

Lattice QCD noise reduction for bosonic correlators through blocking

Luis Altenkort¹, Alexander M. Eller², O. Kaczmarek¹, Lukas Mazur³, Guy D. Moore², and Hai-Tao Shu⁴

¹*Fakultät für Physik, Universität Bielefeld, D-33615 Bielefeld, Germany*

²*Institut für Kernphysik, Technische Universität Darmstadt Schlossgartenstraße 2, D-64289 Darmstadt, Germany*

³*Paderborn Center for Parallel Computing, Paderborn University, D-33098 Paderborn, Germany*

⁴*Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany*



(Received 13 December 2021; accepted 27 April 2022; published 20 May 2022)

We propose a method to substantially improve the signal-to-noise ratio of lattice correlation functions for bosonic operators or other operator combinations with disconnected contributions. The technique is applicable for correlations between operators on two planes (zero momentum correlators) when the dimension of the plane is larger than the separation between the two planes which are correlated. In this case, the correlation arises primarily from points whose in-plane coordinates are close, but noise arises from all pairs of points. By breaking each plane into bins and computing bin-bin correlations, it is possible to capture these short-distance correlators exactly while replacing (small) correlators at large spatial extent with a fit, with smaller uncertainty than the data. The cost is only marginally larger than averaging each plane before correlating, but the improvement in signal-to-noise can be substantial. We test the method on correlators of the gradient-flowed topological charge density and squared field strength, finding noise reductions by a factor of ~ 3 – 7 compared to the conventional approach on the same ensemble of configurations.

DOI: [10.1103/PhysRevD.105.094505](https://doi.org/10.1103/PhysRevD.105.094505)

I. INTRODUCTION

Many problems in quantum field theory can be expressed in terms of the correlation function of two operators as a function of separation along one axis, while averaging over directions transverse to that axis. For example, when the operators are the interpolating operator for a particle, the exponential rate of the falloff of the correlator determines the particle mass. Moreover, zero spatial-momentum correlators of conserved currents, such as the energy-momentum tensor (EMT) or the vector current, encode transport coefficients like shear and bulk viscosity or flavor diffusion coefficients and the electrical conductivity in the small-frequency limiting behavior of their reconstructed spectral functions. For a review see Ref. [1]. Some recent lattice studies using this approach for the calculation of viscosities can be found in Refs. [2–6] and for a recent overview of results for the electrical conductivity see Ref. [7]. Problems like these ultimately come down to the computation of correlation functions of operators

averaged over a transverse plane, evaluated as a function of the separation between two planes.

Usually such studies require that the correlation function be determined very precisely. For some operators built out of fermions which carry nontrivial flavor, there are no disconnected contributions. In this case, the precision is generally good, as there are explicit factors of propagators between the two planes, which cause the configuration-by-configuration value of the correlator to decay. In this case the signal to noise is generally good,¹ and the main limitations are, for example, contamination from higher states, the continuum limit, etc. However, for correlation functions of operators consisting of bosonic fields, or correlators built out of fermions such that there are disconnected contributions, signal-to-noise problems are generically severe. Recently there have been some advances in dealing with this problem, such as covariant coordinate-space methods [9,10], frequency splitting methods [11], and the cluster decomposition method [12].

This paper will introduce and demonstrate a new approach for problems where such disconnected contributions lead to severe signal-to-noise problems. In the next section, we review the origin of the noise in such

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

¹In some cases, e.g., correlators of baryon operators, the signal-to-noise is bad even though there are only connected contributions [8]. This paper will not address this problem.

disconnected correlation functions. We show a numerically efficient way to express the correlation function between operators on two planes as an integral over the transverse separation between the operator locations on the planes in Sec. III. The signal is dominated by small transverse separations; the noise is dominated by large transverse separations. By fitting the region with a strong signal and using the fit where the signal is poor, one can avoid these noisy contributions, improving the overall signal-to-noise significantly. We demonstrate this for EMT correlators in the bulk channel and the topological charge density correlator in Sec. IV. As a byproduct of this work we also update the double extrapolated topological charge density correlators which we provided recently [13].

II. ORIGIN OF THE PROBLEM

The origin of noise in correlators was first explored by Parisi [8], and our analysis follows his pioneering work. Consider an operator \mathcal{O} of dimension Δ and its Hermitian conjugate \mathcal{O}^\dagger . Suppose we are interested in the correlation function

$$G(\tau) \equiv \frac{1}{L^3} \int d^3\vec{x} d^3\vec{y} \langle \mathcal{O}^\dagger(\vec{x}, \tau) \mathcal{O}(\vec{y}, 0) \rangle. \quad (1)$$

Here the transverse space integration $d^3\vec{x}$ is over a transverse space of extent L^3 , and we are primarily concerned with the case where $\tau \ll L$. This is the case for transport coefficients because $\tau < 1/(2T)$ but $L \gg 1/T$ to ensure that we are close to the thermodynamic limit. In the opposite limit, that is $\tau > L$, our approach will be

ineffective. For correlations between planes which include the time direction and are separated instead along a space axis, exchange the label τ for the label of the relevant space direction in what follows.

First let us analyze the expected size of the signal. On dimensional grounds we expect that

$$\langle \mathcal{O}^\dagger(\vec{x}, \tau) \mathcal{O}(\vec{y}, 0) \rangle \sim ((\vec{x} - \vec{y})^2 + \tau^2)^{-\Delta}, \quad (2)$$

$$\begin{aligned} G(\tau) &= \frac{1}{L^3} \int d^3\vec{x} d^3\vec{y} \langle \mathcal{O}^\dagger(\vec{x}, \tau) \mathcal{O}(\vec{y}, 0) \rangle \\ &\propto \frac{1}{L^3} \int d^3\frac{\vec{x} + \vec{y}}{2} \int d^3(\vec{x} - \vec{y}) \frac{1}{((\vec{x} - \vec{y})^2 + \tau^2)^\Delta} \\ &\sim \tau^{3-2\Delta}. \end{aligned} \quad (3)$$

The correlation function is extensive in the transverse area because of the integral over the average coordinate $(\vec{x} + \vec{y})/2$; the integral over the difference coordinate $\vec{x} - \vec{y}$ is dominated by $|\vec{x} - \vec{y}| \lesssim \tau$.

If a large mass gap plays a role in the correlator of interest, the decay will instead be exponential. In general, one expects polynomial decay at short distances and exponential decay at large distances.

