to be profitable to choose n_1 so that the product of the (level-1) mean value of the two colorless spinors is significantly reduced, and possibly proportional to $\exp\{-M_N|y_0 - x_0|\}$, then a good statistical precision is reached with a number of updates of the lattice $(n_0 \cdot n_1) \propto \exp\{(M_N - 3M_\pi/2)|y_0 - x_0|\}$. Notice again that the factor at the exponent is halved with respect to the standard Monte Carlo procedure. The remaining correction can be computed by a two-level algorithm with a succession of simulations with larger and larger $\bar{\Gamma}_0$. Also in this case the real effectiveness of the multi-level can only be quantified by a realistic numerical test, see below.

6   Numerical tests for the disconnected pseudoscalar propagator

We test the ideas discussed in section 3 in the quenched approximation of QCD. We discretize gluons and fermions with the Wilson action, and we impose open and periodic boundary conditions in the time and spatial directions respectively [11, 13]. The inverse coupling constant is fixed to $\beta = 6/g_0^2 = 6.0$, the length of each spatial direction to $L = 24$, and the time extent to $T = 64$. The lattice spacing is $a = 0.093$ fm as fixed

by assuming a physical value of 0.5 fm for the Sommer scale $r_0/a = 5.368$ [14]. The up and down quarks are degenerate. Their masses are fixed by the hopping parameter value $k = 0.1560$, corresponding to a pion of approximatively 455 MeV [15].

Numerical simulations have been carried out with a modified version of the open-QCD code version 1.4 [13, 16]. We have generated $n_0 = 200$ level-0 independent gauge field configurations spaced by 400 molecular-dynamics units (MDUs) with the Hybrid Monte Carlo (HMC). Following section 3, the lattice has been split at $x_0^{\text{cut}} = 32$ in two domains of equal size Γ_0 and Γ_0^*. For all level-0 background gauge fields, $n_1 = 100$ level-1 configurations were generated by updating the two regions independently while keeping fixed the spatial links at $x_0^{\text{cut}} = 32$. Also for these updates we used the HMC by skipping 400 MDUs between measurements, a very conservative choice for which the generation of the level-1 configurations is still cheaper than the computation of the Wick contractions. Within this setup, the correlator in Eq. (3.1) is naturally decomposed as

$$C_{P_d}(y_0, x_0) = C_{P_d}^{(\text{f})}(y_0, x_0) + C_{P_d}^{(\text{r}_1)}(y_0, x_0) + C_{P_d}^{(\text{r}_2)}(y_0, x_0) \, . \tag{6.1}$$

The fully factorized contribution can be written as

$$C_{P_d}^{(\text{f})}(y_0, x_0) = \frac{1}{L^3} \left\langle \left[\sum_{\boldsymbol{x}} \text{tr} \left\{ \gamma_5 D_{\Gamma_0}^{-1}(x, x) \right\} \right] \left[\sum_{\boldsymbol{y}} \text{tr} \left\{ \gamma_5 D_{\Gamma_0^*}^{-1}(y, y) \right\} \right] \right\rangle \, . \tag{6.2}$$

The other two terms are given by

$$C_{P_d}^{(\text{r}_1)}(y_0, x_0) = \frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \left\langle W_{P_d}^{(\text{r}_1)}(y, x) + (\Gamma_0, x) \leftrightarrow (\Gamma_0^*, y) \right\rangle \, , \tag{6.3}$$

$$C_{P_d}^{(\text{r}_2)}(y_0, x_0) = \frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \left\langle W_{P_d}^{(\text{r}_2)}(y, x) \right\rangle \, , \tag{6.4}$$

where $W_{P_d}^{(\text{r}_1)}(y, x)$ and $W_{P_d}^{(\text{r}_2)}(y, x)$ are defined in Eqs. (3.5) and (3.6). If the Wilson–Dirac operator is written as

$$2\kappa D = \mathbb{1} - \kappa D_{\text{hop}} \, , \tag{6.5}$$

the trace can be re-expressed as[5] [17–19].

$$\text{tr} \left\{ \gamma_5 D^{-1} \right\} = \kappa^p \, \text{tr} \left\{ \gamma_5 D_{\text{hop}}^p D^{-1} \right\} \qquad p \leq 8 \, . \tag{6.6}$$

By choosing $p = 8$, all the traces appearing in the contributions (6.2)–(6.4) have been estimated stochastically, e.g.

$$\sum_{\boldsymbol{x}} \text{tr} \left\{ \gamma_5 D^{-1}(x, x) \right\} \longrightarrow \frac{1}{n_{\text{src}}} \sum_{i=1}^{n_{\text{src}}} \sum_{\boldsymbol{x}} \eta_i^\dagger(x) \left[\kappa^8 D_{\text{hop}}^8 D^{-1} \gamma_5 \eta_i \right](x) \, , \tag{6.7}$$

---

[5]For the $O(a)$-improved Wilson-Dirac operator $p \leq 2$.

by inverting the various Dirac operators on the very same $n_{\mathrm{src}} = 100$ Gaussian random sources η_i [20, 21], defined on the whole space-time volume[6], and by contracting the solution with a time-slice of η_i. The hopping parameter expansion, used in Eq. (6.6), reduces the variance of the stochastic estimator significantly. Other techniques [22–25] may further reduce the cost of the computation, but we prefer to keep it simple and focus on factorization.

The $C_{P_d}^{(\mathrm{f})}$ contribution is estimated by first averaging, for each of the level-0 configurations, the two traces independently over the n_1 level-1 background fields. The expectation value of the product of the two means is then computed by averaging over the n_0 configurations. The other two contributions are computed as if the n_1 (subset of) configurations, generated for each of the level-0 boundary fields, were correlated level-0 ones. The measures of $C_{P_d}^{(\mathrm{r}_1)}$ and $C_{P_d}^{(\mathrm{r}_2)}$ are thus grouped in bins of n_1, and the expectation values and their errors are determined as usual by treating the bins as n_0 independent measurements.

6.1 Numerical results

The numerical results for $C_{P_d}^{(\mathrm{f})}$, $C_{P_d}^{(\mathrm{r}_1)}$, and $C_{P_d}^{(\mathrm{r}_2)}$ are plotted in Fig. 4 as a function of the time separation of the pseudoscalar densities. The central values and their errors are shown in the plots on the left and right columns respectively. The best estimate of C_{P_d} (the sum of the three) is also shown in each plot on the left for comparison. In all cases x_0 and y_0 belong to different domains, $y_0 > x_0$, and they are chosen to be as much as possible equidistant from x_0^{cut}.

The statistical error on $C_{P_d}^{(\mathrm{f})}$, top-right plot of Fig. 4, is a flat function of $|y_0 - x_0|$ with sizeable deviations near the boundaries of the domains. Error bars are smaller than the symbols. Up to the largest value that we have, $n_1 = 100$, the error decreases as n_1^{-1}, i.e. the two-level Monte Carlo works at full potentiality. The mean value of $C_{P_d}^{(\mathrm{f})}$, top-left plot, is compatible with zero. The correlation between C_{P_d} and $C_{P_d}^{(\mathrm{f})}$ goes from 0.9 to 1.0 when $|y_0 - x_0|$ varies from 15 to 50, a value which collapses toward zero when the multi-level is switched on.

