



Large-time correlation functions in bosonic lattice field theories

Cagin Yunus^{a,*}, William Detmold^{a,b}

^a Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^b The NSF Institute for Artificial Intelligence and Fundamental Interactions, USA

ARTICLE INFO

Article history:

Received 8 March 2023

Accepted 30 March 2023

Available online 4 April 2023

Editor: A. Schwenk

ABSTRACT

Large-time correlation functions have a pivotal role in extracting particle masses from Euclidean lattice field theory calculations, however little is known about the statistical properties of these quantities. In this work, the asymptotic form of the distributions of the correlation functions at vanishing momentum is determined for bosonic interacting lattice field theories with a unique gapped vacuum. It is demonstrated that the deviations from the asymptotic form at large Euclidean times can be utilized to determine the spectrum of the theory.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Funded by SCOAP³.

1. Introduction

Quantum field theory (QFT) has an essential role in nuclear and particle physics and in condensed matter physics. From the Standard Model of particle physics to topological phase structures in quantum materials, the language of QFT provides a concise and predictive mathematical description. In some cases these descriptions contain a small parameter (coupling) in which an expansion can be performed to derive analytical expressions for relevant physical quantities. However in other cases, the couplings are large and numerical approaches are required to extract physical results; the dominant such approach in many strongly-coupled theories is referred to as lattice QFT (LQFT). The LQFT method involves stochastic Monte-Carlo sampling of the very high-dimensional lattice-regulated path integrals that define the correlation functions of the theory in which physical information is encoded. The large Euclidean-time behavior of correlation functions plays a crucial role in LQFT as it is central to obtaining the energy spectrum of the theory under study. If $\mathcal{O}(x)$ is a localized operator built from combinations of the fundamental fields centered around the spacetime point $x = (t, \vec{x})$ and $\overline{\mathcal{O}}(t) \propto \sum_{\vec{x}} \mathcal{O}(x)$ then, for Euclidean theories satisfying reflection positivity, the bilocal operator $C(t) \equiv \overline{\mathcal{O}}(t) \mathcal{O}^\dagger(0)$ has (vacuum, $|\Omega\rangle$) expectation value

$$\langle C(t) \rangle = \langle \Omega | C(t) | \Omega \rangle \xrightarrow{t \rightarrow \infty} Z_C e^{-mt}, \quad (1)$$

where m is the mass of the lowest energy eigenstate that $\overline{\mathcal{O}}^\dagger(0) |\Omega\rangle$ has an overlap with and Z_C is the overlap factor. Therefore, for large enough¹ t , an estimator for m is given by

$$\hat{m} = \ln \langle C(t) \rangle - \ln \langle C(t+1) \rangle. \quad (2)$$

In practice, to extract m , one calculates the sample mean and sample variance of $\hat{m}$ and assumes the validity of the Central Limit Theorem to construct confidence intervals. However, for most choices of correlation function, as the Parisi-Lepage [1,2] argument shows, the signal-to-noise ratio vanishes exponentially fast as $t \rightarrow \infty$.² This issue has been investigated in detail in lattice calculations in the context of lattice Quantum Chromodynamics (LQCD), given its phenomenological interest as the theory of the strong interaction. Recent works have focused on characterizing the nature of statistical fluctuations [3–10] and on proposing strategies for noise reduction [11–25].

In confronting noise issues in the large-time behavior of correlation functions, it would be useful to have an analytical form for the probability distribution function (PDF) of $C(t)$ even in simple examples. Such a form would allow for statistical tests of empirical

* Corresponding author.

E-mail addresses: cyunus@mit.edu (C. Yunus), wdetmold@mit.edu (W. Detmold).

¹ “Large” implies that t^{-1} is much smaller than the difference between the masses of the ground state and the first excited state with zero spatial momentum.

² A notable exception is the correlation function built from local pseudoscalar interpolating operators.

distributions determined in numerical calculations that may diagnose statistical limitations in the empirical sampling. Examination of higher moments of the PDF may also be useful in improving extractions of physical properties of the system through improved estimators [4,5]. In the present work, it will be demonstrated that in certain cases it is possible to obtain such analytical forms in the large time regime for correlation functions whose source and sink are both constructed to have vanishing spatial momentum. In Sec. 2, the analytical form of correlation function distributions for the free real scalar field theory will be derived. In Sec. 3, it will be shown that the distribution obtained in Sec. 2 is valid for correlation functions constructed from local operators for interacting bosonic theories at large times assuming a unique, gapped vacuum. Deviations from the asymptotic form of the distribution are linked to excited states in the spectrum and are controlled by the correlation lengths (or masses) of the theory. Consequently, analysis of these deviations provides a means of determining correlation lengths. Through their more general approach to statistical distributions, these results can potentially provide a path towards more robust determinations of energies and matrix elements in LQFT.

2. Exact results for free scalar field theory

In this section, the distributions of correlation functions whose source and sink are both constructed to have vanishing spatial momentum are studied in the setting of a free real scalar field theory. The general form of the distribution is derived before a specific lattice discretization is chosen and investigated.

Since the free theory momentum modes decouple, all non-zero spatial momentum modes can be factored out trivially. That is, if $\phi(t)$ is defined as

$$\phi(t) = \frac{1}{\sqrt{V}} \sum_{\vec{x}} \phi(t, \vec{x}), \quad (3)$$

$$Z = \oint_{\phi(0)=\phi(N_t)} [\mathcal{D}\phi] \exp \left(-\frac{1}{2} \sum_{t', t''} \phi(t') D(t', t'') \phi(t'') \right), \quad (4)$$

where N_t is the lattice size in the temporal direction, V is the spatial volume in lattice units, $D(t', t'')$ is a discretization of the continuum Klein-Gordon operator, and $[\mathcal{D}\phi] = \prod_{t=0}^{N_t-1} d\phi(t)$ is the integration measure.

For the following argument, it is assumed that the discretized $D(t', t'')$ is

- positive-definite, in order that the integral in Eq. (4) is convergent,
- real and symmetric,
- translationally invariant, only depending on $t' - t'' \pmod{N_t}$.

The characteristic function for the composite operator $\phi(t)\phi(0)$ is defined as

$$\begin{aligned} \Phi_{\phi(t)\phi(0)}(\omega) &= \langle \exp(-i\omega\phi(t)\phi(0)) \rangle \\ &= \frac{1}{Z} \oint_{\phi(0)=\phi(N_t)} [\mathcal{D}\phi] \exp \left(-\frac{1}{2} \sum_{t', t''} \phi(t') D(t', t'') \phi(t'') - i\omega\phi(t)\phi(0) \right). \end{aligned} \quad (5)$$

If $Q_{t', t''} = \delta_{t', 0} \delta_{t'', t} + \delta_{t', t} \delta_{t'', 0}$ is further defined, then it is easy to show that

$$\Phi_{\phi(t)\phi(0)}(\omega) = \frac{1}{\sqrt{\det R(\omega)}}, \quad (6)$$

where

$$R(\omega) = 1 + i\omega D^{-\frac{1}{2}} Q D^{-\frac{1}{2}}. \quad (7)$$

Note that D and Q are linear operators acting on a vector space of dimension N_t and let $e_{t'}$ for $t' = 0, \dots, N_t - 1$ be the basis of unit vectors on each timeslice and for which matrix elements are given as $D_{t', t''}$ and $Q_{t', t''}$ respectively. Note that in this basis $Q e_0 = e_t$ and $Q e_t = e_0$ and $Q e_{t'} = 0$ for $t' \notin \{0, t\}$. Let $(\cdot, \cdot)$ be the inner product for which the vectors $e_{t'}$ are orthonormal. For the basis $\{v_{t'} = D^{\frac{1}{2}} e_{t'}\}$, $R(\omega)v_{t'} = v_{t'}$ for $t' \notin \{0, t\}$. Therefore, to calculate $\det R(\omega)$, only the subspace spanned by v_0 and v_t needs to be considered. Further, the vectors $v_{\pm} = D^{\frac{1}{2}} e_{\pm}$ can be defined where $e_{\pm} = \frac{1}{\sqrt{2}}(e_t \pm e_0)$. If $R(\omega)v_+ = R_{++}v_+ + R_{+-}v_- + \dots$ and $R(\omega)v_- = R_{--}v_- + R_{-+}v_+ + \dots$, where the ellipsis represents terms that are spanned by v_j for $j \neq 0, t$ and does not contribute to the determinant, then:

$$\det R(\omega) = R_{++}R_{--} - R_{+-}R_{-+}. \quad (8)$$

Further, for $w_{t'} = D^{-\frac{1}{2}} e_{t'}$ and $w_{\pm} = D^{-\frac{1}{2}} e_{\pm}$. Then, $(w_{\pm}, v_{\pm}) = \delta_{\pm, \pm}$. It follows that

$$\begin{aligned} R_{\sigma\sigma'} &= (w_{\sigma}, R(\omega)v_{\sigma'}) \\ &= \delta_{\sigma, \sigma'} + i\omega\sigma'(e_{\sigma}, D^{-1}e_{\sigma'}), \end{aligned} \quad (9)$$

where $\sigma, \sigma' = \pm$. From translational invariance, it also follows that $(e_{\pm}, D^{-1}e_{\mp}) = (e_0, D^{-1}e_0) - (e_t, D^{-1}e_t) \pm (e_t, D^{-1}e_0) - (e_0, D^{-1}e_t) = 0$ and so $R_{+-} = R_{-+} = 0$. Then, by defining³

³ The positivity of $\omega_{\pm}$ follows from the positive-definiteness of D .

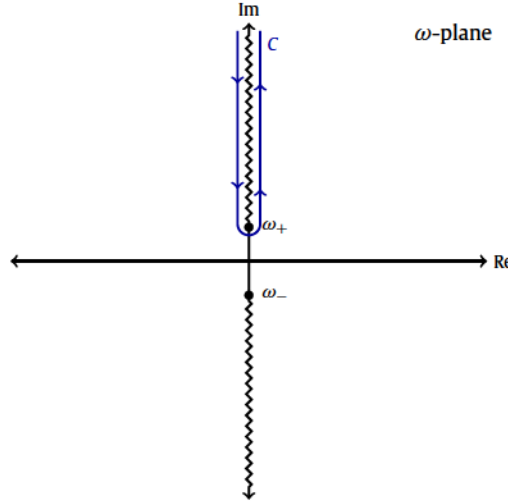


Fig. 1. The integration contour used to evaluate $I(x)$ for $x > 0$.

$$\omega_{\pm} = \frac{1}{(e_{\pm}, D^{-1}e_{\pm})} > 0, \quad (10)$$

it follows that

$$\det(R(\omega)) = \frac{(\omega - i\omega_+)(\omega + i\omega_-)}{\omega_+\omega_-}. \quad (11)$$

Therefore, the probability distribution of $\phi(t)\phi(0)$ is obtained from the characteristic function as:

$$\begin{aligned} P_{\phi(t)\phi(0)}(x) &\equiv \frac{1}{2\pi} \int d\omega e^{i\omega x} \Phi_{\phi(t)\phi(0)}(\omega) \\ &= \frac{\sqrt{\omega_+\omega_-}}{2\pi} \int d\omega \frac{e^{i\omega x}}{\sqrt{(\omega - i\omega_+)(\omega + i\omega_-)}}. \end{aligned} \quad (12)$$

The calculation of $P_{\phi(t)\phi(0)}$ then reduces to the evaluation of

$$I(x) = \int_{-\infty}^{\infty} d\omega \frac{e^{i\omega x}}{\sqrt{(\omega - i\omega_+)(\omega + i\omega_-)}}. \quad (13)$$

Since $\omega_{\pm} > 0$ (Eq. (10)), for $x > 0$ the integration contour in $I(x)$ can be deformed to the contour $C \equiv \{i\infty - \epsilon \rightarrow i\omega_+ - \epsilon\} \cup \{i\omega_+ + \epsilon \rightarrow i\infty + \epsilon\}$ as in Fig. 1. The integral can thus be expressed as:

$$I(x) = \left(-e^{i\frac{3\pi}{4}} + e^{-i\frac{\pi}{4}}\right) i e^{-i\frac{\pi}{4}} I_0(x) = 2\bar{I}(x), \quad (14)$$

where

$$\begin{aligned} \bar{I}(x) &= \int_{\omega_+}^{\infty} dy \frac{e^{-xy}}{\sqrt{(y - \omega_+)(y + \omega_-)}} \\ &= e^{\frac{1}{2}(\omega_- - \omega_+)x} K_0\left(\frac{1}{2}(\omega_+ + \omega_-)x\right). \end{aligned} \quad (15)$$

Here, $K_0(x)$ is a modified Bessel function of the second kind.

These formulas can be summarized as⁴:

$$P_{\phi(t)\phi(0)}(x) = \frac{\sqrt{\omega_+\omega_-}}{\pi} e^{\frac{1}{2}(\omega_- - \omega_+)x} K_0\left(\frac{1}{2}(\omega_+ + \omega_-)|x|\right). \quad (16)$$

The lattice action needs to be specified explicitly to gain more insight into $\omega_{\pm}$. A simple discretization is:

$$(D \cdot \phi)(t) = m^2 \phi(t) - (\phi(t+1) + \phi(t-1) - 2\phi(t)). \quad (17)$$

⁴ Note that the similarity of this equation to the form of the scalar field propagator in position space,

$$C(x) \sim \frac{m^{(d-1)/2}}{|x|^{(d-1)/2}} K_{(d-1)/2}(m|x|) \xrightarrow{d \rightarrow 1} K_0(m|x|),$$

is coincidental for $d = 1$ and does not hold in other numbers of spatial dimensions, d .

Assuming that N_t is odd, the normalized eigenvectors of D are given by:

$$\begin{aligned} E(t) &= \frac{1}{\sqrt{N_t}}, \\ S_j(t) &= \sqrt{\frac{2}{N_t}} \sin\left(\frac{2\pi j}{N_t} t\right), \\ C_j(t) &= \sqrt{\frac{2}{N_t}} \cos\left(\frac{2\pi j}{N_t} t\right), \end{aligned} \quad (18)$$

where $j = 1, \dots, \frac{N_t-1}{2}$ and t indexes the components of the each vector. The corresponding eigenvalues are:

$$\begin{aligned} \lambda(E) &= m^2, \\ \lambda(S_j) &= m^2 + 4 \sin^2\left(\frac{\pi j}{N_t}\right), \\ \lambda(C_j) &= m^2 + 4 \sin^2\left(\frac{\pi j}{N_t}\right). \end{aligned} \quad (19)$$

Expanding in eigenvectors of D , the vector representing an arbitrary field ϕ can be expressed as $\phi = \phi^e E + \sum_{j=1}^{\frac{N_t-1}{2}} \phi_j^s S_j + \sum_{j=1}^{\frac{N_t-1}{2}} \phi_j^c C_j$ and it follows that

$$\begin{aligned} \langle \phi(t) \phi(0) \rangle &= \frac{1}{N_t} \langle (\phi^e)^2 \rangle + \sqrt{\frac{2}{N_t}} \sum_{j=1}^{\frac{N_t-1}{2}} C_j(t) \langle (\phi_j^c)^2 \rangle \\ &= \frac{1}{N_t} \frac{1}{m^2} + \frac{2}{N_t} \sum_{j=1}^{\frac{N_t-1}{2}} \frac{\cos\left(\frac{2\pi j}{N_t} n\right)}{m^2 + 4 \sin^2\left(\frac{\pi j}{N_t}\right)}, \end{aligned} \quad (20)$$

where $\langle \dots \rangle$ indicates integration over the field ϕ where the integration measure is given by $[\mathcal{D}\phi] = \prod_{t=0}^{N_t-1} d\phi_t = d\phi^e \prod_{j=1}^{\frac{N_t-1}{2}} d\phi_j^s d\phi_j^c$.