Next we want to understand the noise. The signal-to-noise achieved from N_{sample} independent gauge field configurations will scale as $1/\sqrt{N_{\text{sample}}}$ times the signal-to-noise from a single configuration. We can estimate this noise by asking about the mean value of $G(\tau)$ and about the mean-squared value of $G(\tau)$. Then the variance of the measurement is determined as usual by

$$\begin{aligned} \sigma_G^2 &= \langle G(\tau) G^*(\tau) \rangle - |\langle G(\tau) \rangle|^2 \\ &= \frac{1}{L^6} \int d^3\vec{x}_1 d^3\vec{x}_2 d^3\vec{y}_1 d^3\vec{y}_2 \langle (\mathcal{O}^\dagger(\vec{x}_1, \tau) \mathcal{O}(\vec{x}_2, \tau) \mathcal{O}(\vec{y}_1, 0) \mathcal{O}^\dagger(\vec{y}_2, 0)) - \langle \mathcal{O}^\dagger(\vec{x}_1, \tau) \mathcal{O}(\vec{y}_1, 0) \rangle \langle \mathcal{O}(\vec{x}_2, \tau) \mathcal{O}^\dagger(\vec{y}_2, 0) \rangle \rangle \\ &\simeq \frac{1}{L^6} \int d^3\vec{x}_1 d^3\vec{x}_2 \langle \mathcal{O}^\dagger(\vec{x}_1, \tau) \mathcal{O}(\vec{x}_2, \tau) \rangle \int d^3\vec{y}_1 d^3\vec{y}_2 \langle \mathcal{O}(\vec{y}_1, 0) \mathcal{O}^\dagger(\vec{y}_2, 0) \rangle \\ &\sim \frac{1}{L^6} \int d^3\frac{\vec{x}_1 + \vec{x}_2}{2} \int d^3(\vec{x}_1 - \vec{x}_2) |\vec{x}_1 - \vec{x}_2|^{-2\Delta} \int d^3\frac{\vec{y}_1 + \vec{y}_2}{2} \int d^3(\vec{y}_1 - \vec{y}_2) |\vec{y}_1 - \vec{y}_2|^{-2\Delta}. \end{aligned} \quad (4)$$

The first term in the second line is the *full* correlator, including both the connected correlator and various disconnected contributions. The second term cancels one of these disconnected contributions, but the disconnected contribution shown in the third line involves large small-separation contributions when $\vec{x}_1 \approx \vec{x}_2$ and $\vec{y}_1 \approx \vec{y}_2$, and is therefore expected to dominate the correlation function.

The variance has two worrying features. First, the integrals over $(\vec{x}_1 - \vec{x}_2)$ and $(\vec{y}_1 - \vec{y}_2)$ are short-distance divergent, presumably cut off by the lattice spacing a . Second, *each* overall integration $\int d^3(\vec{x}_1 + \vec{x}_2)/2$, $\int d^3(\vec{y}_1 + \vec{y}_2)/2$ introduces an overall L^3 factor. Thus one estimates that

$$\sigma_G^2 \sim a^{6-4\Delta}. \quad (5)$$

The signal-to-noise from correlating a single pair of planes on a single lattice is therefore on the order of $G(\tau)/\sigma_G \sim (a/\tau)^{2\Delta-3} \ll 1$.

The gradient flow method [14,15] offers an approach to ameliorate the short-distance divergent behavior in these correlation functions. Rather than evaluating the correlation functions directly on the lattice configuration, one first applies a well-defined procedure to remove UV fluctuations in the fields down to a length scale $\sim \sqrt{8\tau_F}$, where τ_F is the gradient-flow depth. This reduces the divergent

short-distance behavior such that $\int d\vec{x} \langle \mathcal{O}^\dagger(\vec{x}, 0) \mathcal{O}(\vec{0}, 0) \rangle \sim (8\tau_F)^{(3-2\Delta)/2}$. Physical results require an extrapolation to small τ_F , which partially counteracts the gain in signal-to-noise. Nevertheless, it is still necessary to apply some additional kind of noise-reduction technique in order to get a signal in a reasonable amount of computing time.²

This leaves, however, the problem that the signal-to-noise does not improve as one makes L large. One might have hoped for such improvement, because boxes with larger transverse extent L should be generating more statistically independent samples. But we do not see such an improvement in our parametric estimates, nor in simulations. To see why, we look at the role of transverse integrations in the signal and in the noise. In Eq. (2) we see that only $|\vec{x} - \vec{y}| \lesssim \tau$, that is, small transverse separations, contribute to the signal. But Eq. (4) contains *independent* integrations over $d\vec{x}$ and $d\vec{y}$; all values of $\vec{x} - \vec{y}$ contribute equally to the noise. So points with small transverse difference are responsible for signal and noise, but points with large transverse difference contribute to the noise, but not to the signal.

In this work we propose a blocking technique which eliminates the noise contributions from large transverse separations, and restores the expected behavior that the signal-to-noise ratio improves as $(L/\tau)^{3/2}$. The technique is numerically cheap and can be used in conjunction with gradient flow. In addition, it is applied at the analysis level, not as part of the configuration generation, and it is perfectly compatible with unquenching.

Naturally we are far from the first people to confront this particular problem. The issue of rapid falloff in the signal but not the noise has been known for a long time [8], and has been discussed and confronted frequently in the recent literature [24–29]. In particular, Ref. [12] presented an approach which is in some ways similar to what we argue for here. We will discuss the relative advantages of the approaches after giving an exposition of what we propose.

III. BLOCKING METHOD

Let us specialize to Hermitian operators \mathcal{O} and consider the lattice form of the correlation function, where space integrals are replaced by discrete sums:

$$G(\tau_1 - \tau_2) = \frac{a^3}{V} \sum_{\vec{x} \in V} \mathcal{O}(\tau_1, \vec{x}) \sum_{\vec{y} \in V} \mathcal{O}(\tau_2, \vec{y}), \quad (6)$$

²An alternative noise-reduction technique is the use of the multilevel algorithm [16]. The multilevel method has been successfully applied to correlation functions relevant for transport [17–20]. However, this technique is only applicable to pure-gauge theories; it cannot directly be generalized to the unquenched case. Furthermore, it is implemented using Monte Carlo updates of the gauge fields, rather than during the calculation of correlation functions on generated configurations. Ideas on the implementation of multilevel algorithms including dynamical fermions can be found in Refs. [21,22]. Reference [23] presented an approach for combining multilevel and gradient flow techniques.

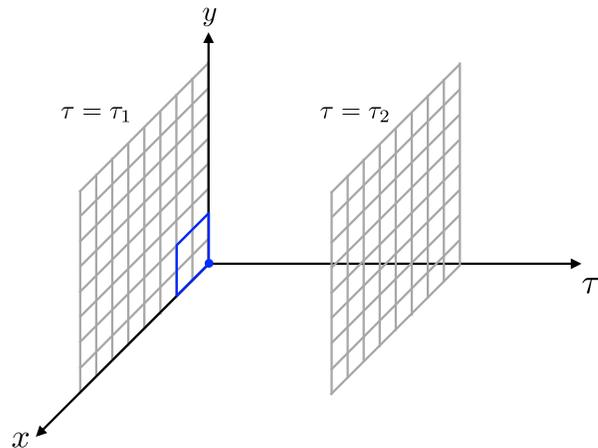


FIG. 1. Illustration of a 3D lattice on which a temporal correlator is measured. Operators at each site of plane $\tau = \tau_1$ are summed and the same is done for $\tau = \tau_2$. Two summed operators will be correlated via Eq. (6).

where $V = N_x N_y N_z$ is the spatial volume of the lattice. To calculate it one first evaluates the operator on each site on the plane at temporal position τ_1 and also those at τ_2 . The two planes are shown as grey squares at $\tau = \tau_1$ and $\tau = \tau_2$ in a simplified 3D sketch in Fig. 1. Then one calculates all site-to-site correlators of two operators: one operator runs over all sites on plane $\tau = \tau_1$ while the other is fixed to one site on plane $\tau = \tau_2$ and the process is repeated for each site on the second plane. The economical and also the most common way to do this is to first compute the sum of the operators on the plane at $\tau = \tau_1$ and then repeat this procedure for all planes. The data on each plane can be reduced to a single number, thus saving memory.

