Euclidean-Monte-Carlo-informed ground-state preparation for quantum simulation of scalar field theory

Navya Gupta,[1,2,3,*] Christopher David White,[2,4,†] and Zohreh Davoudi[1,2,3,‡]

[1]*Maryland Center for Fundamental Physics and Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

[2]*Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, Maryland 20742 USA*

[3]*The NSF Institute for Robust Quantum Simulation, University of Maryland, College Park, Maryland 20742, USA*

[4]*Center for Computational Materials Science, U.S. Naval Research Laboratory, Washington, DC 20375, USA*

(Received 2 July 2025; accepted 5 January 2026; published 12 June 2026)

Quantum simulators offer great potential for investigating dynamical properties of quantum field theories. However, preparing accurate nontrivial initial states for these simulations is challenging. Classical Euclidean-time Monte-Carlo methods provide a wealth of information about states of interest to quantum simulations. Thus, it is desirable to facilitate state preparation on quantum simulators using this information. To this end, we present a fully classical pipeline for generating efficient quantum circuits for preparing the ground state of an interacting scalar field theory in $1+1$ dimensions. The first element of this pipeline is a variational *Ansatz* family based on the stellar hierarchy for bosonic quantum systems. The second element of this pipeline is the classical moment-optimization procedure that augments the standard variational energy minimization by penalizing deviations in selected sets of ground-state correlation functions (i.e., moments). The values of ground-state moments are sourced from classical Euclidean methods. The resulting states yield comparable ground-state energy estimates but exhibit distinct correlations and local non-Gaussianity. The third element of this pipeline is translating the moment-optimized *Ansatz* into an efficient quantum circuit with an asymptotic cost that is polynomial in system size. This work opens the way to systematically applying classically obtained knowledge of states to prepare accurate initial states in quantum field theories of interest in nature.

DOI: 10.1103/3c4m-b4t3

I. INTRODUCTION

Classical-computational methods, rooted in the path-integral formulation of quantum mechanics, continue to advance nonperturbative studies of strongly interacting quantum field theories [1–5]. For example, several properties of hadrons and light nuclei, such as their low-lying spectra, structure, and low-energy reaction rates, and of low-density matter at finite temperatures, can now be computed from the underlying quantum chromodynamics (QCD) [6–9]. These studies are performed in Euclidean spacetime to be amenable to Monte-Carlo sampling methods—methods which only scale polynomially with system size. Two-point correlation functions decay as a function of Euclidean time, and at late times are dominated by the ground-state contribution. Three- and higher-point functions are also accessible via Euclidean Monte-Carlo methods. Using theoretical frameworks such as finite-volume formalisms [10,11], such Euclidean correlation functions can be related to scattering amplitudes in the low-energy, low-inelasticity regimes [12–15]. The Euclidean signature used in these computations, however, hinders access to general dynamical, Minkowski-time observables, and direct Minkowski-time Monte-Carlo computations suffer from a sign problem [16,17]. Sign problems also hinder simulations of systems at finite baryon density, even in Euclidean spacetime [18–20].

Quantum simulators are expected to overcome these hurdles: they enable a sign-problem-free simulation of real-time dynamics, and offer an efficient encoding of the Hilbert space of the simulated quantum system [21–25]. Several algorithms for simulating quantum field theories, including gauge theories of relevance to the Standard Model of particle physics [26–36], have been proposed in recent years. These algorithms have mainly focused on concrete resource analyses for simulating time evolution of

*Contact author: navyag@umd.edu
†Contact author: christopher.d.white117.ctr@us.navy.mil
‡Contact author: davoudi@umd.edu

a generic quantum state. One major bottleneck, however, is the preparation of nontrivial initial states, which are often ground and low-energy excited states or their superpositions [37–39]. In fact, ground-state preparation has been shown to be QMA-complete[1] for certain quantum systems [41,42]. In the context of quantum-field-theory simulations, the adiabatic preparation of the initial state in the scattering algorithms of Refs. [37–39,43] constitutes the most computationally demanding step. One often needs to resort to preparing approximate ground states, and to using the knowledge of states and symmetries to find efficient routes to their preparation.

More precisely, preparing a state involves determining and implementing the quantum circuit which realizes that state on a quantum simulator (assuming some exact or approximate mapping between the theory's and simulator's degrees of freedom). In many cases, the wave function of the desired state is already known, and the main state-preparation task is to approximate this known wave function using a quantum circuit [44–48]. In this work, we are primarily interested in cases where the wave function is not known and in fact, is not expected to possess an analytical form (as is the case for general interacting quantum field theories). Thus, the state-preparation task also includes direct or indirect computation of this state. Different state-preparation algorithms can be classified by the source(s) of their dominant computational cost (quantum, classical, or both) together with their guarantees for accuracy (provably accurate or heuristic).

On the one hand, there are quantum-cost-dominated algorithms, which often offer provable guarantees on the accuracy of the prepared ground state. In such cases, modest classical resources can be used to determine quantum circuits which implement certain quantum routines. For example, real-time evolution is used in adiabatic quantum computing to turn the easily prepared ground state of an initial Hamiltonian into that of a final, desired Hamiltonian by slowly varying Hamiltonian parameters [37–39,43]; quantum imaginary-time evolution is used to suppress the excited-state components of an initial, trial state at long imaginary times [49]; eigenvalue-estimation methods such as quantum phase estimation are used to postselect the low-lying eigenstates of the system by measuring the phase of the state [50]; iterative filtering methods are used to suppress the excited-state components of a trial initial state upon repetitive action of a suitable operator [51,52]; quantum measurements are used in projective ground-state preparations [53–55] and reservoir-based methods [56,57] to approach the ground state. While these methods often enjoy provable success and accuracy guarantees, they require deep coherent quantum circuits. For this reason, such quantum algorithms are often untenable for implementation on near-term quantum hardware. Most algorithms, further, require a suitable, easily prepared trial state, and the depth or the success probability of the algorithm may depend on the overlap between the trial state and the desired state to be prepared [50–53].

In contrast, variational quantum algorithms trade these deeper quantum circuits for shallower, parametrized quantum circuits, which are optimized by a quantum-classical feedback loop: the quantum simulator estimates the value of some loss function while the classical computer optimizes it. In the context of quantum chemistry, nuclear physics, and quantum field theories, the parametrized quantum circuit is typically based upon physics-inspired *Ansätzes* (such as Hartree Fock, unitary coupled cluster, approximate full configuration interaction [58–67], etc.). One popular variational quantum algorithm is the variational quantum eigensolver (VQE), which minimizes the energy of the output state. It has been used to prepare ground states [68–72], scattering wave packets [73–75], and thermal states [76,77] of gauge theories in recent years. The variational algorithms, nonetheless, face several challenges: the classical optimization may run into barren plateaus [59,78,79], and encounter a multitude of local minima [80] (unless the circuit *Ansatz* and its initialization are chosen very carefully); statistical fluctuations can hinder quantum-assisted classical estimates of the loss function [59,78,79]; it is not guaranteed that the optimized circuit represents the target ground state accurately; and VQE typically only optimizes the energy, potentially at the cost of reducing accuracy of other physically relevant quantities. Thus, while such variational algorithms offer promising state-preparation routes on near-term quantum computers, their success often hinges upon a good choice for the *Ansatz*.

When additional information about the target wave function is available, one may want to perform a more fine-tuned optimization depending upon subsequent tasks. We desire an *Ansatz* that is *accurate*, i.e., it can get sufficiently close to the ground state, is *circuit-translatable*, i.e., it can be efficiently mapped to a quantum circuit, and is *circuit-efficient*, i.e., the size of the resulting quantum circuit scales with a low-degree polynomial in system size. We will demonstrate an *Ansatz* possessing these qualities for an example bosonic quantum field theory. In fact, we require that both the computation and optimization of the loss function for this *Ansatz* be performed *classically*. Thus, we will fully determine the quantum circuit for preparing ground states in this theory using classical computing, and eliminate any quantum-computing costs beyond implementing this fixed predetermined quantum circuit. We will refer to this property of the *Ansatz* as *classical tractability*. This property is desirable since it avoids the statistical noise associated with quantum computations of the loss function.

[1]QMA-complete complexity class is the quantum counterpart of classical Merlin-Arthur (MA)-complete (i.e., probabilistic NP-complete) class [40].

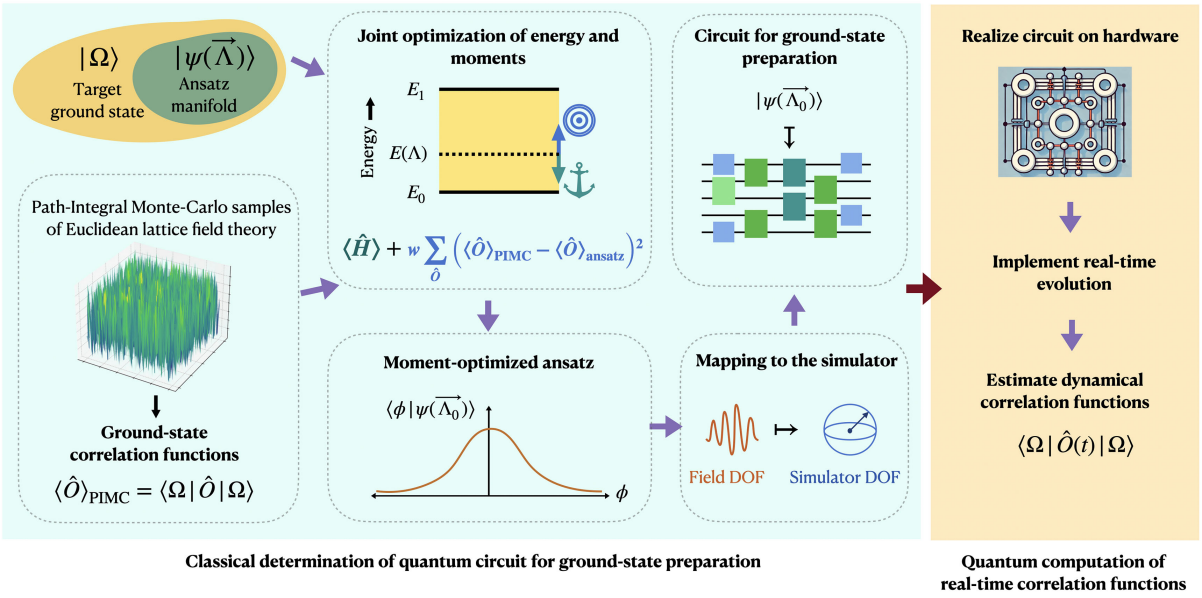


FIG. 1. Schematic overview of classically informed ground-state preparation. Using ground-state correlation functions in an interacting scalar field theory, sourced from Euclidean path-integral Monte Carlo, a joint optimization of the energy and moments of the *Ansatz* wave function is performed. Based on a mapping between the field and simulator degrees of freedom, the optimized *Ansatz* is translated into a quantum circuit using a classical algorithm. This classically determined circuit can thereafter be implemented on quantum hardware, and be used as the starting point for performing other tasks, such as simulating real-time dynamics and estimating dynamical correlation functions. The goal of this paper is to demonstrate the classical determination of the quantum circuit for preparing the ground state.

We would further like to use classically computed properties of ground states to inform the optimization of our variational *Ansatz*. Typically, variational algorithms minimize the *Ansatz* energy and, therefore, assume minimal knowledge about the target ground state. However, one often has access to some special information about ground states. Notably, while Euclidean path-integral Monte-Carlo (PIMC) computations do not provide direct access to the state's wave function, they yield static ground-state correlation functions, or "moments." As depicted in Fig. 1, we will incorporate this knowledge in our loss function to jointly optimize, on a classical computer, the energy and a set of prespecified ground-state moments of the *Ansatz*. Once such a "moment-optimized" *Ansatz* is obtained, we map the degrees of freedom of the theory to those of the quantum simulator, and determine the quantum circuit for preparing this optimized wave function. This classically determined ground-state quantum circuit could then be implemented on quantum hardware. For example, such a circuit can be used as the starting point for the quantum computation of dynamical properties of the state, which are often classically intractable but are quantumly accessible. In this paper, we will focus on only the classical pipeline for determining the ground-state circuit in a quantum field theory, and will leave studies of real-time dynamics to future work.

We demonstrate this procedure using the simplest nontrivial quantum field theory (QFT), a ϕ^4 theory in $1+1$ dimensions (D). Despite their simplicity, scalar fields in $(3+1)$D describe, among other things, pions, the Higgs boson, and scalar inflation fields in nature. Furthermore, scalar field theories are important testing grounds for the development of quantum-simulation algorithms [37–39,43,45–47,81–91]. These theories have also been widely studied using sign-problem-free lattice techniques rooted in Euclidean PIMC; see, e.g., [92]. Hence, static ground-state correlation functions can, in principle, be computed using classical computing. As befits a bosonic theory, we will use bosonic *Ansätze*—states with finite stellar rank [93,94]—to first estimate and then circuitize the ground-state wave function. We show that the structure of this *Ansatz* enables an efficient translation to a quantum circuit on a discrete-variable, i.e., qubit-based platform, and the resulting quantum circuit is efficient, i.e., scales polynomially in the system size. We also comment on the feasibility of preparing this state on a continuous-variable, i.e., bosonic quantum simulator.

To summarize, the two main pillars of our work are: a classically tractable, circuit-translatable, and circuit-efficient *Ansatz* for interacting scalar field theory, and the PIMC-informed moment-optimization procedure for optimizing this *Ansatz*. The efficiency of our *Ansatz* stems from the fact that it is directly expressed in terms of the degrees of freedom of the target field theory, making it straightforward to encode features such as correlations and non-Gaussianity in the *Ansatz*. This is in contrast to the more hardware-oriented variational-circuit *Ansätze*. Our *Ansatz*

is further amenable to optimization on a classical computer, i.e., is classically tractable. Given that any *Ansatz* will only express limited features of a nontrivial interacting ground state, we use PIMC-informed moment optimization to tune the *Ansatz* to reproduce the most relevant features of the ground state. Such features depend on the subsequent quantum-simulation goals, e.g., studies of certain excitations and their dynamics.

This paper is organized as follows. In Sec. II, we introduce the optimization procedure—Euclidean-Monte-Carlo-informed moment optimization—which is independent of the choice of theory and *Ansatz*. In Secs. III and IV, we describe the *Ansatz* and optimization procedure for the $(0+1)$D and $(1+1)$D ϕ^4 ground states, respectively. The discussion in the above sections is independent of quantum simulations, and is applicable to Hamiltonian simulations at large. In Sec. V, we specialize the discussion to quantum simulation by describing the translation of these optimized *Ansätze* to discrete variable, i.e., qubit-based quantum simulators. We conclude in Sec. VI with a summary and outlook. Several appendices are presented to provide further details on our methodology, and to supplement our discussions, including a discussion of the viability of continuous variable, i.e., bosonic quantum simulators, for our state-preparation scheme.

II. EUCLIDEAN-MONTE-CARLO-INFORMED MOMENT OPTIMIZATION

Consider an *Ansatz* family $|\psi(\vec{\Lambda})\rangle$ parametrized by some finite-dimensional vector $\vec{\Lambda}$. In general, a finite-dimensional *Ansatz* family cannot exactly represent the ground state $|\Omega\rangle$ (assumed to be nondegenerate) of a (lattice-regularized Hilbert-space truncated) quantum-field-theory Hamiltonian $\hat{H}$. Therefore, the aim is to determine $\vec{\Lambda}$ such that $|\psi(\vec{\Lambda})\rangle$ is a good proxy for $|\Omega\rangle$. Broadly, we seek a value of $\vec{\Lambda}$ for which the *Ansatz* can be expressed as

$$|\psi(\vec{\Lambda})\rangle = e^{i\phi(\vec{\Lambda})}\sqrt{1-|\epsilon(\vec{\Lambda})|^2}|\Omega\rangle + \epsilon(\vec{\Lambda})|\Omega^{\perp}(\vec{\Lambda})\rangle, \quad (1)$$

where $\epsilon(\vec{\Lambda})$ has a sufficiently small magnitude, and $|\Omega^{\perp}(\vec{\Lambda})\rangle$ is orthogonal to the ground state. Assuming a nonzero mass gap, ΔE, the ratio of the energy discrepancy to the spectral gap is bounded as

$$\begin{aligned}\delta_E(\vec{\Lambda}) &\coloneqq \frac{\langle\psi(\vec{\Lambda})|\hat{H}|\psi(\vec{\Lambda})\rangle - \langle\Omega|\hat{H}|\Omega\rangle}{\Delta E} \\ &= |\epsilon(\vec{\Lambda})|^2\,\frac{\langle\Omega^{\perp}(\vec{\Lambda})|\hat{H}|\Omega^{\perp}(\vec{\Lambda})\rangle - \langle\Omega|\hat{H}|\Omega\rangle}{\Delta E} \\ &\geq |\epsilon(\vec{\Lambda})|^2. \end{aligned} \quad (2)$$

This relation leads to a lower bound on the fidelity $F(\vec{\Lambda}) \coloneqq |\langle\Omega|\psi(\vec{\Lambda})\rangle|^2 \leq 1$

$$F(\vec{\Lambda}) = 1 - |\epsilon(\vec{\Lambda})|^2 \geq 1 - \delta_E(\vec{\Lambda}). \quad (3)$$

Thus, a natural strategy is to minimize $\delta_E(\vec{\Lambda})$ [and hence the *Ansatz* energy $\langle\psi(\vec{\Lambda})|\hat{H}|\psi(\vec{\Lambda})\rangle$], yielding the familiar variational principle. While foundational, this is not a unique prescription. For instance, the relation (3) only provides a lower bound on the fidelity, and the maximum-fidelity *Ansatz* can differ from the minimum-energy *Ansatz*, as demonstrated through an analytically solvable example in Appendix A. Furthermore, for many applications, achieving an energy discrepancy $\delta_E(\vec{\Lambda})$ below a certain tolerance δ_E^{tol} is sufficient. This is especially true in the context of numerical simulations, where systematic errors (due to, e.g., finite volume, discretization, truncation cutoffs) naturally define such a nonzero threshold. This threshold then identifies a *region* of acceptable parameter space rather than a single point. Within this region, *Ansätze* can exhibit varied characteristics despite having similar values of energy. What constitutes a good choice of $\vec{\Lambda}$ within this region, and how one arrives at it, depend upon various factors: the information available about $|\Omega\rangle$, the intended use for $|\psi(\vec{\Lambda})\rangle$, and the computational resources available for optimizing $|\psi(\vec{\Lambda})\rangle$.

When no information about the target ground state is available (except for the Hamiltonian $\hat{H}$), the variational principle provides an upper bound on the ground-state energy upon minimizing $\langle\psi(\vec{\Lambda})|\hat{H}|\psi(\vec{\Lambda})\rangle$. The state $|\psi(\vec{\Lambda}_0)\rangle$ which achieves this minimum energy can then serve as a proxy for the target ground state $|\Omega\rangle$. In some cases, $|\Omega\rangle$ may be known exactly, and one may instead maximize the fidelity F. Oftentimes, one operates at neither extreme of minimal or maximal information about $|\Omega\rangle$. In particular, PIMC techniques efficiently determine (static) ground-state correlation functions $\langle\Omega|\hat{O}|\Omega\rangle$ and the spectral gap ΔE (when the computation is not subject to a sign problem). Thus, one could optimize the values of specific correlation functions of the *Ansatz* depending upon the physics one wants to study. In the context of Hamiltonian simulations, the ground-state representative $|\psi(\vec{\Lambda})\rangle$ serves as an initial state for real-time evolution. Based on the preexisting knowledge about the correlation functions of $|\Omega\rangle$ (sourced from methods like PIMC) and the dynamical properties one wishes to investigate with $|\psi(\vec{\Lambda})\rangle$, one can discriminate between the values of $\vec{\Lambda}$ in regions of low energy discrepancy. This is unlike energy minimization and fidelity maximization, which offer limited control over the features of $|\Omega\rangle$ that can be reproduced by the *Ansatz*.

There are several ways in which one could incorporate preexisting knowledge about ground-state correlation functions, i.e., moments, in the search for $\vec{\Lambda}$. Here, we will work with a loss function which is close to the energy loss function. Suppose that it is important for $|\psi(\vec{\Lambda})\rangle$ to

accurately reproduce the ground-state expectation values for a given set of operators $\mathcal{T} = \{\hat{O}_i\}$, which we refer to as the "target set." We propose to find the appropriate value of $\vec{\Lambda}$ by minimizing the loss function

$$\vec{\Lambda}_0 = \mathrm{argmin}_{\vec{\Lambda}}[\langle\psi(\vec{\Lambda})|\hat{H}|\psi(\vec{\Lambda})\rangle + \sum_{\hat{O}\in\mathcal{T}} w_{\hat{O}}|\langle\psi(\vec{\Lambda})|\hat{O}|\psi(\vec{\Lambda})\rangle - \langle\Omega|\hat{O}|\Omega\rangle|^2], \quad (4)$$

where the real weights $w_{\hat{O}} \geq 0$ are determined by physical intuition and experimentation.[2] We will refer to this procedure of minimizing the above loss function as "moment optimization." When $w_{\hat{O}} = 0$, one recovers energy minimization. As the values of the $w_{\hat{O}}$ are increased, the *Ansatz* better captures the expectation values $\langle\hat{O}\rangle$—but the energy of the optimized *Ansatz* grows. In an effective instance of moment optimization, the behavior of the *Ansatz* target moments will improve at sufficiently small values of the weights, such that the energy penalty paid is small with respect to the spectral gap. In a less favorable attempt, target-moment behavior would only improve at high weight values, implying that a bigger energy penalty is paid, which in turn may lower the fidelity with the ground state. As will become clear through examples of this work, the efficacy of this procedure will depend upon the characteristics of the *Ansatz*, the Hamiltonian parameters, and the chosen set of target moments. We will refer to the resulting *Ansatz* as the "moment-optimized *Ansatz*."

In summary, we will quantify the success of the moment-optimization procedure by comparing the behavior of target moments, and by ensuring that $\delta_E \ll 1$. In what follows, we focus on the case of an interacting scalar field theory. Since fidelity is only efficiently computable for systems with small Hilbert spaces, we will compute both energy discrepancy and fidelity for the single-mode, i.e., the $(0+1)$D case, but restrict the analysis to just energy discrepancy for the multimode, i.e., the $(1+1)$D case.

III. THE SINGLE-MODE CASE: $(0+1)$D ϕ^4 GROUND STATES

Consider a single bosonic mode governed by the dimensionless Hamiltonian

$$\hat{H}_\sigma^{(0+1)} = \frac{\hat{\pi}^2}{2} + \sigma\frac{\hat{\phi}^2}{2} + \lambda\hat{\phi}^4, \quad (5)$$

where $\lambda \geq 0$ and $\sigma = \pm 1$. This Hamiltonian is invariant under the parity, $(\hat{\phi}, \hat{\pi}) \mapsto (-\hat{\phi}, -\hat{\pi})$, and antiunitary time-reversal, $(\hat{\phi}, \hat{\pi}) \mapsto (\hat{\phi}, -\hat{\pi})$, transformations. When $\sigma = 1$, Eq. (5) describes a quantum anharmonic oscillator which possesses a nondegenerate ground state. When $\sigma = -1$, the Hamiltonian can be written in terms of a double-well potential, $V(\hat{\phi})$

$$\hat{H}_{-1}^{(0+1)} = \frac{\hat{\pi}^2}{2} + \lambda\left(\hat{\phi}^2 - \frac{1}{4\lambda}\right)^2 - \frac{1}{16\lambda} \equiv \frac{\hat{\pi}^2}{2} + V(\hat{\phi}) - \frac{1}{16\lambda}. \quad (6)$$

The minima for the two wells are located at $\phi = \pm\frac{1}{\sqrt{4\lambda}}$ while the height of the well is given by $V(\phi = 0) = \frac{1}{16\lambda}$. Thus, as λ becomes smaller, the barrier separating the two wells becomes taller. However, as long as λ is nonzero, there is a finite gap between the lowest-energy parity-symmetric and parity-antisymmetric eigenstates, and hence the Hamiltonian is nondegenerate.