The statistical error on $C_{P_d}^{(\mathrm{r}_1)}$, middle-right plot of Fig. 4, shows a strong dependence on $|y_0 - x_0|$. It is compatible with an exponential behavior of the form $\exp\{-M|y_0 - x_0|/2\}$ with an effective mass $M = 0.14$, i.e. lighter than expected and roughly 2/3 of the pion mass[7]. It decreases as $n_1^{-1/2}$ up to $n_1 = 100$ and, at fixed time distance, it becomes the dominant contribution to the error of C_{P_d} once a large enough number n_1 of level-1 updates have been carried out. The mean value of $C_{P_d}^{(\mathrm{r}_1)}$ is roughly 2/3 of the full correlator at $|y_0 - x_0| = 15$, and it becomes the dominant contribution up to

---

[6]For the factorized contribution the η_i acts effectively as two independent random sources, one for each domain. The estimate of the two traces is thus obtained with a single global inversion per random source.

[7]We did not attempt to study the dependence of this parameter on the finite size or other sources of systematics.
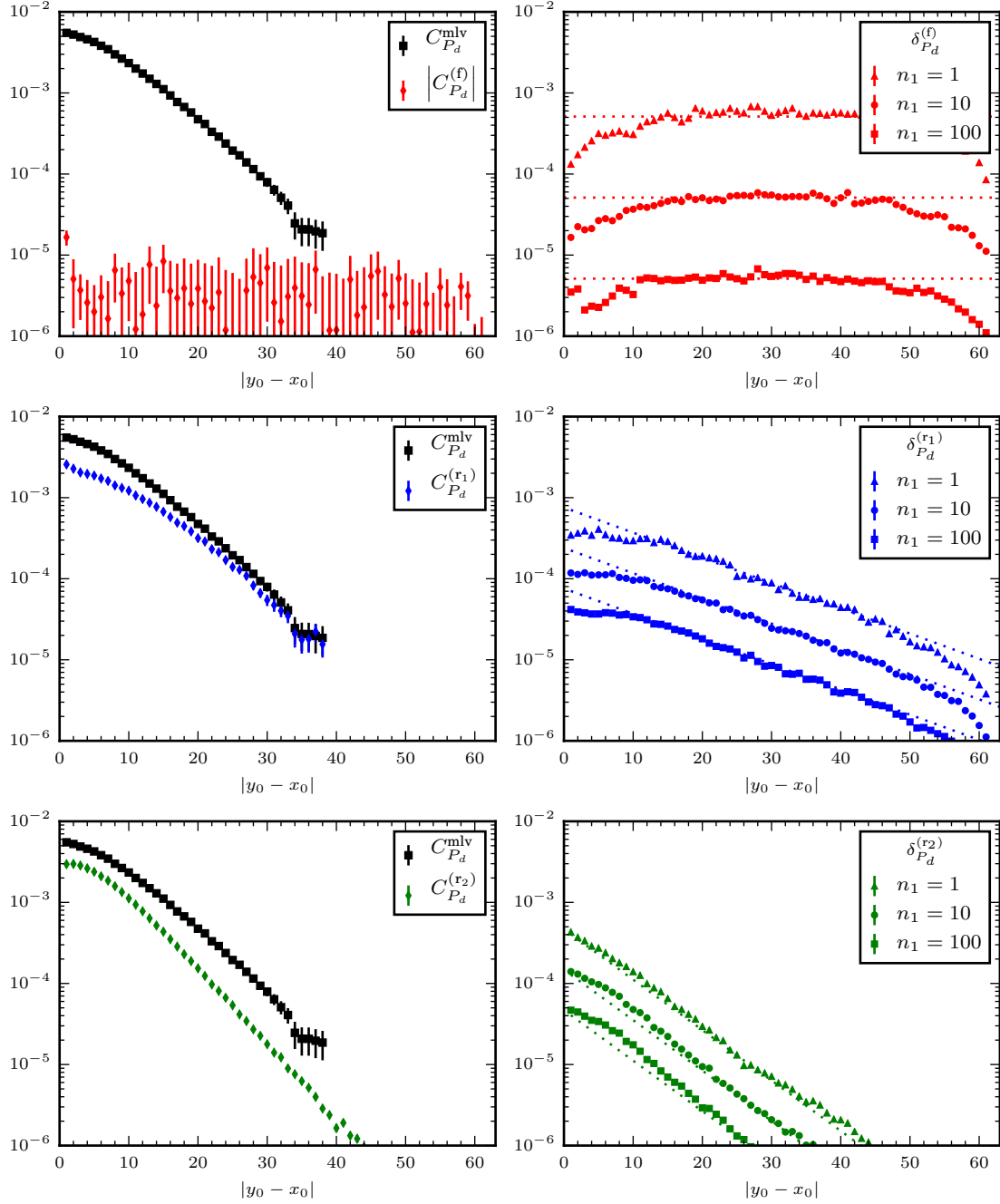
Figure 4: Left-column plots: the three contributions on the r.h.s. of Eq. (6.1) are shown, together with the best estimate of the full correlator (the sum of the three), as a function of the time separation $|y_0 - x_0|$. Right-column plots: the errors of the various contributions are shown as a function of the time distance for various values of n_1.
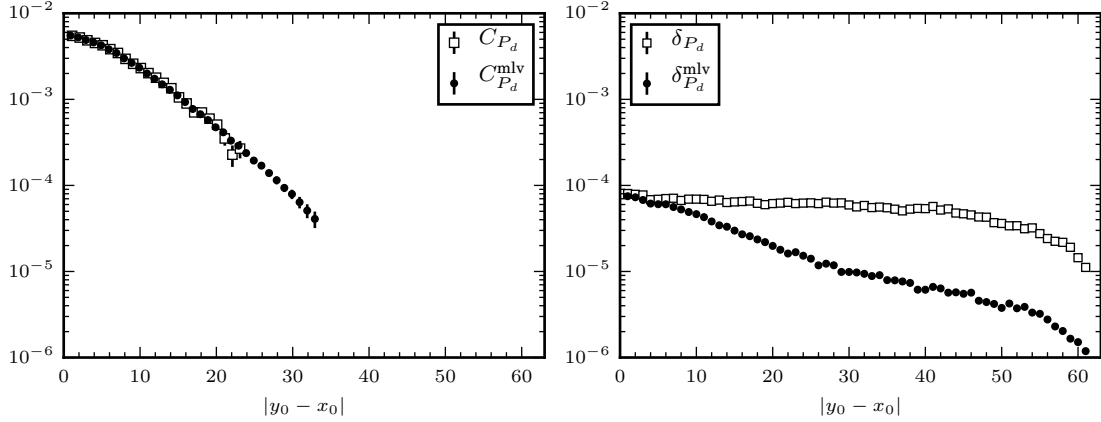
14

Figure 5: The best estimate of $C_{P_d}(y_0, x_0)$ (left) and of its error (right) are shown as a function of the time distance, with and without two-level integration of the factorized contribution. In the latter case the n_1 (subset of) configurations, generated for each of the level-0 boundary fields, are treated as if they were correlated level-0 ones. The n_1 measures are thus binned together, and the mean and its error are computed as usual by treating the bins as independent.

$|y_0 - x_0| = 33$, after which the signal is lost. The statistical errors of $C_{P_d}^{(r_2)}$ decreases exponentially as $\exp\{-M|y_0 - x_0|\}$, and it scales as $n_1^{-1/2}$.

A clear picture emerges from the above analysis. At large time distances, the statistical error on the standard estimate of the disconnected pseudoscalar propagator is dominated by the one on $C_{P_d}^{(f)}$. The second largest contribution is the statistical error on $C_{P_d}^{(r_1)}$ which, however, is exponentially suppressed as $\exp\{-M|y_0 - x_0|/2\}$. Once the two-level integration is switched on, the error on $C_{P_d}^{(f)}$ decreases as n_1^{-1}, while the one on $C_{P_d}^{(r_1)}$ continues to scale as[8] $n_1^{-1/2}$. The parameter n_1 can thus be tuned, up to a prefactor of $\mathcal{O}(1)$, so that $n_1 \sim \exp\{Md\}$ with d being the maximum temporal distance in which one is interested in[9]. This way the error on the factorized contribution is reduced to the level of (or below) the uncertainty on $C_{P_d}^{(r_1)}$ at the same cost of generating $n_0 \cdot n_1$ global configurations. The net computational gain is therefore $\sim n_1$, and a good statistical precision is reached with a number of updates $(n_0 \cdot n_1) \propto \exp\{M_\pi|y_0 - x_0|\}$. Notice that the factor at the exponent is halved with respect to the standard Monte Carlo.