The method mentioned above includes contributions from all possible spatial distances ($s = \sqrt{(\vec{x} - \vec{y})^2}$). But as we have seen, the signal is dominated by small $s \lesssim \tau_2 - \tau_1$. Therefore, we want to obtain differential information, namely, how the correlation function varies as a function of s . A complete differential measurement involves correlating each \vec{x} position with each \vec{y} position. The numerical cost of this scales as V^2 , which is prohibitive, so we must seek a numerically less costly alternative, which we now present.

In the blocking method, each plane, for instance the $\tau = \tau_1$ plane in Fig. 1, is split into equal-sized bins. In each bin, we measure the operator on all sites belonging to this bin and save the sum at one corner. For instance, the blue square in Fig. 1 encloses one bin of size 2×2 and the sum is denoted as a dot at the origin. The relative position of this dot inside of the bin should be the same for all bins. Covering the whole lattice with bins leads to Fig. 2, in which the original lattice has effectively been compressed into a smaller lattice denoted by the blue dots. On this smaller lattice a calculation of the point-to-point correlators at all possible distances is feasible. One can see this

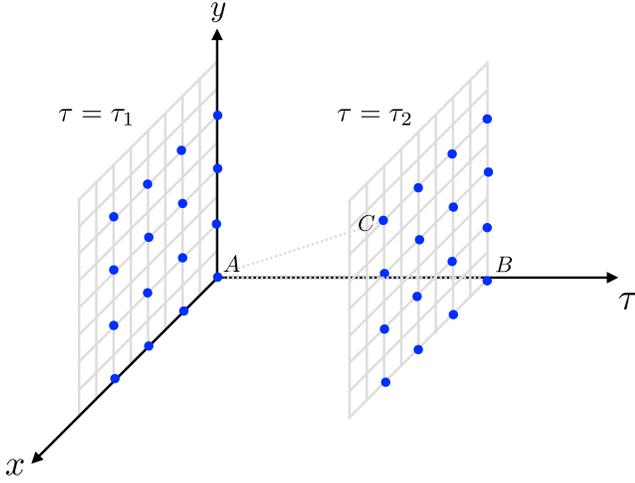


FIG. 2. Illustration of how the blocking method works: each plane is divided into bins. The operators in each bin are summed and saved to the site denoted by the blue dot on the corner. Then, for each pair of planes, one computes the correlators of all pairs of blue dots, one on each plane.

procedure is nothing but rewriting Eq. (6) as

$$G(\tau) = \frac{a^3}{V} \sum_{\{v_1\}} \left[\sum_{\vec{m} \in v_1} \mathcal{O}(\tau_1, \vec{m}) \right] \sum_{\{v_2\}} \left[\sum_{\vec{n} \in v_2} \mathcal{O}(\tau_2, \vec{n}) \right] \quad (7)$$

where $\tau = \tau_1 - \tau_2$, v_1, v_2 are the individual bins in each plane, and $\vec{m} \in v_1$ are all points in the bin v_1 . We advocate the use of cubic bins, that is, each bin contains $V_b = n_b \times n_b \times n_b$ lattice sites. Defining

$$\begin{aligned} \mathcal{O}_{v_1}(\tau_1) &= \sum_{\vec{m} \in v_1} \mathcal{O}(\tau_1, \vec{m}), \\ G_{v_1 v_2}(\tau) &= \langle \mathcal{O}_{v_1}(\tau_1) \mathcal{O}_{v_2}(\tau_2) \rangle \end{aligned} \quad (8)$$

we can reexpress Eq. (7) as

$$G(\tau) = \frac{a^3}{V} \sum_{v_1, v_2} G_{v_1 v_2}(\tau). \quad (9)$$

It is useful to rearrange this sum in terms of the transverse separation between the corner points of the bins $s = |\vec{v}_1 - \vec{v}_2|$. For $s = 0$ there are V/V_b contributions (equal to the number of bins). For any other s value there is an additional degeneracy factor d_s . For instance, for $s = n_b$ we can have $\vec{v}_1 - \vec{v}_2 \in \{[\pm n_b, 0, 0], [0, \pm n_b, 0], [0, 0, \pm n_b]\}$ for a total of $d_s = 6$ degenerate choices.³ Introducing the sum of all correlations between bins with separation s , normalized to contain V/V_b contributions,

³Normalizing s to $s^2/n_b^2 = 0, 1, 2, 3, \dots$, the d_s with $s^2 \leq (L/2n_b)^2$ are the OEIS integer sequence A005875; see <https://oeis.org/A005875>.

$$G(\tau, s) = a^3 d_s^{-1} \sum_{v_1, v_2} G_{v_1 v_2}(\tau) \delta(|\vec{v}_1 - \vec{v}_2| - s), \quad (10)$$

we can write the total correlation function as

$$G(\tau) = \frac{1}{V_b} \sum_{s=0}^{s_{\max}} d_s G(\tau, s), \quad (11)$$

where $s_{\max}^2 = 3(L/2)^2$.

This representation will be particularly practical in the following. Specifically, $G(\tau, s)$ should be a smooth function of s , which allows us to fit its behavior in a range of s where the signal-to-noise is good, and to use this fit to estimate its behavior at large s , where the signal-to-noise is bad.

A similar decomposition can be achieved using Fourier techniques, as described in Ref. [12]. The authors show how fast Fourier transform techniques can determine the transverse-separation-by-separation correlator with of order $V \log V$ operations. (In comparison, our approach requires of order V^2/n_b^6 operations; in practice, either scaling renders the required compute time small compared to the time required to update and gradient-flow gauge configurations, except perhaps on extremely large lattices.) To get something analogous to Eq. (10), one could histogram the resulting separation-by-separation correlator into transverse-separation ranges, which allows a transverse-separation-differential analysis of the data which is analogous to what we achieve here. Therefore we consider the Fourier method of Ref. [12], together with some histogramming, to be an alternative to our blocking technique, with very similar advantages. The main difference between our approach and the approach of Ref. [12] will be in how we use this differential information. The next section will show how we extract the integrated correlator, applying it to two specific physical problems.

IV. APPLICATIONS OF THE BLOCKING METHOD

Let us illustrate how to take advantage of blocked-correlator information in a real calculation which can be used to study transport phenomena.