Partially following Ref. [26], in order to take the continuum limit, the infinitesimal time interval ϵ is introduced and the limit $t, N_t \rightarrow \infty$ is considered such that $\beta = \epsilon N_t$ and t/N_t are fixed. Further, $\tau \equiv t\epsilon$ is defined. The quantity m depends on ϵ (equivalently on N_t) and $m(\epsilon)$ should be chosen such that the correlation function decays as $\exp(-m_R \tau)$ as $\epsilon \rightarrow 0$ for large τ and $\beta \rightarrow \infty$ (the $\beta \rightarrow \infty$ limit must be taken first), where m_R is the renormalized mass. The renormalized field $\phi_R = \sqrt{Z_\phi(\epsilon)} \phi$, where $Z_\phi(\epsilon)$ will be chosen such that correlation functions of ϕ_R have non-singular behavior as the continuum limit is taken. Setting $k = \frac{2\pi}{\beta} j$, the renormalized correlation function is:

$$\begin{aligned} \langle \phi_R(\tau) \phi_R(0) \rangle &\equiv Z_\phi(\epsilon) \langle \phi(t) \phi(0) \rangle \\ &= \frac{Z_\phi(\epsilon) \epsilon}{\pi} \sum_{k=\frac{2\pi}{\beta}}^{\frac{\pi}{\beta}(N_t-1)} \frac{\Delta k \cos(k\tau)}{m(\epsilon)^2 + 4 \sin^2\left(\frac{k}{2}\epsilon\right)} + \frac{Z_\phi(\epsilon)}{N_t m(\epsilon)^2} \\ &= \frac{Z_\phi(\epsilon) \epsilon}{\pi} \sum_{k=0}^{\frac{\pi}{\beta}(N_t-1)} \frac{\Delta k \cos(k\tau)}{m(\epsilon)^2 + 4 \sin^2\left(\frac{k}{2}\epsilon\right)} - \frac{Z_\phi(\epsilon) \epsilon}{\beta m(\epsilon)^2}. \end{aligned} \quad (21)$$

Taking the limit $\beta \rightarrow \infty$ (or equivalently $N_t \rightarrow \infty$) one therefore obtains:

$$\begin{aligned} \lim_{\beta \rightarrow \infty} \langle \phi_R(\tau) \phi_R(0) \rangle &= \frac{Z_\phi(\epsilon) \epsilon}{\pi} \int_0^{\frac{\pi}{\epsilon}} dk \frac{\cos(k\tau)}{m(\epsilon)^2 + 4 \sin^2\left(\frac{k}{2}\epsilon\right)} \\ &= \frac{Z_\phi(\epsilon) \epsilon}{2\pi} \int_{-\frac{\pi}{\epsilon}}^{\frac{\pi}{\epsilon}} dk \frac{\exp(ik\tau)}{m(\epsilon)^2 + 4 \sin^2\left(\frac{k}{2}\epsilon\right)}. \end{aligned} \quad (22)$$

The integration can be performed by defining $z = \exp(ik\epsilon)$ that maps onto the unit circle, giving⁵:

$$\langle \phi_R(\tau) \phi_R(0) \rangle = -\frac{Z_\phi(\epsilon)}{2\pi i} \oint dz \frac{z^{\frac{\tau}{\epsilon}}}{z^2 - z(m(\epsilon)^2 + 2) + 1}. \quad (23)$$

⁵ Note that $\frac{\tau}{\epsilon}$ is a non-negative integer, so the integrand is meromorphic.

The poles of the integrand are at $z_{\pm} = 1 + \frac{m(\epsilon)^2}{2} \pm m(\epsilon)\sqrt{1 + \frac{1}{4}m(\epsilon)^2}$, with only z_- occurring inside the unit circle. Therefore:

$$\begin{aligned} \langle \phi_R(\tau) \phi_R(0) \rangle &= \frac{Z_{\phi}(\epsilon)}{2m(\epsilon)\sqrt{1 + \frac{1}{4}m(\epsilon)^2}} \\ &\times \left(1 + \frac{m(\epsilon)^2}{2} - m(\epsilon)\sqrt{1 + \frac{1}{4}m(\epsilon)^2} \right)^{\frac{\tau}{\epsilon}}. \end{aligned} \quad (24)$$

To observe the expected exponential decay $\propto \exp(-m_R \tau)$, one is forced to set $m(\epsilon) = m_R \epsilon + \mathcal{O}(\epsilon^2)$. This also shows that one must chose $Z(\epsilon) = \text{constant} \times \epsilon + \mathcal{O}(\epsilon^2)$ for the above equation to be finite for $\epsilon \rightarrow 0$. Given these constraints, $m(\epsilon) = m_R \epsilon$ and $Z_{\phi}(\epsilon) = \epsilon$ are chosen, leading to a renormalized correlation function that is finite in the $\epsilon \rightarrow 0$ limit.

It is clear from Eqs. (18) and (19) that⁶

$$\begin{aligned} D^{-1}e_t &= \frac{1}{m_R^2 \epsilon^2 \sqrt{N_t}} E \\ &+ \sqrt{\frac{2}{N_t}} \sum_{j=1}^{N_t-1} \frac{\sin\left(\frac{2\pi j t}{N_t}\right) S_j + \cos\left(\frac{2\pi j t}{N_t}\right) C_j}{m_R^2 \epsilon^2 + 4 \sin\left(\frac{\pi j}{N_t}\right)^2}, \end{aligned} \quad (25)$$

so it is straightforward to show that:

$$\begin{aligned} (e_+, D^{-1}e_+) &= \frac{2}{N_t m_R^2 \epsilon^2} + \frac{2}{N_t} \sum_{j=1}^{N_t-1} \frac{1 + \cos\left(\frac{2\pi j t}{N_t}\right)}{m_R^2 \epsilon^2 + 4 \sin\left(\frac{\pi j}{N_t}\right)^2}, \\ (e_-, D^{-1}e_-) &= \frac{2}{N_t} \sum_{j=1}^{N_t-1} \frac{1 - \cos\left(\frac{2\pi j t}{N_t}\right)}{m_R^2 \epsilon^2 + 4 \sin\left(\frac{\pi j}{N_t}\right)^2}, \\ (e_+, D^{-1}e_-) &= (e_-, D^{-1}e_+) \\ &= 0. \end{aligned} \quad (26)$$

The probability density function of the renormalized two-point function is given as:

$$P_{\phi_R(\tau)\phi_R(0)}(x) = \langle \delta(x - \phi_R(\tau)\phi_R(0)) \rangle \quad (27)$$

and since $Z(\epsilon) = \epsilon$ is chosen, this may be written equivalently as

$$P_{\phi_R(\tau)\phi_R(0)}(x) = \langle \delta(x - \epsilon \phi(t) \phi(0)) \rangle. \quad (28)$$

The factor ϵ leads to the modification of Eqs. (7) and (10) as

$$R(\omega) = 1 + i\epsilon \omega D^{-\frac{1}{2}} Q D^{-\frac{1}{2}}, \quad (29)$$

and

$$\omega_{\pm} = \frac{1}{\epsilon (e_{\pm}, D^{-1}e_{\pm})}. \quad (30)$$

In the limit $\epsilon \rightarrow 0$:

$$\omega_+^{-1} \rightarrow \frac{2}{\beta m_R^2} + \frac{\cosh\left(\frac{\beta m_R}{2}\right) + \cosh\left(\frac{\beta m_R}{2} - m_R \tau\right)}{2m_R \sinh\left(\frac{\beta m_R}{2}\right)}, \quad (31)$$

and

$$\omega_-^{-1} \rightarrow \frac{\cosh\left(\frac{\beta m_R}{2}\right) - \cosh\left(\frac{\beta m_R}{2} - m_R \tau\right)}{2m_R \sinh\left(\frac{\beta m_R}{2}\right)}, \quad (32)$$

where the following identity⁷ has been used:

⁶ Note that this is a vector equation: both $D^{-1}e_t$ and E , S_j and C_j have N_t components.