Our best estimate of the disconnected pseudoscalar propagator is shown in Fig. 5, where also the result without the multi-level is reported for comparison. Using the two-

---

[8] A two-level algorithm can be used to further reduce the statistical error on $C_{P_d}^{(r_1)}$ by a domain decomposition of the exact inverse D^{-1} in Eq. (3.5) with the cut at, for instance, $x_0 = 40$. This is an improvement which goes beyond the exploratory numerical study of this work.

[9] The contributions $C_{P_d}^{(f)}$ and $C_{P_d}^{(r_1)}$ can be computed with different number of sources, different value of p for the HPE, etc. The prefactor of this estimate can thus change depending on the details of the computation.

level algorithm, the signal-to-noise ratio remains larger than 1 for ten additional time slices. This is better seen in the right plot, where the statistical error is shown in the two cases. With the standard Monte Carlo the error is approximatively flat, while for the two-level algorithm decreases exponentially. The reduction is evident from distance 15, and becomes approximatively $n_1^{1/2} = 10$ at the point $x_0 = 30$ that was taken to fix n_1. For $n_1 = 100$, the overall gain in the computational cost is approximatively 50 since we have to invert two time the Wilson–Dirac operator on each random source.

7   Numerical tests for the pion propagator

We have tested the factorized approximation of the quark propagator within the same lattice setup of the previous section. The number of independent gauge configurations have been increased to 1000. For each of them 10 Gaussian random sources have been generated on the time-slice at $x_0 = 4$, and the exact and the approximated quark propagators, as defined in Eqs. (5.1) and (5.3), have been computed on the sources. The corresponding pion propagators have then been calculated by contracting the indices and averaging over the random sources as usual. The very same sources have been used for $C_{P_c}(y_0, x_0)$ in Eq. (5.4), for

$$C_{P_c}^{(X)}(y_0, x_0) = -\frac{1}{L^3} \sum_{\boldsymbol{x}, \boldsymbol{y}} \langle W_{P_c}^{(X)}(y, x) \rangle \qquad X = \mathrm{f, r} \;, \tag{7.1}$$

and for the analogous correlators $\tilde{C}_{P_c}^{(\mathrm{f})}$ and $\tilde{C}_{P_c}^{(\mathrm{r})}$. Similarly to Eq. (1.1), the pion propagator can be expanded as

$$C_{P_c} = C_{P_c}^{(0)} + \left[C_{P_c}^{(1)} - C_{P_c}^{(0)} \right] + \left[C_{P_c}^{(2)} - C_{P_c}^{(1)} \right] + C_{P_c}^{(\mathrm{rest})} \;. \tag{7.2}$$

where $C_{P_c}^{(i)}$ for $i = 0, 1, 2$ is a succession of factorized propagators $C_{P_c}^{(\mathrm{f})}$ as defined in Eq. (7.1) for $x_0^{\mathrm{cut}} = 24$ and for $\Delta = 8, 12, 16$ respectively, while $C_{P_c}^{(\mathrm{rest})} = C_{P_c}^{(\mathrm{r})}$ for $\Delta = 16$.

The correlation between $C_{P_c}(y_0, 4)$ and $C_{P_c}^{(\mathrm{f})}(y_0, 4)$ turns out to be practically 1 for all y_0 and for all three values of Δ. In the plots on the top of Fig. 6 we show the central values and the statistical errors of the five terms appearing on both sides of Eq. (7.2) as a function of y_0. The statistical errors are normalized to $C_{P_c}(y_0, 4)$. If we pick up a typical point, $y_0 = 40$, we get

$$C_{P_c}^{(0)} = (101.1 \pm 1.3) \cdot 10^{-7} \;,$$

$$\left[C_{P_c}^{(1)} - C_{P_c}^{(0)} \right] = (5.22 \pm 0.15) \cdot 10^{-7} \;,$$

$$\left[C_{P_c}^{(2)} - C_{P_c}^{(1)} \right] = (1.59 \pm 0.07) \cdot 10^{-7} \;, \tag{7.3}$$

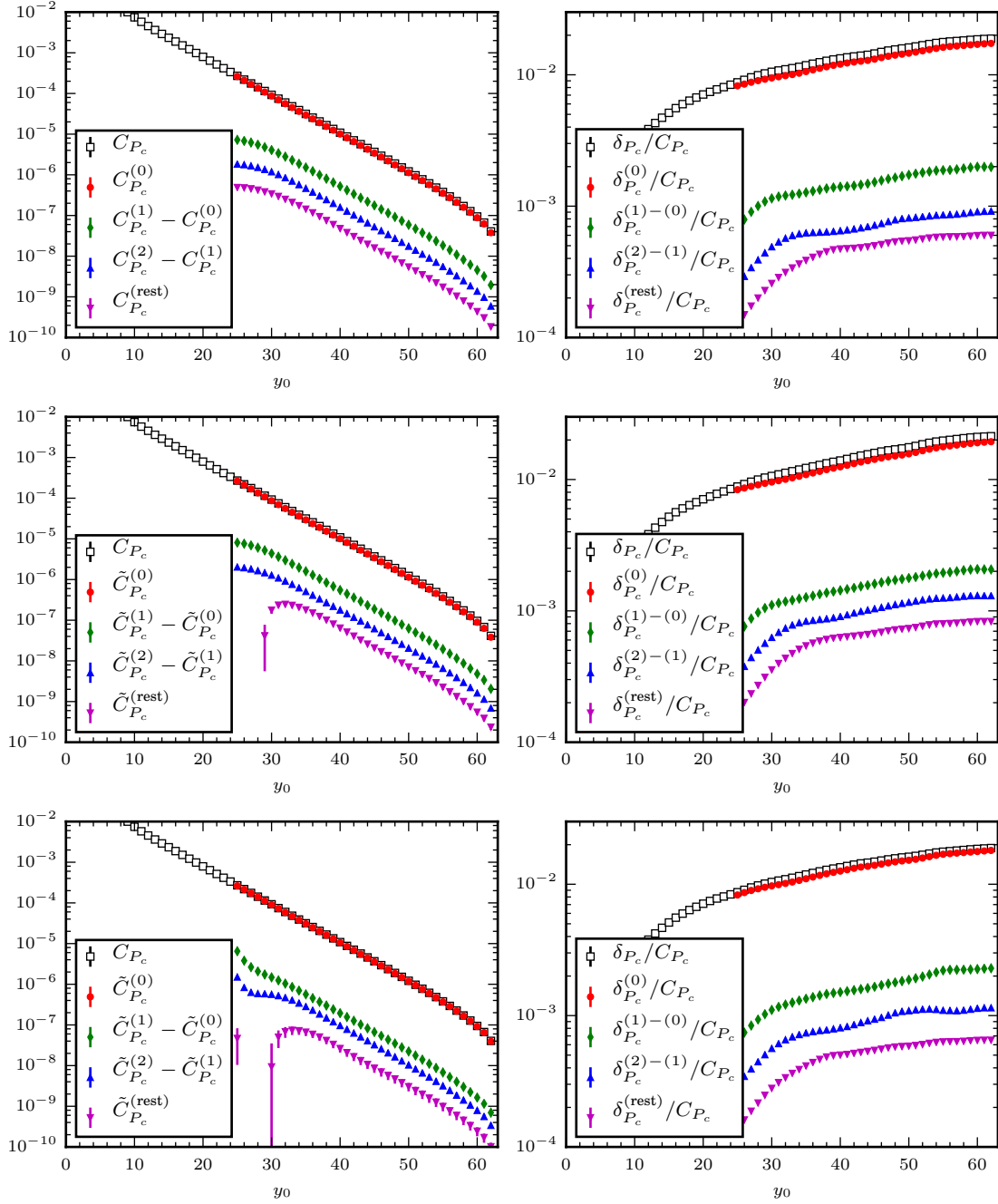$$C_{P_c}^{(\mathrm{rest})} = (0.48 \pm 0.05) \cdot 10^{-7} \;.$$