Bulk viscosity can be determined from the small-frequency behavior of the spectral function for the squared field strength operator

$$E(x) = \frac{1}{4} F_{\rho\sigma}^a(x) F_{\rho\sigma}^a(x). \quad (12)$$

To compute bulk viscosity on the lattice, we first need the zero-momentum correlation function

$$G(\tau) \equiv \int d^3\vec{x} \langle \delta E(0, \vec{0}) \delta E(\tau, \vec{x}) \rangle \quad (13)$$

as a function of τ . Here $\delta E(x) \equiv E(x) - \langle E(x) \rangle$ is the field strength with its expectation value subtracted off to remove

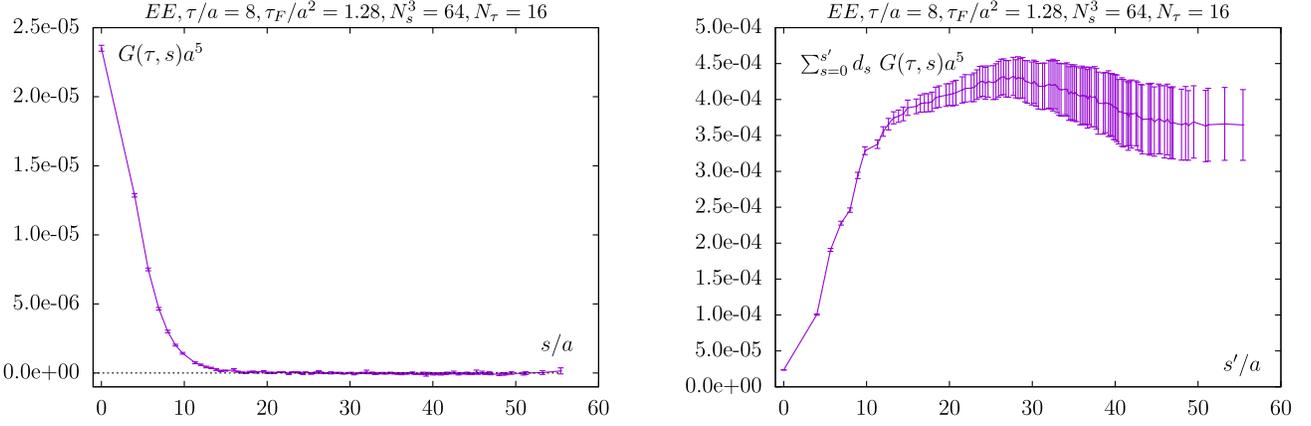


FIG. 3. Left: (bare) energy-momentum tensor correlator as a function of spatial distance s on a 16×64^3 lattice at temperature $T = 1.5T_c$ in the bulk channel at temporal separation $\tau/a = 8$ and flow depth $\tau_F/a^2 = 1.28$. The correlator is measured using the blocking method on 10 000 quenched configurations. The bin size is $4 \times 4 \times 4$. Right: sum of the correlator (sum of all terms $s \leq s'$) as a function of the maximal spatial distance s' . The last data point (at the largest s') is equal to the correlator calculated in the conventional way (without blocking) using Eq. (6).

the disconnected contributions. In our implementation, we construct E using the clover definition of the field strength tensor. In this work we will focus on determining this correlator with good signal-to-noise. The issues of correctly normalizing the operator, continuum and zero-flow extrapolating, and extracting the bulk viscosity from $G(\tau)$ are left for a separate study. We measure this correlator using the blocking method on a $64^3 \times 16$ quenched lattice over 10 000 configurations under gradient flow, determining errors using the bootstrap method. The bin size is $4 \times 4 \times 4$. The lattice setup and gradient flow setup are the same as used in Ref. [13]. We will also revisit the correlation function of topological charge density q which we first addressed in Ref. [13], using the same q definition introduced there and the same lattice setup as just described.

We illustrate the EMT correlators calculated using the blocking method in Fig. 3. The left panel shows the correlator $G(s, \tau)$ as a function of s at a fixed flow time $\tau_F/a^2 = 1.28$ and temporal separation $\tau/a = 8$, where a is the lattice spacing. We can see that $G(s, \tau)$ is a relatively smooth function of s , and that it falls off fast such that only the first few data points ($s/a \lesssim 15$) contribute significantly to $G(\tau)$. At distances $s/a > 17$ the correlators cannot be statistically distinguished from zero. If one sums the correlator over all $s \leq s'$, including the degeneracy factor, one obtains the data shown in the right panel of Fig. 3. The data point at the largest s' recovers the result and errors calculated in the usual (nonblocking) way. It can be seen that at $s'/a \sim 15$ the integrated correlator reaches a plateau but the error size becomes larger and larger as s' increases. From this it is clear that the bin-to-bin correlators with small s are contributing most of the signal while the large- s ones mainly introduce noise. The key idea of the blocking method is to use only the reliable lattice data coming from small s and to estimate the contribution from the long tail

by fitting the data that has good signal-to-noise using some theoretically inspired *Ansatz*.

For each (τ, τ_F) pair, we break the s range into three regions based on two cut points, s_0 and s_{cut} . The first region, $s < s_0$, is characterized by very high signal-to-noise ratios in the data $G(\tau, s)$. The point s_0 is chosen as the largest s value where the signal-to-noise is better than 10. This point is very easy to find from the data, and its value is stable across different bootstrap samples. We will justify this choice *a posteriori*, after describing the rest of the fitting procedure. The middle region $s_0 \leq s \leq s_{\text{cut}}$, is characterized by signal-to-noise between 10 and 2. s_{cut} is determined in a self-consistent way which we will explain soon. Finally, there is the region $s > s_{\text{cut}}$, where the signal-to-noise is very poor. Our procedure is to perform a fit of the data, and to replace a direct evaluation of $G(\tau)$ with an evaluation which takes into account the fit, as follows:

$$\begin{aligned}
 G(\tau) &= G_{\text{dom}}(\tau) + G_{\text{mid}}(\tau) + G_{\text{tail}}(\tau), \\
 G_{\text{dom}}(\tau) &\equiv \frac{a^3}{V_b} \sum_{s=0}^{s_0-1} d_s G(\tau, s), \\
 G_{\text{mid}}(\tau) &\equiv \frac{a^3}{V_b} \sum_{s=s_0}^{s_{\text{cut}}} d_s (x G_{\text{fit}}(\tau, s) + (1-x) G(\tau, s)), \\
 G_{\text{tail}}(\tau) &\equiv \frac{a^3}{V_b} \sum_{s>s_{\text{cut}}} d_s G_{\text{fit}}(\tau, s). \tag{14}
 \end{aligned}$$

Here $x = (s - s_0)/(s_{\text{cut}} - s_0)$ is the fraction of the way from s_0 to s_{cut} ; that is, in the first region we purely use the data, in the middle region we vary linearly from purely using the data at $s = s_0$ to purely using the fit at $s = s_{\text{cut}}$, and in the final region we purely use the fit.