⁷ The sum $f(\tau) = \sum_k \frac{e^{ik\tau}}{m_R^2 + k^2}$ is to be calculated where $k = \frac{2\pi}{\beta} j$ with $j \in \mathbb{Z}$. One observes that $f''(\tau) = -\sum_k e^{ik\tau} + m_R^2 f(\tau)$. Using the fact that $\sum_k e^{ik\tau} = \beta \delta(\tau)$ under the restriction $0 \leq \tau < \beta$, we obtain $f''(\tau) = -\beta \delta(\tau) + m_R^2 f(\tau)$. Therefore, one has the solution $f(\tau) = A_+ e^{-m_R \tau} + A_- e^{m_R \tau}$, except at $\tau = 0$. The symmetries of the summation in $f(\tau)$ imply that $f(\beta - \tau) = f^*(\tau) = f(\tau)$, and from this condition, one obtains $A_- = e^{-m_R \beta} A_+$. Finally, using the boundary condition $f'(\beta - \epsilon) - f'(\epsilon) = \beta$ as required to satisfy the inhomogeneous differential equation for $f(\tau)$, one can also fix $A_+ = \frac{\beta}{4m_R} \frac{e^{\frac{\beta m_R}{2}}}{\sinh\left(\frac{\beta m_R}{2}\right)}$, resulting in Eq. (33).

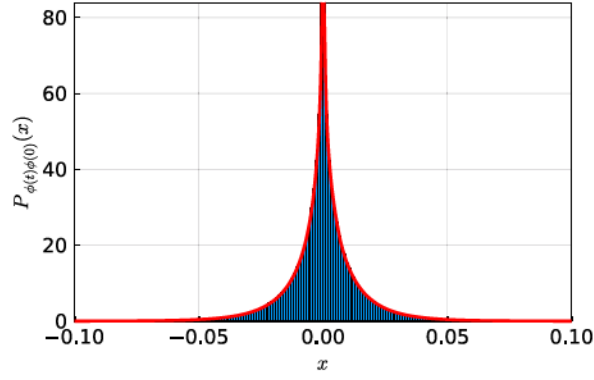


Fig. 2. The empirical distribution of $\phi(t)\phi(0)$ for free scalar field theory with the discretization given by Eq. (17). The red curve is given by Eq. (16) with $\omega_{\pm}$ from Eq. (37).

$$\sum_{k=\frac{2\pi j}{\beta}; j \in \mathbb{Z}} \frac{e^{ik\tau}}{m_R^2 + k^2} = \frac{\beta}{2m_R} \frac{\cosh\left(\frac{\beta m_R}{2} - m_R \tau\right)}{\sinh\left(\frac{\beta m_R}{2}\right)}. \quad (33)$$

In the limit $\beta \rightarrow \infty$,

$$\omega_{\pm}^{-1} \rightarrow \frac{1 \pm e^{-m_R \tau}}{2m_R}, \quad (34)$$

and further in the limit $\tau \rightarrow \infty$,

$$\omega_{\pm}^{-1} \rightarrow \frac{1}{2m_R}. \quad (35)$$

In the large τ, β limit (at non-zero lattice spacing) with $\omega_+ = \omega_-$, Eq. (16) will be shown to valid for a much larger class of theories in the next section. First however, the distribution in Eq. (16) is compared to the numerical distribution of $\phi(t)\phi(0)$ for the two-dimensional free real scalar field theory discretized as in Eq. (17). Computations are performed for a lattice of size $N_s \times N_t = 20 \times 40$ for $t = 6$ and a sample size of $\mathcal{N} = 10^6$ and the resulting distribution is shown in Fig. 2.

The parameters $\omega_{\pm}$ to use in Eq. (16) for comparison are determined as follows. The lowest moments of the distribution in Eq. (16) are given as:

$$\begin{aligned} \langle x \rangle &= \frac{\omega_- - \omega_+}{2\omega_+ \omega_-}, \\ \langle x^2 \rangle &= \frac{3\omega_-^2 - 2\omega_- \omega_+ + 3\omega_+^2}{4\omega_-^2 \omega_+^2}, \end{aligned} \quad (36)$$

where $\langle x^n \rangle = \int_{-\infty}^{\infty} dx P_{\phi(t)\phi(0)}(x) x^n$. After some algebra, these relations can be inverted to give

$$\omega_{\pm} = \frac{\sqrt{\langle x^2 \rangle - 2 \langle x \rangle^2} \mp \langle x \rangle}{\langle x^2 \rangle - 3 \langle x \rangle^2}. \quad (37)$$

As well as the empirical distribution from the numerical calculations, Fig. 2 shows the distribution of Eq. (16) with the above values of $\omega_{\pm}$ determined from moments of the empirical distribution. The choice of $t = 6$ and lattice size are completely arbitrary and equivalently good agreement is seen for all t and for various lattice geometries. The numerical data are clearly well-represented by the expected behavior and the empirical cumulative distribution function converges to the exact cumulative distribution function given by

$$F_{\phi(t)\phi(0)}(x) = \int_{-\infty}^x du P_{\phi(t)\phi(0)}(u) \quad (38)$$

as the sample size is increased, as seen in Fig. 3 where sample sizes $\mathcal{N} = 10^2$ and 10^6 are used.

3. Large-time correlators

The above results for a free real scalar field theory can approximately describe the statistical behavior of correlation functions for many operators at large time in a broad class of bosonic interacting lattice field theories. Consider a theory in $D = d + 1$ dimensions and a local operator⁸ $\mathcal{O}(t, \vec{x})$ such that:

- The Euclidean action of the theory is real,

⁸ This operator does not have to be an elementary field but can be a composite operator constructed from elementary fields.

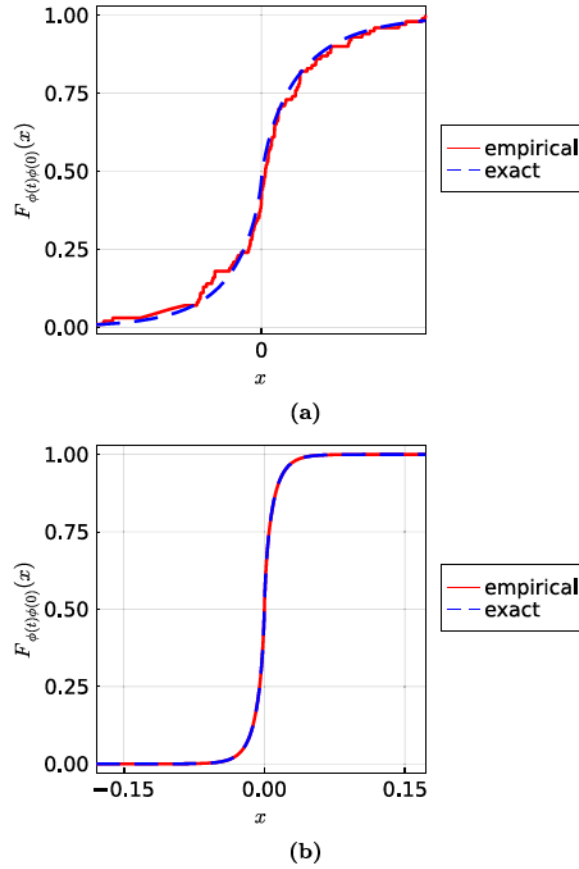


Fig. 3. (a) shows the empirical distribution function and the exact cumulative distribution function for the sample size 10^2 . (b) shows the same data and curves for sample size 10^6 .

- There is a unique, translationally invariant, gapped vacuum $|\Omega\rangle$,
- $\mathcal{O}(t, \vec{x})$ is covariant under temporal and spatial translations,
- The vacuum expectation value of $\mathcal{O}(t, \vec{x})$, $\langle\Omega|\mathcal{O}(t, \vec{x})|\Omega\rangle$, vanishes.

If $\bar{\mathcal{O}}(t)$ is defined as:

$$\bar{\mathcal{O}}(t) = \sqrt{\frac{1}{V}} \sum_{\vec{x}} \mathcal{O}(t, \vec{x}), \quad (39)$$

where $\vec{x}$ runs over all lattice sites on the given time slice, then as $t, N_t \rightarrow \infty$:

$$\lim_{V \rightarrow \infty} P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(q) = \frac{\omega}{\pi} K_0(\omega|q|) + \mathcal{O}\left(e^{-m \min(t, N_t - t)}\right), \quad (40)$$

with $\omega > 0$. Here, $|m\rangle$ is the zero-momentum eigenstate with the smallest energy such that

$$\langle\Omega|\delta(\bar{\mathcal{O}} - u)|m\rangle \neq 0 \quad (41)$$

for some $u \in \mathbb{R}$ and m is its energy.