Figure 6: Top-line plots: central values (left) and their statistical errors (right) of the five terms appearing on both sides of Eq. (7.2). Middle-line plots: analogous results but for a factorized approximation where a projector on the deflation subspace has been inserted to cut the fermion lines. Bottom-line plots: the same but with fermion lines cut by a projector defined via 120 modes computed by the inverse iteration technique, see main text.

17

All these results show that, for this quark mass, the factorized correlator approximates the exact one at the level of 5% already for $\Delta = 8$, a precision which increases by one order of magnitude for $\Delta = 16$. The reduction of the central value of $[C_{P_c}^{(i)} - C_{P_c}^{(i-1)}]$ is in line (even a bit faster) with the expectations from Eq. (4.13), while the decreasing of its statistical error is a bit slower.

The two plots in the middle of Fig. 6 show analogous results but with the factorized propagators computed by inserting $\tilde{S}^{(f)}(y, x)$, as defined in Eq. (5.3), in the contraction. The set of orthonormal vectors ϕ_i are chosen to be those which form the deflation subspace generated from $N_s = 60$ global modes as defined in Ref. [12]. The cut is again at $x_0^{\text{cut}} = 24$, and $\Delta = 8, 12, 16$. It is clear that the contribution from the deflation subspace saturates nicely the exact pion correlator, and that the factorization combines well with deflation provided the number of modes is large enough.

Finally in the plots on the bottom of Fig. 6 we show analogous data but with the ϕ_i being $N_m = 120$ orthonormal vectors computed by applying 10 inverse iterations of the block Wilson–Dirac operator, defined in the thick time-slice $\Omega_{i_c-1}^* = \Lambda_{i_c-1} \cup \Lambda_{i_c}$, on randomly chosen vectors. In this case the leading contribution $\tilde{C}_{P_c}^{(f)}$ can be written as a sum of a manageable number of terms defined as in Eqs. (5.7)–(5.8). Since the ϕ_i depend on the gauge links belonging to $\Omega_{i_c-1}^*$ only, each term in the sum could have been computed with a two-level Monte Carlo by keeping frozen the links in $\Omega_{i_c-1}^*$ during the level-1 updates. It is our experience, however, that for the pion propagator this is not cost effective as for baryons, see below. It must also be said that, while $N_m = 120$ is good enough for the lattice setup we have chosen, this number may have to be increased significantly at larger volumes.

8   Tests of two-level integration for the baryon propagator

We have computed the baryon propagator on the same $n_0 = 1000$ configurations used for the pion in section 7. Also in this case we have calculated the exact correlator $C_N(y_0, x_0)$, and the various contributions defined as

$$C_N^{(X)}(y_0, x_0) = \sum_{\mathbf{y}} \left\{ \langle W_{N1}^{(X)}(y, x) \rangle - \langle W_{N2}^{(X)}(y, x) \rangle \right\}, \qquad X = \text{f}, \text{r} \qquad (8.1)$$

by using the exact quark propagator and the factorized approximations as defined in Eqs. (5.1) and (5.3). All of them have been determined starting from 4 local sources located at a randomly chosen spatial position on the time-slice at $x_0 = 4$.

As for the pion, we expand the nucleon propagator as

$$C_N = C_N^{(0)} + \left[C_N^{(1)} - C_N^{(0)} \right] + \left[C_N^{(2)} - C_N^{(1)} \right] + C_N^{(\text{rest})} . \qquad (8.2)$$

where $C_N^{(i)}$ for $i = 0, 1, 2$ is a succession of factorized correlators $C_N^{(f)}$ as defined in Eq. (8.1) for $x_0^{\text{cut}} = 24$ and for $\Delta = 8, 12, 16$ respectively, while $C_N^{(\text{rest})} = C_N^{(r)}$ for
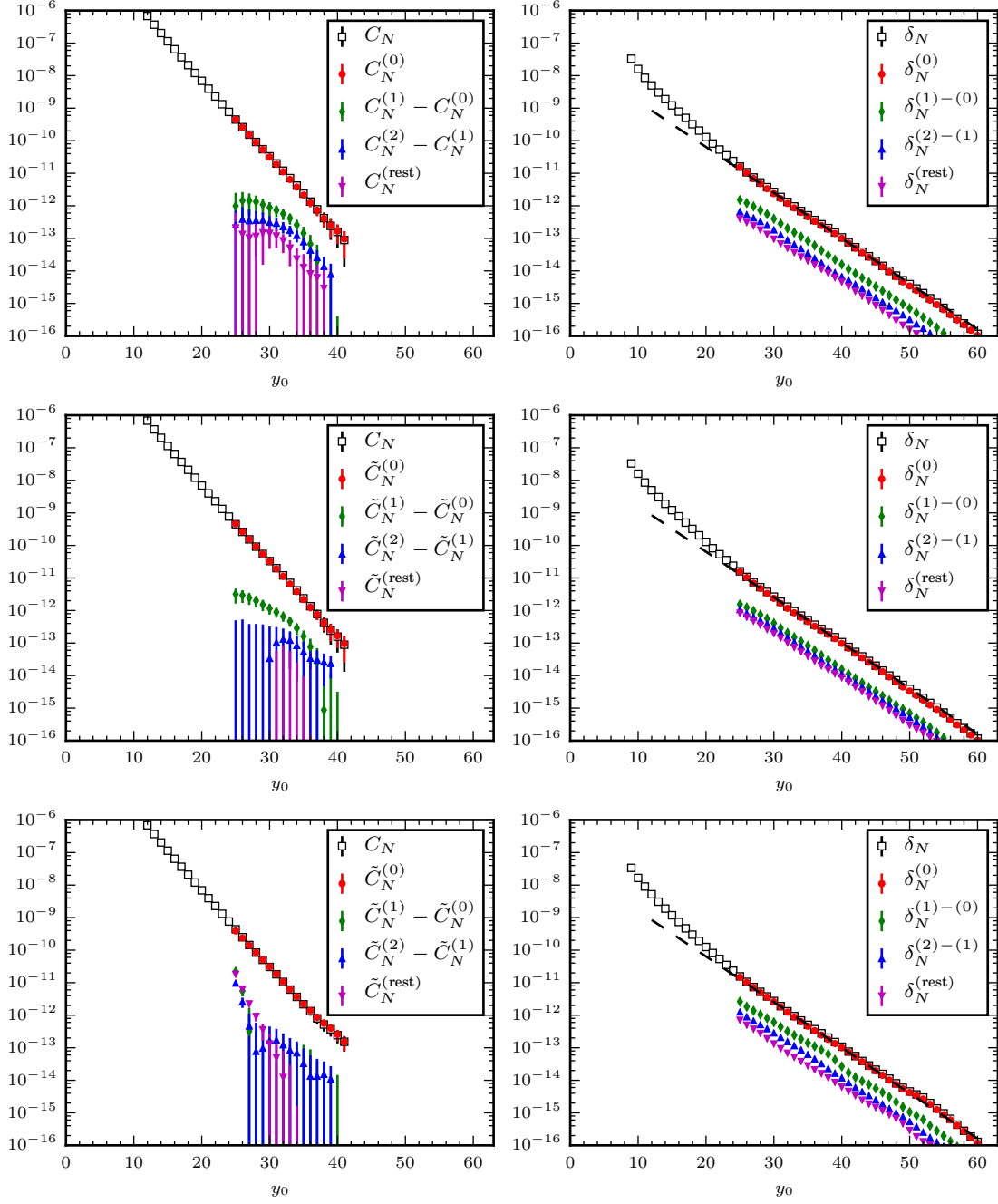
Figure 7: Top-line plots: central values (left) and their statistical errors (right) of the five terms appearing on both sides of Eq. (7.2). For clarity in data are shown only up to $x_0 = 41$, after which the signal for the correlator is lost. Middle-line plots: analogous results but for a factorized approximation where a projector on the deflation subspace has been inserted to cut the fermion lines. Bottom-line plots: the same but with fermion lines cut by a projector defined via 120 modes computed by the inverse iteration technique, see main text.