To perform a fit of the data we need two things: an *Ansatz*, and a data range to use in the fit. We will return to the *Ansatz* momentarily. First we need to emphasize what s range the fit needs to be precise in. It is not important to find a fit function which describes the whole s range. As we see in Eq. (14), we only use the fit for $s > s_0$. We know on physical grounds that $G(\tau, s)$ falls rapidly at large s , and this should be reflected in our *Ansatz*. Therefore, the data range where $G_{\text{fit}}(\tau, s)$ is most important is the range around s_0 and s_{cut} . This fact needs to be reflected both in our choice of *Ansatz*, and in the data range we use to fit the *Ansatz*. In particular, using data with $s < s_0$ in our fitting procedure is actually not a good idea. The high signal-to-noise tends to control the fit, but it gives information about the functional form too far away from the region where we need the fit to work. Therefore, we choose instead to fit the data with $s \geq s_0$. We could cut off the s range used in the fitting procedure, for instance, at the not-yet-established value s_{cut} , but in practice the fit is always dominated by the first few points above s_0 , even if we use *all* data with $s \geq s_0$. This is in fact what we do: we fit a physically well-motivated *Ansatz* to all data with $s \geq s_0$. We have checked that the fit, χ^2 , and errors in the fit parameters are almost unaffected by introducing an upper cutoff on the s range used in the fit.

With a fit in hand, we can then estimate the signal-to-noise ratio as $\text{STN}(s) = G_{\text{fit}}(\tau, s)/\sigma(\tau, s)$, where the noise is determined from the fluctuations in the data and the signal is estimated from the fit. We choose s_{cut} to be the s value above which this estimated signal-to-noise ratio is always worse than 2. That is, we fully replace the data with the result of our fit starting where the signal-to-noise is consistently below 2. We estimate the errors in G_{dom} , G_{mid} and G_{tail} , whose sum gives the total correlator, using the bootstrap method.

Next we discuss the *Ansatz* (*Ansätze*) used in the fits. The choice of *Ansatz* is clearly dependent on the specifics of the theory under consideration; here we specialize to correlation functions of the action density E and of the topological charge density q .

The first *Ansatz* is a simple power law in s , $a^5 G(s, \tau) = A(s/s_{\text{pivot}})^{-B}$, where A and B are fit parameters and s_{pivot} is the third s value larger than s_0 .⁴ The motivation is that a rapidly falling function is generically fit

over a narrow range by a power law. Also, Eq. (2) implies that the true falloff could be a power law, with an unknown coefficient due to operator anomalous dimensions.⁵

The second *Ansatz* is based on the leading-order perturbative value for the correlator, accounting for time periodicity, gradient flow, and our blocking procedure. In vacuum, the leading-order correlator of two field strength tensors is

$$\begin{aligned} & \langle F_{\mu\nu}^a(r) F_{\alpha\beta}^b(0) \rangle \\ &= \frac{g^2 \delta_{ab}}{\pi^2 r^4} \left[\delta_{\mu\alpha} \delta_{\nu\beta} - \delta_{\mu\beta} \delta_{\nu\alpha} \right. \\ & \quad \left. - \frac{2}{r^2} (r_\mu r_\alpha \delta_{\nu\beta} - r_\mu r_\beta \delta_{\nu\alpha} - r_\nu r_\alpha \delta_{\mu\beta} + r_\nu r_\beta \delta_{\mu\alpha}) \right]. \end{aligned} \quad (15)$$

Applying gradient flow to a depth τ_F modifies this expression to [30]

$$\begin{aligned} & \langle G_{\mu\nu}^a(r) G_{\alpha\beta}^b(0) \rangle_{\tau_F} \\ &= \frac{g^2 \delta_{ab}}{\pi^2 r^4} \left[A(r, \tau_F) (\delta_{\mu\alpha} \delta_{\nu\beta} - \delta_{\mu\beta} \delta_{\nu\alpha}) \right. \\ & \quad \left. + \frac{B(r, \tau_F)}{r^2} (r_\mu r_\alpha \delta_{\nu\beta} - r_\mu r_\beta \delta_{\nu\alpha} - r_\nu r_\alpha \delta_{\mu\beta} + r_\nu r_\beta \delta_{\mu\alpha}) \right], \end{aligned} \quad (16)$$

$$A(r, \tau_F) = 1 - \left(1 + \frac{r^2}{8\tau_F} \right) e^{-r^2/8\tau_F}, \quad (17)$$

$$B(r, \tau_F) = -2 + \left[2 - 2 \frac{r^2}{8\tau_F} + \left(\frac{r^2}{8\tau_F} \right)^2 \right] e^{-r^2/8\tau_F}. \quad (18)$$

Note that this is a continuum, not lattice, expression; but when $\tau_F/a^2 > 0.5$ the lattice-continuum difference for flowed correlators is small, and the use of a continuum limit at fixed flow depth based only on data which satisfies this criterion should avoid the need to include lattice spacing corrections as well.

Using these expressions, the leading-order $\langle EE \rangle$ correlator at finite τ_F , τ , $|\vec{r}|$ and with periodic boundaries in the time direction is

$$\langle E(\vec{r}, \tau) E(0, 0) \rangle_{\tau_F} \propto \sum_{n_1, n_2 \in \mathbb{Z}} \frac{A(r_1)A(r_2)}{r_1^4 r_2^4} + \frac{A(r_1)B(r_2) + A(r_2)B(r_1)}{2r_1^4 r_2^4} + \frac{B(r_1)B(r_2)}{6r_1^6 r_2^6} (2(r_1 \cdot r_2)^2 + r_1^2 r_2^2), \quad (19)$$

⁴ s_{pivot} is introduced to suppress the correlation between A and B , which stabilizes the fit.

⁵ If the falloff is with distance, one should really fit to $(s^2 + \tau^2)^{-B/2}$. Since s_{cut} is almost always significantly larger than τ , this turns out to make very little difference.

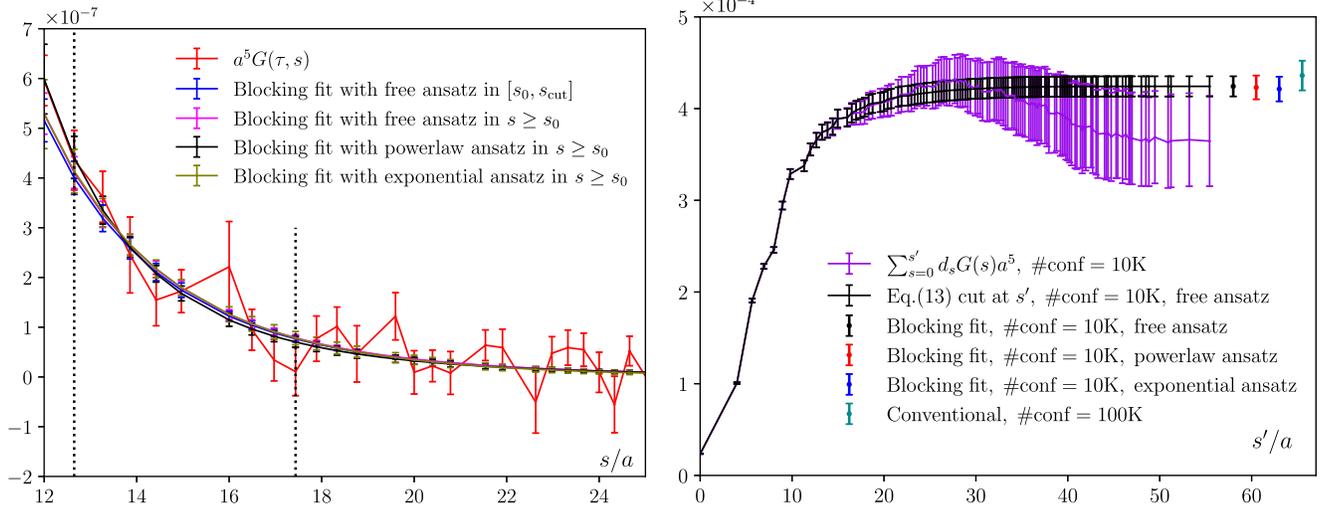


FIG. 4. Left: the same data as in the left panel of Fig. 3, but zoomed in around the value where the signal-to-noise becomes poor, together with three fits of the data, representing fitting functions and fit ranges as described in the text. Right: the same as the right plot in Fig. 3, but also showing the result summed over $s \leq s'$ using the mix of data and *Ansatz* defined in Eq. (14). Also, the fully summed result from Eq. (14) (using power law and free *Ansatz*) and from the conventional approach using 10 times as many configurations are shown (with a slight horizontal offset) as single data points on the far right.