When $\lambda \neq 0$, the theory is interacting, and its ground state does not have a known analytical solution in either regime. In the following, we introduce the *finite-stellar-rank Ansatz* to approximate the ground state of this theory.

A. Single-mode stellar hierarchy and the finite-rank *Ansatz*

In terms of the dimensionless quadrature operators introduced in Eq. (5), one can define the ladder operators

$$\hat{a} \coloneqq \frac{\hat{\phi} + i\hat{\pi}}{\sqrt{2}}, \qquad \hat{a}^\dagger \coloneqq \frac{\hat{\phi} - i\hat{\pi}}{\sqrt{2}}, \quad (7)$$

which satisfy the commutation $[\hat{a}, \hat{a}^\dagger] = 1$. The Fock states, i.e., the eigenstates of the number operator $\hat{a}^\dagger\hat{a}$, will be denoted by $|n\rangle$, where $n \in \mathbb{N} \cup \{0\}$. Unitary transformations acting on the bosonic Hilbert space can either be Gaussian or non-Gaussian. Gaussian unitary transformations are generated by Hamiltonians which are at most quadratic in the bosonic ladder operators. All other unitary transformations are non-Gaussian, and are generated by Hamiltonians with a greater than quadratic degree in the ladder operators.

Our choice of ground-state *Ansatz* is based on the *stellar hierarchy* of Ref. [93]. A single-mode bosonic state $|\psi\rangle$ is said to have a finite stellar rank $R_\psi \in \mathbb{N} \cup \{0\}$ if it admits the decomposition

$$|\psi\rangle = \hat{U}_{G_\psi}\hat{C}_\psi|0\rangle \equiv \hat{U}_{G_\psi}|C_\psi\rangle, \quad (8)$$

where

$$\hat{C}_\psi \coloneqq C_\psi(\hat{a}^\dagger) = \sum_{n=0}^{R_\psi} c'_{\psi,n}(\hat{a}^\dagger)^n. \quad (9)$$

[2] We either work with Hermitian target operators, or non-Hermitian target operators whose expectation value is guaranteed to be real by symmetry constraints. Thus, our loss function is real-valued even in the absence of the absolute value surrounding the terms proportional to $w_{\hat{O}}$.

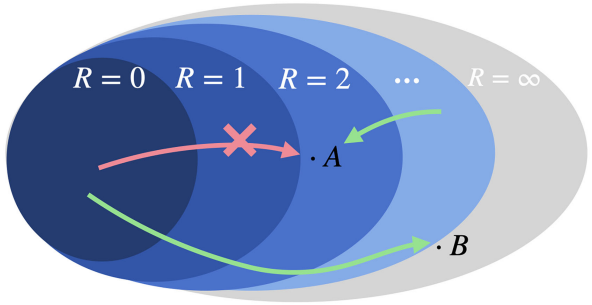


FIG. 2. Schematic representation of the stellar hierarchy inspired by Fig. 1 of Ref. [94]. Stellar rank partitions the single-mode bosonic Hilbert space into disjoint subspaces with different ranks. Finite-rank states are dense in this Hilbert space. Thus, an infinite-rank state, such as the one marked by the point B in this figure, can be approached by a sequence of finite-rank states (depicted by the green arrow), i.e., it can be approximated with arbitrary precision in trace distance using finite-rank states. Similarly, a state with finite-rank R, such as the one marked by point A, can be approached by a sequence of states which have a rank greater than or equal to R (indicated by the green arrow). However, the state A cannot be approached by a sequence of states with rank strictly less than R (indicated by the orange arrow). In other words, there is a finite trace distance between states with finite rank R, and all lower-rank states.

Here, $\hat{U}_{G_\psi}$ is a Gaussian unitary transformation, $C_\psi(\cdot)$ is a polynomial of degree R_ψ with complex coefficients $c'_{\psi,n}$, and $\hat{C}_\psi$ is the operator obtained by evaluating this polynomial with the argument $\hat{a}^\dagger$. The coefficients of the polynomial are chosen so that the state is normalized. The state $|C_\psi\rangle \equiv \hat{C}_\psi|0\rangle$ is called the *core state* corresponding to $|\psi\rangle$. This decomposition of a finite-rank state into a Gaussian unitary acting on a core state is unique. If a single-mode bosonic state does not admit the decomposition in Eq. (8), it is said to have an infinite stellar rank [93].

As shown in Fig. 2, the stellar rank partitions the single-mode Hilbert space into disjoint subspaces with different ranks $R_\psi \in \bar{\mathbb{N}} \cup \{0\}$ ($\bar{\mathbb{N}} \equiv \mathbb{N} \cup \{\infty\}$), and this is referred to as stellar hierarchy. Finite-rank states, i.e., states with $R_\psi \in \mathbb{N} \cup \{0\}$ form a dense subset of the Hilbert space under the trace norm. In other words, infinite-rank states (such as the one marked by B in Fig. 2) are not isolated from finite-rank states in trace distance. Thus, one can find a sequence of finite-rank states which gets arbitrarily close to any given infinite-rank state. Finite-rank states, on the other hand, are isolated from states with lower stellar ranks, and this property is referred to as *stellar robustness*. Thus, there is a finite trace distance between any specific rank-R_ψ state (such as the one marked by A in Fig. 2) and all the states with rank less than R_ψ. A rank-R_ψ state, nonetheless, can be approached by a sequence of states with rank greater than or equal to R_ψ. Thus, for a finite-rank state, the rank indicates the minimal cost of boson additions needed to engineer that state with arbitrary precision in trace distance.

The ground state of the Hamiltonian in Eq. (5) is expected to have an infinite rank [95]. Thus, we choose to approximate this state by finite-rank R_ψ states which obey the symmetries of the Hamiltonian. The ψ subscript on quantities will be dropped from now on for simplicity. Instead of demanding the full rank-R state, i.e., $\hat{U}_G|C\rangle$, to possess the desired symmetries, we will impose the potentially stronger condition that both $|C\rangle$ and $\hat{U}_G$ possesses these symmetries. An arbitrary single-mode Gaussian unitary $\hat{U}_G$ can be expressed as [96] $\hat{U}_G = \hat{S}(\xi)\hat{D}(\alpha)$, up to a phase. Here, $\hat{S}(\xi) \coloneqq e^{\frac{1}{2}(\xi(\hat{a}^\dagger)^2 - \xi^*\hat{a}^2)}$ with $\xi \in \mathbb{C}$ is the squeezing operator, and $\hat{D}(\alpha) \coloneqq e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$ with $\alpha \in \mathbb{C}$ is the displacement operator. $\hat{S}(\xi)$ is invariant under a parity transformation while under the same transformation, $\hat{D}(\alpha) \mapsto \hat{D}(-\alpha)$. This means that a parity-invariant Gaussian is just the squeezing operator $\hat{S}(\xi)$. Under time reversal, $\hat{S}(\xi) \mapsto \hat{S}(\xi^*)$, so time-reversal invariance further restricts the squeezing operator to have a real-valued parameter $\xi = r \in \mathbb{R}$. This results in the symmetric Gaussian unitary transformation

$$\hat{U}_G = \hat{S}(r) = e^{\frac{1}{2}r(\hat{a}^{\dagger 2} - \hat{a}^2)}, \tag{10}$$

with $r \in \mathbb{R}$. Similarly, the core state is constrained to have the decomposition

$$|C\rangle_R \coloneqq \sum_{n=0,2,\ldots,R} c'_n (\hat{a}^\dagger)^n |0\rangle, \tag{11}$$

with $c'_n \in \mathbb{R}$ and even rank R. Defining $c_n \coloneqq c'_n\sqrt{n!}$, the symmetric rank-$R$ single-mode *Ansatz* can finally be expressed as

$$|\psi\rangle_R = \hat{S}(r)(c_0|0\rangle + c_2|2\rangle + \cdots + c_R|R\rangle). \tag{12}$$

This *Ansatz* is parametrized by the $2 + \frac{R}{2}$ real parameters $\vec{\Lambda} = (r, c_0, c_2, \ldots, c_R)$. These $2 + R/2$ real parameters are not independent due to the normalization constraint. During optimization, we will allow all these parameters to vary freely within the normalized manifold of states.

B. Moment optimization

Consider the Weyl-ordered basis $\{\hat{\mathcal{O}}_{p,q} \coloneqq \hat{\phi}^p\hat{\pi}^q\}$ with $p, q \in \mathbb{N} \cup \{0\}$ for the operator Hilbert space. Then:

(i) The lowest-order expectation value of the form $\langle\hat{\phi}^p\hat{\pi}^q\rangle$ with $p = 2m + 1$ and $q = 2n + 1$, is $\langle\hat{\phi}\,\hat{\pi}\rangle$. Time-reversal invariance fixes its value as follows. First, one notes that

$$\langle\hat{\phi}\,\hat{\pi}\rangle = -\langle\hat{\pi}\,\hat{\phi}\rangle = -\langle\hat{\phi}\,\hat{\pi}\rangle + i. \tag{13}$$

In the first equality, the time-reversal transformation is applied. In the second equality, the canonical

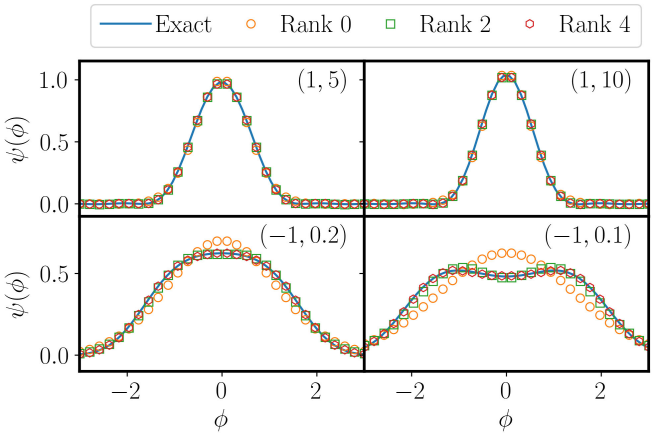


FIG. 3. Minimum-energy wave functions for the $(0+1)$D ϕ^4 Hamiltonian starting from the *Ansatz* in Eq. (12) for various values of rank and Hamiltonian parameters, (σ, λ) (indicated on the top right of each plot). All wave functions show relatively good agreement with the exact wave function except for the $R = 0$ *Ansatz* deep in the double-well regime, i.e., for $(\sigma, \lambda) = (-1, 0.1)$.

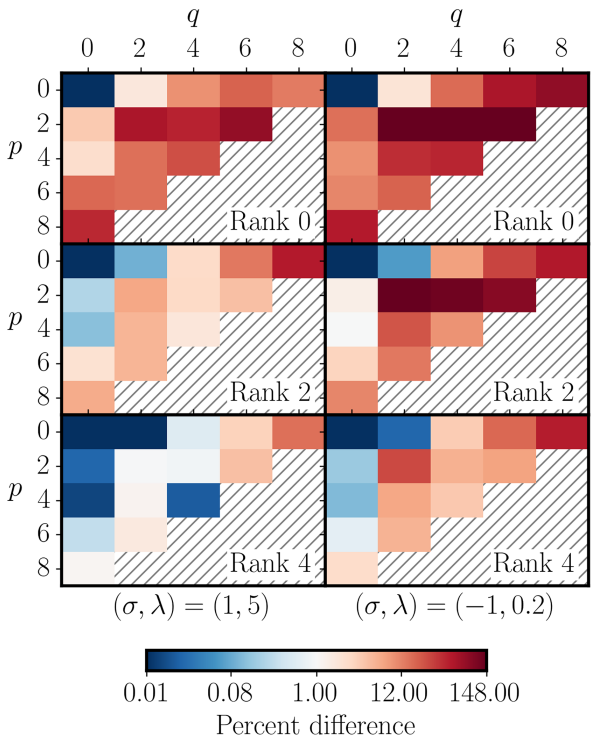


FIG. 4. Percent errors in the moments $\langle \hat{\mathcal{O}}_{p,q} \rangle = \langle \hat{\phi}^p \hat{\pi}^q \rangle$ with $p + q \leq 8$ for minimum-energy states with various *Ansatz* ranks in Eq. (12), and for $(\sigma, \lambda) = (1, 5)$ and $(-1, 0.2)$ in the left and right columns, respectively. In all cases, moment errors are the least for the operators that feature in the Hamiltonian, i.e., $\hat{\phi}^2, \hat{\pi}^2, \hat{\phi}^4$ [corresponding to $(p, q) = (2, 0), (0, 2), (4, 0)$, respectively], and they increase as the value of $p + q$ increases. Errors decrease as the *Ansatz* rank increases, and are worse for the double-well case compared with the anharmonic-oscillator case.

bosonic commutation relation is used. Therefore, $\langle \hat{\phi}\, \hat{\pi} \rangle = \frac{i}{2}$.

(ii) All other expectation values of the form $\langle \hat{\phi}^{2m+1} \hat{\pi}^{2n+1} \rangle$ with $m, n \neq 0$ can be expressed as a sum of expectation values of lower-order basis elements. To see this, one notes that

$$\begin{aligned} \langle \hat{\phi}^{2m+1} \hat{\pi}^{2n+1} \rangle &= -\langle \hat{\pi}^{2n+1} \hat{\phi}^{2m+1} \rangle \\ &= -\langle \hat{\phi}^{2m+1} \hat{\pi}^{2n+1} \rangle + \cdots . \end{aligned} \tag{14}$$

Time-reversal invariance is used in the first line to equate the two expectation values. In the second line, the canonical bosonic commutation relation is employed. The ellipses represent a linear combination of the expectation values of lower-order basis elements, i.e., expectation values of the form $\langle \hat{\phi}^{p'} \hat{\pi}^{q'} \rangle$ with $p' < 2m + 1$ and $q' < 2n + 1$, considering the commutation relation $[\hat{\phi}, \hat{\pi}] = i$. Thus, $\langle \hat{\phi}^{2m+1} \hat{\pi}^{2n+1} \rangle$ can be expressed as a sum of expectation values of lower-order basis elements, which are eventually related to the trivial value $\langle \hat{\phi}\, \hat{\pi} \rangle = \frac{i}{2}$ and to expectation values with p' and q' being both even.

(iii) When p and q are not both even or odd, $\langle \hat{\mathcal{O}}_{p,q} \rangle$ vanishes as a result of parity invariance.

Thus, an arbitrary (parity and time-reversal) symmetric state is characterized by independent correlation functions of the form $\langle \hat{\mathcal{O}}_{p,q} \rangle = \langle \hat{\phi}^p \hat{\pi}^q \rangle$, where both p and q are even.[3] As expected, this set of independent correlation functions is infinite-dimensional (unless further constraints such as Gaussianity or stationarity are imposed [97–99]). Thus, we will restrict our attention to bosonic operators $\hat{\mathcal{O}}_{p,q}$ with some fixed maximum degree.

1. Energy minimization

We first minimize energy alone, i.e., we set $w_{\hat{\mathcal{O}}} = 0$ in Eq. (12). The resulting wave functions are shown in Fig. 3 while the errors in the moments $\langle \hat{\mathcal{O}}_{p,q} \rangle$ with $p + q \leq 8$ are presented in Fig. 4. The values of the fidelity F and the energy-discrepancy to spectral-gap ratio δ_E are provided in Table I. For a fixed value of (σ, λ), the behavior of all three metrics—F, δ_E, and moment errors—improves with increasing the *Ansatz* rank. In the double-well regime $(\sigma = -1)$, the behavior of *Ansätze* of all three lowest ranks is worse compared with that in the $\sigma = 1$ case. This is reflected in the values of δ_E, F, the moment errors, as well as the shape of the wave function. The rank-0 *Ansatz* obviously fails to reproduce the bimodal form of the ground-state wave function. Even though the rank-2 and

[3]Note that $\langle \hat{\phi}^{2p} \hat{\pi}^{2q} \rangle$ is real-valued. This is because the complex conjugate $\langle \hat{\phi}^{2p} \hat{\pi}^{2q} \rangle^* = \langle \hat{\pi}^{2p} \hat{\phi}^{2q} \rangle$ is equal to $\langle \hat{\phi}^{2p} \hat{\pi}^{2q} \rangle$ due to time-reversal invariance.

TABLE I. Fidelity and the energy-discrepancy to spectral-gap ratio for the minimum-energy *Ansatz* for various (σ, λ) and ranks R.

σ	λ	Rank	Fidelity	δ_E (%)
1	5	0	0.9986896	0.6415
		2	0.9999647	0.0315
		4	0.9999981	0.0024
1	10	0	0.9985955	0.6885
		2	0.9999620	0.0343
		4	0.9999979	0.0027
−1	0.2	0	0.9885943	5.8026
		2	0.9996527	0.4166
		4	0.9999771	0.0418
−1	0.1	0	0.9557221	26.5263
		2	0.9983677	2.5773
		4	0.9998871	0.3021

rank-4 *Ansätze* are able to achieve this bimodality, the fidelities obtained in this case are not as high as those obtained for the anharmonic-oscillator case. Increasing the *Ansatz* rank would, therefore, be necessary to achieve a better approximation of the ground-state wave function.

The moment-error plots in Fig. 4 show that minimizing the energy results in some natural distribution of moment errors across different values of p and q. In particular, the errors are small for $(p, q) = (2, 0), (4, 0), (0, 2)$, i.e., for the moments $\langle\hat{\phi}^2\rangle$, $\langle\hat{\phi}^4\rangle$, and $\langle\hat{\pi}^2\rangle$ (corresponding to operators that are present in the Hamiltonian). However, the moment errors increase as p and q increase. It may be desirable to improve the behavior of higher-order moments. In the following, we show how this can be done at a minimal cost in F and δ_E.

2. Column-moment optimization

Suppose one wishes to accurately reproduce the following "column" moments

$$\mathcal{T}_q = \{\hat{\mathcal{O}}_{p,q} \text{ for even } p \text{ satisfying } p + q \leq 8\}. \quad (15)$$

We take $\mathcal{T}_q$ to be the target set for the moment-optimization loss function in Eq. (4), and the weight accompanying the target operator $\hat{O}$ to be

$$w_{\hat{O}} = \frac{1}{(\langle\Omega|\hat{O}|\Omega\rangle)^2}. \quad (16)$$

Thus, the loss function in Eq. (4) becomes the sum of the *Ansatz* energy and the squared fractional discrepancies in the target moments.

The errors in the moments $\langle\hat{\mathcal{O}}_{p,q}\rangle$ with $p + q \leq 8$ resulting from this optimization are shown in Fig. 5 for $(\sigma, \lambda) = (1, 5)$ with the rank-4 *Ansatz*. We observe that the moment optimization redistributes moment errors such that

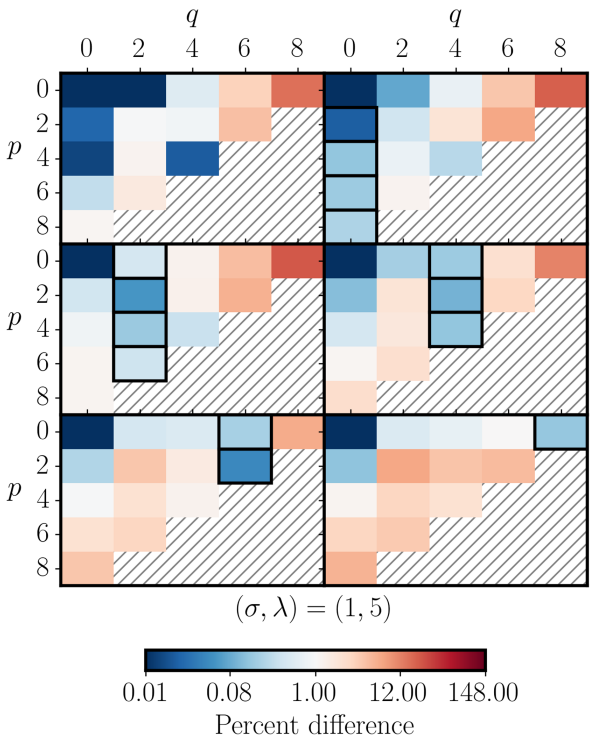


FIG. 5. Percent errors in the moments $\langle\hat{\mathcal{O}}_{p,q}\rangle = \langle\hat{\phi}^p\hat{\pi}^q\rangle$ with $p + q \leq 8$ resulting from column-moment optimization for $(\sigma, \lambda) = (1, 5)$ with the rank-4 *Ansatz* in Eq. (12). The top left plot shows the moment errors for the minimum-energy *Ansatz* (also shown as a subplot in Fig. 4), and the rest of the plots show the errors with the boxed column moments chosen as target moments. For the columns $q = 6$, 8, all errors in the boxed moments are significantly lowered when the corresponding columns moments are optimized (as compared to mere energy minimization). For the columns $q = 0$, 2, 4, the error in some of the boxed moments is raised (relative to the minimum-energy case) while the errors in other boxed moments is lowered. However, the errors in all boxed moments are less than 1% (unlike the minimum-energy case for the same columns). Thus, there is a redistribution of moment errors so that the overall behavior of target moments is improved.

the overall behavior of target moments is improved. More quantitatively, for the minimum-energy *Ansatz*, the average percent errors in each column from left to right (i.e., from $q = 0$ to $q = 8$) are given by 0.35%, 0.94%, 0.46%, 3.8%, and 15.43%, respectively. When the corresponding column moments are optimized, these errors go down to 0.13%, 0.24%, 0.13%, 0.12%, and 0.14%, respectively. While the errors in the first three columns, i.e., $q = 0$, 2, 4, were already small for the minimum-energy *Ansatz*, dramatic improvements are seen in the behavior of the last two columns $q = 6$, 8. These improvements only occur at the cost of increasing nontarget moment errors. For example, when moments in the columns corresponding to $q = 2$, 4, 6, 8 are optimized, the average percent errors in the moments in the $q = 0$ column are 0.87%, 1.00%, 1.83%, and 2.5%, respectively. The redistribution of errors

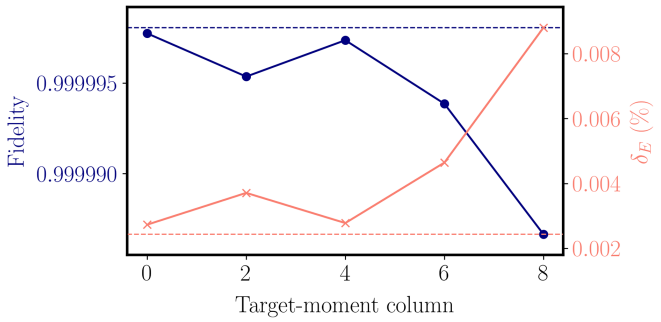


FIG. 6. Fidelity F and the energy-discrepancy to spectral-gap ratio δ_E for the column-moment-optimized *Ansatz* plotted as a function of q, where q is the column labeling the target set. The optimization is performed for a rank-4 *Ansatz* in Eq. (12) with $(\sigma, \lambda) = (1, 5)$. Dashed horizontal lines mark the value of F and δ_E corresponding to the minimum-energy *Ansatz*.

is achieved at an insignificant cost in F and δ_E, as plotted in Fig. 6. The improvement in higher-order target moments $q \geq 4$ is much more pronounced than that in lower-order moments. This could be due to higher-order moments being more *independent* of the lower-order moments featuring in the expression for energy, and also the simpler fact that the size of the target set grows smaller as q increases, imposing less demand on the optimization.