19
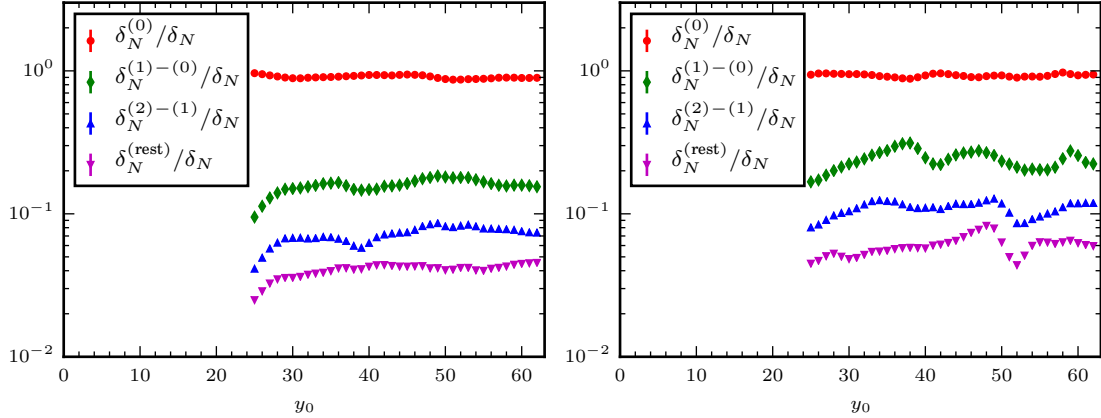
Figure 8: Left: statistical errors of the four terms appearing on the r.h.s. side of Eq. (7.2) normalized to the error of the exact propagator on the l.h.s. of the same equation. Right: analogous results but for a factorized approximation where the fermion lines are cut by the projector defined via the 120 modes computed by the inverse iteration technique, see main text.

$\Delta = 16$. The correlation between $C_N(y_0, 4)$ and $C_N^{(f)}(y_0, 4)$ is practically 1 for all y_0 and for all three values of Δ also for the nucleon. In the plots on the top of Fig. 7 we show the central values (left) and the statistical errors (right) of the five terms appearing on both sides of Eq. (8.2) as a function of y_0. Due to the exponential suppression of the signal with respect to the noise the data on the left plot has large statistical errors, especially the smallest three contributions. On the right plot the statistical error of the exact correlator (and of all the others) follows the expected exponential behavior $\propto e^{-3M_\pi|y_0-4|/2}$ (black line) with $M_\pi = 0.215$ from $y_0 \gtrsim 25$, fully confirming the analysis in Ref. [1]. The hierarchy among the statistical errors of the various terms is evident in the left plot of Fig. 8, where the errors are normalized to the one of the exact correlator. If we pick up a typical point, $y_0 = 35$ not to lose the signal for the correlator, we get

$$C_N^{(0)} = (21.3 \pm 4.6) \cdot 10^{-13} \,,$$

$$\left[C_N^{(1)} - C_N^{(0)}\right] = (1.44 \pm 0.83) \cdot 10^{-13} \,,$$

$$\left[C_N^{(2)} - C_N^{(1)}\right] = (0.78 \pm 0.35) \cdot 10^{-13} \,, \tag{8.3}$$

$$C_N^{(\text{rest})} = (0.13 \pm 0.20) \cdot 10^{-13} \,.$$

At this quark mass, the factorized correlator approximates the exact one at the level of $5 - 10\%$ already for $\Delta = 8$, a precision which increases which $\Delta = 16$ even though the statistical errors are too large to justify a more precise statement. The reduction of the statistical error from top to bottom in Eqs. (8.3) is more than a factor 20, and it seems to decrease a bit slower than the expectation from Eq.(4.13).

The other four plots in Fig. 7 are analogous to the ones for the pions. The two plots in the middle show the results for the factorized propagators computed by inserting in the contraction $\tilde{S}^{(\mathrm{f})}(y, x)$ as defined in Eq. (5.3). The set of orthonormal vectors ϕ_i are chosen to be those which form the deflation subspace. The cut is again at $x_0^{\mathrm{cut}} = 24$, and $\Delta = 8, 12, 16$. The remarkable fact is that the contribution from the deflation subspace saturates nicely the exact nucleon correlator, provided the number of modes is large enough. The factorization combines well with deflation.

Finally in the plots on the bottom of Fig. 7 we show analogous data but with the ϕ_i being $N_m = 120$ orthonormal vectors computed as for the pions. The statistical errors normalized to the one of the full correlator are shown on the right plot of Fig. 8.

8.1 Two-level Monte Carlo

The colorless spinors $\mathbb{B}$ in Eqs. (5.12) depend on the gauge field belonging either to $\bar{\Gamma}_0$ or Γ_0^*. When the links in $\Omega_{i_c}^*$ are kept frozen, the dependence of the action and of the approximated contractions on them is factorized. We can thus compute $C_N^{(\mathrm{f})}(y_0, x_0)$ with a two-level Monte Carlo. For $n_0 = 50$ level-0 configurations (the first 50 of the 1000 generated previously), we generate $n_1 = 20$ level-1 gauge fields[10] by freezing the link variables in $\Omega_{i_c}^*$. The thick time-slice averages of the colorless spinors are then performed independently on the level-1 configurations, and the matrix elements of $\gamma_5 P_-$ between the spinors is averaged over the 50 level-0 boundary fields.

In the top-left plot of Fig. 9 we show the total reduction achieved for the statistical error on $C_N^{(\mathrm{f})}(y_0, x_0)$. In particular what is plotted in this graph is the ratio of (a) the statistical error on the factorized correlator – computed by averaging over the 1000 level-0 configurations at our disposal and the 4 local sources and finally multiplied by $\sqrt{1000/50}$ – over (b) the statistical error achieved with the two-level integration. For $y_0 \gtrsim 30$ we observe a sharp reduction of the error by a factor which becomes 40–50 for $y_0 \gtrsim 40$. The cost per level-0 configuration, without counting scalar products and updates, is roughly 70 times higher. For a given target statistical error, this results in a net reduction of the cost of the simulation of 20–40 times.