To see this, the joint probability distribution $P_{\mathcal{O}(t), \mathcal{O}(0)}(u, v)$ that $\mathcal{O}(t)$ takes value u and $\mathcal{O}(0)$ takes value v will be considered. This is defined as

$$\begin{aligned} P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(u, v) &= \oint_{\phi(0)=\phi(N_t)} [\mathcal{D}\phi] e^{-S[\phi]} \delta(\bar{\mathcal{O}}(t) - u) \delta(\bar{\mathcal{O}}(0) - v) \\ &= \text{Tr} \left(e^{-(N_t - t)H} \delta(\bar{\mathcal{O}}(t) - u) e^{-tH} \delta(\bar{\mathcal{O}}(0) - v) \right) \\ &\xrightarrow{t, N_t \rightarrow \infty} \langle\Omega|\delta(\bar{\mathcal{O}}(0) - u)|\Omega\rangle \langle\Omega|\delta(\bar{\mathcal{O}}(0) - v)|\Omega\rangle + \mathcal{O}\left(e^{-m \min(t, N_t - t)}\right), \end{aligned} \quad (42)$$

where ϕ is the set of elementary fields in the theory and $[\mathcal{D}\phi]e^{-S[\phi]}$ is interpreted as a probability measure. The third line expresses the decoupling at large time separations consistent with cluster decomposition. The factor $\langle\Omega|\delta(\bar{\mathcal{O}}(0) - u)|\Omega\rangle$ can be calculated as $\lim_{N_t \rightarrow \infty} \text{Tr} \left(e^{-N_t H} \delta(\bar{\mathcal{O}}(0) - u) \right)$. Therefore, one may write

$$\langle \Omega | \delta(\bar{\mathcal{O}}(0) - u) | \Omega \rangle = \lim_{N_t \rightarrow \infty} \oint_{\phi(0)=\phi(N_t)} [\mathcal{D}\phi] e^{-S[\phi]} \delta(\bar{\mathcal{O}}(0) - u). \quad (43)$$

To evaluate this, N d -dimensional spatial boxes $B_1, \dots, B_N$ centered at an equispaced set of points $\{\vec{x}_i\}$ are defined, each encompassing $\frac{V}{N}$ sites. Coarse-grained quantities $\bar{\mathcal{O}}_i(t)$ are defined as:

$$\bar{\mathcal{O}}_i(t) = \sqrt{\frac{N}{V}} \sum_{\vec{x} \in B_i} \mathcal{O}(t, \vec{x}), \quad (44)$$

such that

$$\bar{\mathcal{O}}(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \bar{\mathcal{O}}_i(t). \quad (45)$$

By integrating over all remaining variables, Eq. (43) can be expressed as:

$$\begin{aligned} \langle \Omega | \delta(\bar{\mathcal{O}}(0) - u) | \Omega \rangle &= \int \left[\prod_{i=1}^N du_i \right] P_{\bar{\mathcal{O}}_1, \dots, \bar{\mathcal{O}}_N}(u_1, \dots, u_N) \\ &\times \delta\left(\frac{u_1 + \dots + u_N}{\sqrt{N}} - u\right), \end{aligned} \quad (46)$$

where here and henceforth the abbreviation $\bar{\mathcal{O}}_i \equiv \bar{\mathcal{O}}_i(0)$ will be used and $P_{\bar{\mathcal{O}}_1, \dots, \bar{\mathcal{O}}_N}(u_1, \dots, u_N)$ is the joint probability density of events in which each $\bar{\mathcal{O}}_i$ takes the value u_i and is normalized such that

$$\int \left[\prod_{i=1}^N du_i \right] P_{\bar{\mathcal{O}}_1, \dots, \bar{\mathcal{O}}_N}(u_1, \dots, u_N) = 1. \quad (47)$$

If the box sizes are larger than a few correlation lengths, the $\bar{\mathcal{O}}_i$ are independent of each other up to exponentially small effects proportional to e^{-ml} , where m is the mass gap specified after Eq. (41) and l is the distance between centers of neighboring boxes. Consequently, $P_{\bar{\mathcal{O}}_1, \dots, \bar{\mathcal{O}}_N}(u_1, \dots, u_N) \approx \prod_{i=1}^N P_{\bar{\mathcal{O}}_i}(u_i)$. Such an approximation becomes exact in the infinite volume limit, and with N and l both taken to infinity⁹:

$$\begin{aligned} \lim_{V \rightarrow \infty} \langle \Omega | \delta(\bar{\mathcal{O}} - u) | \Omega \rangle &= \lim_{N, l \rightarrow \infty} \int \left[\prod_{i=1}^N du_i \right] P_{\bar{\mathcal{O}}_1, \dots, \bar{\mathcal{O}}_N}(u_1, \dots, u_N) \delta\left(\frac{u_1 + \dots + u_N}{\sqrt{N}} - u\right) \\ &= \lim_{N \rightarrow \infty} \int \left[\prod_{i=1}^N du_i P_{\bar{\mathcal{O}}_i}(u_i) \right] \delta\left(\frac{u_1 + \dots + u_N}{\sqrt{N}} - u\right) \\ &= \lim_{N \rightarrow \infty} \frac{1}{(2\pi)^{N+1}} \int_{-\infty}^{\infty} d\lambda \int \left[\prod_{i=1}^N du_i dk_i \tilde{P}_{\bar{\mathcal{O}}_i}(k_i) \right] e^{i(k_1 u_1 + \dots + k_N u_N)} e^{i\lambda \left(u - \frac{u_1 + \dots + u_N}{\sqrt{N}}\right)} \\ &= \lim_{N \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda e^{i\lambda u} \left[\tilde{P}_{\bar{\mathcal{O}}_1}\left(\frac{\lambda}{\sqrt{N}}\right) \right]^N, \end{aligned} \quad (48)$$

where $\tilde{P}_{\bar{\mathcal{O}}_i}(k_i)$ is the Fourier transform of $P_{\bar{\mathcal{O}}_i}(u_i)$. In the fourth line, that fact that $\tilde{P}_{\bar{\mathcal{O}}_i}(k) = \tilde{P}_{\bar{\mathcal{O}}_i}(k)$ by translational invariance has been used. Note that $\tilde{P}_{\bar{\mathcal{O}}_1}(0) = 1$ as $P_{\bar{\mathcal{O}}_1}(\cdot)$ is a normalized probability density function. Similarly, $\tilde{P}'_{\bar{\mathcal{O}}_1}(0) = 0$ as $\langle \bar{\mathcal{O}}_1 \rangle = 0$ by assumption, and setting $\tilde{P}''_{\bar{\mathcal{O}}_1}(0) = -\sigma_{\bar{\mathcal{O}}_1}^2$, $\tilde{P}_{\bar{\mathcal{O}}_1}(\lambda)$ can be expressed as:

$$\tilde{P}_{\bar{\mathcal{O}}_1}(\lambda) = e^{-\frac{1}{2}\sigma_{\bar{\mathcal{O}}_1}^2 \lambda^2 + \mathcal{O}(\lambda^3)}.$$

This implies that as $N \rightarrow \infty$:

$$\lim_{N \rightarrow \infty} \left[\tilde{P}_{\bar{\mathcal{O}}_1}\left(\frac{\lambda}{\sqrt{N}}\right) \right]^N = e^{-\frac{1}{2}\sigma_{\bar{\mathcal{O}}_1}^2 \lambda^2}, \quad (49)$$

and by integrating over λ it is clear that:

⁹ The following set of manipulations follows results of Andrey Andreyevich Markov, see Ref. [27] for further details.

$$\lim_{V \rightarrow \infty} \langle \Omega | \delta(\bar{\mathcal{O}} - u) | \Omega \rangle = \frac{1}{\sqrt{2\pi}\sigma_{\bar{\mathcal{O}}_1}} e^{-\frac{u^2}{2\sigma_{\bar{\mathcal{O}}_1}^2}}. \quad (50)$$

Therefore in the infinite-volume limit, the joint probability density $P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(u, v)$ is given as:

$$\lim_{V \rightarrow \infty} P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(u, v) = \frac{1}{2\pi\sigma_{\bar{\mathcal{O}}_1}^2} e^{-\frac{u^2+v^2}{2\sigma_{\bar{\mathcal{O}}_1}^2}} + \mathcal{O}\left(e^{-m \min(t, N_t-t)}\right). \quad (51)$$