When column-moment optimization is performed for the rank-2 and rank-0 *Ansätze*, the result is qualitatively similar: When the corresponding column's moments are chosen as the target set, there is a systematic reduction in the average percent errors in that column. Furthermore, the improvements for the last two columns are much higher than those in the first three columns. However, the size of the variations in energy and fidelity are higher as compared to the rank-4 case, which indicates that greater energy and fidelity penalties have to be paid to optimize these lower-rank *Ansätze*. In particular, for the rank-2 *Ansatz*, variations in fidelity occur at the fourth decimal place (as opposed to the sixth place for the rank-4 *Ansatz*), while variations in δ_E occur at the second decimal place (as opposed to the third decimal place for the rank-4 *Ansatz*). For the rank-0 *Ansatz*, the variations occur at the third (first) decimal place for fidelity (δ_E). For the rank-4 *Ansatz*, column-moment optimization brings down the mean percent error in the target column to less than 0.25% for each column. However, for the rank-2 and rank-0 *Ansätze*, while the behavior of the average percent errors in the target column improves, the improved errors are still quite high. For instance the average percent error in the $q = 2$ ($q = 6$) column is reduced from about 4.5% (50%) to about 1.4% (29%) for the rank-2 (rank-0) case when the corresponding column is chosen as the target. This shows that lowering the rank reduces the overall capacity for the *Ansatz* to express various ground-state moments, but moment optimization succeeds in improving target-moment behavior in each instance.

IV. THE MULTIMODE CASE: (1+1)D ϕ^4 GROUND STATES

Our aim in this section is to construct and constrain approximate wave functions for the ground state of the $(1+1)$D lattice ϕ^4 theory. The continuum Hamiltonian for this theory is

$$\hat{H}_{\text{cont}} = \int dx \left[\frac{\hat{\pi}(x)^2}{2} + \frac{\partial_x \hat{\phi}(x)^2}{2} + \frac{1}{2} m_0^2 \hat{\phi}(x)^2 + \frac{\lambda_0}{4} \hat{\phi}(x)^4 \right]. \tag{17}$$

Here, m_0^2 and λ_0 are the bare mass and coupling constant, respectively. $\hat{\phi}(x)$ and $\hat{\pi}(x)$ are the field and its conjugate, respectively, obeying the canonical bosonic commutation relation $[\hat{\phi}(x), \hat{\pi}(x')] = i\delta(x - x')$. We discretize this theory on a finite periodic lattice with lattice spacing a and size $L = Na$, to obtain the dimensionless lattice Hamiltonian

$$\hat{H} \equiv a\hat{H}_{a,N} = \sum_{j=0}^{N-1} \left[\frac{\hat{\pi}_j^2}{2} + \frac{(\hat{\phi}_{j+1} - \hat{\phi}_j)^2}{2} + \frac{1}{2} m^2 \hat{\phi}_j^2 + \frac{\lambda}{4} \hat{\phi}_j^4 \right]. \tag{18}$$

Here, we have introduced the dimensionless lattice mass $m^2 \equiv m_0^2 a^2$, the dimensionless coupling $\lambda \equiv \lambda_0 a^2$, and dimensionless lattice-field operators $\hat{\phi}_n \equiv \hat{\phi}(na)$ and $\hat{\pi}_n \equiv a\hat{\pi}(na)$. These field operators obey the commutation relations $[\hat{\phi}_j, \hat{\pi}_k] = i\delta_{j,k}$, where $\delta_{j,k}$ is the Kronecker delta function. Due to periodicity, $\phi_{j+N} = \phi_j$ and similarly for the conjugate field.

This Hamiltonian in Eq. (18) possesses ($\mathbb{Z}_2$) parity [i.e., $(\hat{\phi}, \hat{\pi}) \mapsto (-\hat{\phi}, -\hat{\pi})$], time reversal, lattice inversion (i.e., spatial parity), and ($\mathbb{Z}_N$) lattice translation symmetries. In the thermodynamic limit and for $m^2 < 0$, the system undergoes a phase transition, in which the parity symmetry is spontaneously broken beyond a critical value λ_c. The bare mass m_0 diverges logarithmically $m_0^2 \sim m_r^2 \log(\frac{1}{a})$ [92], where m_r is some finite renormalized mass, while the coupling λ_0 stays finite as the lattice spacing is reduced. Thus, the continuum theory can be characterized by the dimensionless coupling constant λ_0/m_r^2 with a suitable choice for mass renormalization (as discussed in Ref. [92]). The continuum limit is achieved upon taking $N \to \infty$ followed by $a \to 0$, i.e., $(m^2, \lambda) = (m_0^2(a^2)a^2, \lambda_0(a)a^2) \to (0, 0)$ while holding the coupling λ_0/m_r^2 fixed.

To keep the subsequent PIMC computations manageable, we work with a fixed small value of $N = 10$. This value suffices to demonstrate the conclusions of this section. The values of (m^2, λ) are chosen such that $N\Delta E > 1$ and $\Delta E < \pi$, where ΔE is the (dimensionless)

spectral gap. The first condition ensures that the Compton wavelength of the scalar excitation is contained within the system volume while the second condition ensures that the scalar particle is light compared to the UV cutoff imposed by the lattice spacing. To qualitatively study the effects toward the continuum limit, we choose the values of (m^2, λ) to be on a straight line approaching the origin with $m^2 > 0$. Thus, we are restricted to the symmetric regime. We will discuss how our strategy can be generalized to the full phase diagram, and argue that the qualitative features of our approach, including its efficiency, carry through to the symmetry-broken phase with minor modifications.

Using Euclidean PIMC techniques, we estimate the values of various ground-state correlation functions. In particular, we will study the two-point function $\langle\hat{\phi}_{j'}\hat{\phi}_j\rangle$ with $j' \neq j$, which characterizes the strength of nonlocal correlations in the theory. We will also study the non-Gaussianity of the ϕ_j^{2n}-moments resulting from the quartic coupling. While the raw moments $\langle\hat{\phi}_j^{2n}\rangle$ are useful for examining the scale of fluctuations, information on shape (and hence non-Gaussianity) can be more cleanly extracted through relations among different moments. In particular, we will compute the moment ratio[4]

$$R_{2n}(j) \coloneqq \frac{1}{(2n-1)!!}\frac{\langle\hat{\phi}_j^{2n}\rangle}{\langle\hat{\phi}_j^{2}\rangle^n}. \tag{19}$$

For Gaussian states, $R_{2n}(j) = 1$. Thus, any deviations from Gaussianity can be readily seen in the moment ratio. Note that due to translational invariance, $\langle\hat{\phi}_{j'}\hat{\phi}_j\rangle = \langle\hat{\phi}_0\hat{\phi}_{j-j'}\rangle$. Thus, we will simply refer to the two-point functions by $\langle\hat{\phi}_0\hat{\phi}_j\rangle$. Also, one single quantity $R_{2n} \equiv R_{2n}(j)$ characterizes the moment ratio across the translationally invariant lattice. More details about the PIMC calculations of these quantities can be found in Appendix B.

A. Multimode stellar hierarchy and finite-rank *Ansätze*

Consider the ladder operators

$$\hat{a}_j \coloneqq \frac{\hat{\phi}_j + i\hat{\pi}_j}{\sqrt{2}}, \qquad \hat{a}_j^\dagger \coloneqq \frac{\hat{\phi}_j - i\hat{\pi}_j}{\sqrt{2}}, \tag{20}$$

for each mode $j \in \{0, \ldots, N-1\}$, satisfying the commutation relations $[\hat{a}_j, \hat{a}_k^\dagger] = \delta_{j,k}$ and $[\hat{a}_j, \hat{a}_k] = [\hat{a}_j^\dagger, \hat{a}_k^\dagger] = 0$. We will collectively denote all the mode operators with the notation $\hat{\boldsymbol{a}}^\dagger \coloneqq (\hat{a}_0^\dagger, \ldots, \hat{a}_{N-1}^\dagger)$, and a multimode Fock state with $|\boldsymbol{n}\rangle \coloneqq (n_0!\ldots n_{N-1}!)^{-1/2}(\hat{\boldsymbol{a}}^\dagger)^{\boldsymbol{n}}|\boldsymbol{0}\rangle \coloneqq |n_0\rangle_0 \otimes \ldots \otimes |n_{N-1}\rangle_{N-1}$, where $|\boldsymbol{0}\rangle$ is the Fock vacuum and $(\hat{\boldsymbol{a}}^\dagger)^{\boldsymbol{n}} \equiv (\hat{a}_0^\dagger)^{n_0}\ldots(\hat{a}_{N-1}^\dagger)^{n_{N-1}}$. Just as in the single-mode case, multimode Gaussian unitary transformations are defined to be those which can be generated by Hamiltonians quadratic in the ladder operators.

A multimode state $|\psi\rangle$ with rank $R \in \mathbb{N} \cup \{0\}$ can be expressed as

$$|\psi\rangle = \hat{U}_G\hat{C}|\boldsymbol{0}\rangle = \hat{U}_G|C\rangle. \tag{21}$$

$\hat{U}_G$ is a multimode Gaussian operation, while $|C\rangle \equiv C(\hat{\boldsymbol{a}}^\dagger)|\boldsymbol{0}\rangle$ is the multimode core state. $C(\cdot)$ is the degree R N-variate polynomial

$$C(\hat{\boldsymbol{a}}^\dagger) = \sum_{R'=0}^{R}\sum_{\substack{\boldsymbol{n}\\ \mathrm{sum}(\boldsymbol{n})=R'}} c'_{\boldsymbol{n}}(\hat{\boldsymbol{a}}^\dagger)^{\boldsymbol{n}}, \tag{22}$$

where $\mathrm{sum}(\boldsymbol{n}) \equiv n_0 + n_1 + \cdots + n_{N-1}$, and the complex coefficients $c'_{\boldsymbol{n}}$ are chosen to ensure $|\psi\rangle$ is normalized to unity. The core state is a unique state with bounded support over multimode Fock states which can be rotated into $|\psi\rangle$ by means of a Gaussian unitary transformation.

States which do not admit the decomposition presented in Eq. (21) are said to have an infinite rank. The properties of the single-mode stellar hierarchy are applicable in the multimode case too: finite-rank states can get arbitrarily close to infinite-rank states in trace distance [94]. The $(1+1)$D ϕ^4 ground state is expected to have an infinite rank, and thus, we will approximate it using finite-rank states.

Similar to the single-mode case, we demand that both the multimode core state $|C\rangle$ and Gaussian unitary transformation $\hat{U}_G$ obey the symmetries of the lattice ϕ^4 Hamiltonian in the symmetric phase, i.e., parity, time-reversal, lattice translations, and lattice inversion. As argued later in this section, these restrictions are not enough to make this *Ansatz* classically tractable, i.e., amenable to efficient classical computation. To ensure classical tractability, we will introduce two additional restrictions to our *Ansatz*. First, we will demand the Gaussian operation to not entangle the modes—thus, all cross-mode entanglement in the *Ansatz* will arise from the core state. Second, we will introduce an additional truncation in the core-state polynomial. These restrictions are discussed in more detail below. Eventually, it will be argued that these restrictions do not merely ensure classical tractability, but are also physically inspired.

An arbitrary multimode Gaussian can be expressed as [96]

$$\hat{U}_G = \hat{R}(\Phi_1) \otimes_{j=0}^{N-1} [\hat{S}_j(\xi_j)\hat{D}_j(\alpha_j)]\hat{R}(\Phi_2), \tag{23}$$

[4]This measure is inspired by Ref. [95], which performs a detailed characterization of the non-Gaussianity of single-mode bosonic systems with a quasisolvable sextic potential. Nonetheless, despite adopting the same terminology and notation, our definition of moment ratio in Eq. (19) is different from that introduced in Ref. [95].

where $\hat{S}(\xi_j)$ and $\hat{D}(\alpha_j)$ are the single-mode squeezing and displacement operators introduced earlier, while $\hat{R}(\Phi) \coloneqq e^{i\sum_{j,j'=0}^{N-1}\hat{a}_j^\dagger \Phi_{j,j'}\hat{a}_{j'}}$ is a passive rotation parametrized by the Hermitian rotation matrix Φ. The passive rotations are responsible for entangling different modes, and we will drop them from our *Ansatz*. This will result in a tensor product of single-mode Gaussian operations. Due to translation invariance, and following the arguments in the single-mode section, this would leave us with the simplified symmetric Gaussian operation

$$\hat{U}_G = \otimes_{j=0}^{N-1}\hat{S}_j(r), \tag{24}$$

that is, a tensor product of single-mode squeezing operators. When working in the symmetry-broken phase, in addition to the single-mode squeezers, one needs to include parity-violating single-mode displacements.

Next, we impose the Hamiltonian symmetries on the coefficients $c'_{\boldsymbol{n}}$ of the polynomial $C(\hat{\boldsymbol{a}}^\dagger)$, which generates the core state $|C\rangle$. As a result of the translation and inversion symmetries, $c'_{\boldsymbol{n}}$ is invariant under cyclic and anticylic permutations of $\boldsymbol{n}$. Parity constrains $c'_{\boldsymbol{n}} = 0$ for $\boldsymbol{n}$ such that sum$(\boldsymbol{n})$ is odd (and thus the rank is constrained to be even). Finally, $c'_{\boldsymbol{n}} \in \mathbb{R}$ due to time-reversal invariance. Despite the symmetry constraints imposed above, the space of symmetric multimode rank-R core states grows rapidly with R. Thus, we will introduce an additional truncation in the *Ansatz*. Define the *span* of a Fock state $|\boldsymbol{n}\rangle$ (on a translation- and inversion-symmetric lattice) to be the smallest integer Q' such that all bosonic excitations are contained within the sites $\{j, j+1, \ldots, j+Q'\}$ for some site j. Consequently, define the span of a superposition of such Fock states to be the largest value of span associated with the Fock states in the superposition. We will restrict our *Ansatz* to rank-R, span-$Q \leq \lfloor \frac{N}{2} \rfloor$ core states generated by

$$\hat{C}_{R,Q} \equiv C_{R,Q}(\hat{\boldsymbol{a}}^\dagger) = \sum_{j=0}^{N-1}\hat{c}_{R,Q}^{(j)}, \tag{25}$$

where

$$\begin{aligned}\hat{c}_{R,Q}^{(j)} &\equiv c_{R,Q}^{(j)}(\boldsymbol{a}^\dagger) \\ &= \sum_{R'=0,2,\ldots,R}\sum_{Q'=0}^{Q}\sum_{\substack{n'_0,\ldots,n'_{Q'} \\ n'_0+\cdots+n'_{Q'}=R' \\ n'_0,n'_{Q'}\geq 1}} d_{n'_0,\ldots,n'_{Q'}} \\ &\quad \times (a_j^\dagger)^{n'_0}(a_{j+1}^\dagger)^{n'_1}\ldots(a_{j+Q'}^\dagger)^{n'_{Q'}},\end{aligned} \tag{26}$$

where $d_{n'_0,\ldots,n'_{Q'}} \in \mathbb{R}$ and $d_{n'_0,\ldots,n'_{Q'}} = d_{n'_{Q'},\ldots,n'_0}$. Using straightforward combinatorics, the number of terms in the polynomial $C_{R,Q}(\cdot)$ is $|C_{R,Q}| = N|c_{R,Q}|$, where

$$|c_{R,Q}| = 1 + \sum_{R'=2,4,\ldots,R}\sum_{Q'=0}^{Q}\binom{R'+Q'-2}{Q'}. \tag{27}$$

In the symmetry-broken phase, odd-rank terms will also be included in the above polynomial. Thus, the number of terms in the core-state polynomial will increase. The number of terms will be given by an expression similar to the one in Eq. (27) with the only difference being that R' would take all integer values from 2 to R (including odd values).

Incorporating both the simplification in the Gaussian unitary transformation and the Q-truncation of the core state, the complete form for our simplified *Ansatz* reads

$$|\psi\rangle_{R,Q} \coloneqq [\otimes_{j=0}^{N-1}\hat{S}_j(r)]|C\rangle_{R,Q}, \tag{28}$$

where $|C\rangle_{R,Q} \equiv C_{R,Q}(\hat{\boldsymbol{a}}^\dagger)|\boldsymbol{0}\rangle$. We shall refer to this as the (R,Q) *Ansatz*. Since the squeezing operations are purely on site, any nonlocal correlations in the (R,Q) *Ansatz* must come from the core state. The Q-truncation, therefore, imposes a hard cutoff on two-point correlations unless $Q \geq \lceil \frac{N}{4} \rceil$. Consider two local operators $\hat{A}_{j_1}$ and $\hat{B}_{j_2}$ with support on sites j_1 and j_2 respectively. Then, the *Ansatz* expectation value $\langle \hat{A}_{j_1}\hat{B}_{j_2}\rangle - \langle \hat{A}_{j_1}\rangle\langle \hat{B}_{j_2}\rangle = 0$ if $\min(|j_1 - j_2|, N - |j_1 - j_2|) > 2Q$. Since $\min(|j_1 - j_2|, N - |j_1 - j_2|) \leq \lfloor \frac{N}{2} \rfloor$, the *Ansatz* can yield nonvanishing two-point correlation functions regardless of operators' distance, provided that $Q \geq \lceil \frac{N}{4} \rceil$.

In order to compute local expectation values involving the operators $(\hat{\phi}_j, \hat{\pi}_j)$, it will be useful to express the core state using a polynomial generator with the argument $\hat{\boldsymbol{\phi}}$, i.e., $|C\rangle_{R,Q} = C_{R,Q}(\hat{\boldsymbol{a}}^\dagger)|\boldsymbol{0}\rangle \equiv C^{\phi}_{R,Q}(\hat{\boldsymbol{\phi}})|\boldsymbol{0}\rangle$, with $\hat{\boldsymbol{\phi}} \coloneqq (\hat{\phi}_0, \ldots, \hat{\phi}_{N-1})$. It is always possible to deduce a unique polynomial $C^{\phi}_{R,Q}(\cdot)$ given some core state $|C\rangle_{R,Q}$, such that there is a one-to-one mapping between the terms and the number of independent coefficients of the polynomials $C_{R,Q}(\cdot)$ and $C^{\phi}_{R,Q}(\cdot)$; see Appendix C. For simplicity, we will drop the ϕ superscript in $C^{\phi}_{R,Q}(\cdot)$; the argument of the polynomial clarifies whether $C_{R,Q}(\hat{\boldsymbol{\phi}})$ or $C_{R,Q}(\hat{\boldsymbol{a}}^\dagger)$ is assumed.

Our subsequent numerical study assumes maximum values of $R = 4$ and $Q = 2$ (and the number of modes $N = 10 \geq 2Q$). Thus, it is instructive to explicitly enumerate all the terms in core-state polynomial associated with these maximum R and Q values

$$C_{4,2}(\hat{\boldsymbol{\phi}}) = \sum_{j=0}^{N-1} \{A + B_0\hat{\phi}_j^2 + B_1\hat{\phi}_j\hat{\phi}_{j+1} + B_2\hat{\phi}_j\hat{\phi}_{j+2} + C_0\hat{\phi}_j^4 + C_{11}(\hat{\phi}_j^3\hat{\phi}_{j+1} + \hat{\phi}_j\hat{\phi}_{j+1}^3) + C_{12}\hat{\phi}_j^2\hat{\phi}_{j+1}^2$$
$$+C_{21}(\hat{\phi}_j^3\hat{\phi}_{j+2} + \hat{\phi}_j\hat{\phi}_{j+2}^3) + C_{22}(\hat{\phi}_j^2\hat{\phi}_{j+1}\hat{\phi}_{j+2} + \hat{\phi}_j\hat{\phi}_{j+1}\hat{\phi}_{j+2}^2) + C_{23}\hat{\phi}_j\hat{\phi}_{j+1}^2\hat{\phi}_{j+2} + C_{24}\hat{\phi}_j^2\hat{\phi}_{j+2}^2\}$$
$$= \sum_{j=0}^{N-1} c_{4,2}^{(j)}(\hat{\phi}_j, \hat{\phi}_{j+1}, \hat{\phi}_{j+2}). \tag{29}$$

Here, all coefficients are real. Recall the cyclic translation invariance of the theory which imposes $\phi_j = \phi_{j+N}$. There are $|c_{4,2}| = 14$ independent terms in the $c_{4,2}^{(j)}(\cdot)$ polynomial, consistent with Eq. (27). There are, however, only 11 distinct polynomial coefficients due to inversion symmetry. All these 11 coefficients are allowed to vary during the optimization within the constrained manifold of normalized states.

To determine the complexity of computing expectation values with respect to the (R, Q) *Ansatz*, consider the expectation value of some operator $\hat{\mathcal{O}} = \mathcal{O}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\pi}})$

$$_{R,Q}\langle\psi|\hat{\mathcal{O}}|\psi\rangle_{R,Q}$$
$$= {}_{R,Q}\langle C|[\otimes_{j=0}^{N-1}\ \hat{S}_j^\dagger(r)]\hat{\mathcal{O}}[\otimes_{j=0}^{N-1}\ \hat{S}_j(r)]|C\rangle_{R,Q}. \tag{30}$$

Since $\hat{S}(r)\hat{\boldsymbol{\phi}}\hat{S}^\dagger(r) = e^{-r}\hat{\boldsymbol{\phi}}$ and $\hat{S}(r)\hat{\boldsymbol{\pi}}\hat{S}^\dagger(r) = e^{r}\hat{\boldsymbol{\pi}}$, one can apply the squeezing operators to $\hat{\mathcal{O}}$ to express the expectation value in Eq. (30) as

$$_{R,Q}\langle\psi|\hat{\mathcal{O}}|\psi\rangle_{R,Q}$$
$$= \langle\mathbf{0}|C_{R,Q}(\hat{\boldsymbol{\phi}})\mathcal{O}(e^r\hat{\boldsymbol{\phi}}, e^{-r}\hat{\boldsymbol{\pi}})C_{R,Q}(\hat{\boldsymbol{\phi}})|\mathbf{0}\rangle. \tag{31}$$

Upon writing the terms in the $C_{R,Q}(\hat{\boldsymbol{\phi}})$ polynomial explicitly and defining $\hat{\mathcal{O}}_r \coloneqq \mathcal{O}(e^r\hat{\boldsymbol{\phi}}, e^{-r}\hat{\boldsymbol{\pi}})$, the full expectation value will break into a sum of $|c_{R,Q}|^2$ expectation values of the form

$$\sum_{j,j'=0}^{N-1}\langle\mathbf{0}|\hat{\phi}_j^{n_0}\cdots\hat{\phi}_{j+Q}^{n_Q}\hat{\mathcal{O}}_r\hat{\phi}_{j'}^{n'_0}\cdots\hat{\phi}_{j'+Q}^{n'_Q}|\mathbf{0}\rangle. \tag{32}$$

We will be interested in computing expectation values of operators that feature in the Hamiltonian, i.e., the local operators $\hat{\phi}_j^2, \hat{\phi}_j^4, \hat{\pi}_j^2$, and the nonlocal operators $\hat{\phi}_0\hat{\phi}_j$. Additionally, we will choose local operators such as $\hat{\phi}_j^{2n}$ with $n > 2$, and nonlocal operators such as $\hat{\phi}_0\hat{\phi}_j$ with $j \geq 2$ to feature in the target-moment set. In all these cases, only $O(Q^2)$ expectations value in the above double-sum need to be evaluated independently, and each such expectation value breaks into a trivial product of local expectation values. Thus, expectation values with respect to the (R, Q) *Ansatz* can be computed with classical time complexity $O(|c_{R,Q}|^2Q^2)$. Notably, this complexity is independent of the system size N. A similar asymptotic scaling is expected in symmetry-broken phase except that the number of terms in the polynomial will be greater because of the inclusion of parity-odd terms. Since we are interested in expectation values of low-order observables, the inclusion of displacements in the product Gaussian would only lead to changes in prefactors in the complexity: In the single-mode case, under a displacement by some $b \in \mathbb{C}$, an expectation value of the form $\langle\hat{\phi}^n\rangle$ would map to $\langle(\hat{\phi} + b)^n\rangle$. Thus, the original expectation value breaks into a linear combination of expectation values of lower orders of $\hat{\phi}$. Since n is independent of N, R, and Q, and in fact, for this work is no more than four, this just changes the total complexity by an $O(1)$ prefactor, and leaves the asymptotic complexity unchanged. A similar argument works for the multimode case when only a product of such single-mode displacements is included in the *Ansatz*. In either phase, if the *Ansatz* consisted of a general multimode Gaussian instead of the product Gaussian, the complexity would be $O(|c_{R,Q}|^2N^2)$. This is a primary reason for choosing a product Gaussian.