The origin of the gain is due to various factors: sum over all the points on the time-slice at $x_0 = 4$, averaging over the n_1 level-1 configurations, two-level averaging. On the top-right plot we show the ratio of the errors of the local estimator, averaged over the 4 local sources and multiplied by $\sqrt{1000/50}$, over the one on $C_N^{(0)}(y_0, x_0)$ obtained with the 50 level-0 configurations only. After a few transient time-slices, the gain is in the range 2–3. On the bottom-left plot we show the gain due to the averaging over the n_1 gauge configurations but without two-level integration. We just bin the n_1 measurements for each level-0 configurations, and we compute the errors considering the bins as independent. After few time-slices were no gain is observed due to the freezing of the links belonging to $\Omega_{i_c}^*$, the gain scales approximatively as $\sqrt{n_1}$. In the bottom-right

---

[10]When n_1 is not a large number, it is feasible to avoid the cut of the fermion lines by computing the factorized propagator on the n_1^2 combinations of level–1 configurations.
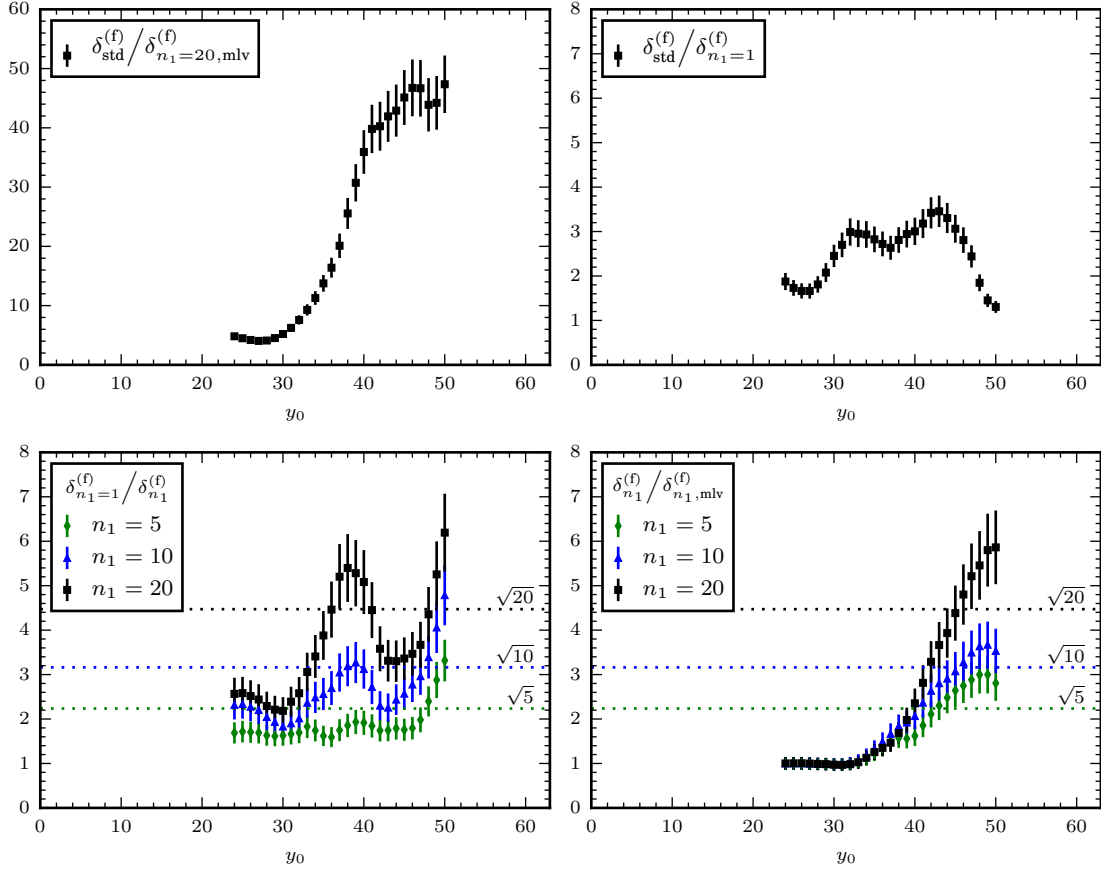
Figure 9: The total gain on the statistical error of $C_N^{(f)}(y_0, x_0)$ in the two-level Monte Carlo (top-left) is due to various factor: the sum over all the points on the time-slice at $x_0 = 4$ (top-right), the averaging over the n_1 level-1 configurations (bottom-left), the two-level averaging of the $\mathbb{B}$-spinors (bottom-right). See main text for more details.

plot the reduction of the error due to the two-level independent averaging is shown. For $y_0 \gtrsim 35$ the gain is clearly visible, and at larger distances saturates the expected $\sqrt{n_1}$ factor up to $n_1 = 20$.

The final results for the correlator with and without the two-level integration are shown in the left plot of Fig. 10. In the two-level Monte Carlo the signal-to-noise ratio for the factorized contribution remains larger than one for 10 additional time-slices with respect to the standard evaluation. When we add the rest of the correlator, $C_N^{(r)}$, the gain reduces to 5 additional points. The effectiveness of the two-level integration is better seen on the right plot of Fig. 10, which is a replica of the right graph in Fig. 8 but with the error on $C_N^{(0)}$ coming from the two-level integration (error with standard Monte Carlo also shown with shadow red points). For completeness we report also the statistical error on our best two-level estimate of the exact correlator, normalized to the
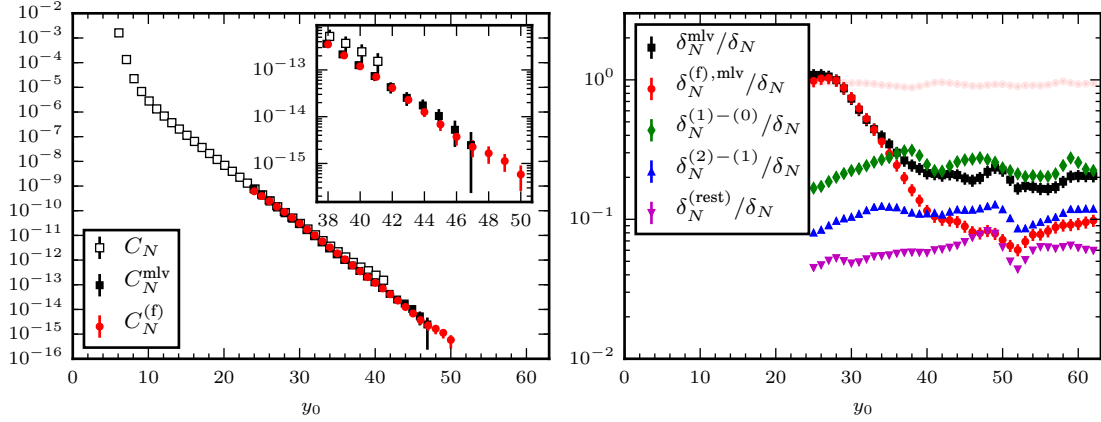
Figure 10: Left: best results for $C_N(y_0, x_0)$ with and without two-level integration, and for $C_N^{(0)}(y_0, x_0)$. Right: same as in the right plot of Fig. 8 but with the error on $C_N^{(0)}(y_0, x_0)$ from the two-level Monte Carlo. For completeness we show also the statistical error on our best two-level estimate of the exact correlator

one obtained with the standard Monte Carlo.