where $r_1 = (\tau + n_1\beta, \vec{r})$ and $r_2 = (\tau + n_2\beta, \vec{r})$ are the 4-displacement with the temporal displacement shifted by independent integer multiples of the inverse temperature β . Similarly, when we compute the correlation function of two topological charge density operators $q = F_{\mu\nu}\tilde{F}_{\mu\nu}/32\pi^2$ (see below), the leading-order correlation function after flow is

$$\langle q(\vec{r}, \tau)q(0, 0) \rangle_{\tau_F} \propto - \sum_{n_1, n_2 \in \mathbb{Z}} \frac{A(r_1)A(r_2)}{r_1^4 r_2^4} + \frac{A(r_1)B(r_2) + A(r_2)B(r_1)}{2r_1^4 r_2^4} + \frac{B(r_1)B(r_2)}{6r_1^6 r_2^6} (r_1^2 r_2^2 - (r_1 \cdot r_2)^2). \quad (20)$$

Our second *Ansatz* is, for each s value, to integrate these expressions over the relative coordinates in the two bins whose centers are separated by s to determine what the time-periodic, flowed, bin-averaged correlation function would be at leading perturbative order. We then fit to a single overall normalization. This fit gives a poor description of the *whole* correlator $G(s, \tau)$, because it does not get the ratio of the peak to the tail accurately. However, it gives a reasonable description of the tail shape with a single fitting parameter.

Our third *Ansatz* will be based on the generic expected long-distance behavior for correlation functions in a nontrivial interacting theory. Consider a correlator as a function of space separation s along the z axis. If we think about the z axis as “Euclidean time,” then this corresponds to a zero-temperature system in a space with one compact periodic direction $\mathcal{R}^2 \times S^1$. This theory is expected to have a gapped spectrum, and the large-distance correlation behavior is controlled by the lightest state in the symmetry channel under investigation. One then expects that the correlation function decays

at asymptotically large distances as⁶ $a^5 G(s, \tau) = A(s/s_{\text{pivot}})^{-1} \exp(-B(s - s_{\text{pivot}}))$, where A, B are two fitting parameters. This *Ansatz* makes sense if one assumes that s is large enough to suppress the contributions of all higher excited states.

Some readers may be concerned about the *ad hoc* nature of our fitting functions. It is important to emphasize three things:

- (i) The function $G(s, \tau)$ falls very fast with s . Perturbatively, $G(s, \tau) \propto s^{-8}$. We see this rapid falloff explicitly in Fig. 3, both in the data and in the fits. Therefore, it is only important that our fit to $G(s, \tau)$ be reasonable in a fairly narrow range above s_0 .
- (ii) The fit is stable when we introduce an upper s cutoff on the fit range, and it is stable when we adjust (somewhat) the starting point s_0 for the fit.

⁶In $\mathcal{R}^3 \times \mathcal{R}$, $a^5 G \propto (s/s_{\text{pivot}})^{-3/2} \exp(-B(s - s_{\text{pivot}}))$ (see for instance Ref. [29]), but the presence of a small-radius S^1 makes the large-distance $s \gg 1/T$ decay behave like that in $\mathcal{R}^2 \times \mathcal{R}$.

TABLE I. s_0 and s_{cut} values for the correlator defined in Eq. (13) using our three fitting *Ansätze*, based on 10 000 independent configurations at the indicated values of lattice size, block size, and τ, τ_F values.

$N_\sigma^2 \times N_\tau$	n_b	τ/a	τ_F/a^2	<i>Ansatz</i>	s_0/a	s_{cut}/a
$64^3 \times 16$	4	4	1.28	Free	13.9	17.9
$64^3 \times 16$	4	8	1.28	Free	12.6	17.4
$64^3 \times 16$	4	8	0.605	Free	8.0	12.0
$64^3 \times 16$	4	4	1.28	Power law	13.9	18.3
$64^3 \times 16$	4	8	1.28	Power law	12.6	17.4
$64^3 \times 16$	4	8	0.606	Power law	8.0	12.0
$64^3 \times 16$	4	4	1.28	Exponential	13.9	18.8
$64^3 \times 16$	4	8	1.28	Exponential	12.6	17.9
$64^3 \times 16$	4	8	0.605	Exponential	8.0	12.6

(iii) The influence of the fit *Ansatz* is small, as determined by comparing these three rather different *Ansätze*.

To illustrate these points, consider Fig. 4. The left panel shows the same data as in Fig. 3, but zoomed in around the region where the signal-to-noise becomes poor. The vertical bars indicate s_0 and s_{cut} , the red points are the data, and the indicated lines are three fits to the data. We see that all three fit *Ansätze* give nearly the same results, and the free-*Ansatz* fit is almost unchanged when we fit to the data in the range $[s_0, s_{\text{cut}}]$ rather than all data with $s \geq s_0$. The fit indicates that the correlation function becomes very small for s a little higher than s_{cut} , consistent with the data but without the large error bars.

In the right panel of Fig. 4 we show a comparison of the correlators measured in the conventional way on 100 000 configurations (green data point), fitted correlators based on blocking data on 10 000 configurations with the free *Ansatz* (black points), and the same using the power-law *Ansatz* (red data point) and exponential-decaying *Ansatz* (blue point). These are compared to the same partial sums as in Fig. 3. We can see that if we fit the tail of our data to a proper *Ansatz*, we can reduce the error by a significant factor (in this case by a factor of ~ 4). Comparison with a much larger data set shows that this is achieved without corrupting the value; the blocking method gives a result which is consistent with that achieved by the conventional method using 10 times more data.

Now we return to discussing our choice of s_0, s_{cut} . Table I lists the values of s_0 and s_{cut} determined at some

typical separations τ and flow times τ_F . We see that s_0 is primarily determined by the gradient flow depth, and secondarily by τ . s_{cut} is additionally weakly dependent on our *Ansatz* choice. Since s_0, s_{cut} are determined from the data signal-to-noise, they will also change as we increase/decrease the computational resources and hence the amount of fit data. Specifically, more data means better signal-to-noise which means larger s_0, s_{cut} . Our approach always balances s_0 such that the fit is only used where the signal-to-noise has become problematic.