The ground state of an interacting ϕ^4 scalar field theory with bare mass m can also be approximated by that of a free scalar field theory, but with a *different* mass μ. This constitutes another finite-rank *Ansatz*, namely, the Gaussian effective potential (GEP) [86,100,101]. Here, μ is a single parameter specifying the *Ansatz*.[5] For the sake of comparison, we will also consider this *Ansatz*. Recall that the free scalar-field-theory Hamiltonian can be diagonalized as

$$\hat{H}_{\text{free}} = \sum_{k=0}^{N-1}\omega_k\left(\hat{A}_k^\dagger\hat{A}_k + \frac{1}{2}\right), \tag{33}$$

where $\omega_k^2 \coloneqq \mu^2 + 4\sin^2(\frac{\pi k}{N})$ and

$$\hat{A}_k \coloneqq \sqrt{\frac{\omega_k}{2}}\hat{\tilde{\phi}}_k + \frac{i}{\sqrt{2\omega_k}}\hat{\tilde{\pi}}_k$$
$$= \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}(\cosh(\chi_k)\hat{a}_j + \sinh(\chi_k)\hat{a}_j^\dagger)e^{-i\frac{2\pi jk}{N}}$$
$$\equiv \hat{U}_{\text{GEP}}\hat{a}_k\hat{U}_{\text{GEP}}^\dagger, \tag{34}$$

[5]The full definition of the GEP introduced in Ref. [100] includes a displacement. The displacement is dropped here because we are only interested in states that respect parity.

with $\hat{\tilde{O}}_k \coloneqq \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}\hat{O}_j e^{-i\frac{2\pi jk}{N}}$ and $\chi_k \coloneqq \frac{1}{2}\log(\omega_k)$. The equality in the second line arises from replacing the $\hat{\phi}_j$ and $\hat{\pi}_j$ operators with the $\hat{a}_j$ and $\hat{a}_j^\dagger$ operators using Eq. (20). We have then used the standard form of a Gaussian unitary transformation in Eq. (G3) to define $\hat{U}_{\mathrm{GEP}}$ with $\alpha_k = 0, u^*_{kj} = \frac{\cosh(\chi_k)e^{-i\frac{2\pi jk}{N}}}{\sqrt{N}}$, and $v^*_{kj} = \frac{\sinh(\chi_k)e^{-i\frac{2\pi jk}{N}}}{\sqrt{N}}$, which yield the equivalence in the third line. The ground state $|\psi\rangle_{\mathrm{GEP}}$ is given by the vacuum of the $\hat{A}_k$ modes. Using $\hat{A}_k|\psi\rangle_{\mathrm{GEP}} = \hat{U}_{\mathrm{GEP}}\hat{a}_k\hat{U}^\dagger_{\mathrm{GEP}}|\psi\rangle_{\mathrm{GEP}} = 0$, we deduce that $\hat{U}^\dagger_{\mathrm{GEP}}|\psi\rangle_{\mathrm{GEP}}$ is the vacuum of the original modes $\hat{a}_k$. Thus, $|\psi\rangle_{\mathrm{GEP}} = \hat{U}_{\mathrm{GEP}}|0\rangle$, and this is a rank-0 *Ansatz*. Due to translation invariance, two-point correlation functions in momentum space have the form ${}_{\mathrm{GEP}}\langle\psi|\hat{\tilde{\phi}}_{k_1}\hat{\tilde{\phi}}_{k_2}|\psi\rangle_{\mathrm{GEP}} = f(\omega_{k_1})\delta_{k_1+k_2,0}$. In this instance, $f(\omega_k) = \frac{1}{2\omega_k}$. Thus, the cost of computing the (low-order) position-space correlation functions is equivalent to the cost of taking a discrete Fourier transform of $f(\omega_k)$. For example, for the two-point function,

$$\langle\hat{\phi}_j\hat{\phi}_{j'}\rangle = \frac{1}{N}\sum_k f(\omega_k)e^{i\frac{2\pi(j-j')k}{N}}. \tag{35}$$

This computation has, therefore, an $O(N\log N)$ classical complexity.

B. Euclidean-Monte-Carlo-informed moment optimization

We aim to optimize the multimode *Ansätze* introduced in Sec. IV A by minimizing the loss function in Eq. (4). Ground-state moments $\langle\Omega|\hat{O}|\Omega\rangle$ are sourced from Euclidean PIMC computations. Monte-Carlo data exhibit statistical uncertainties; these uncertainties need to be propagated to the optimized parameters $\vec{\Lambda}_0$. Appendix B presents details of this uncertainty propagation.

We first minimize the energy of various *Ansätze*, and observe the behavior of the two-point functions and local non-Gaussianity of the resulting optimized wave functions. Thereafter, we demonstrate moment optimization for target sets consisting of local and nonlocal $\hat{\phi}$ moments.

1. Energy minimization

The top row in Fig. 7 plots the energy obtained upon minimizing the Hamiltonian expectation value using the (R, Q) and GEP *Ansätze* for various values of (m^2, λ). The bands show the PIMC estimates for the ground and first excited-state energies. For each parameter value, the *Ansätze* can be separated into two groups based on their energy. The first group is the $Q = 0$ (R, Q) *Ansätze*. These exhibit similar values of minimum energies which are higher than the PIMC value. The second group is the $Q \neq 0$ (R, Q) *Ansätze* and the GEP *Ansatz*. They exhibit minimum energies that are close to the PIMC value. The separation in the minimum-energy values associated with these two groups grows, both in absolute terms and relative to the spectral gap, as (m^2, λ) approaches (0,0). This underscores the growing importance of nonlocal correlations in these ground-states in the $(m^2, \lambda) \to (0, 0)$ limit. The $Q = 0$ *Ansätze* have vanishing nonlocal correlations, and increasing their rank, R, is unable to lower the value of minimum energy. However, increasing R *does* improve the behavior of the local non-Gaussianity of these $Q = 0$ *Ansätze*, as will be discussed below. Likewise, once Q is tuned to the minimum nonzero value of $Q = 1$, the minimum-energy *Ansatz* already achieves the PIMC value. Further improvements in R and Q leave this minimum *Ansatz* energy unchanged, but they *do* improve the behavior of the *Ansatz*'s local non-Gaussianity and nonlocal correlations respectively, which will be discussed below. The fact that the GEP leads to a good energy estimate illustrates that the quartic coupling behaves like an effective mass. This can also be seen in the behavior of the two-point function shown in the center row of Fig. 7. The GEP does a good job at reproducing the two-point function for all values of j, while the (R, Q) *Ansätze* do a worse job with increasing j. The Q truncation in the (R, Q) *Ansatz* results in $\langle\hat{\phi}_0\hat{\phi}_j\rangle = 0$ for $j > 2$ when $Q = 1$ and for $j > 4$ when $Q = 2$.

The quartic coupling, however, induces non-Gaussianity in the ground state. As mentioned previously, we inspect non-Gaussianity by studying the ϕ^{2n}-moments and their ratios. The bottom row of Fig. 7 compares the PIMC values of the ϕ^{2n}-moments with the corresponding *Ansatz* values. The $\langle\hat{\phi}^{2n}\rangle$ values for all *Ansätze* gradually deviate from the Monte-Carlo values as n increases, and this deviation gets more pronounced as $(m^2, \lambda) \to (0, 0)$. The values of these moments by themselves do not clearly convey any information on the states' non-Gaussianity. However, when one computes moment ratios, even small discrepancies among different *Ansätze*'s ϕ^2-moments (not visible in Fig. 7 given the scale of the moment plots) contribute significantly to the ratio in Eq. (19), and a clear hierarchy emerges: The moment ratios for the GEP ansatz are pinned to one. For the $R = 2$ *Ansätze*, the value only deviates slightly from one. In contrast, the $R = 4$ *Ansätze* get much closer to reproducing the Monte-Carlo values of the moment ratios. Thus, while the purely Gaussian effect of readjusting the mass can lead the GEP close to the ground state energetically, the GEP fails to reproduce the non-Gaussian moment ratios. The (R, Q) *Ansatz*, on the other hand, can reproduce the expected values of the non-Gaussian moment ratios for sufficiently large rank R. This plot does not show the $Q = 0$ moment-ratio values to avoid crowding. However, the qualitative behavior is similar: the $(R, Q) = (0, 0)$ moment ratio is pinned to one, the $(R, Q) = (2, 0)$ moment ratios bend only slightly away from one, while the $(R, Q) = (4, 0)$ moment ratios get much closer to the corresponding PIMC estimates.

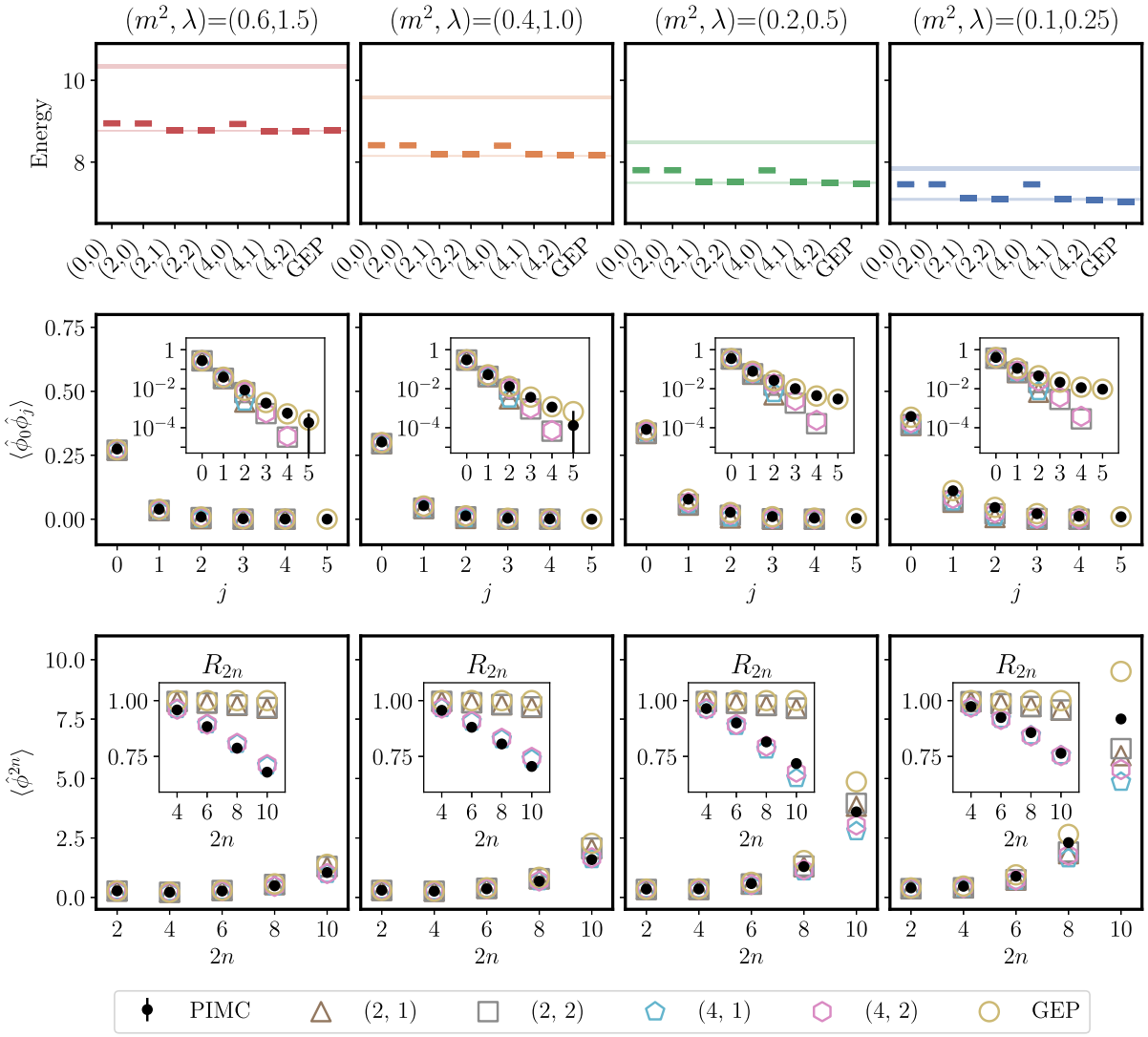


FIG. 7. Minimum Hamiltonian expectation values (top panels), two-point functions (middle panels), and local $\hat{\phi}$ moments and moment ratios (bottom panels) of various *Ansätze* compared with the Monte-Carlo estimates for several values of (m^2,λ). The two bands in the top panels indicate the corresponding Monte-Carlo energy estimate for the ground state and first excited state. Black points with error bars in the middle and bottom panels show the values from PIMC. A log scale is adopted for the inset plots in the middle panels. Error bars and error bands are obtained from bootstrap resampling of the Monte-Carlo estimates, as explained in Appendix B.

Importantly, even though energy minimization leads to comparable values of energy for different *Ansatz* families, the resulting minimum-energy *Ansätze* behave differently in terms of their nonlocal correlations and non-Gaussianity. Given this distinctive behavior of moments for different *Ansatz* families, in the next section, we will explore how moment errors can be varied within a *fixed Ansatz* family using moment optimization.

Before moving on, we plot the values of the squeezing parameter r for the minimum-energy (R,Q) *Ansätze* in Fig. 8. The value of r drops as the continuum limit $(m^2,\lambda)\to(0,0)$ is approached. When the squeezing operator acts on the core state, it extends its support from a finite domain in the local Hilbert space to the entire local Hilbert space. The value of the squeezing parameter r determines how the state's Fock-basis amplitudes are distributed across the Hilbert space. The r value, therefore, becomes important when considering Hilbert-space truncations for implementing the state on qubit-based platforms. We will discuss this point in more detail in Sec. V C.

2. Moment-ratio optimization

We observed that the minimum-energy $(R,Q)=(2,2)$ *Ansatz* captures the ground-state energy rather accurately. However, its moment ratios are quite Gaussian. Thus, we will employ moment optimization to look for alternative low-energy solutions which reproduce the state's non-Gaussianity more accurately. We take the target set to be $\mathcal{T}=\{\hat{\phi}_j^6,\hat{\phi}_j^8,\hat{\phi}_j^{10}\}$ (*any* value of j works because

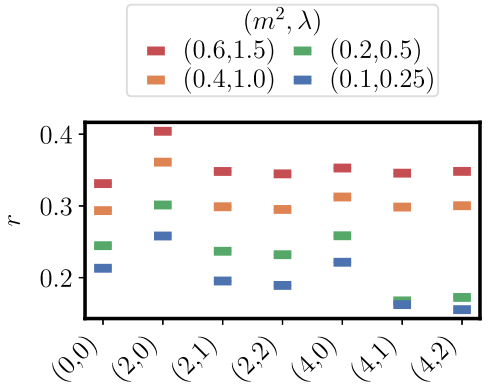


FIG. 8. Optimized value of the squeezing parameter r for the minimum-energy (R, Q) *Ansätze*.

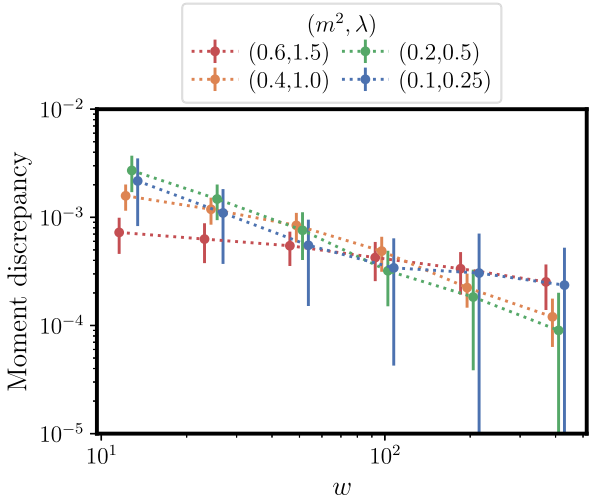


FIG. 9. Sum of squared discrepancies in the target moments plotted as a function of the weight w for moment-ratio optimization. For comparison, the corresponding values for the minimum-energy *Ansatz* for $(m^2, \lambda) = (0.6, 1.5)$, (0.4, 1.0), (0.2, 0.5), and (0.1, 0.25) are 0.07, 0.25, 0.19, and 2.10, respectively. The error bars on the moment discrepancy are obtained by propagating the Monte-Carlo uncertainties using the procedure described in Appendix B. The values associated with different (m^2, λ) are slightly offset in the horizontal direction to improve visibility of the error bars.

of translation invariance) with uniform weights $w_{\hat{O}} = w \in \{12.5, 25, 50, 100, 200, 400\}$, see Eq. (4). These values of $w_{\hat{O}}$ are selected empirically: starting from a small value, $w_{\hat{O}}$ is increased until the target moments shift by a large-enough fraction relative to their minimum-energy values. Such shifts are deduced via inspecting plots shown in Fig. 10, as will be explained shortly.

Figure 9 shows the value of the optimized target-moment discrepancy, i.e., the term proportional to w in the loss function, $\sum_{\hat{O}\in\mathcal{T}} (\langle\psi(\vec{\Lambda})|\hat{O}|\psi(\vec{\Lambda})\rangle - \langle\Omega|\hat{O}|\Omega\rangle)^2$, as a function of w. One observes that the value of this moment discrepancy decreases as a function of w, indicating a steady improvement in the behavior of the target moments $\mathcal{T} = \{\hat{\phi}_j^6, \hat{\phi}_j^8, \hat{\phi}_j^{10}\}$.

Crucially, this improvement happens at a negligible cost in energy relative to the spectral gap, as shown in the top panel of Fig. 10. Additionally, the middle panel in Fig. 10 shows that the behavior of the two-point functions is not altered significantly. As a consequence of the improvement in the expectation value of $\hat{\phi}_j^6, \hat{\phi}_j^8$, and $\hat{\phi}_j^{10}$, there is an initial improvement in the behavior of their moment-ratios as shown in the bottom panel of Fig. 10. However, as w becomes larger, the accuracy of the values of individual moments $\langle\hat{\phi}_j^n\rangle$ is prioritized over their ratios. This results in the moment ratios for $(m^2, \lambda) = (0.4, 1.0), (0.2, 0.5)$, and $(0.1, 0.25)$ overshooting their PIMC value at some value of w. In other words, the value of w at which the optimal behavior of moment ratios is recovered decreases as $(m^2, \lambda) \to (0, 0)$.

Finally, Fig. 11 shows how the optimized parameters evolve as a function of the weight w. Interestingly, for all values of (m^2, λ), the moment optimization guides the *Ansatz* in a roughly similar direction in parameter space—B_2 stays mostly constant, r and B_1 increase slowly, A decreases slowly, and there is a more dramatic increase in B_0 with w.

Since the ground-state moments in the moment-optimization loss function are sourced from PIMC, they carry statistical uncertainty. A greater uncertainty is observed in the moment-optimized quantities corresponding to smaller values of (m^2, λ), as seen in all plots in Figs. 9–11. Furthermore, statistical errors on moment-optimized quantities increase with w, as the moments carrying statistical error are weighed more.

3. Two-point function optimization

Next, we explore alternative low-energy solutions with improved behavior for the two-point function of the $(R, Q) = (2, 2)$ *Ansatz*. Toward this goal, we take the target set to be $\mathcal{T} = \{\hat{\phi}_0\hat{\phi}_4\}$ with the weight $w \in \{1.25, 2.5, 5, 10, 20, 50\} \times 10^4$, see Eq. (4). These weight values are chosen based on similar considerations as described in the previous section. It is observed that much bigger values of weight are required to shift the two-point functions as compared to the moment ratios. This is simply due to the fact that the magnitude of the ground-state moment in this target set is much smaller than that in the previous case.[6] The results of the consequent optimization are shown in Figs. 12–14.

Figure 12 shows that the value of the optimized target-moment discrepancy, i.e., the term proportional to w in the loss function, $\sum_{\hat{O}\in\mathcal{T}} (\langle\psi(\vec{\Lambda})|\hat{O}|\psi(\vec{\Lambda})\rangle - \langle\Omega|\hat{O}|\Omega\rangle)^2$, decreases as a function of w.

[6]For instance, the PIMC estimate for $\langle\hat{\phi}_0\hat{\phi}_4\rangle$ varies from 5.6×10^{-4} for $(m^2, \lambda) = (0.6, 1.5)$ to 0.011 for $(m^2, \lambda) = (0.1, 0.25)$. In contrast, the value of $\langle\hat{\phi}^{10}\rangle$ varies from 1.05 for $(m^2, \lambda) = (0.6, 1.5)$ to 7.5 for $(m^2, \lambda) = (0.1, 0.25)$.