As a consequence, the distribution of the product $\bar{\mathcal{O}}(t)\bar{\mathcal{O}}(0)$ in the same limit is:

$$\begin{aligned} \lim_{V \rightarrow \infty} P_{\bar{\mathcal{O}}(t)\bar{\mathcal{O}}(0)}(q) &= \int du dv \left[\lim_{V \rightarrow \infty} P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(u, v) \right] \delta(q - uv) \\ &\xrightarrow{t, N_t \rightarrow \infty} \frac{1}{2\pi\sigma_{\bar{\mathcal{O}}_1}^2} \int d\omega \frac{e^{i\omega q}}{\sqrt{\left(\omega - \frac{i}{\sigma_{\bar{\mathcal{O}}_1}^2}\right)\left(\omega + \frac{i}{\sigma_{\bar{\mathcal{O}}_1}^2}\right)}} + \mathcal{O}\left(e^{-m \min(t, N_t-t)}\right). \end{aligned} \quad (52)$$

This expression reduces to Eq. (12) with $\omega_{\pm} = \frac{1}{\sigma_{\bar{\mathcal{O}}_1}^2}$, so finally from Eq. (16) one obtains¹⁰:

$$\lim_{t \rightarrow \infty} \lim_{N_t \rightarrow \infty} \left[\lim_{V \rightarrow \infty} P_{\bar{\mathcal{O}}(t)\bar{\mathcal{O}}(0)}(q) \right] = \frac{1}{\pi\sigma_{\bar{\mathcal{O}}_1}^2} K_0\left(\frac{|q|}{\sigma_{\bar{\mathcal{O}}_1}^2}\right). \quad (53)$$

To test the validity of Eq. (53), interacting ϕ^4 theory in two dimensions is investigated numerically following Ref. [28]. The action for this theory is:

$$S = \sum_i \left[-\frac{b}{2} \phi_i^2 + \frac{u}{4} \phi_i^4 + \frac{K}{2} \sum_{\hat{\mu}=1}^2 (\phi_{i+\hat{\mu}} - \phi_i)^2 \right], \quad (54)$$

where i labels the sites, b, u, K are couplings and $\hat{\mu}$ labels the directions. Through the rescalings

$$\begin{aligned} \phi_i &= \frac{1}{\sqrt{K}} \varphi_i, \\ b &= \theta K, \\ u &= \chi K^2, \end{aligned} \quad (55)$$

the action be rewritten as

$$S = \sum_i \left[\left(2 - \frac{\theta}{2}\right) \varphi_i^2 + \frac{\chi}{4} \varphi_i^4 \right] - \sum_{\langle ij \rangle} \varphi_i \varphi_j, \quad (56)$$

where $\sum_{\langle ij \rangle}$ indicates summation over all pairs of neighboring points. A schematic illustration of the phase diagram corresponding to the above action is given in the Fig. 4 in terms of exponentials of the couplings θ and χ .

Defining the one-parameter path through the coupling space

$$\begin{aligned} e^{-\chi} &= s, \\ e^{-\theta} &= 1 - s, \end{aligned} \quad (57)$$

calculations have been performed for a lattice size of $N_s \times N_t = 40 \times 40$ and sample size $\mathcal{N} = 10^6$ for $s = 0.05 \times k$ where $k \in \{2, \dots, 10\}$ using a publicly available code [29]. For these values of s , the system is in the disordered phase, and larger s values correspond to smaller renormalized mass values, m , closer to the critical line. The parameter $\sigma_{\bar{\mathcal{O}}}^2$ in Eq. (53) is determined through Eq. (37). To quantify how well Eq. (53) describes the numerical data, the total variation between the empirical distribution $E_t(q)$ at time t and the expected asymptotic distribution of Eq. (53) is calculated:

$$\mathcal{T}(t) \equiv \frac{1}{2} \int dq \left| E_t(q) - \frac{1}{\pi\sigma_{\bar{\mathcal{O}}_1}^2} K_0\left(\frac{|q|}{\sigma_{\bar{\mathcal{O}}_1}^2}\right) \right|. \quad (58)$$

It is expected that the total variation vanishes in the large t, N_t limit. The logarithm of the total variation versus t is shown in Fig. 5 where it is seen that the total variation decreases as t is increased until it reaches a plateau value that appears to be independent of s . Note that, as is seen in Eq. (53), $\frac{1}{\pi\sigma_{\bar{\mathcal{O}}_1}^2} K_0\left(\frac{|q|}{\sigma_{\bar{\mathcal{O}}_1}^2}\right)$ is the dominant term in the expansion of $P_{\bar{\mathcal{O}}(t)\bar{\mathcal{O}}(0)}(q)$ at large times and arises from the

¹⁰ It must be stressed that the limit $\beta \rightarrow \infty$ must be taken before the limit $t \rightarrow \infty$ while the limit $V \rightarrow \infty$ can be interchanged with the limits $t \rightarrow \infty$ and $\beta \rightarrow \infty$.

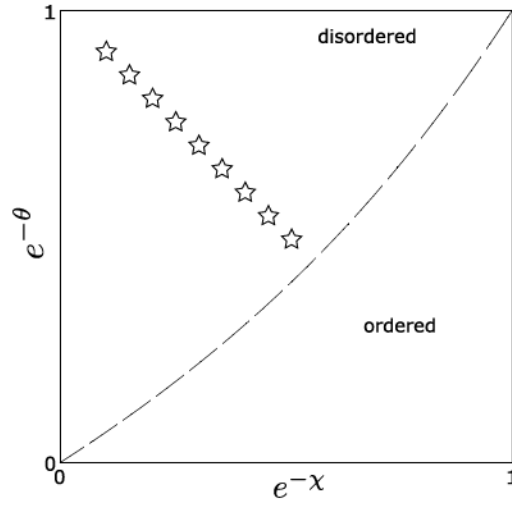


Fig. 4. Schematic phase diagram corresponding to the action given in Eq. (56). The dashed line separates the disordered and ordered phases and is known to be a second order transition. The stars correspond to the values of the couplings at which simulations are performed. A numerical determination of this phase diagram is given by Fig. 8 of Ref. [28].

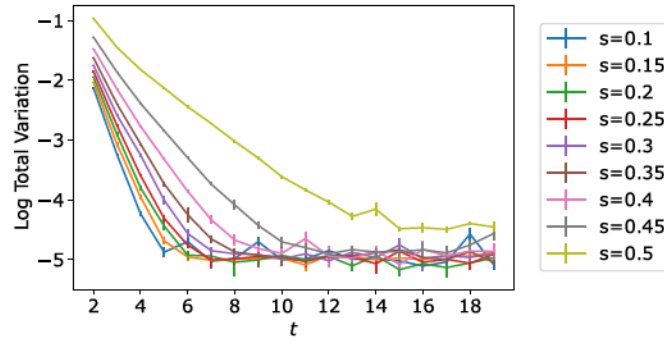


Fig. 5. The logarithm of the total variation as a function of time separation for the ϕ^4 theory in the disordered phase for various values of the parameter s in Eq. (57). Calculations are performed with $N_s \times N_t = 40 \times 40$, and a sample size $\mathcal{N} = 10^6$.