Our choice for s_0 was based on the somewhat arbitrary criterion of a signal-to-noise of 10. We have explored what happens when we either loosen or sharpen this criterion. If we choose a smaller value for the signal-to-noise, the fit ceases to be very constraining, and we fail to get a good description of the tail. But what if we start the fit where the signal-to-noise is much higher, say, 50? We explored this possibility, and our choice, over the full range of τ, τ_F values which we use in our data analysis. We consistently find that the $\chi^2/\text{d.o.f.}$ of our fits, using the data in the range $[s_0, s_{\text{cut}}]$, is close to 1 for the criterion we use. But when we set s_0 at the point where the S/N is 50, we find that the smallest- τ_F results—where the data is the noisiest—produce a poor (large) $\chi^2/\text{d.o.f.}$, apparently because the fit relies on data at a small s value where none of our fitting functions are good descriptions. This χ^2 analysis provides an *a posteriori* justification for our s_0 choice: it gives fits which are constraining but which are also internally consistent.

To understand our procedure better, Table II shows a decomposition of the total determined $G(\tau)$ based on the three regions defined in Eq. (14), together with the fitted parameters for each *Ansatz*. It can be seen that the first two *Ansätze* give almost the same contribution for each part. The exponential *Ansatz* has a somewhat larger s_{cut} , such that the middle and tail regions contain a different number of points. But the sum of these two regions is approximately the same in all cases. The difference between the results using different *Ansätze* is about 10 times smaller than the overall statistical uncertainty.

An alternative approach, advocated in Ref. [12], is to use physical arguments to determine the s value where almost all of the signal has been included, and to discard the data at higher s values. In Fig. 4, this would correspond to using the purple data point at an s' value somewhere above $s'/a = 20$. We see that this approach would be consistent with ours, but with larger errors.

TABLE II. A comparison of the three fitting *Ansätze*, showing the fit coefficients and the decomposition of the correlator shown in Eq. (14) for each *Ansatz*. The other values are the same as in each middle line of Table I.

<i>Ansatz</i>	$G_{\text{dom}} \times 10^6$	$G_{\text{mid}} \times 10^6$	$G_{\text{tail}} \times 10^6$	A	B
Free	355.5 (6.6)	44.6 (11.8)	24.3 (1.9)	5.5×10^{-5} (4.3×10^{-6})	...
Power law	355.5 (6.6)	43.7 (11.5)	23.9 (6.7)	2.1×10^{-7} (1.7×10^{-8})	5.7 (0.56)
Exponential	355.5 (6.6)	46.2 (12.1)	19.3 (6.9)	2.2×10^{-7} (1.8×10^{-8})	0.27 (0.049)

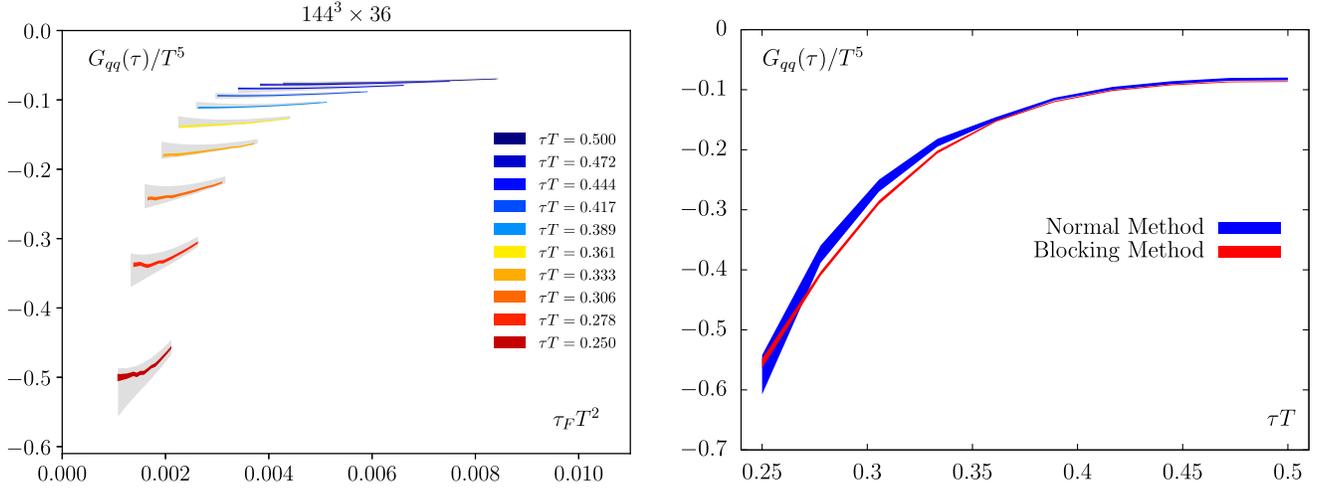


FIG. 5. Left: a comparison of the topological charge density correlators measured in the conventional way on 10 000 configurations taken from Ref. [13] (grey bands) and those from blocking fits (colorful bands) on the same configurations on the $144^3 \times 36$ lattice. Right: the same as in the left panel, but after continuum and flow-time-to-zero extrapolation.

As an application of our technique, we reanalyze the topological charge density correlators which we originally explored in Ref. [13]. The correlation function under study is

$$G_{qq}(\tau) = \int d^3\vec{x} \langle q(\vec{0}, 0) q(\vec{x}, \tau) \rangle, \quad (21)$$

where the topological charge density is defined as

$$q(x) = \frac{g^2}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{Tr}\{F_{\mu\nu}(x) F_{\rho\sigma}(x)\}. \quad (22)$$

Our implementation constructs this operator using an improved field strength tensor $F_{\mu\nu}(x)$; see Ref. [13] for details. We repeat the analysis of Ref. [13] carried out on five lattices— $64^3 \times 16$, $80^3 \times 20$, $96^3 \times 24$, $120^3 \times 30$ and $144^3 \times 36$ —but now applying the blocking method. The bin size is $4^3, 4^3, 4^3, 6^3, 8^3$ for each lattice respectively. The number of configurations is 10 000 for all lattices. Other details about the lattice setup and gradient flow setup can be found in Ref. [13].

In Fig. 5 we show a comparison of the correlators measured in the conventional way to those from the blocking method on the same configurations. Only correlators in the flow time range valid for the $\tau_F \rightarrow 0$ extrapolation are shown. In the left panel we take the finest available lattice as an example. In the right panel we compare the correlators after continuum extrapolation and flow time extrapolation. From the left panel we can see that the two ways of calculating the correlators give consistent results but with significantly reduced statistical uncertainty at the cost of introducing tiny systematic uncertainty when using the blocking method. In the right panel a discrepancy between two methods occurs in the range $\tau T \in [0.27, 0.35]$. This is mostly due to two discrepant points in the original

data analysis (not shown) at $\tau T = 0.3, 0.35$ on the $80^3 \times 20$ lattice, which pull the original continuum extrapolation in this region. The smaller-error results obtained with the blocking method are in better agreement with the other lattice spacings, suggesting that the problem lies in the results obtained without blocking. Besides these two points, the results using the new approach generally lie within the error bars of the previous determinations. Using this higher-precision data, we repeat the spectral analysis carried out in Ref. [13] with the updated correlators, and find that the spectral function *Ansätze* we considered there cannot describe our data well any more. All the fits have $\chi^2/\text{d.o.f.} > 10$. This indicates that more sophisticated and physically motivated *Ansätze* for the spectral function are needed. We leave this for future work.