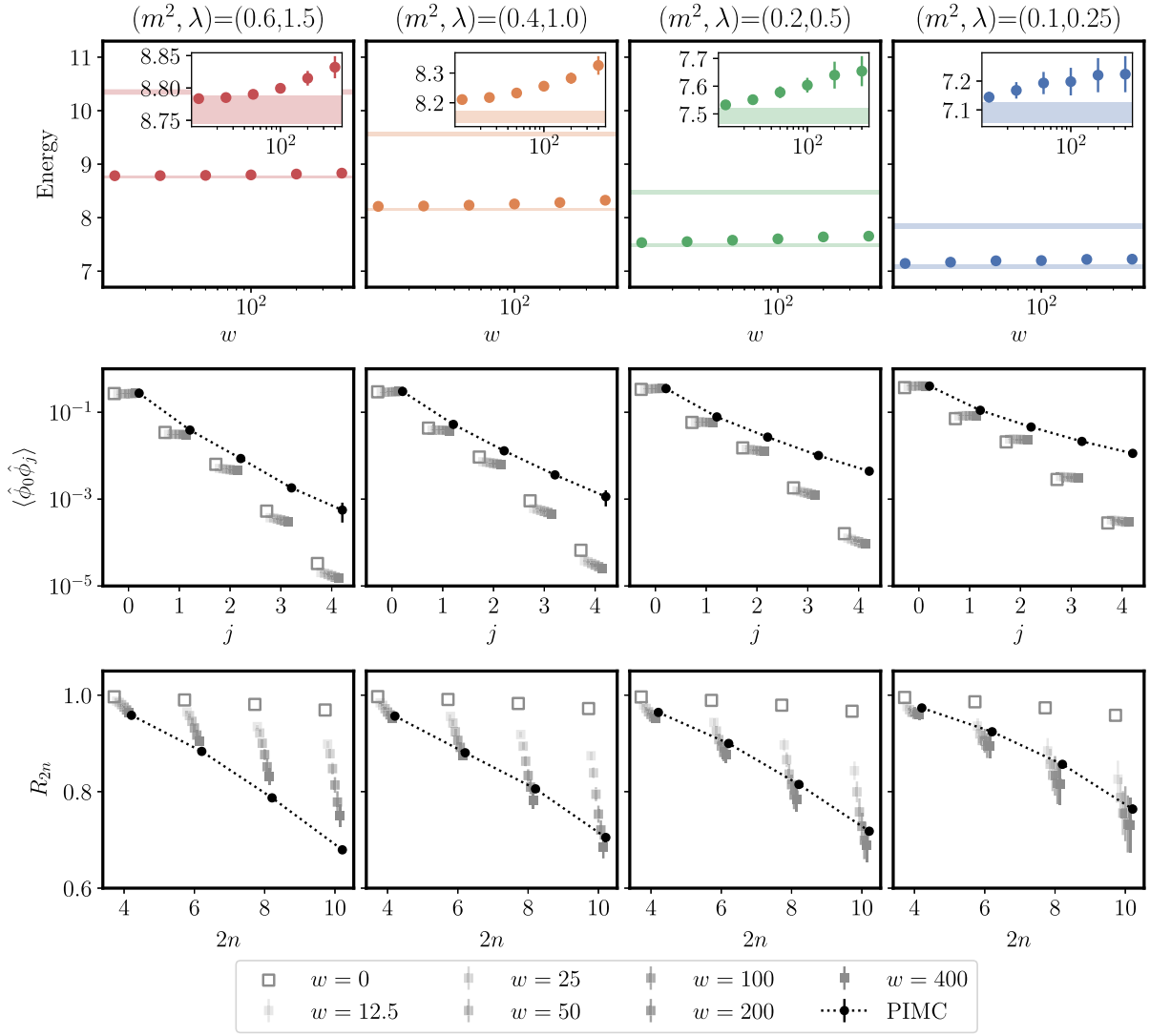


FIG. 10. Optimization of the moment ratio for the $(R, Q) = (2, 2)$ *Ansatz* for various values of (m^2, λ). Moment optimization is performed for the target set $\mathcal{T} = \{\hat{\phi}_j^6, \hat{\phi}_j^8, \hat{\phi}_j^{10}\}$ for various values of w. The bands in the plots on the top panels represent the PIMC values of the ground- and first excited-state energies and their associated uncertainties. In the central and bottom panels, the values associated with different weights and PIMC are slightly offset in the horizontal direction to improve visibility of the error bars. Error bars and error bands on PIMC values are obtained from bootstrap resampling of the Monte-Carlo estimates, as explained in Appendix B. The error bars on the expectation values from moment optimization are obtained by propagating the Monte-Carlo uncertainties using the procedure described in Appendix B.

Even though only the $\hat{\phi}_0\hat{\phi}_4$ moment belongs to the target set, the behavior of both $\langle\hat{\phi}_0\hat{\phi}_3\rangle$ and $\langle\hat{\phi}_0\hat{\phi}_4\rangle$ improves while that of $\langle\hat{\phi}_0^2\rangle$ and $\langle\hat{\phi}_0\hat{\phi}_1\rangle$ stays intact as shown in the bottom panel of Fig. 13. The $\langle\hat{\phi}_0\hat{\phi}_2\rangle$ value hits its PIMC value at a smaller value of w, and overshoots the PIMC value at bigger values of w (i.e., at values where $\hat{\phi}_0\hat{\phi}_j$ with $j = 3, 4$ hit their PIMC values). This peculiar behavior of the *Ansatz* could be potentially fixed by incorporating $\hat{\phi}_0\hat{\phi}_2$ into the target set. The moment ratios stay largely intact in all cases. However, the energy penalty paid rises steadily as $(m^2, \lambda) \to (0, 0)$, and it is significant for $(m^2, \lambda) = (0.1, 0.25)$. Thus, this example illustrates the limits of performing moment optimization with this *Ansatz*. As one approaches the continuum limit, larger values of R and Q are needed to accommodate the correlations.

Figure 14 displays the values of the optimized parameters as a function of the weight w. The largest shifts and errors are associated with the parameter B_2 (which accompanies the $\hat{\phi}_j\hat{\phi}_{j+2}$ term in the core-state polynomial). This is perhaps not surprising because this parameter controls the strength of nonlocal correlations. However, the parameter B_1 (which accompanies $\hat{\phi}_j\hat{\phi}_{j+1}$) hardly shifts.

In contrast with the previous section, the propagated errors in the moment-optimized quantities now decrease as

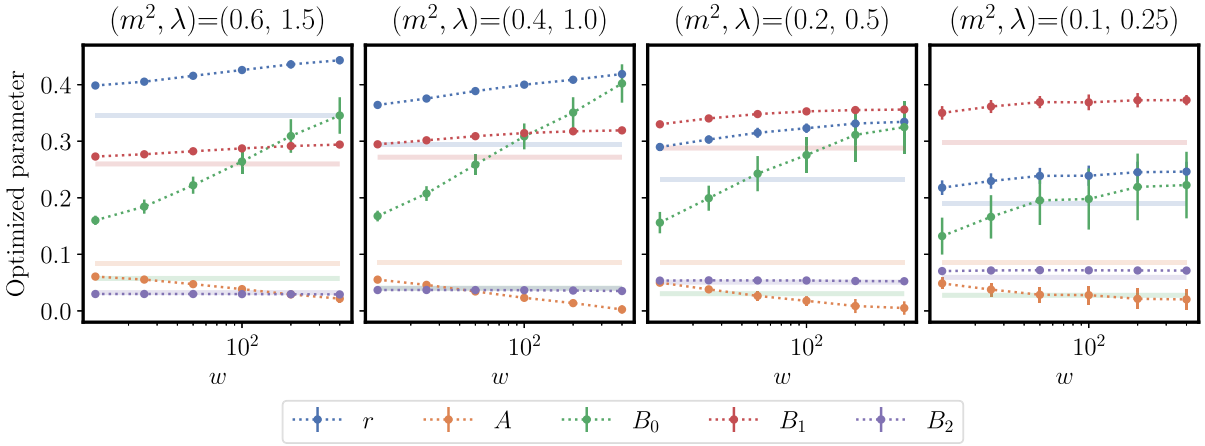


FIG. 11. Optimized values of the $(R, Q) = (2, 2)$ *Ansatz* parameters [with its core state polynomial given by Eq. (29) when all parameters except A, B_0, B_1, and B_2 are set to zero] plotted as a function of w for moment-ratio optimization. The dotted lines passing through the data points are to guide the eye and do not represent a fit. The error bars on the moment-optimized parameters are obtained by propagating the Monte-Carlo uncertainties using the procedure described in Appendix B. The thick light-colored lines show the parameter values for the minimum-energy *Ansatz*. These lines are thickened to improve visibility and do not imply error bars.

$(m^2, \lambda) \to (0, 0)$. Also in contrast with the previous section, while the sizes of the errors increase with weight in the moment-optimized energy, such a dependence is not observed for the moment-optimized two-point functions and moment discrepancy.

To summarize, one observes several distinct behaviors of *Ansatz* moments in the above examples. First, the examples

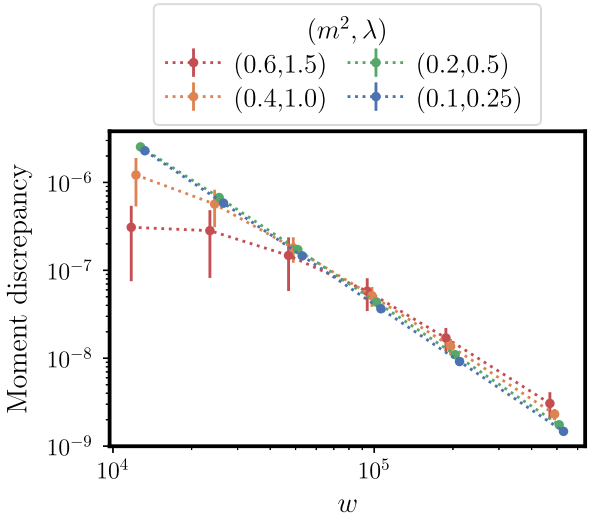


FIG. 12. Sum of squared discrepancies in the target moments plotted as a function of the weight w for two-point function optimization. For comparison, the corresponding values for the minimum-energy *Ansatz* for $(m^2, \lambda) = (0.6, 1.5)$, (0.4, 1.0), (0.2, 0.5), and (0.1, 0.25) are 7.7×10^{-7}, 7.8×10^{-7},1.7×10^{-5}, and 1.4×10^{-4}, respectively. The error bars on the moment discrepancy are obtained by propagating the Monte-Carlo uncertainties using the procedure described in Appendix B. The values associated with different (m^2, λ) are slightly offset in the horizontal direction to improve visibility of the error bars.

reveal completely rigid moments, e.g., for the GEP *Ansatz*, the moment ratio is pinned to one while for the (R, Q) *Ansatz*, $\langle \hat{\phi}_0 \hat{\phi}_j \rangle$ is pinned to be zero for $j > 2Q$ (when $Q < N/2$). Minimizing the *Ansatz* energy imposes a natural distribution of moment errors in the remaining nonrigid moments. When one tries to control these moment errors by using moment optimization, one observes more or less rigid-moment behaviors. For example, the local $\hat{\phi}$ moments for the (R, Q) *Ansatz* appear to be less rigid as they can be continuously improved with w with only a steadily but slowly rising energy penalty. The $\hat{\phi}_0 \hat{\phi}_j$ moments, on the other hand, behave more rigidly as $(m^2, \lambda) \to (0, 0)$ because they are improved at a greater energy penalty.

As mentioned previously, the Q-truncation imposes a hard cutoff on the correlations of the *Ansatz*: $\langle \hat{A}_{j_1} \hat{B}_{j_2} \rangle - \langle \hat{A}_{j_1} \rangle \langle \hat{B}_{j_2} \rangle = 0$, if $2Q < \mathrm{dist}(j_1, j_2) \coloneqq \min(|j_1 - j_2|, N - |j_1 - j_2|) \leq \lfloor \frac{N}{2} \rfloor$. Consequently, we do not expect Q to scale with the size N, but rather with physical theory's correlation length ΔE^{-1} (ΔE is the spectral gap). For distances $\mathrm{dist}(j_1, j_2) < 2Q$, the *Ansatz* state can (potentially) represent correlations correctly. For $\mathrm{dist}(j_1, j_2) \geq 2Q \gg \Delta E^{-1}$ the *Ansatz*'s correlations are identically zero, while the physical state's correlations are nonzero but small. In some cases, e.g., when attempting to distinguish a power-law-correlated state from one with a large but finite correlation length, or preparing a state near a critical point, the *Ansatz* may not be sufficient because one wishes to precisely resolve correlations beyond a correlation length. This is also the reason why we expect parity-violating version of the *Ansatz* to work well in the parity-broken phase because this phase is still a gapped phase with a finite correlation length.

In summary, this section equipped us with an *Ansatz* for the $(1 + 1)$D interacting scalar field theory that possesses at

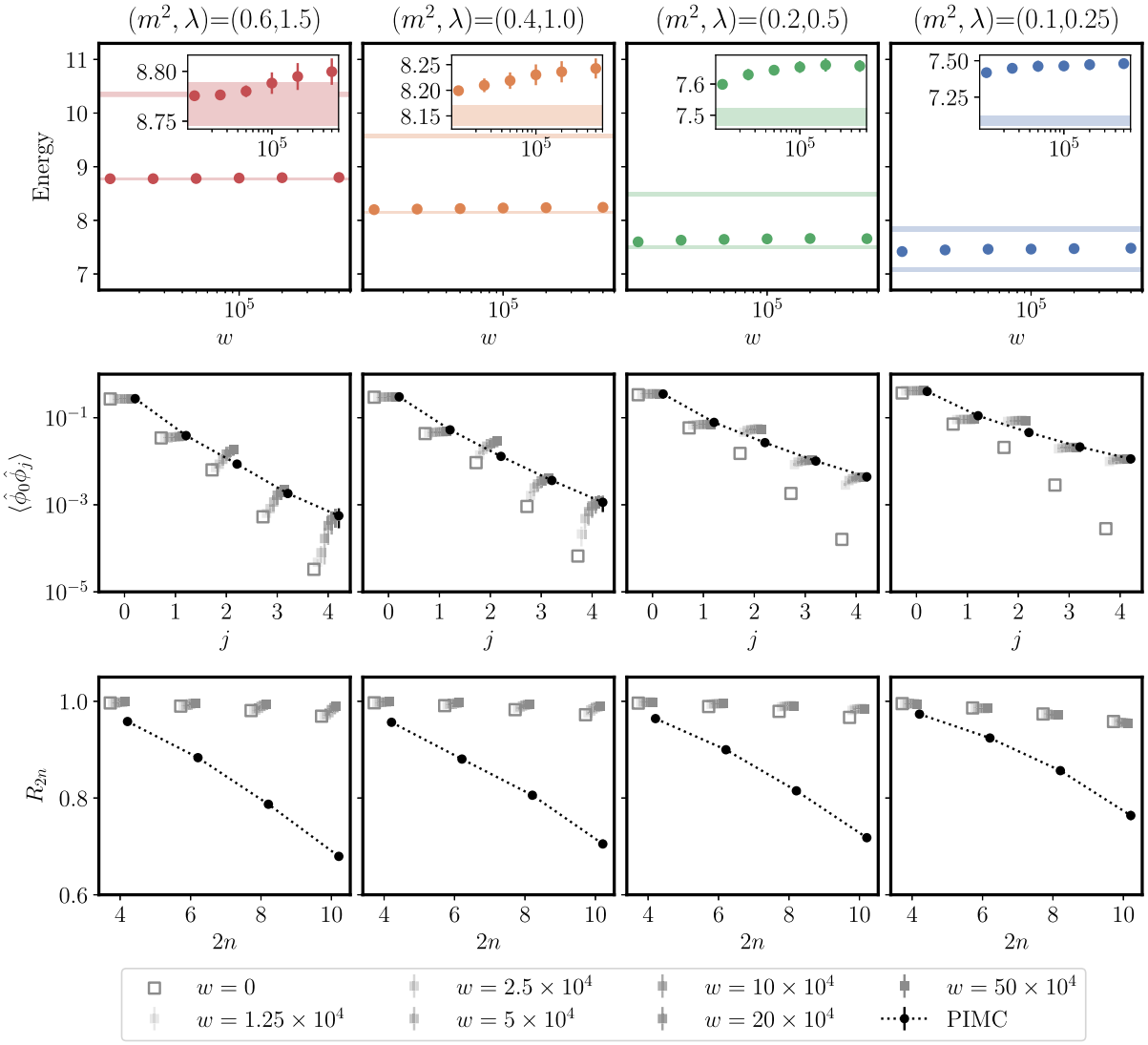


FIG. 13. Optimization of the two-point function for the $(R, Q) = (2, 2)$ *Ansatz* for various values of (m^2, λ). Moment optimization is performed for the target set $\mathcal{T} = \{\hat{\phi}_0\hat{\phi}_4\}$ for various values of w. The bands in the plots on the top panels represent the PIMC values of the ground- and first excited-state energies and their associated uncertainty. In the central and bottom panels, the values associated with different weights and PIMC are slightly offset in the horizontal direction to improve visibility of the error bars. Error bars and error bands on PIMC values are obtained from bootstrap resampling of the Monte-Carlo estimates, as explained in Appendix B. The error bars on the expectation values from moment optimization are obtained by propagating the Monte-Carlo uncertainties using the procedure described in Appendix B.

most $|c_{R,Q}|$ independent parameters, and is systematically improvable. Importantly, it can be optimized with the aid of Euclidean-Monte-Carlo-informed moment optimization, to exhibit accurate energies, nonlocal correlation functions, and local non-Gaussianity. The efficiency of this *Ansatz* for quantum simulation will be discussed in the next section.

V. QUANTUM-CIRCUIT REPRESENTATION

In this section, we encode the bosonic wave functions obtained in the previous sections into efficient quantum circuits. We consider a circuit design for discrete-variable, i.e., qubit-based, platforms, and discuss algorithms for core-state preparation and Gaussian-unitary implementation, as well as their associated complexities by analyzing the theoretical and algorithmic errors. In Appendix G, we further discuss opportunities for continuous-variable, i.e., bosonic, platforms.

A. Encoding bosonic states on a discrete-variable platform

To represent bosonic degrees of freedom using qubits, one must cut off the infinite-dimensional bosonic Hilbert space. Thus, in addition to thermodynamic and continuum-limit extrapolations, a discrete-variable quantum simulation

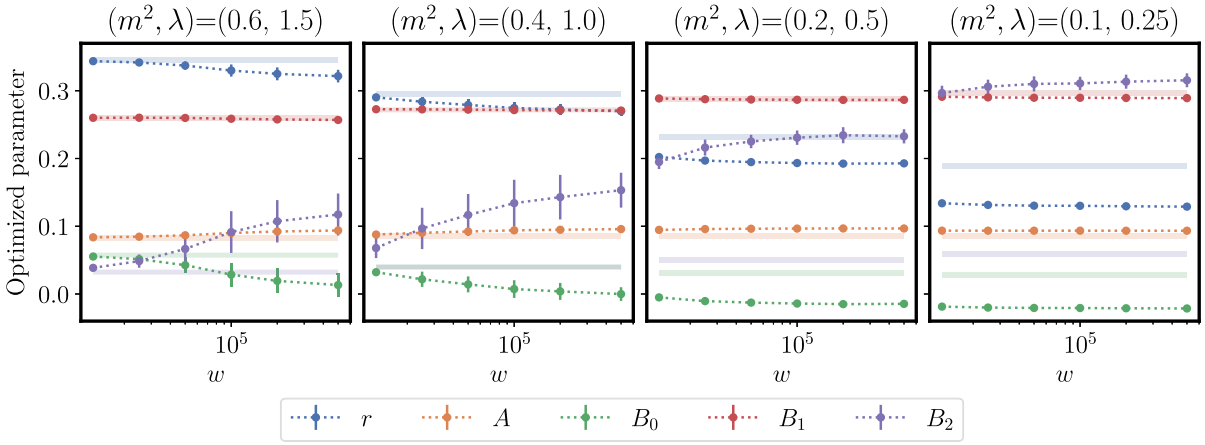


FIG. 14. Optimized values of the $(R, Q) = (2, 2)$ *Ansatz* parameters [with its core state polynomial given by Eq. (29) when all parameters except A, B_0, B_1, and B_2 are set to zero] plotted as a function of w for two-point function optimization. The dotted lines passing through the data points are to guide the eye and do not represent a fit. The error bars on the moment-optimized parameters are obtained by propagating the Monte-Carlo uncertainties using the procedure described in Appendix B. The thick light-colored lines show the parameter values for the minimum-energy *Ansatz*. These lines are thickened to improve visibility and do not imply error bars.

must first remove the dependence on this cutoff in observables. Several digitization and truncation schemes for bosons have been considered in the literature [38,44,81,83,102–105]. Here, we present one straightforward choice, and describe its quantum and classical computing costs. For the finite-rank *Ansätze* considered in this work, a natural choice is to encode bosonic Fock states in the qubits' computational basis states, and to introduce a cutoff in the total number of bosons or the number of bosons per mode. Such Fock-space Hamiltonian truncation has been employed for the two-dimensional Euclidean ϕ^4 theory in Ref. [106] to make predictions about the theory's spectral and critical properties, and for bosonic systems in Refs. [107,108] in the context of quantum simulations.

We assume a cutoff Λ on the number of bosons per mode. Thus, the truncated single-mode Fock space is defined as the span of the Fock states $|n\rangle_j$ with $n = 0, \ldots, \Lambda$, for each mode $j \in \{0, \ldots, N-1\}$. The full truncated multimode Hilbert space is the tensor product of these Hilbert spaces. Following Ref. [107], we introduce the ladder operators

$$\hat{a}_j^{\Lambda} \coloneqq \sum_{n=1}^{\Lambda} \sqrt{n}\, |n-1\rangle_j \langle n|_j, \tag{36}$$

$$\hat{a}_j^{\Lambda\dagger} \coloneqq \sum_{n=0}^{\Lambda-1} \sqrt{n+1}\, |n+1\rangle_j \langle n|_j, \tag{37}$$

on the truncated Fock space, with the commutation relation

$$[\hat{a}_j^{\Lambda}, \hat{a}_k^{\Lambda\dagger}] = \delta_{j,k} \left(1 - \frac{\Lambda+1}{\Lambda!} (\hat{a}_j^{\Lambda\dagger})^{\Lambda} (\hat{a}_j^{\Lambda})^{\Lambda}\right). \tag{38}$$

$(\cdot)^{\Lambda}$ denotes the Λth powers of the operator enclosed. The canonical bosonic commutation relation is recovered in the limit $\Lambda \to \infty$. In the following, when working with a single mode, the mode index j will be dropped for brevity.

Denote the binary ($q = \mathrm{b}$) or unary ($q = \mathrm{u}$) representation of n by $q^{\Lambda}(n)$. Both of these representations can be encoded in the computational basis of a register of $n_q(\Lambda)$ qubits given by $|q^{\Lambda}(n)\rangle$, where

$$n_q(\Lambda) = \begin{cases} \Lambda + 1 & \text{if } q = \mathrm{u}, \\ \lceil \log(\Lambda+1) \rceil & \text{if } q = \mathrm{b}. \end{cases} \tag{39}$$

Label the qubits in each register by $i \in \{0, 1, \ldots, n_q(\Lambda) - 1\}$. Further denote the ladder operators on the ith qubit as

$$\hat{L}_i^{\pm} \coloneqq \frac{1}{2} (\hat{X}_i \mp i \hat{Y}_i), \tag{40}$$

with Pauli operators $\hat{X}_i$ and $\hat{Y}_i$ acting on the ith qubit. Using these ladder operators, the mapping from Fock states to the qubits' computational basis states can be defined as

$$|n\rangle \mapsto |\mathrm{u}^{\Lambda}(n)\rangle = \hat{L}_n^{+} |0\rangle^{\otimes n_{\mathrm{u}}(\Lambda)} \tag{41}$$

for the unary representation, and

$$|n\rangle \mapsto |\mathrm{b}^{\Lambda}(n)\rangle = \sum_{i=0}^{n_{\mathrm{b}}(\Lambda)-1} (\hat{L}_i^{+})^{b_i} |0\rangle^{\otimes n_{\mathrm{b}}(\Lambda)} \tag{42}$$

for the binary representation. For the latter, we we have assumed the 0th qubit to represent the least significant digit and introduced the binary notation $n = \sum_{n=0}^{n_{\mathrm{b}}(\Lambda)-1} b_i 2^i$.

An N-mode Fock state $|\boldsymbol{n}\rangle$ can be encoded into a collection of N registers of the form in Eqs. (41) or (42) with $n_q(\Lambda)$ qubits each. We define

$$|\boldsymbol{n}\rangle \mapsto |q^\Lambda(\boldsymbol{n})\rangle \coloneqq |q^\Lambda(n_0)\rangle \otimes \ldots \otimes |q^\Lambda(n_{N-1})\rangle. \quad (43)$$

In the following, we describe the preparation of the (R, Q) *Ansatz* $|\psi\rangle_{R,Q} = \hat{U}_G|C\rangle_{R,Q}$, where $\hat{U}_G = \bigotimes_{j=0}^{N-1} \hat{S}_j(r)$. We assume that the qubit registers are initialized in the state $|0\rangle^{\otimes n_q(\Lambda)}$.