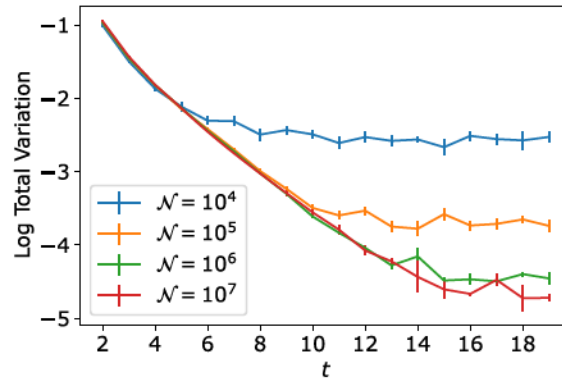


Fig. 6. The logarithm of the total variation as a function of time separation for the ϕ^4 theory in the disordered phase. Calculations are performed for $N_s \times N_t = 40 \times 40$, with $s = 0.5$, and results for two different sample sizes $\mathcal{N} = \{10^4, 10^5, 10^6, 10^7\}$ are shown.

contributions of the vacuum state. The subleading terms in this expansion are due to excited states with vanishing spatial momentum.¹¹ It follows that:

$$\lim_{V \rightarrow \infty} P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(q) = \frac{1}{\pi \sigma_{\bar{\mathcal{O}}_1}^2} K_0 \left(\frac{|q|}{\sigma_{\bar{\mathcal{O}}_1}^2} \right) + e^{-mt} A_m(q) + \int_{\vec{m}} d\mu e^{-\mu t} \rho(\mu) A_\mu(q), \quad (59)$$

¹¹ Since $\bar{\mathcal{O}}(0)$ is invariant under spatial translations, $\delta(\bar{\mathcal{O}}(0) - u)$ is also invariant. Therefore $\langle \Psi | \delta(\bar{\mathcal{O}}(0) - u) | \Omega \rangle$ is non-vanishing only if $|\Psi\rangle$ has vanishing spatial momentum. This proves that $P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(u, v)$ can be expanded in eigenstates with vanishing spatial momentum and same is the true for $P_{\bar{\mathcal{O}}(t), \bar{\mathcal{O}}(0)}(q)$.

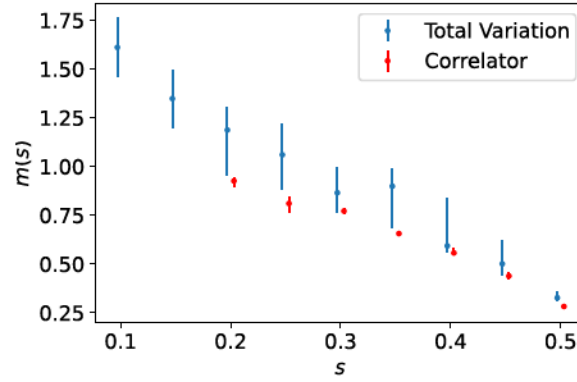


Fig. 7. Estimates of $m(s)$ from the time derivative of the logarithm of total variation are shown in blue, while those from the time-dependence of the correlation function ($C(t)$) are shown in red. Error bars represent the standard deviations calculated using 250 bootstrap resamplings.

where $\tilde{m}$ is the mass of the second excited state with zero spatial momentum, $\rho(\mu)$ is the density of states, and A_m and A_μ are t -independent quantities. In this expression, it is assumed that $0 \ll t \ll \beta/2$ so that effects of the finite temporal extent can be ignored. Considering the total variation as a function of time, it is seen that

$$\log(\mathcal{T}_{th}(t)) = -\log(2) - mt + \log \left(\int dq \left| A_m(q) + \int_{\tilde{m}}^{\infty} d\mu e^{-(\mu-\tilde{m})t} \rho(\mu) A_\mu(q) \right| \right), \quad (60)$$

where $\mathcal{T}_{th}(t)$ is the infinite sample size limit of $\mathcal{T}(t)$. At values of t for which $A_m(q)$ dominates the remaining contributions, one expects to find linear behavior with slope $-m$. Such behavior is seen in Fig. 5 for some values of t , but at larger t , a constant behavior is seen. The above expansion assumes the distribution $\lim_{V \rightarrow \infty} P_{\tilde{O}(t)} \tilde{O}(0)(q)$, while numerical calculations determine $E_t(q)$, that deviates from the exact distribution at finite statistical sampling. Since $e^{-mt} A_m(q)$ vanishes as t increases, the deviation of $E_t(q)$ from the exact distribution will be larger than $e^{-mt} A_m(q)$ at large times, invalidating the above expansion. The main contribution to this deviation is expected to be due to the finite sample size, since convergence to normality is generically very robust if $mN_s \gg 1$. Additionally, since $t \sim N_t/2$ in the figure, effects of the finite temporal extent may need to be accounted for. This expectation is numerically confirmed for the system under consideration in Fig. 6 where results of calculations are shown for $s = 0.5$ for sample sizes $\mathcal{N} \in \{10^4, 10^5, 10^6, 10^7\}$.

Since Eq. (60) depends on the mass of the lowest energy state with the correct quantum numbers (Eq. (41)), the time-dependence of the total variation between the empirical distribution and the asymptotic expectation can be used to extract the corresponding mass, m . Estimates of m from the exponential falloffs seen in Fig. 5 are shown in Fig. 7 as a function of s . Estimates from fits to the time dependence of the correlation function ($C(t)$) are also shown. The fits to extract these masses are discussed in Appendix A. As can be seen in Fig. 7, the masses extracted from the total variation and the time-dependence of the correlation function itself are consistent although the total variation provides a less precise estimate for the parameters used in this study.

4. Summary and outlook

In this work, the statistical behavior of correlation functions of bosonic lattice field theories at large Euclidean time-separations has been investigated. The exact distribution of bilocal correlation functions in free scalar field theory was determined and numerical Monte-Carlo calculations were seen to converge to this distribution. It was also shown that the distributions of many correlation functions in a general class of bosonic lattice field theories approach the same universal distribution in the large-time limit. Numerical tests also confirmed this behavior.

The existence of analytic expressions for distributions even in simplified theories provides an interesting direction for future work. Extension of these results to other phenomenologically relevant theories such as lattice Quantum Chromodynamics is possible and may help in diagnosing the signal-to-noise problem that plagues calculations of many quantities. A more thorough understanding of noise in Monte-Carlo sampling of lattice field theories along directions analogous to those pursued here, may lead to improved strategies for extracting physical information. In particular, tests of empirical distributions against the expected asymptotic distribution of correlation functions at large time may build confidence that a given level of sampling is sufficient for robust physical results to be determined. A deeper exploration of this direction is left to future work.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: William Detmold reports financial support was provided by US Department of Energy. William Detmold reports financial support was provided by National Science Foundation. Cagin Yunus reports financial support was provided by US Department of Energy.

Data availability

The data is available on request.

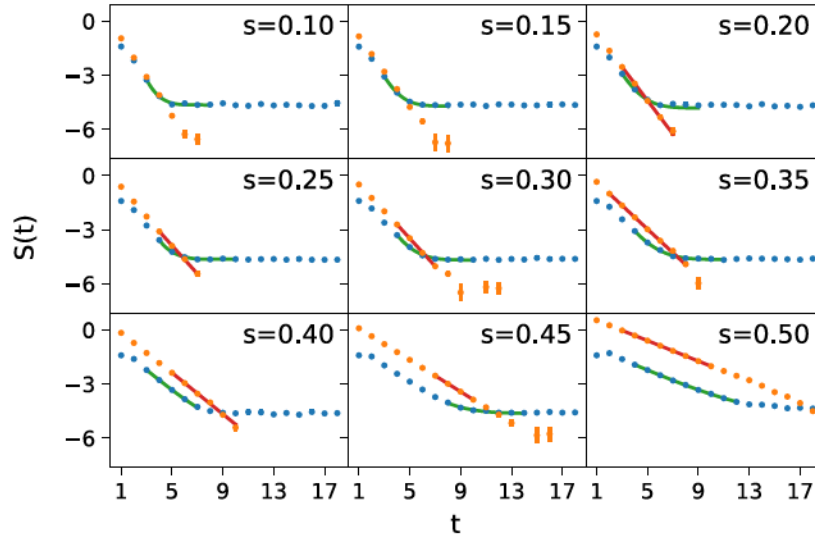


Fig. 8. The behavior of $S(t) = \log \mathcal{T}(t)$ (blue) and $S(t) = \log \langle C(t) \rangle$ (orange) for $t = 1, \dots, 18$ and $s = 0.1, 0.15, \dots, 0.5$ is shown. Uncertainties are determined through the bootstrap method and the best fits $\hat{f}$ (defined in Appendix A) are shown as smooth curves over the time range of the best fit.

Acknowledgements

We are grateful for insightful discussions with P. Shanahan and M. L. Wagman. This work is supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, <http://iafi.org/>) and by the U.S. Department of Energy Office of Science, Office of Nuclear Physics under grant Contract Number DE-SC0011090. WD is also supported by the SciDAC4 award DE-SC0018121.