A rather clear picture emerges, similar to what we discussed for the disconnected correlator. At large time distances, the statistical error on the standard estimate of C_N is dominated by the one on $C_N^{(0)}$. Once the two-level integration is switched on, the error on $C_N^{(0)}$ decreases roughly as n_1^{-1}, while the one on the rest of the correlator continues to scale as $n_1^{-1/2}$. We thus tuned n_1 so that the error on $C_N^{(0)}$ is smaller (roughly $1/2$) of the uncertainty on $\left[C_N^{(1)} - C_N^{(0)} \right]$ at the same cost of generating $n_0 \cdot n_1$ global configurations. If one wants to gain further (beyond the scope of this paper), the leading correction $\left[C_N^{(1)} - C_N^{(0)} \right]$ needs also to be integrated with a two-level algorithm with $\Delta = 12$. This way the error on this contribution can be reduced at the level or below the one on $C_N^{(0)}(y_0, x_0)$ and $\left[C_N^{(2)} - C_N^{(1)} \right]$, and so on. A good statistical precision can thus be reached with a number of updates of the lattice $(n_0 \cdot n_1) \propto \exp\{(M_N - 3M_\pi/2)|y_0 - 4|\}$, i.e. a factor at the exponent which is halved with respect to the one in the standard Monte Carlo procedure.

9 Conclusions

The numerical computation of many interesting hadronic correlation functions in lattice QCD suffers from signal-to-noise ratios which decrease exponentially with the time distance of the sources. Notable examples are meson correlators with disconnected contributions, static-light correlators, baryonic correlation functions with and without disconnected Wick contractions, etc. Based on the experience in purely bosonic theories, our physics intuition would suggest that multi-level algorithms would lead to an

impressive acceleration of those computations, opening new perspectives in lattice QCD.

Formulating multi-level integration schemes in systems with fermions, however, is not as straight forward as for bosons. The gauge-field dependence of the fermion determinant and of the propagator need to be judiciously factorized before integrating the Wick contractions. Here we have shown that this can be achieved in (quenched) QCD by properly combining the ideas of multi-level integration and domain decomposition of the quark propagator.

The numerical tests that we have carried out for the disconnected correlator of two flavor-diagonal pseudoscalar densities and for a nucleon two-point function show indeed that the signal-to-noise ratio increases exponentially with the time distance of the sources when a two-level integration is at work. In the very simple setup that we have implemented, the number of configurations needed to reach a given statistical precision is proportionally to the square root of those required in the standard case.

For the strategy to be useful in full QCD, the next step is to devise a similar domain decomposition for the quark determinant.

10 Acknowledgments

A   Wilson–Dirac operator

The massive Wilson–Dirac operator is defined as [26]

$$D = D_{\mathrm{w}} + m_0 \ , \tag{A.1}$$

where m_0 is the bare quark mass, D_{w} is the massless operator

$$D_{\mathrm{w}} = \frac{1}{2} \left\{ \gamma_\mu (\nabla_\mu^* + \nabla_\mu) - \nabla_\mu^* \nabla_\mu \right\} \ , \tag{A.2}$$

γ_μ are the Dirac matrices, and the summation over repeated indices is understood. The covariant forward and backward derivatives ∇_μ and ∇_μ^* are defined to be

$$\nabla_\mu \psi(x) = U_\mu(x)\psi(x + \hat{\mu}) - \psi(x) \ , \quad \nabla_\mu^* \psi(x) = \psi(x) - U_\mu^\dagger(x - \hat{\mu})\psi(x - \hat{\mu}) \ , \tag{A.3}$$

where $U_\mu(x)$ are the link variables. The Wilson–Dirac operator satisfies the γ_5-hermiticity relation

$$D = \gamma_5 D^\dagger \gamma_5 \ . \tag{A.4}$$

B   LU factorization of the block-banded Wilson-Dirac operator

The LU factorization for block banded matrices leads to the simple result for the Wilson–Dirac operator [27]

$$
\begin{pmatrix}
D_{\Lambda_{0,0}} & D_{\Lambda_{0,1}} & 0 & \cdots \\
D_{\Lambda_{1,0}} & D_{\Lambda_{1,1}} & D_{\Lambda_{1,2}} & \cdots \\
0 & D_{\Lambda_{2,1}} & D_{\Lambda_{2,2}} & \\
\vdots & & & \ddots
\end{pmatrix}
=
\begin{pmatrix}
1 & B_0 & 0 & \cdots \\
0 & 1 & B_1 & \cdots \\
0 & 0 & 1 & \\
\vdots & & & \ddots
\end{pmatrix}
\begin{pmatrix}
A_0 & 0 & 0 & \cdots \\
D_{\Lambda_{1,0}} & A_1 & 0 & \cdots \\
0 & D_{\Lambda_{2,1}} & A_2 & \\
\vdots & & & \ddots
\end{pmatrix},
\quad \text{(B.1)}
$$

where the block matrices $D_{\Lambda_{i,j}}$ are defined in Eqs. (4.1)–(4.2), while the A_i and the B_i are defined uniquely in terms of the $D_{\Lambda_{i,j}}$ by the following recursion relations

$$
A_{T-1} = D_{\Lambda_{T-1,T-1}} \,,
$$

$$
A_i = D_{\Lambda_{i,i}} - D_{\Lambda_{i,i+1}} A_{i+1}^{-1} D_{\Lambda_{i+1,i}} \qquad (i = 0, \ldots, T-2)\,,
\qquad \text{(B.2)}
$$

$$
B_i = D_{\Lambda_{i,i+1}} A_{i+1}^{-1} \qquad\qquad\qquad (i = 0, \ldots, T-2)\,.
$$

Using the factorization (B.1), the linear system $D\,\psi = \eta$ can be easily solved, again leading to recursion relations. Let us consider the case where the source η is non-zero only on one thick time-slice Λ_k. Solutions for sources on multiple time slices can be obtained by superposition. The system

$$
\begin{pmatrix}
1 & B_0 & 0 & \cdots & \\
0 & 1 & B_1 & \cdots & \\
0 & 0 & 1 & & \\
\vdots & & & \ddots & \\
0 & & & & 1
\end{pmatrix}
\begin{pmatrix}
\chi_0 \\ \chi_1 \\ \vdots \\ \\ \chi_{T-1}
\end{pmatrix}
=
\begin{pmatrix}
0 \\ \vdots \\ \eta_k \\ \vdots \\ 0
\end{pmatrix}
\qquad \text{(B.3)}
$$

is solved by

$$
\chi_i =
\begin{cases}
\left[\displaystyle\prod_{j=i}^{k-1} (-B_j) \right] \eta_k & i < k \\[2.2em]
\eta_k & i = k \\[0.6em]
0 & i > k
\end{cases}
\qquad \text{(B.4)}
$$

where in the last line the obvious ordered product has to be taken. Using Eq. (B.2) we can rewrite

$$
\prod_{j=i}^{k-1} (-B_j)\,\eta_k = (-)^{k-i}\, (D_{\Lambda_{i,i+1}} A_{i+1}^{-1}) \,\cdots\, (D_{\Lambda_{k-1,k}} A_k^{-1})\,\eta_k \,,
\qquad \text{(B.5)}
$$

where for $i < k$ the χ_i have support only on the boundaries. By solving the system

$$
\begin{pmatrix}
A_0 & 0 & 0 & \cdots \\
D_{\Lambda_{1,0}} & A_1 & 0 & \cdots \\
0 & D_{\Lambda_{1,2}} & A_2 & \\
\vdots & & & \ddots
\end{pmatrix}
\begin{pmatrix}
\psi_0 \\
\vdots \\
\\
\psi_{T-1}
\end{pmatrix}
=
\begin{pmatrix}
\chi_0 \\
\vdots \\
\\
\chi_{T-1}
\end{pmatrix} ,
\tag{B.6}
$$

we get the final result

$$
\begin{aligned}
\psi_0 &= A_0^{-1} \chi_0 , \\
\psi_i &= A_i^{-1}(\chi_i - D_{\Lambda_{i,i-1}} \psi_{i-1}) \qquad (i = 1, \ldots, T - 1) .
\end{aligned}
\tag{B.7}
$$

As for the χ_i, the second term in the parentheses on the r.h.s. of Eq. (B.7) live on the boundaries. The matrix A_i^{-1} propagates these two contributions into the center of the thick time-slice. The Eq. (B.7) is the basis of the so-called Thomas algorithm for the solution of (block) banded linear systems [27].