B. Core-state preparation

Since the core state has a finite support over the Fock space, it can be represented exactly in a qubit register provided that the cutoff is set to $\Lambda \geq R$. The core state is mapped to the qubit state

$$|C\rangle_{R,Q} = \bar{\sum_{\boldsymbol{n}}} c_{\boldsymbol{n}}|\boldsymbol{n}\rangle \mapsto |C\rangle_{R,Q}^{q,\Lambda} = \bar{\sum_{\boldsymbol{n}}} c_{\boldsymbol{n}}|q^\Lambda(\boldsymbol{n})\rangle, \quad (44)$$

where the notation $\bar{\sum}$ indicates that the summation is restricted such that it obtains the symmetric rank-R and Q-truncated core state described in the previous section.[7] Since the state is only nontrivial over the first $n_q(R)$ qubits in each register, we focus our attention to this part of the qubit register. Recall from Eq. (27) that the core state is mapped to a superposition of $N|c_{R,Q}|$ multiqubit computational basis states, the number of which is much smaller than the total Hilbert-space dimension of $2^{Nn_q(R)}$ (provided that $R, Q \ll N$). Thus, one can invoke sparse state-preparation methods such as the ones proposed in Refs. [58,109–113].

In particular, the method presented in Ref. [110] can prepare the core state using $O(N^2 n_q(R)|c_{R,Q}|)$ two-qubit entangling (CNOT) gates and $O(Nn_q(R) + N|c_{R,Q}|\log(N|c_{R,Q}|))$ single-qubit rotation gates. The key idea is to determine the circuit decomposition of a unitary transformation $\hat{\mathcal{U}}_{R,Q}^{q,\Lambda}$ which maps $|C\rangle_{R,Q}^{q,\Lambda}$ to the multiqubit computational basis state $|q^\Lambda(\boldsymbol{0})\rangle$, i.e., $|C\rangle_{R,Q}^{q,\Lambda} = \hat{\mathcal{U}}_{R,Q}^{q,\Lambda\dagger}|q^\Lambda(\boldsymbol{0})\rangle$. The transformation works by successively rotating two elements in the superposition $|C\rangle_{R,Q}^{q,\Lambda}$ into one. This operation is performed with suitable controls so that the merging of two computational basis states is not accompanied by introducing additional basis states in the superposition. Thus, by performing these merges $N|c_{R,Q}| - 1$ times, $|C\rangle_{R,Q}^{q,\Lambda}$ is mapped to a single computational basis state, which can be converted to $|q^\Lambda(\boldsymbol{0})\rangle$ using NOT gates. The decomposition of $\hat{\mathcal{U}}_{R,Q}^{q,\Lambda}$ into elementary quantum gates is obtained using a classical algorithm with an $O(N^3 n_q(R)|c_{R,Q}|^2\log(N|c_{R,Q}|))$ runtime. More details about this algorithm, together with an example, can be found in Appendix D.

Once the *Ansatz*, i.e., the values of R and Q as well as the system size N, are fixed, the circuit template that prepares the core state is known. The values of the couplings (m^2, λ) only affect the magnitudes of the rotation angles in the single-qubit rotation gates within this fixed template.

C. Gaussian unitary transformation

Once the core state $|C\rangle_{R,Q}$ is prepared, the next step is to the apply Fock-space-truncated squeezing operators [recall Eq. (28)]. This step introduces two kinds of error: truncation error from restriction to a finite Fock space, and error in the implementation of the resulting truncated operator.[8]

The core state $|C\rangle_{R,Q}$ has a bounded support over Fock space, i.e., the maximum number of bosons in any mode is the rank R. Thus, as already mentioned, it can be represented exactly in the truncated Fock space with a cutoff of R bosons per mode, which is in turn isomorphic to the multiqubit Hilbert space (or to a subspace of it in case of the unary map). In contrast, the Gaussian unitary transformation, $\hat{U}_G = \bigotimes_{j=0}^{N-1} \hat{S}_j(r)$, extends the support of the (R, Q) *Ansatz* $|\psi\rangle_{R,Q} = \hat{U}_G|C\rangle_{R,Q}$ to the entire Hilbert space; thus, it can only be represented approximately. Following Refs. [108,114], an approximate squeezing operator on the truncated bosonic Hilbert space can be obtained as follows:

$$\hat{S}(r) = e^{\frac{r}{2}(\hat{a}^{\dagger 2} - \hat{a}^2)} \mapsto \hat{S}^\Lambda(r) \coloneqq e^{\frac{r}{2}[(\hat{a}^{\Lambda\dagger})^2 - (\hat{a}^\Lambda)^2]} \coloneqq e^{-ir\hat{s}^\Lambda}. \quad (45)$$

Here, the original ladder operators in the squeezing Hamiltonian are replaced by their truncated counterparts, i.e., $\hat{a}^{\Lambda\dagger}$ and $\hat{a}^\Lambda$ defined in Eqs. (36)–(37). Furthermore, we have defined

$$\hat{s}^\Lambda \coloneqq \sum_{n=0}^{\Lambda-2} \hat{T}_n, \quad (46)$$

with

$$\hat{T}_n \coloneqq \frac{i\ell_n}{2}(|n+2\rangle\langle n| - |n\rangle\langle n+2|) \quad (47)$$

and $\ell_n \coloneqq \sqrt{(n+1)(n+2)}$, which can be deduced upon replacing definitions of the truncated ladder operators in the form $(\hat{a}^{\Lambda\dagger})^2 - (\hat{a}^\Lambda)^2$.

The application of this approximate squeezing operator to the core state results in the approximate (R, Q) state $|\psi^\Lambda\rangle_{R,Q} \coloneqq \bigotimes_{j=1}^{N-1} \hat{S}_j^\Lambda(r)|C\rangle_{R,Q}$, which has support only on the truncated Hilbert space. As shown in Appendix E 1, a

[7] This preparation method, nonetheless, works for arbitrary core states.

[8] Experimental errors arising from environmental noise, imperfections in gate implementations, etc. will not be considered here.

suitable value of the cutoff, Λ, can be determined by upper-bounding the distance between $|\psi\rangle_{R,Q}$ and $|\psi^{\Lambda}\rangle_{R,Q}$

$$\|(\otimes_{j=0}^{N-1}\hat{S}_j(r)-\hat{\otimes}_{j=0}^{N-1}S_j^{\Lambda}(r))|C\rangle_{R,Q}\| \leq \epsilon_{\text{trunc}}^{(N)}(r,R,\Lambda,Q), \tag{48}$$

where

$$\epsilon_{\text{trunc}}^{(N)}(r,R,\Lambda,Q) \coloneqq N\sqrt{N|c_{R,Q}|}[G(r,R,\Lambda)+g^{\text{leak}}(r,R,\Lambda)]. \tag{49}$$

Here, the term proportional to $g^{\text{leak}}(r,R,\Lambda)$ bounds the error resulting from discarding the weight of $|\psi\rangle_{R,Q}$ outside the truncated Hilbert space, with

$$g^{\text{leak}}(r,R,\Lambda) \coloneqq \text{argmax}_{n_1\leq R}\left[\sum_{n_2=\Lambda+1}^{\infty}|\langle n_2|\hat{S}(r)|n_1\rangle|^2\right]^{1/2}. \tag{50}$$

For the small values of the squeezing parameter r in this work, the squeezing operator $\hat{S}(r)$ distributes most of the Fock-basis amplitude among the lower Fock states, and this limits the value of the above leakage parameter; see Fig. 32 in Appendix E 1. The term proportional to $G(r,R,\Lambda)$ bounds the discrepancy between the matrix elements of the exact and approximate squeezing operators within the truncated Hilbert space, and the expression for $G(r,R,\Lambda)$ can be found in Appendix E 1. Once a truncation-error tolerance is set, Eq. (49) determines the smallest value of Λ that achieves such an error tolerance.

It is important to note that the above analysis focuses only on the accuracy of the initial-state representation. However, as this state evolves in time, its support on higher Fock states is expected to grow, and thus, eventually larger cutoff values are needed. As a result, additional qubits need to be appended to the state register to account for this expansion of the state's domain within the Hilbert space. One could use the strategy of Refs. [115,116] to achieve an analytical bound on the error accumulated throughout the evolution due to constraining the evolved state to a truncated Hilbert space.

The implementation error depends on the capabilities of the device and the computation mode. The most common mode of quantum computation, which we used for core-state preparation as well, is the digital mode. This mode relies on applying a discrete set of universal single and two-qubit gates, hence requiring Trotterization of, or other digitization schemes for, the squeezing operator. The squeezing operator in Eq. (45) can be expressed as

$$\hat{S}^{\Lambda}(r) = e^{-ir(\hat{s}_0^{\Lambda}+\hat{s}_2^{\Lambda})}e^{-ir(\hat{s}_1^{\Lambda}+\hat{s}_3^{\Lambda})}, \tag{51}$$

where

$$s_m^{\Lambda} \coloneqq \sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor}\hat{T}_{4n+m}, \tag{52}$$

with $m\in\{0,1,2,3\}$.[9] With this splitting of the squeezing operator, we define the first-order Trotter approximation to $\hat{S}^{\Lambda}$ as

$$\hat{S}^{\Lambda,K}(r) \coloneqq \left(e^{-i\frac{r}{K}\hat{s}_0^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_2^{\Lambda}}\right)^K\left(e^{-i\frac{r}{K}\hat{s}_1^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_3^{\Lambda}}\right)^K. \tag{53}$$

As shown in Appendix E 2, this results in the following bound on the distance between the state $|\psi^{\Lambda}\rangle_{R,Q}$ defined previously, and the state, $|\psi^{\Lambda,K}\rangle_{R,Q} \coloneqq \otimes_{j=0}^{N-1}\hat{S}_j^{\Lambda,K}(r)|C\rangle_{R,Q}$,

$$\|(\otimes_{j=0}^{N-1}\hat{S}_j^{\Lambda}(r)-\hat{\otimes}_{j=0}^{N-1}S_j^{\Lambda,K}(r))|C\rangle_{R,Q}\| \leq \epsilon_{\text{trott}}^{(N)}(r,\Lambda,K), \tag{54}$$

where

$$\epsilon_{\text{trott}}^{(N)}(r,\Lambda,K) \coloneqq \frac{Nr^2}{2K}(\|[\hat{s}_0^{\Lambda},\hat{s}_2^{\Lambda}]\|+\|[\hat{s}_1^{\Lambda},\hat{s}_3^{\Lambda}]\|). \tag{55}$$

Using triangle inequality, a looser but analytical bound $\tilde{\epsilon}_{\text{trott}}^{(N)}(r,\Lambda,K) > \epsilon_{\text{trott}}^{(N)}(r,\Lambda,K)$ can also be obtained

$$\tilde{\epsilon}_{\text{trott}}^{(N)}(r,\Lambda,K) \coloneqq \frac{Nr^2\beta(\Lambda)}{2K}, \tag{56}$$

where

$$\beta(\Lambda) \coloneqq \frac{1}{2}\sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor}\ell_{4n+m}(\ell_{4n+m-2}+\ell_{4n+m+2}), \tag{57}$$

resulting in an $O(\frac{Nr^2\Lambda^3}{K})$ error. Unlike the truncation error bound, this bound is obtained using a *state-independent* first-order product-formula error bound [117], which is quite loose.

It is useful to enumerate the minimum value of the cutoff, Λ, and the number of Trotter layers, K, which guarantee a minimum fidelity F_0 between $|\psi\rangle_{R,Q}$ and $|\psi^{\Lambda,K}\rangle_{R,Q}$ based on these bounds (we use the tighter commutator-norm bound for the Trotter error). Assuming an equal value for the maximum permissible error resulting from truncation and Trotterization, we list these minimum Λ and K values in Table II for the minimum-energy $(R,Q)=(4,2)$ *Ansatz* studied previously.

In the above example, the number of Trotter layers needed to guarantee high fidelity in the implementation of the squeezing operator is larger for larger values of (m^2,λ).

[9] It is assumed that $s_m^{\Lambda}=0$ when $\lfloor\frac{\Lambda-2-m}{4}\rfloor<0$.

TABLE II. Minimum cutoff Λ and number of Trotter layers K which guarantee a minimum fidelity, F_0, for the Trotterized and truncated multimode product squeezing operator with $N = 10$ and various mass and coupling values (m^2, λ). r is the squeezing parameter obtained upon minimizing the energy of the $(R, Q) = (4, 2)$ *Ansatz*.

		$F_0 = 0.9$		$F_0 = 0.95$	
(m^2, λ)	r	Λ	K	Λ	K
(0.6, 1.5)	0.348	33	2375	35	3876
(0.4, 1)	0.300	27	1102	27	1569
(0.2, 0.5)	0.172	15	87	15	123
(0.1, 0.25)	0.155	14	59	15	100

This is because of the larger value of the associated squeezing parameter, which in turn leads to a higher value for the Fock-space cutoff. The value of the optimized squeezing parameter decreases as $(m^2, \lambda) \to (0, 0)$, resulting in the much smaller number of required Trotter layers in this regime. One should, however, note that the accuracy with which a fixed (R, Q) *Ansatz* represents the ground state declines as $(m^2, \lambda) \to (0, 0)$ (where the gap becomes smaller while the correlations grow stronger). This feature can be captured, for instance, by a growing value of δ_E. To maintain a constant value of δ_E as one takes $(m^2, \lambda) \to (0, 0)$, the value of both R and Q must be increased. Consequently, the complexity of implementing the core state would increase. The complexity of implementing the squeezing operator, on the other hand, would depend upon the values of R, Q, and the squeezing parameter r, and could be determined empirically.

One can now faithfully map each of the single-mode Trotterized squeezing operators on the truncated Hilbert space to an operator acting on a register of $n_q(\Lambda)$ qubits

$$\hat{S}^{q,\Lambda,K}(r) \coloneqq \left(e^{-i\frac{r}{K}\hat{s}_0^{q,\Lambda}} e^{-i\frac{r}{K}\hat{s}_2^{q,\Lambda}}\right)^K \left(e^{-i\frac{r}{K}\hat{s}_1^{q,\Lambda}} e^{-i\frac{r}{K}\hat{s}_3^{q,\Lambda}}\right)^K. \tag{58}$$

Here,

$$s_m^{q,\Lambda} \coloneqq \sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor} \hat{T}_{4n+m}^{q,\Lambda}, \tag{59}$$

with

$$\hat{T}_n^{q,\Lambda} \coloneqq \frac{i\ell_n}{2}\left(|q^\Lambda(n+2)\rangle\langle q^\Lambda(n)| - |q^\Lambda(n)\rangle\langle q^\Lambda(n+2)|\right). \tag{60}$$

In the unary representation, to implement each Trotter layer $e^{-i\frac{r}{K}\hat{s}_m^{u,\Lambda}}$ with $m = 0, 1, 2, 3$, one first notes that

$$\begin{aligned} e^{-i\frac{r}{K}\hat{s}_m^{u,\Lambda}} &= \prod_{n=0}^{\lfloor(\Lambda-2-m)/4\rfloor} e^{-i\frac{r}{K}\hat{T}_{4n+m}^{u,\Lambda}} \\ &= \prod_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor} e^{\frac{r\ell_{4n+m}}{2K}\left(|u^\Lambda(4n+m+2)\rangle\langle u^\Lambda(4n+m)| - \text{H.c.}\right)} \\ &= \prod_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor} \left(e^{i\frac{ir\ell_{4n+m}}{4K}\hat{X}_{4n+m+2}\hat{Y}_{4n+m}} e^{-i\frac{r\ell_{4n+m}}{4K}\hat{Y}_{4n+m+2}\hat{X}_{4n+m}}\right). \end{aligned} \tag{61}$$

Here, we have used Eq. (41) for the unary representation of the qubitized bosonic states. Thus, for the single-mode operator, each Trotter layer can be implemented with $O(\Lambda)$ CNOT and single-qubit rotation gates. For the multimode tensor-product operator, this gate count becomes $O(N\Lambda)$. Switching to the binary map, and recalling the representation of the qubitized bosonic states in Eq. (42), one readily observes a gate complexity $O(\Lambda \log \Lambda)$ for each Trotter layer of the single-mode squeezing operator, following steps similar to those outlined for the unary map. This implementation, however, introduces additional Trotter error and, thus is less preferred.

Alternatively, we rely upon the singular-value-decomposition (SVD) approach of Ref. [118] which works both for the binary and unary representations, as detailed in Appendix F 1. Importantly, this approach avoids any additional Trotter error beyond that from Eq. (53). SVD enables the implementation of the single-mode squeezing operator in the binary representation using $O(\Lambda \log \Lambda)$ CNOT and single-qubit rotation gates. However, the SVD representation for the single-mode operator in the unary representation relies upon $O(\Lambda 2^\Lambda)$ CNOT and single-qubit rotation gates. Thus, the direct implementation discussed in the previous paragraph is favored for the unary representation.

With a hybrid analog-digital platform, such as that in Ref. [119], the squeezing operator $\hat{S}^{q,\Lambda}(r)$ can be implemented without incurring a Trotter error. The operator $\hat{s}^{q,\Lambda}$ in the exponent of Eq. (45) can be put into the form of a nearest-neighbor XY Hamiltonian by rearranging the qubits in the unary representation, and upon performing a local basis change. Such a Hamiltonian can be natively generated on this platform, upon careful calibration and basis choice [119]. Details of such a scheme are provided in Appendix F 2. The accuracy of this implementation is controlled by spurious next-to-nearest-neighbor and farther terms in the effective spin Hamiltonian implemented on the device, characterization of which is beyond the scope of our work. One may still need a fully digital scheme to implement the core state on such a platform.

Even more powerful is the use of continuous-variable quantum platforms, e.g., those taking advantage of bosonic degrees of freedom. While the discrete-variable approach is more standard and is commonly practiced, the

continuous-variable approach is more natural for bosonic theories and may ultimately present practical benefits. As shown in Appendix G, our single-mode finite-rank *Ansatz* for the ground state of an interacting scalar field theory can be implemented using standard operations in bosonic platforms, namely single-mode Gaussian (squeezing $\hat{S}(\xi) = e^{\frac{1}{2}[\xi(\hat{a}^\dagger)^2 - \xi^*\hat{a}^2]}$ and displacement $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$) operations, and single-boson addition ($\hat{a}_j^\dagger$). However, more work needs to be done to arrive at a bosonic preparation recipe for the multimode (R, Q) *Ansatz*; see Appendix G for more details.

VI. DISCUSSION AND OUTLOOK

In this work, we demonstrated a classical Euclidean-Monte-Carlo-informed method for determining quantum circuits that prepare states close to the ground state of the $(1+1)$D ϕ^4 theory. The crucial components for translating Euclidean ground-state information to Minkowski wave functions are the finite stellar-rank *Ansatz*, which we specialize to the (R, Q) *Ansatz*, and a Euclidean-Monte-Carlo-informed moment-optimization procedure. The (R, Q) *Ansatz* is an N-mode bosonic *Ansatz* consisting of a tensor-product of single-mode squeezing operations acting on a superposition of Fock states, which is referred to as the (R, Q) core state. The rank, R, specifies the maximum total number of bosons while Q specifies the largest span of occupied modes featuring in this superposition. Roughly, the rank R is associated with the level of non-Gaussianity in the *Ansatz*, while the maximum span Q is associated with its level of locality. Given a fixed (R, Q) *Ansatz*, we showed how the moment-optimization procedure can systematically improve the *Ansatz*'s estimates of the ground-state's two-point correlation functions as well as its non-Gaussianity. This procedure works by augmenting the familiar variational energy minimization by penalizing deviations in selected target sets of ground-state moments (whose values are sourced from Euclidean path-integral Monte-Carlo computations).

The extent to which various target-moment sets can be optimized depends upon the structure and expressiveness of the *Ansatz*, together with the properties of the theory itself (such as its spectral gap). Based on this, we conclude that the (R, Q) *Ansatz* possesses the following properties for the tested range of couplings and masses:

(i) *Classical tractability:* We showed that expectation values with respect to the (R, Q) *Ansatz* can be computed in $O(|c_{R,Q}|^2 Q^2)$ classical-computing time, with $|c_{R,Q}|$ defined in Eq. (27). Away from the continuum limit, this cost is independent of the system size N. However, close to the continuum limit, it is expected that greater values of R and Q would be needed to reproduce the diverging correlation length, and in fact, R and Q may no longer be small compared to N, increasing the classical-computing cost.

(ii) *Circuit translatability:* For a discrete-variable quantum simulator, the core state can be translated into a quantum circuit with classical time complexity $O(N^3 R|c_{R,Q}|^2 \log(N|c_{R,Q}|))$ using a sparse quantum-state preparation algorithm (and assuming a unary mapping of the bosons). The circuit for the squeezing operator, on the other hand, can be determined using $O(1)$ classical computing time (for the direct Trotterization recipe with the unary map).

(iii) *Circuit efficiency:* With a near-term implementation in mind, we characterized circuits based on their CNOT and single-qubit gate counts. The core-state circuit consists of $O(N^2 R|c_{R,Q}|)$ CNOT gates while the squeezing circuit consists of $O(N\Lambda)$ CNOT gates per Trotter layer (assuming a unary encoding of bosons). Here, Λ is the maximum number of bosons in each mode in the truncated Hilbert space. While the core state is implemented exactly, the qubitized squeezing operator is subject to two sources of error, namely, the truncation of the bosonic Hilbert space and Trotterization. We show how suitable values for the cutoff and the number of Trotter layers can be chosen to meet specified fidelity goals.

We ascribe the above features (i.e., classical tractability, circuit translatability, and circuit efficiency) to an *Ansatz* suitable for the purpose of classical determination of quantum circuits that prepare ground states.

A comprehensive field-theory study necessitates a detailed examination of the thermodynamic and continuum limits. As these limits are approached, the spectral gap closes and correlation lengths diverge, imposing significant demands on the *Ansatz*, and implying system-dependent scalings for both R and Q. Critical slowing down also poses challenges for sourcing ground-state correlation functions from Monte-Carlo simulations. However, since practical computations often operate at points slightly removed from the continuum limit, extrapolations using finite-lattice-spacing results remain a feasible strategy. Furthermore, as the *Ansatz* becomes less classically tractable toward the continuum limit, one could potentially switch to quantum-computing *Ansatz* moments, as in more conventional variational-quantum-eigensolver settings. In particular, it would be interesting to use ADAPT-VQE methods to systematically grow the (R, Q) *Ansatz* [60,70].

Optimizing the choice of weights in the moment-optimization loss function is another area needing further study. In our current work, these weights were selected through experimentation. Furthermore, we chose the same value of weight accompanying the squared discrepancy in each target moment. One may want to choose moment-dependent weights for a more fine-tuned expression of ground-state correlations. This strategy would also help control errors because higher-order moments have larger associated statistical PIMC errors, and hence these could be assigned smaller weights. Thus, developing a principled

approach for selecting moment-dependent weights could enhance the effectiveness of the optimization process. Since the choice of weight is an optimization problem in itself, it would be crucial to quantify how the complexity of this optimization varies with the number and type of moments in the target set. Moreover, since ground-state correlation functions are sourced from Monte-Carlo simulations, the moment-optimized parameters inherently carry statistical errors. Understanding the implications of these errors for quantum simulations is crucial. Specifically, one may need to perform multiple runs of the quantum simulation to propagate these errors to subsequent real-time quantities.

Our introduction of the moment-optimization procedure was aimed at obtaining *Ansätze* that are more dynamics-aware. A deeper and more comprehensive understanding of the relationship between dynamics and target-moment sets, utilizing both analytical and numerical tools, remains an open question. Studying dynamics using these moment-optimized *Ansätze* as initial states would be particularly exciting. Interestingly, the construction of the (R, Q) *Ansatz* is reminiscent of matrix product states. Making this relationship more concrete could facilitate the development of more efficient *Ansatz*-optimization methods, or even a density matrix renormalization group (DMRG) analog. Additionally, finite-rank *Ansätze* hold relevance for the far-term fault-tolerant era. The reason is that the efficiency of near-optimal ground-state preparation strategies [51,52] is often related to the overlap of an initial trial state with the ground state. Our method, therefore, could be used to prepare such a suitable trial state, and subsequent quantum computation could evolve this state even closer to the ground state.