Appendix A. Fits to correlation functions and total variation

The fits to extract masses from the correlation functions and total variation functions are performed as follows. For each value of s , $B = 250$ bootstrap samples are generated, each having 10^6 samples. For each bootstrap sample, $\mathcal{T}(t)$ and $C(t) \equiv \bar{\mathcal{O}}(t)\bar{\mathcal{O}}(0)$ are calculated. In order to determine the mass, fits of the form $f_{\mathcal{T}}(t) = A + Be^{-E_{\mathcal{T}}t}$ must be performed to $\mathcal{T}(t)$ and fits of the form $f_C(t) = De^{-E_C t}$ must be performed to $C(t)$. In what follows, only the extracted energies E_C and $E_{\mathcal{T}}$ are meaningful and referred to generically as E . For each s , bounds of acceptable fit ranges $t_{<}^{\mathcal{T}}, t_{>}^{\mathcal{T}}, t_{<}^C, t_{>}^C$ are set as follows: $t_{<}^C$ is always chosen to be $t_{<}^C = 1$; $t_{<}^{\mathcal{T}}$ is chosen to be minimum t that satisfies $\log \mathcal{T}(t') < \log \mathcal{T}(t)$ for all $t' > t$; $t_{>}^C$ is chosen to be minimum value of t that satisfies $\langle C(t) \rangle < 0$; $t_{>}^{\mathcal{T}}$ is chosen to be t value where $\log \mathcal{T}(t)$ becomes consistent with a constant. Fits to the two quantities are performed over subranges $(t_{\min}, t_{\max})$ within these bounds that satisfy the inequalities $t_{<} \leq t_{\min} \leq \max(2, t_{\max} - 3)$ and $5 \leq t_{\max} \leq \max(t_{>}, 7)$. Fully correlated fits are performed for both $\mathcal{T}(t)$ and $C(t)$ with covariance matrices calculated through optimal shrinkage [30]. Only the fits that satisfy $\frac{\chi^2}{d.o.f.} < 1.5$ are deemed acceptable and kept for further analysis. These surviving fits are labeled by $f \in \{1, \dots, F\}$ and are assigned a weight $W_f \propto \frac{1}{\delta E_f^2} e^{-\frac{\chi^2}{d.o.f.}}$.

where δE_f^2 is a measure of the statistical variance of a given fit. Defining the deviations $\Delta E_f^b = E_f^b - E_f$ on each bootstrap b for each fit range f (E_f is the bootstrap mean for fit range f), this statistical uncertainty is defined by

$$\delta E_f = \frac{1}{2} \left[Q_{\frac{5}{8}}(\{\Delta E_f\}) - Q_{\frac{1}{8}}(\{\Delta E_f\}) \right] \quad (\text{A.1})$$

where $Q_q(\cdot)$ is the quantile function, whose argument is a bootstrap set. A 68% confidence interval is given as $(Q_{\frac{1}{8}}(\{E\}) - \delta E_f, Q_{\frac{5}{8}}(\{E\}) + \delta E_f)$ where $\{E\} = \{E_f | f = 1, \dots, F\}$ and $\hat{f} = \text{argmax}_f W_f$. In Fig. 8, the calculated values of $\log \langle C(t) \rangle$ and $\mathcal{T}(t)$ are shown as a function of t for each $s \in \{0.10, 0.15, \dots, 0.50\}$. The resulting masses extracted from the two data sets are shown in Fig. 7 as a function of s .

References

- [1] G. Parisi, *Phys. Rep.* 103 (1984) 203.
- [2] G.P. Lepage, in: *Theoretical Advanced Study Institute in Elementary Particle Physics*, 1989.
- [3] S.R. Beane, W. Detmold, T.C. Luu, K. Orginos, A. Parreno, M.J. Savage, A. Torok, A. Walker-Loud, *Phys. Rev. D* 80 (2009) 074501, arXiv:0905.0466 [hep-lat].
- [4] M.G. Endres, D.B. Kaplan, J.-W. Lee, A.N. Nicholson, *Phys. Rev. Lett.* 107 (2011) 201601, arXiv:1106.0073 [hep-lat].
- [5] M.G. Endres, D.B. Kaplan, J.-W. Lee, A.N. Nicholson, *PoS LATTICE2011* (2011) 017, arXiv:1112.4023 [hep-lat].
- [6] T. DeGrand, *Phys. Rev. D* 86 (2012) 014512, arXiv:1204.4664 [hep-lat].
- [7] D. Grabowska, D.B. Kaplan, A.N. Nicholson, *Phys. Rev. D* 87 (2013) 014504, arXiv:1208.5760 [hep-lat].
- [8] A.N. Nicholson, D. Grabowska, D.B. Kaplan, *J. Phys. Conf. Ser.* 432 (2013) 012032, arXiv:1210.7250 [hep-lat].
- [9] J.E. Drut, W.J. Porter, *Phys. Rev. E* 93 (2016) 043301, arXiv:1508.04375 [cond-mat.str-el].
- [10] M.L. Wagman, M.J. Savage, *Phys. Rev. D* 96 (2017) 114508, arXiv:1611.07643 [hep-lat].
- [11] M. Della Morte, L. Giusti, *Comput. Phys. Commun.* 180 (2009) 813.
- [12] M. Della Morte, L. Giusti, *Comput. Phys. Commun.* 180 (2009) 819, arXiv:0806.2601 [hep-lat].
- [13] M. Della Morte, L. Giusti, *J. High Energy Phys.* 05 (2011) 056, arXiv:1012.2562 [hep-lat].

- [14] W. Detmold, M.G. Endres, Phys. Rev. D 90 (2014) 034503, arXiv:1404.6816 [hep-lat].
- [15] P. Majumdar, N. Mathur, S. Mondal, Phys. Lett. B 736 (2014) 415, arXiv:1403.2936 [hep-lat].
- [16] M. Cè, L. Giusti, S. Schaefer, Phys. Rev. D 95 (2017) 034503, arXiv:1609.02419 [hep-lat].
- [17] M. Cè, L. Giusti, S. Schaefer, Phys. Rev. D 93 (2016) 094507, arXiv:1601.04587 [hep-lat].
- [18] M.L. Wagman, M.J. Savage, arXiv:1704.07356 [hep-lat], 2017.
- [19] M.L. Wagman, Statistical Angles on the Lattice QCD Signal-to-Noise Problem, Ph.D. thesis, U. Washington, Seattle (main), 2017, arXiv:1711.00062 [hep-lat].
- [20] W. Detmold, G. Kanwar, M.L. Wagman, Phys. Rev. D 98 (2018) 074511, arXiv:1806.01832 [hep-lat].
- [21] W.J. Porter, J.E. Drut, Phys. Rev. A 95 (2017) 053619, arXiv:1609.09401 [cond-mat.quant-gas].
- [22] M. Dalla Brida, L. Giusti, T. Harris, M. Pepe, Phys. Lett. B 816 (2021) 136191, arXiv:2007.02973 [hep-lat].
- [23] W. Detmold, G. Kanwar, M.L. Wagman, N.C. Warrington, Phys. Rev. D 102 (2020) 014514, arXiv:2003.05914 [hep-lat].
- [24] G. Kanwar, Machine Learning and Variational Algorithms for Lattice Field Theory, Ph.d. Thesis, 2021, arXiv:2106.01975 [hep-lat].
- [25] W. Detmold, G. Kanwar, H. Lamm, M.L. Wagman, N.C. Warrington, Phys. Rev. D 103 (2021) 094517, arXiv:2101.12668 [hep-lat].
- [26] J. Smit, Introduction to Quantum Fields on a Lattice, Cambridge Lecture Notes in Physics, Cambridge University Press, 2002.
- [27] M. Kac, Statistical Independence in Probability, Analysis and Number Theory, vol. 12, 1st ed., Mathematical Association of America, 1959.
- [28] R. Toral, A. Chakrabarti, Phys. Rev. B 42 (1990) 2445.
- [29] A. Romero, **Phi-4-model**, <https://github.com/Physics-Simulations/Phi-4-Model>, 2020.
- [30] O. Ledoit, M. Wolf, J. Multivar. Anal. 88 (2004) 365.