B.1   The 2×2 case

The previous derivation for the 2×2 block-banded Wilson-Dirac operator

$$
D = \begin{pmatrix} D_{\Lambda_{0,0}} & D_{\Lambda_{0,1}} \\ D_{\Lambda_{1,0}} & D_{\Lambda_{1,1}} \end{pmatrix} = \begin{pmatrix} I & D_{\Lambda_{0,1}} D_{\Lambda_{1,1}}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} S_{\Lambda_{0,0}} & 0 \\ D_{\Lambda_{1,0}} & D_{\Lambda_{1,1}} \end{pmatrix} ,
\tag{B.8}
$$

where the Schur complement is defined as

$$
S_{\Lambda_{0,0}} = D_{\Lambda_{0,0}} - D_{\Lambda_{0,1}} D_{\Lambda_{1,1}}^{-1} D_{\Lambda_{1,0}} ,
\tag{B.9}
$$

leads to

$$
D^{-1} = \begin{pmatrix} S_{\Lambda_{0,0}}^{-1} & -S_{\Lambda_{0,0}}^{-1} D_{\Lambda_{0,1}} D_{\Lambda_{1,1}}^{-1} \\ -D_{\Lambda_{1,1}}^{-1} D_{\Lambda_{1,0}} S_{\Lambda_{0,0}}^{-1} & D_{\Lambda_{1,1}}^{-1} + D_{\Lambda_{1,1}}^{-1} D_{\Lambda_{1,0}} S_{\Lambda_{0,0}}^{-1} D_{\Lambda_{0,1}} D_{\Lambda_{1,1}}^{-1} \end{pmatrix} .
\tag{B.10}
$$

It is worth noting that $S_{\Lambda_{0,0}}^{-1}$ is the exact block in the block inverse of D. By putting the Schur complement in the bottom-right block, the analogous formula can be written as

$$
D^{-1} = \begin{pmatrix} D_{\Lambda_{0,0}}^{-1} + D_{\Lambda_{0,0}}^{-1} D_{\Lambda_{0,1}} S_{\Lambda_{1,1}}^{-1} D_{\Lambda_{1,0}} D_{\Lambda_{0,0}}^{-1} & -D_{\Lambda_{0,0}}^{-1} D_{\Lambda_{0,1}} S_{\Lambda_{1,1}}^{-1} \\ -S_{\Lambda_{1,1}}^{-1} D_{\Lambda_{1,0}} D_{\Lambda_{0,0}}^{-1} & S_{\Lambda_{1,1}}^{-1} \end{pmatrix} .
\tag{B.11}
$$

with $S_{\Lambda_{1,1}}$ defined as in Eq. (B.9) but with $1 \leftrightarrow 0$.

B.2 Approximate factorization

The formulas (B.4) and (B.7) factorize the solution vector in terms of thick time-slice matrix products. Only the matrices A_i carry the dependence on the links that belong to several slices, via (nested) Schur complements. The Eq. (4.12) can be derived from Eqs. (B.4) and (B.7) by a systematic approximation of the LU decomposition. As in section 4 we choose the source point $x \in \Lambda_m$ and the sink $y \in \Lambda_l$ with $l > m$, see Fig. 2. By using Eq. (B.7) it is easy to show that

$$\psi_l = (-1)^{m-l} A_l^{-1} D_{l,l-1} \dots A_{m+1}^{-1} D_{m+1,m} \psi_m \tag{B.12}$$

By approximating A_i with $D_{\Omega_i^*}$ and ψ_m with $D_{\Omega_{m+2}}^{-1} \eta_m$ we arrive at Eq. (4.12).

References

[1] G. Parisi, Phys. Rept. 103 (1984) 203.

[2] G.P. Lepage, TASI 89 Summer School, Boulder, CO, Jun 4-30, 1989.

[3] M. Albanese *et al.* [APE Collaboration], Phys. Lett. B **192** (1987) 163.

[4] M. Lüscher and P. Weisz, JHEP **0109** (2001) 010 [hep-lat/0108014].

[5] H. B. Meyer, JHEP **0301** (2003) 048 [hep-lat/0209145].

[6] M. Della Morte and L. Giusti, Comput. Phys. Commun. **180** (2009) 813.

[7] M. Della Morte and L. Giusti, Comput. Phys. Commun. **180** (2009) 819 [arXiv:0806.2601 [hep-lat]].

[8] M. Della Morte and L. Giusti, JHEP **1105** (2011) 056 [arXiv:1012.2562 [hep-lat]].

[9] Y. Saad, Iterative methods for sparse linear systems, 2nd ed. (SIAM, Philadelphia, 2003); see also http://www-users.cs.umn.edu/~saad/

[10] M. Lüscher, Comput. Phys. Commun. **156** (2004) 209 [hep-lat/0310048].

[11] M. Lüscher and S. Schaefer, JHEP **1107** (2011) 036 [arXiv:1105.4749 [hep-lat]].

[12] M. Lüscher, JHEP **0707** (2007) 081 [arXiv:0706.2298 [hep-lat]].

[13] M. Lüscher and S. Schaefer, Comput. Phys. Commun. **184** (2013) 519 [arXiv:1206.2809 [hep-lat]].

[14] M. Guagnelli *et al.* [ALPHA Collaboration], Nucl. Phys. B **535** (1998) 389 [hep-lat/9806005].

[15] C. R. Allton, V. Gimenez, L. Giusti and F. Rapuano, Nucl. Phys. B **489** (1997) 427 [hep-lat/9611021].

[16] http://luscher.web.cern.ch/luscher/openQCD/

[17] C. Thron, S. J. Dong, K. F. Liu and H. P. Ying, Phys. Rev. D **57** (1998) 1642 [hep-lat/9707001].

[18] C. McNeile *et al.* [UKQCD Collaboration], Phys. Rev. D **63** (2001) 114503 [hep-lat/0010019].

[19] G. S. Bali, S. Collins and A. Schäfer, Comput. Phys. Commun. **181** (2010) 1570 [arXiv:0910.3970 [hep-lat]].

[20] R. Sommer, Nucl. Phys. Proc. Suppl. **42** (1995) 186 [hep-lat/9411024].

[21] M. Foster *et al.* [UKQCD Collaboration], Phys. Rev. D **59** (1999) 074503 [hep-lat/9810021].

[22] T. A. DeGrand and S. Schaefer, Comput. Phys. Commun. **159** (2004) 185 [hep-lat/0401011].

[23] L. Giusti, P. Hernandez, M. Laine, P. Weisz and H. Wittig, JHEP **0404** (2004) 013 [hep-lat/0402002].

[24] J. Foley, K. Jimmy Juge, A. O'Cais, M. Peardon, S. M. Ryan and J. I. Skullerud, Comput. Phys. Commun. **172** (2005) 145 [hep-lat/0505023].

[25] T. Blum, T. Izubuchi and E. Shintani, Phys. Rev. D **88** (2013) 9, 094503 [arXiv:1208.4349 [hep-lat]].

[26] K.G. Wilson, New Phenomena In Subnuclear Physics. Part A. Proceedings of the First Half of the 1975 International School of Subnuclear Physics, Erice, Sicily, July 11 - August 1, 1975, ed. A. Zichichi, Plenum Press, New York, 1977, p. 69, CLNS-321.

[27] A. Quarteroni, R. Sacco, F. Saleri, Numerical Mathematics (Springer-Verlag, New York, 2000).