Another approach to leveraging Euclidean PIMC information has been proposed in Refs. [120,121]: One stochastically samples the density matrix classically, and then passes the basis states with proper weights to the quantum computer for subsequent computation of real-time expectation values. This method, therefore, avoids preparing the state on the quantum computer but is inherently stochastic. Our approach, on the other hand, constrains an *Ansatz* wave function with the aid of static correlation functions obtained from Euclidean PIMC. It would be worthwhile to explore other alternative routes to leveraging Euclidean PIMC results in the quantum computation of quantum field theories.

There exist quantum-state preparation strategies that rely upon the knowledge of the exact ground-state wave function. For example, Refs. [45–47,81] introduce a quantum-circuit *Ansatz* for preparing the scalar-field-theory ground state. The locality of the digitized free lattice-scalar-field theory enables a systematic truncation of the circuit elements. The circuit elements for preparing large-scale ground states can then be determined using classically computed ground states in small spatial volumes. It would be interesting to apply the results of Refs. [45–47,81] to the interacting wave functions computed in our work. Notably, the concept of fixed-point circuits introduced in Ref. [47] could be used to trade off system-size dependence in our translated quantum circuits with correlation-length dependence. More generally, it would be interesting to see how the structure imposed by the (R, Q) *Ansatz* on digital quantum circuits compares with the structures obtained in these works.

Since the (R, Q) *Ansatz* is primarily designed to enable efficient computation of expectation values using classical computers, its translatability to bosonic quantum simulators is compromised. Investigating alternative modifications to the finite-rank *Ansatz* that retain classical tractability, and can also be efficiently prepared on bosonic quantum simulators (hence bypassing truncation errors), is an intriguing direction for future work. Alternatively, one could develop variational-quantum-eigensolver methods based upon the (R, Q) *Ansatz*.[10] Future work can explore how our *Ansatz*'s features, as well as the moment-optimization principle, could be applied to existing VQE setups and to tensor networks.

In this work, we adopted a single measure, namely the ϕ^{2n}-moment ratio, to assess deviations from Gaussianity. This quantity can be easily computed using only the $\langle\hat{\phi}^{2n}\rangle$ and $\langle\hat{\phi}^{2}\rangle$ values, and cleanly isolates shape effects from the overall scale. One can also study alternative probes, such as connected cumulants, their variance-normalized versions (e.g., binder cumulants), or variants thereof as introduced in Ref. [95]. Furthermore, it may be interesting to study non-Gaussianity in other operator bases, and for nonlocal correlation functions as well. The choice of the probe may ultimately be informed by the type of phenomenon to be studied. For example, non-Gaussianity in scalar fields is physically relevant in the study of cosmic microwave background (CMB) anisotropies arising from slow-roll inflation driven by scalar fields [123,124]. Future work could explore how the (R, Q) *Ansatz* fares on measures of non-Gaussianity in these cosmological settings.

Extending our methods to higher spatial dimensions, gauge theories, and fermionic degrees of freedom represents another set of clear directions for future exploration. For example, it will be interesting to investigate if the finite-rank bosonic *Ansatz* of this work has any relevance to bosonic states in gauge theories, and what extensions of this *Ansatz* best describe fermionic and coupled fermionic-bosonic ground states.

ACKNOWLEDGMENTS

The authors would like to thank Jacob Bringewatt, Mohammad Hafezi, Chung-Chun Hsieh, Matthew Johnson, Saurabh Kadam, En-Jui Kuo, Yannick Meurice, Christine Muschik, David Roberts, Sarah Shandera, Brayden Ware, and Jing-Chen Zhang for discussions about various

[10]Recently, a VQE computation using a multimode finite-stellar-rank *Ansatz* for the Hubbard model was demonstrated in Ref. [122].

aspects of this work. N. G., C. D. W., and Z. D. were supported by the National Science Foundation (NSF) Quantum Leap Challenge Institutes (QLCI) (Award No. OMA-2120757). Z. D. and N. G. further acknowledge funding by the Department of Energy (DOE), Office of Science, Early Career Award (Award No. DESC0020271), as well as by the Department of Physics; Maryland Center for Fundamental Physics; and College of Computer, Mathematical, and Natural Sciences at the University of Maryland, College Park. Z. D. further acknowledges support from the U.S. DOE, Office of Science, Office of Advanced Scientific Computing Research (ASCR), program in Accelerated Research in Quantum Computing, Fundamental Algorithmic Research toward Quantum Utility (FAR-Qu). Z. D. and N. G. are also grateful for the hospitality of Perimeter Institute where part of this work was carried out. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science, and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. Z. D. was also supported in part by the Simons Foundation through the Simons Foundation Emmy Noether Fellows Program at Perimeter Institute. Z. D. is further grateful for the hospitality of Nora Brambilla, and of the Excellence Cluster ORIGINS at the Technical University of Munich, where part of this work was carried out. The research at ORIGINS is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (No. EXC-2094–390783311). C. D. W. further acknowledges support from the U.S. DOE, Office of Science, Office of Advanced Scientific Computing Research (ASCR) Quantum Computing Application Teams program, under fieldwork proposal number ERKJ347, as well as DOE Quantum Systems Accelerator program (Award No. DE-AC02-05CH11231), AFOSR MURI FA9550-22-1-0339, ARO Grant No. W911NF-23-1-0242, and ARO Grant No. W911NF-23-1-0258; the work was completed while C. D. W. held an NRC Research Associateship award at the United States Naval Research Laboratory.

DATA AVAILABILITY

The data that support the findings of this article are openly available [125].

APPENDIX A: ENERGY MINIMIZATION VERSUS FIDELITY MAXIMIZATION

In Sec. III, we argued that the minimum-energy and maximum fidelity *Ansätze* may not coincide. To make this point more transparent, we provide in this appendix a simple example for which both the energy and fidelity are analytically computable.

Consider the following squeezed harmonic-oscillator Hamiltonian. (Section III introduces some of the notation used in this Appendix and Ref. [126] discusses the squeezed harmonic-oscillator states.)

$$\begin{aligned}\hat{H}_S &:= \hat{S}(r)\hat{a}^\dagger\hat{a}\,\hat{S}(r)^\dagger \\ &= \cosh(2r)\hat{a}^\dagger\hat{a} - \frac{1}{2}\sinh(2r)((\hat{a}^\dagger)^2 + \hat{a}^2) + \sinh^2(r),\end{aligned} \tag{A1}$$

where in the second line, we have used the relation $\hat{S}(r)\hat{a}\,\hat{S}(r)^\dagger = \cosh(r)\hat{a} - \sinh(r)\hat{a}^\dagger$. This relation can be derived using the Baker-Campbell-Hausdorff formula, and the definition of the squeezing operator $\hat{S}(r) := e^{\frac{r}{2}(\hat{a}^{\dagger 2} - \hat{a}^2)}$, with $r \in \mathbb{R}$. The ground state of this Hamiltonian is given by

$$|\Omega\rangle = \hat{S}(r)|0\rangle = \frac{1}{\sqrt{\cosh(r)}}\sum_{n=0}^{\infty}\tanh(r)^n\frac{\sqrt{(2n)!}}{2^n n!}|2n\rangle. \tag{A2}$$

The above expansion can be obtained by noting that $\hat{S}(r)\hat{a}\,\hat{S}(r)^\dagger\hat{S}(r)|0\rangle = (\cosh(r)\hat{a} - \sinh(r)\hat{a}^\dagger)\hat{S}(r)|0\rangle = 0$. By inserting the expansion $\hat{S}(r)|0\rangle = \sum_{n=0}^{\infty} c_n|n\rangle$, the resulting recursive relation between the coefficients c_n gives the desired result. The ground-state energy is $\langle\Omega|\hat{H}_S|\Omega\rangle = 0$, and the spectral gap is $\Delta E = 1$.

Let us now take the rank-2 symmetric core state

$$|C\rangle := c_0|0\rangle + c_2|2\rangle; \qquad c_0, c_2 \in \mathbb{R}; \qquad c_0^2 + c_2^2 = 1, \tag{A3}$$

to be the *Ansatz* for the ground state. This leads to the energy and fidelity loss functions

$$\begin{aligned}\langle C|\hat{H}|C\rangle &= 2c_2^2\cosh(2r) - \sqrt{2}c_0c_2\sinh(2r) + \sinh^2(r), \\ |\langle C|\Omega\rangle|^2 &= \frac{1}{\cosh(r)}\left[c_0 + \frac{c_2}{\sqrt{2}}\tanh(r)\right]^2.\end{aligned} \tag{A4}$$

Figure 15 depicts the result of minimizing (maximizing) the energy (fidelity) loss functions. When the value of r is small, $|\Omega\rangle$ has substantial weight on the states $|0\rangle$ and $|2\rangle$ (as opposed to higher occupation states). Thus, in this regime, the *Ansatz* is powerful/expressive enough to yield small values of energy together with high values of fidelity. However, as r becomes larger, the minimum-energy and maximum-fidelity *Ansätze* diverge. It is in this regime that the optimal values of the coefficients are not well constrained, and which optimization to choose may need to be informed by the subsequent use of the *Ansatz*. In practice, computing fidelity quickly becomes intractable due to the large Hilbert-space sizes associated with systems of practical interest. In such cases, moment matching allows one to access such alternative choices for optimal parameters based on the knowledge of moments.

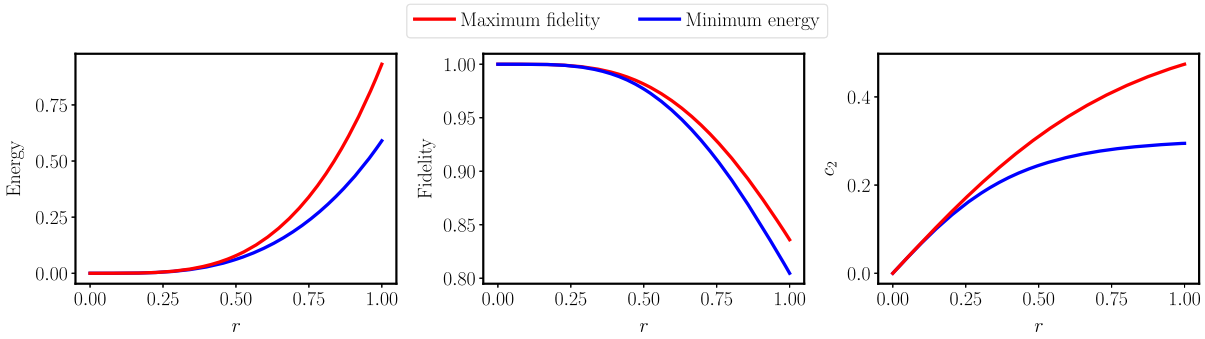


FIG. 15. Behavior of the minimum-energy and maximum-fidelity rank-2 *Ansatz* for the ground state of the squeezed Hamiltonian in Eq. (A1). The exact value of ground-state energy is zero while the spectral gap is 1.

APPENDIX B: EUCLIDEAN PATH-INTEGRAL MONTE CARLO

This appendix details the Euclidean path-integral Monte-Carlo computations used in Sec. IV. Consider the dimensionless lattice Hamiltonian describing the lattice $\hat{\phi}^4$ theory in $(1+1)$D

$$\hat{H} = \sum_{j=0}^{N-1} \left[\frac{\hat{\pi}_j^2}{2} + \frac{(\hat{\phi}_{j+1} - \hat{\phi}_j)^2}{2} + \frac{1}{2} m^2 \hat{\phi}_j^2 + \frac{\lambda}{4} \hat{\phi}_j^4 \right]. \quad \text{(B1)}$$

We introduce the compact notation $\vec{\phi} \equiv (\phi_0, \phi_1, \ldots, \phi_{N-1})$ and $\vec{\pi} \equiv (\pi_0, \pi_1, \ldots, \pi_{N-1})$ for the field values on a spatial slice, and split the Hamiltonian into kinetic and potential terms $\hat{H} = K(\hat{\vec{\pi}}) + V(\hat{\vec{\phi}})$. The partition function $Z_T := \mathrm{Tr}(e^{-T\hat{H}})$ can be approximated with the second-order Trotter decomposition $Z_{T,\theta}$ with (dimensionless) inverse temperature T,

$$\begin{aligned} Z_{T,\theta} &:= \int \mathcal{D}\vec{\phi}_0 \langle \vec{\phi}_0 | (e^{-\theta V/2} e^{-\theta K} e^{-\theta V/2})^M | \vec{\phi}_0 \rangle \\ &= \int \mathcal{D}\vec{\phi}_0 \mathcal{D}\vec{\phi}_1 \cdots \mathcal{D}\vec{\phi}_{M-1} e^{-\frac{\theta}{2} V(\vec{\phi}_0)} \langle \vec{\phi}_0 | e^{-\theta K} | \vec{\phi}_1 \rangle e^{-\theta V(\vec{\phi}_1)} \ldots \langle \vec{\phi}_{M-1} | e^{-\theta K} | \vec{\phi}_0 \rangle e^{-\frac{\theta}{2} V(\vec{\phi}_0)} \\ &= (2\pi\theta)^{-NM/2} \int \mathcal{D}\vec{\phi}_0 \mathcal{D}\vec{\phi}_1 \cdots \mathcal{D}\vec{\phi}_{M-1} e^{-S_{T,\theta}[\vec{\phi}_0, \ldots, \vec{\phi}_{M-1}]}. \end{aligned} \quad \text{(B2)}$$

Here, $\theta := T/M$ is the imaginary-time lattice spacing expressed in units of the spatial lattice spacing a, $\vec{\phi}_t := (\phi_{0,t}, \phi_{1,t}, \ldots, \phi_{N-1,t})$ denotes a spatial field vector at the imaginary time $\tau := t\theta$ with $t \in \{0, 1, \ldots, M-1\}$, and $\mathcal{D}\vec{\phi}_t := \Pi_{j=0}^{N-1} (d\phi_{j,t})$. Finally, the action $S_{T,\theta}$ is given by

$$S_{T,\theta}[\vec{\phi}_0, \ldots, \vec{\phi}_{M-1}] \equiv S_{T,\theta}(\{\phi_{j,t}\}) = \sum_{t=0}^{M-1} \sum_{j=0}^{N-1} \frac{\theta}{2} \left[m^2 \phi_{j,t}^2 + \frac{\lambda}{2} \phi_{j,t}^4 + \frac{(\phi_{j,t+1} - \phi_{j,t})^2}{\theta^2} + (\phi_{j+1,t} - \phi_{j,t})^2 \right]. \quad \text{(B3)}$$

To compute expectation values, one obtains sample configurations $\{\phi_{j,t}^{(\alpha)}\}_{T,\theta}$ with $\alpha \in \{1, 2, \ldots, N_{\text{samples}}\}$, with respect to the weight $e^{-S_{T,\theta}(\{\phi_{j,t}\})}$. The sample autocorrelation is estimated using the $(j,t) = (0,0)$ component of the samples as follows:

$$\mathrm{AC}(\Delta\alpha) := \frac{N_{\text{samples}}}{N_{\text{samples}} - \Delta\alpha} \frac{\sum_{\alpha=1}^{N_{\text{samples}} - \Delta\alpha} (\phi_{0,0}^{(\alpha)} \phi_{0,0}^{(\alpha+\Delta\alpha)})}{\sum_{\alpha=1}^{N_{\text{samples}}} (\phi_{0,0}^{(\alpha)})^2} \quad \text{(B4)}$$

Using this, the integrated autocorrelation time is estimated as follows:

$$\mathrm{IAC} := \frac{1}{2} + \sum_{\Delta\alpha=0}^{100} \mathrm{AC}(\Delta\alpha). \quad \text{(B5)}$$

The sum above is not taken all the way to N_{samples} because the autocorrelation estimates tend to be dominated by noise at large values of $\Delta\alpha$; we verify that the autocorrelation decays sufficiently close to zero for $\Delta\alpha < 100$. The statistical independence of the samples $\{\phi_{j,t}^{(\alpha)}\}_{T,\theta}$ is ensured by arranging their integrated autocorrelation time to be close to 1. This original ensemble of configurations is then resampled to

obtain $N_{\text{bootstrap}}$ sets of samples $\{\phi_{j,t}^{I_b(\alpha)}\}_{T,\theta}$ with $b \in \{1, 2, \ldots, N_{\text{bootstrap}}\}$. The values of $I_b(\alpha)$ are chosen uniformly at random from the set $\{1, \ldots, N_{\text{samples}}\}$. For each computation, the value of $N_{\text{bootstrap}}$ is chosen to ensure the stabilization of the bootstrap mean and standard error within a reasonable threshold. We also verify the normality of the bootstrap means by checking their skewness and kurtosis.

The field operators on a spatial slice are denoted as $\hat{\vec{\phi}} \coloneqq (\hat{\phi}_0, \ldots, \hat{\phi}_{N-1})$. Thus, the operator at Euclidean time τ is obtained as $\hat{\vec{\phi}}(\tau = t\theta) \equiv \hat{\vec{\phi}}_t = e^{\tau \hat{H}} \hat{\vec{\phi}} e^{-\tau \hat{H}}$. The ground-state expectation value of an observable $\mathcal{O}(\hat{\vec{\phi}}_{t_1}, \ldots, \hat{\vec{\phi}}_{t_k})$ can be expressed as

$$\begin{aligned}\langle \Omega | \mathcal{O}(\hat{\vec{\phi}}_{t_1}, \ldots, \hat{\vec{\phi}}_{t_k}) | \Omega \rangle &= \lim_{T\to\infty} \lim_{\substack{\theta\to 0 \\ M\theta = T}} \frac{\int \mathcal{D}\vec{\phi}_0 \mathcal{D}\vec{\phi}_1 \cdots \mathcal{D}\vec{\phi}_{M-1} \mathcal{O}(\vec{\phi}_{t_1}, \ldots, \vec{\phi}_{t_k}) e^{-S_{T,\theta}[\vec{\phi}_0, \ldots, \vec{\phi}_{M-1}]}}{\int \mathcal{D}\vec{\phi}_0 \mathcal{D}\vec{\phi}_1 \cdots \mathcal{D}\vec{\phi}_{M-1} e^{-S_{T,\theta}[\vec{\phi}_0, \ldots, \vec{\phi}_{M-1}]}} \\ &\approx \lim_{T\to\infty} \lim_{\substack{\theta\to 0 \\ M\theta = T}} \frac{1}{N_{\text{bootstrap}}} \sum_{b=1}^{N_{\text{bootstrap}}} \left[\frac{1}{N_{\text{samples}}} \sum_{\alpha=1}^{N_{\text{samples}}} \mathcal{O}(\vec{\phi}_{t_1}^{I_b(\alpha)}, \ldots, \vec{\phi}_{t_k}^{I_b(\alpha)}) \right], \end{aligned} \tag{B6}$$

where the statistical error in the Monte-Carlo estimate of the above integral can be measured by the standard deviation of the bootstrap means contained inside the square brackets. To take the $\theta \to 0$ limit, we perform the Monte-Carlo computation at various values of temporal lattice spacing θ and for a fixed value of inverse temperature T. For the temporal discretization θ to be meaningful, one needs to ensure $\theta < \Delta E^{-1}$, where ΔE is the energy gap measured in units of spatial lattice spacing, and then suitably take the $\theta \to 0$ limit at fixed T. Furthermore, to extract ground-state properties, sampling needs to take place in the limit $T \to \infty$. Instead of taking this limit explicitly, it suffices to work with a single value of $T \gg \Delta E^{-1}$. Since the energy gap is not known *a priori*, one can pick an arbitrary value for this fixed T and for the temporal lattice spacings θ, and check if the conditions $\theta \Delta E < 1$ and $T \Delta E \gg 1$ are satisfied at the end of the computation. In this study, we use $T = 10$ and $\theta \in \{0.4, 0.2, 0.1\}$. We further set $N = 10$, and consider $(m^2, \lambda) \in \{(0.6, 1.5), (0.4, 1.0), (0.2, 0.5), (0.1, 0.25)\}$.