V. CONCLUSION

In this paper we proposed a novel blocking method to improve the signal-to-noise ratio of Euclidean two-point correlators calculated on the lattice. Taking the bulk channel energy-momentum tensor correlators as an example, we demonstrated a factor of 3–7 improvement in the signal-to-noise ratio, with almost no additional cost in numerical effort. Equivalently, this is a factor of 10–50 reduction in computational cost to achieve a given precision goal. We then applied the blocking method to the topological charge density correlators that we studied in a previous publication, finding that the *Ansatz* for the spectral function which we previously considered no longer gave a good fit to the data. Our blocking method can be easily implemented on the lattice and used to study various correlators. It is applied at the analysis level, and does not have to be integrated into the configuration-generating procedure. There is also no obstacle to using it on

unquenched lattices for various bosonic correlators with any physically justified model.

Let us briefly address our choice of bin size. We chose to use bins somewhat smaller than the largest τ difference to be considered, in order to get sufficiently differential information about $G(s)$. As the bin size is made smaller, the numerical cost to correlate all bins eventually becomes significant. For the bin sizes considered here, this was not yet a problem. Also, as we make the bin size smaller, we increase the *relative* error in each individual bin, which might affect our procedure for choosing s_0 . If the bins are chosen smaller than the gradient-flow radius, then data at neighboring s values also becomes correlated, and auto-correlations in $G(\tau, s)$ at nearby s values must be handled carefully. In the opposite direction, if the bins are too large then we get an insufficiently refined determination of the

s dependence of the correlation function. It might be useful to systematically investigate how bin-size choice affects our procedure, but we leave this for future investigation.

All data from our calculations, presented in the figures of this paper, can be found in Ref. [31].

ACKNOWLEDGMENTS

All authors acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the CRC-TR 211 “Strong-interaction matter under extreme conditions”—Project No. 315477589—TRR 211. The computations in this work were performed on the GPU cluster at Bielefeld University using SIMULATEQCD suite [32,33]. We thank the Bielefeld HPC.NRW team for their support.

-
- [1] H. B. Meyer, *Eur. Phys. J. A* **47**, 86 (2011).
 - [2] N. Astrakhantsev, V. Braguta, and A. Kotov, *J. High Energy Phys.* **04** (2017) 101.
 - [3] M. Kitazawa, T. Iritani, M. Asakawa, and T. Hatsuda, *Phys. Rev. D* **96**, 111502 (2017).
 - [4] Sz. Borsányi, Zoltan Fodor, Matteo Giordano, Sandor D. Katz, Attila Pasztor, Claudia Ratti, Andreas Schäfer, Kalman K. Szabo, and Balint C. Tóth, *Phys. Rev. D* **98**, 014512 (2018).
 - [5] Y. Taniguchi, A. Baba, A. Suzuki, S. Ejiri, K. Kanaya, M. Kitazawa, T. Shimojo, H. Suzuki, and T. Umeda, *Proc. Sci., LATTICE2018* (2019) 166 [arXiv:1901.01666].
 - [6] E. Itou and Y. Nagai, *J. High Energy Phys.* **07** (2020) 007.
 - [7] G. Aarts and A. Nikolaev, *Eur. Phys. J. A* **57**, 118 (2021).
 - [8] G. Parisi, *Phys. Rep.* **103**, 203 (1984).
 - [9] H. B. Meyer, *Eur. Phys. J. C* **77**, 616 (2017).
 - [10] M. Cè, A. Gérardin, K. Ottnad, and H. B. Meyer, *Proc. Sci., LATTICE2018* (2018) 137 [arXiv:1811.08669].
 - [11] L. Giusti, T. Harris, A. Nada, and S. Schaefer, *Eur. Phys. J. C* **79**, 586 (2019).
 - [12] K.-F. Liu, J. Liang, and Y.-B. Yang, *Phys. Rev. D* **97**, 034507 (2018).
 - [13] L. Altenkort, A. M. Eller, O. Kaczmarek, L. Mazur, G. D. Moore, and H.-T. Shu, *Phys. Rev. D* **103**, 114513 (2021).
 - [14] R. Narayanan and H. Neuberger, *J. High Energy Phys.* **03** (2006) 064.
 - [15] M. Lüscher and P. Weisz, *Commun. Math. Phys.* **97**, 59 (1985); **98**, 433(E) (1985).
 - [16] M. Lüscher and P. Weisz, *J. High Energy Phys.* **09** (2001) 010.
 - [17] H. B. Meyer, *Phys. Rev. D* **76**, 101701 (2007).
 - [18] H. B. Meyer, *Phys. Rev. Lett.* **100**, 162001 (2008).
 - [19] A. Francis, O. Kaczmarek, M. Laine, T. Neuhaus, and H. Ohno, *Phys. Rev. D* **92**, 116003 (2015).
 - [20] N. Brambilla, V. Leino, P. Petreczky, and A. Vairo, *Phys. Rev. D* **102**, 074503 (2020).
 - [21] M. Cè, L. Giusti, and S. Schaefer, *Phys. Rev. D* **93**, 094507 (2016).
 - [22] L. Giusti, M. Cè, and S. Schaefer, *EPJ Web Conf.* **175**, 01003 (2018).
 - [23] M. García Vera and S. Schaefer, *Phys. Rev. D* **93**, 074502 (2016).
 - [24] G. Bali, S. Collins, and A. Schafer (QCDSF Collaboration), *Proc. Sci., LAT2009* (2009) 149 [arXiv:0911.2407].
 - [25] W. Sun, A. Alexandru, Y. Chen, T. Draper, Z. Liu, and Y.-B. Yang (χ QCD Collaboration), *Chin. Phys. C* **42**, 063102 (2018).
 - [26] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, and C. Lehner, *Phys. Rev. D* **93**, 014503 (2016).
 - [27] M. Cè, L. Giusti, and S. Schaefer, *Phys. Rev. D* **95**, 034503 (2017).
 - [28] E. Shintani, T. Blum, T. Izubuchi, and A. Soni, *Phys. Rev. D* **93**, 094503 (2016).
 - [29] M. Lüscher, *EPJ Web Conf.* **175**, 01002 (2018).
 - [30] A. M. Eller and G. D. Moore, *Phys. Rev. D* **97**, 114507 (2018).
 - [31] L. Altenkort, A. M. Eller, O. Kaczmarek, L. Mazur, G. D. Moore, and H.-T. Shu, *University of Bielefeld* (2022), 10.4119/unibi/2962501.
 - [32] L. Altenkort, D. Bollweg, D. A. Clarke, O. Kaczmarek, L. Mazur, C. Schmidt, P. Scior, and H.-T. Shu, arXiv: 2111.10354.
 - [33] L. Mazur, Ph.D. thesis, Bielefeld University, 2021, 10.4119/unibi/2956493.