We will now describe the results for the computation of various ground-state moments, followed by nonlinear functions of these moments, i.e., the moment ratio and energy gap. Figures 16–19 show the $\theta \to 0$ extrapolations

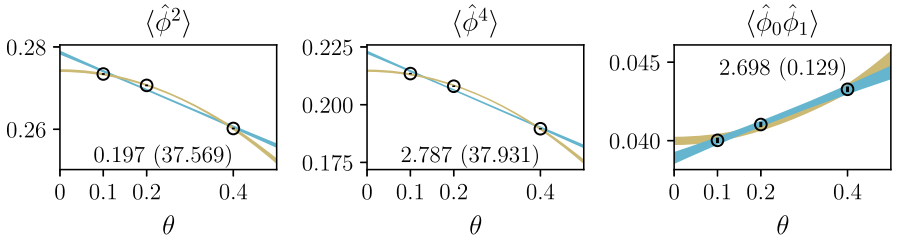


FIG. 16. Extrapolations of the one- and two-point functions of ϕ that feature in the Hamiltonian to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.6, 1.5)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

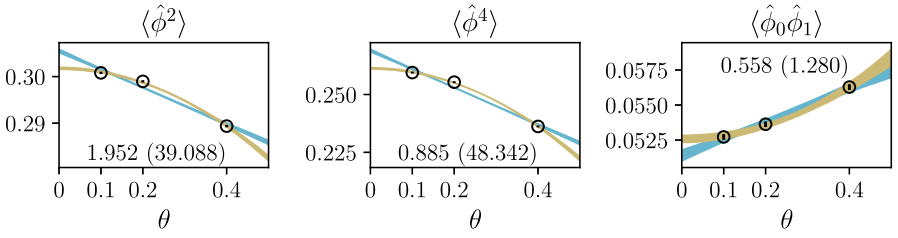


FIG. 17. Extrapolations of the one- and two-point functions of ϕ that feature in the Hamiltonian to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.4, 1.0)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

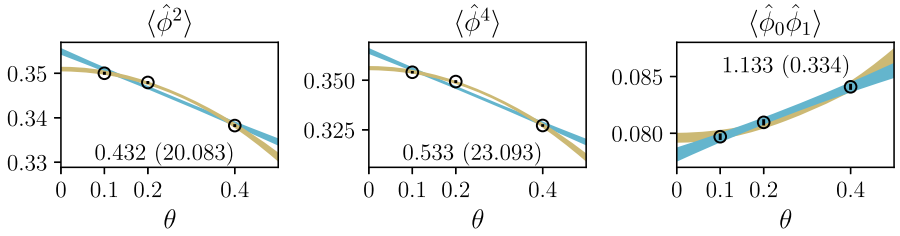


FIG. 18. Extrapolations of the one- and two-point functions of ϕ that feature in the Hamiltonian to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.2, 0.5)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

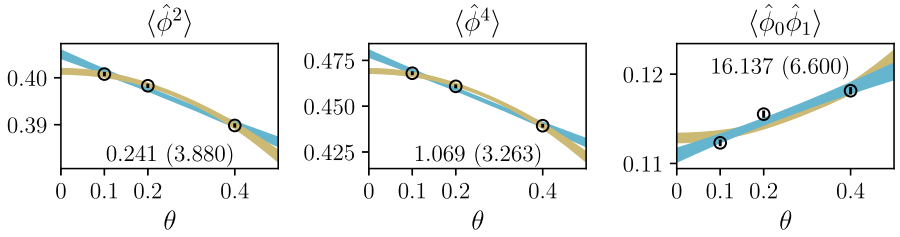


FIG. 19. Extrapolations of the one- and two-point functions of ϕ that feature in the Hamiltonian to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.1, 0.25)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

for the one- and two-point ϕ-moments that feature in the Hamiltonian, i.e., $\langle\hat{\phi}^2\rangle$, $\langle\hat{\phi}^4\rangle$, and $\langle\hat{\phi}_0\hat{\phi}_1\rangle$. Using these extrapolated values, the momentum variance $\langle\hat{\pi}_j^2\rangle$ is estimated using the virial theorem, which together with translation invariance implies that

$$\langle\hat{\pi}_j^2\rangle = \left\langle \hat{\phi}_j \frac{dV(\hat{\vec{\phi}})}{d\hat{\phi}_j} \right\rangle = (2 + m^2)\langle\hat{\phi}_j^2\rangle + \lambda\langle\hat{\phi}_j^4\rangle. \quad \text{(B7)}$$

All of these values are then combined together to get the $(\theta \to 0)$ extrapolated value of energy. Furthermore, Figs. 20–23 show the extrapolation plots for the two-point

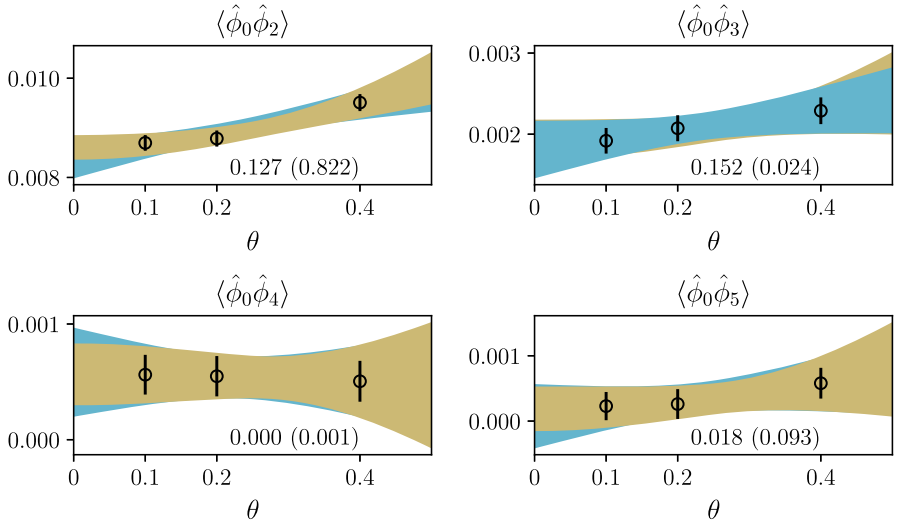


FIG. 20. Extrapolations of the two-point correlation functions to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.6, 1.5)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

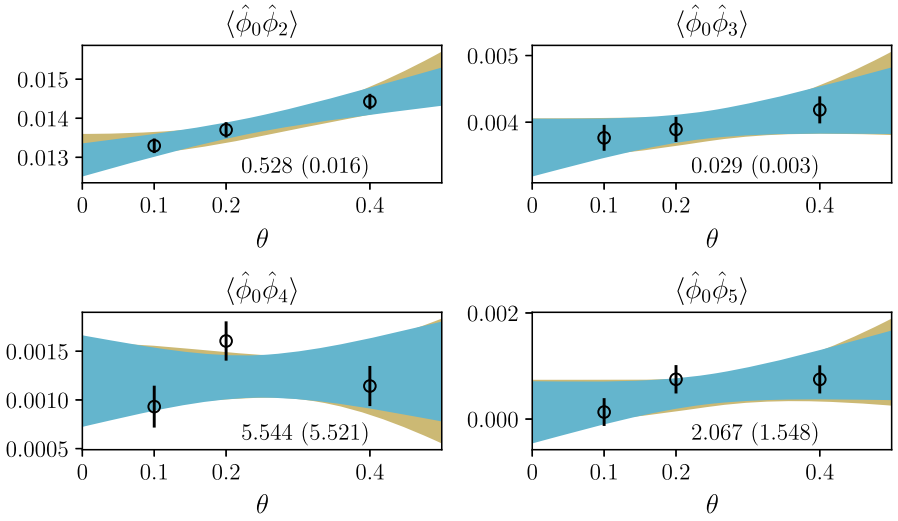


FIG. 21. Extrapolations of the two-point correlation functions to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.4, 1.0)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

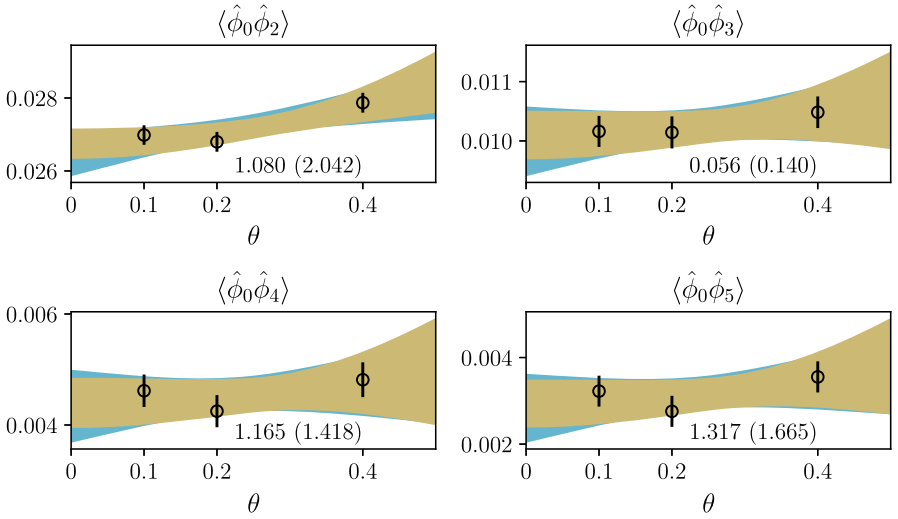


FIG. 22. Extrapolations of the two-point correlation functions to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.2, 0.5)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

correlation functions $\langle\hat{\phi}_0\hat{\phi}_j\rangle$ with $j = 2, 3, 4, 5$, while Figs. 24–27 show the extrapolation for the moment ratios R_n with $n = 4, 6, 8, 10$.

The energy gap can be estimated using the auto-correlation function. Suppose that $\hat{H}$ has a non-degenerate ground state $|\Omega\rangle$, and denote its eigenstates by $|E_n\rangle$, with $|E_0\rangle = |\Omega\rangle$, and the gap $\Delta E = E_1 - E_0$. Recalling that $\hat{\phi}_j(\tau = t\theta) \coloneqq \hat{\phi}_{j,t} = e^{\tau\hat{H}}\hat{\phi}_j(0)e^{-\tau\hat{H}}$, one can write

$$\langle\Omega|\hat{\phi}_j(\tau)\hat{\phi}_j(0)|\Omega\rangle = \sum_{n=0}^{\infty} e^{-\tau(E_n-E_0)}|\langle\Omega|\hat{\phi}_j(0)|E_n\rangle|^2 = e^{-\tau(E_1-E_0)}|\langle\Omega|\hat{\phi}_j(0)|E_1\rangle|^2 + \cdots, \quad \text{(B8)}$$

for $\tau < \frac{T}{2}$. The first term in the sum, $|\langle\Omega|\hat{\phi}_j(0)|\Omega\rangle|^2$, has been set to zero due to the $(\mathbb{Z}_2)$ parity-invariance of the state. Due to the periodic nature of this lattice, the dominant contribution to this correlator in the long Euclidean time limit is, in fact, given by

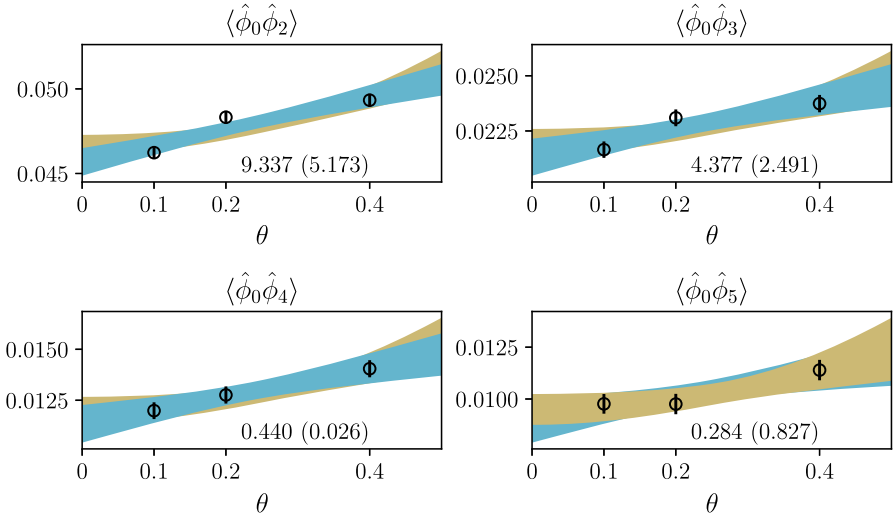


FIG. 23. Extrapolations of the two-point correlation functions to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.1, 0.25)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

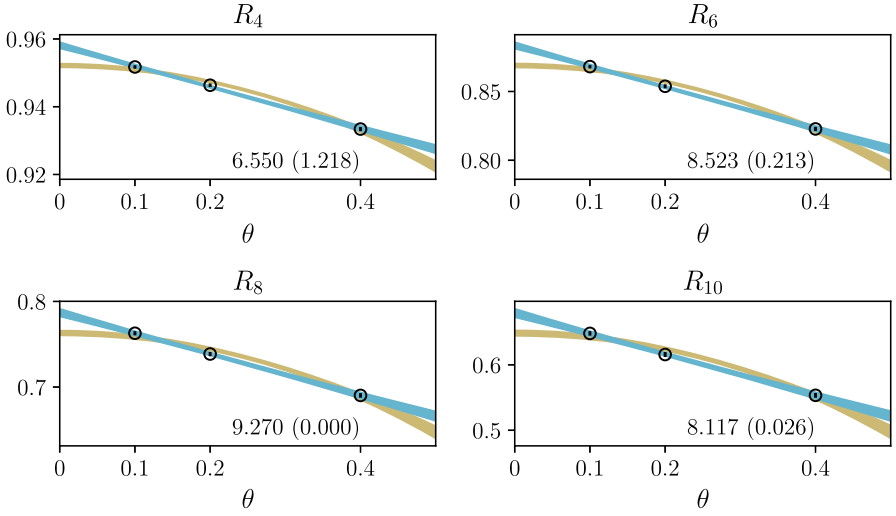


FIG. 24. Extrapolations of the moment ratios to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.6, 1.5)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

$$\begin{aligned}
&\langle\Omega|\hat{\phi}_j(\tau)\hat{\phi}_j(0)|\Omega\rangle \\
&= \frac{1}{2}[\langle\Omega|\hat{\phi}_j(\tau)\hat{\phi}_j(0)|\Omega\rangle + \langle\Omega|\hat{\phi}_j(T-\tau)\hat{\phi}_j(0)|\Omega\rangle] \\
&= |\langle\Omega|\hat{\phi}_j(0)|E_1\rangle|^2 \frac{e^{-\tau\Delta E} + e^{-(T-\tau)\Delta E}}{2} + \cdots \\
&= |\langle\Omega|\hat{\phi}_j(0)|E_1\rangle|^2 e^{-\frac{T\Delta E}{2}} \cosh\left[\Delta E\left(\frac{T}{2} - \tau\right)\right] + \cdots . \qquad \text{(B9)}
\end{aligned}$$

To better identify a range of τ where the above form is the dominant contribution, one can define the effective mass $m_{\text{eff}}(\tau)$ at each point $\tau = t\theta$ on the discrete and periodic Euclidean lattice through the relation

$$\frac{\langle\Omega|\hat{\phi}_{j,t}\hat{\phi}_{j,0}|\Omega\rangle}{\langle\Omega|\hat{\phi}_{j,t+1}\hat{\phi}_{j,0}|\Omega\rangle} = \frac{\cosh\left[m_{\text{eff}}(t\theta)(\frac{T}{2} - t\theta)\right]}{\cosh\left[m_{\text{eff}}(t\theta)(\frac{T}{2} - (t+1)\theta)\right]}. \qquad \text{(B10)}$$

The resulting plots for m_{eff} are shown in Figs. 28–31. The shaded window in yellow represents the time period used to average over for the final gap estimate, ΔE. The gap estimates for different (m^2, λ) values vary from $\Delta E \approx 0.7$ to $\Delta E \approx 1.6$. They are all consistent with the requisite conditions stated

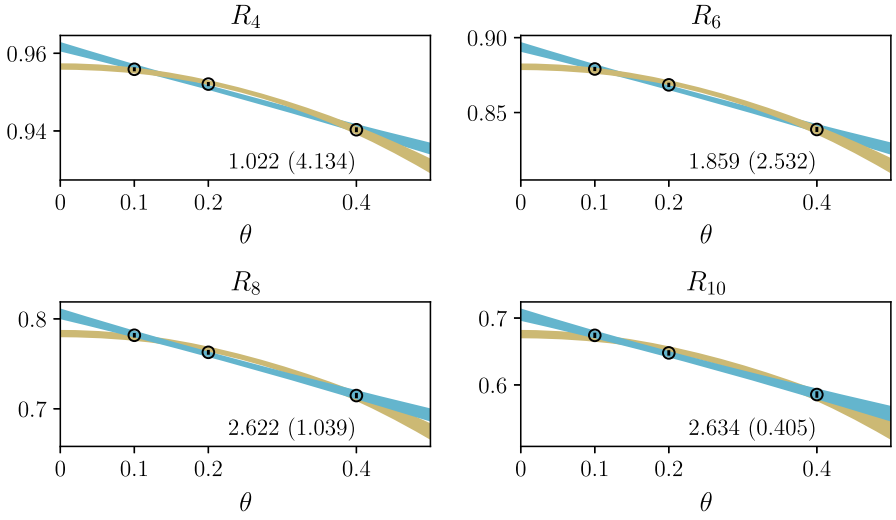


FIG. 25. Extrapolations of the moment ratios to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.4, 1.0)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

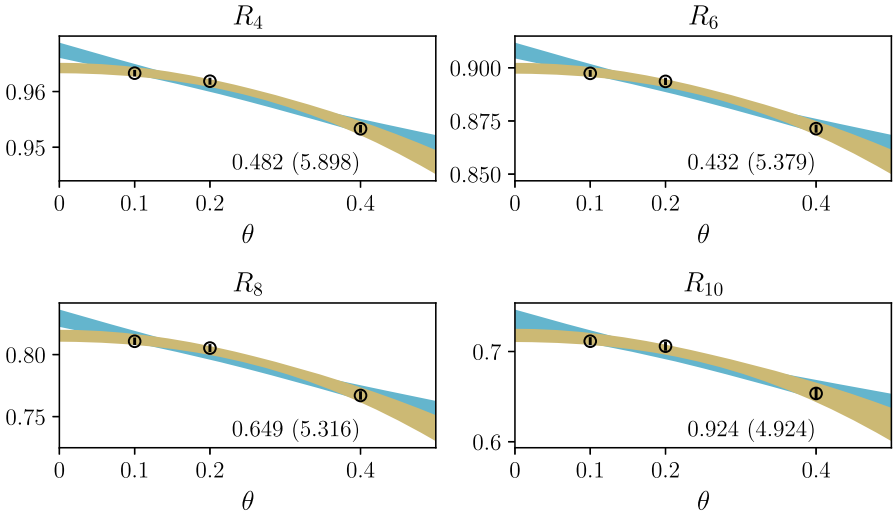


FIG. 26. Extrapolations of the moment ratios to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.2, 0.5)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

before, i.e., $\frac{1}{T} = 0.1 < \Delta E < \frac{1}{\theta_{\max}} = 2.5$, where $\theta_{\max}$ is the largest lattice spacing used in our computations.

To propagate the above errors through moment optimization, we perform the optimization procedure multiple times: each time, we source a random value for the ground-state moments by sampling from a normal distribution specified by the continuum extrapolated values and their associated errors. The plots in the main text show the mean and 68% confidence bands obtained from these samples. In particular, all plots in this work are produced using 100 samples.

APPENDIX C: CORE-STATE POLYNOMIAL WITH QUADRATURE ARGUMENT

In Sec. IV A of the main text, we asserted that the core state $|C\rangle \coloneqq C(\hat{\boldsymbol{a}}^\dagger)|\mathbf{0}\rangle$ can equivalently be expressed as $|C\rangle \coloneqq C^\phi(\hat{\boldsymbol{\phi}})|\mathbf{0}\rangle$. There is a one-to-one mapping between

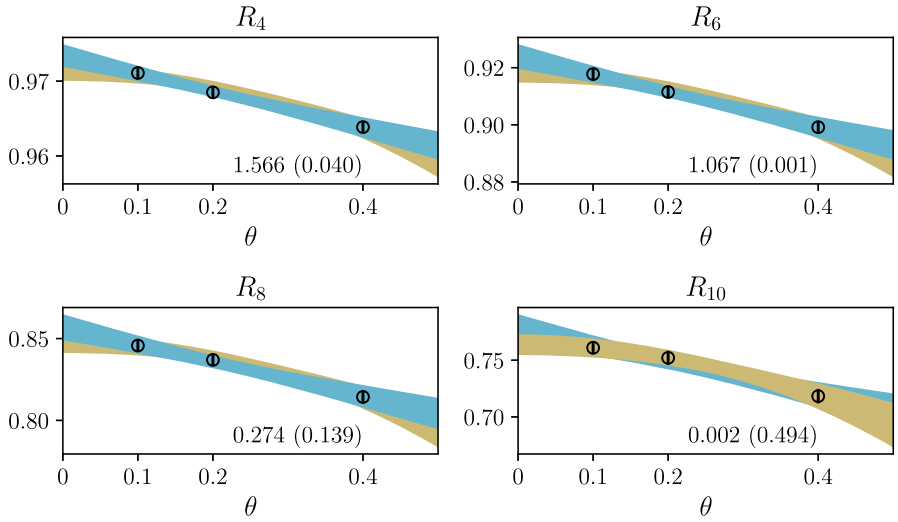


FIG. 27. Extrapolations of the moment ratios to zero temporal lattice spacing θ for $(m^2, \lambda) = (0.1, 0.25)$. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared value.

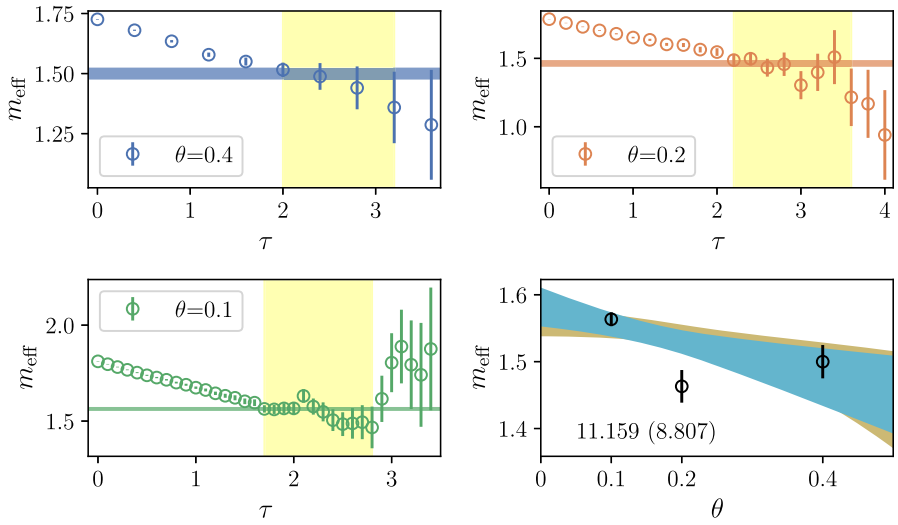


FIG. 28. Effective-mass plots corresponding to $(m^2, \lambda) = (0.6, 1.5)$. The bottom right plot shows the m_{eff} value from each plot averaged over the yellow window as a function of θ. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit with a lower χ-squared.

the terms and independent coefficients of these two polynomials. While this mapping is easy to establish, for completeness, we provide details of the mapping in this appendix.

Consider an N-mode core state $|C\rangle$ with

$$C^{\phi}(\hat{\boldsymbol{\phi}}) \coloneqq \sum_{\substack{\boldsymbol{n} \\ \text{sum}(\boldsymbol{n}) \le R}} d_{\boldsymbol{n}} \hat{\boldsymbol{\phi}}^{\boldsymbol{n}} |\mathbf{0}\rangle = \sum_{\substack{\boldsymbol{n} \\ \text{sum}(\boldsymbol{n}) \le R}} d_{n_0,\ldots,n_{N-1}} \hat{\phi}_0^{n_0} \ldots \hat{\phi}_{N-1}^{n_{N-1}} |\mathbf{0}\rangle. \tag{C1}$$

Now since $\hat{a}_j|\mathbf{0}\rangle = 0$ and $\hat{a}_j \hat{a}_j^\dagger = \hat{a}_j^\dagger \hat{a}_j + 1$,

$$\hat{\phi}_j^{n_j}|\mathbf{0}\rangle = \frac{1}{2^{n_j}} [(\hat{a}_j^\dagger)^{n_j} + A_{n_j,2}(\hat{a}_j^\dagger)^{n_j-2} + A_{n_j,4}(\hat{a}_j^\dagger)^{n_j-4} + \cdots + A_{n_j,n_j-2\lfloor \frac{n_j}{2} \rfloor}(\hat{a}_j^\dagger)^{n_j-2\lfloor \frac{n_j}{2} \rfloor}]|\mathbf{0}\rangle. \tag{C2}$$

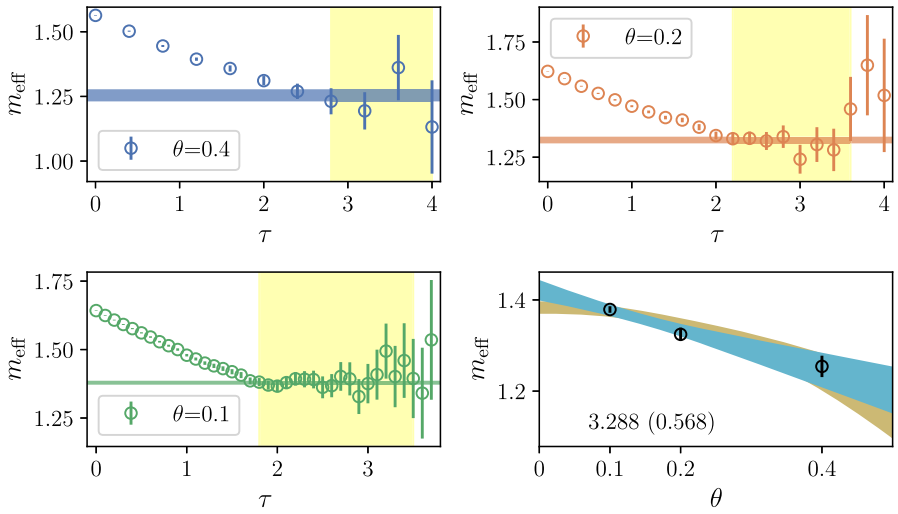


FIG. 29. Effective-mass plots corresponding to $(m^2, \lambda) = (0.4, 1.0)$. The bottom right plot shows the m_{eff} value from each plot averaged over the yellow window as a function of θ. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit which has a lower χ-squared.

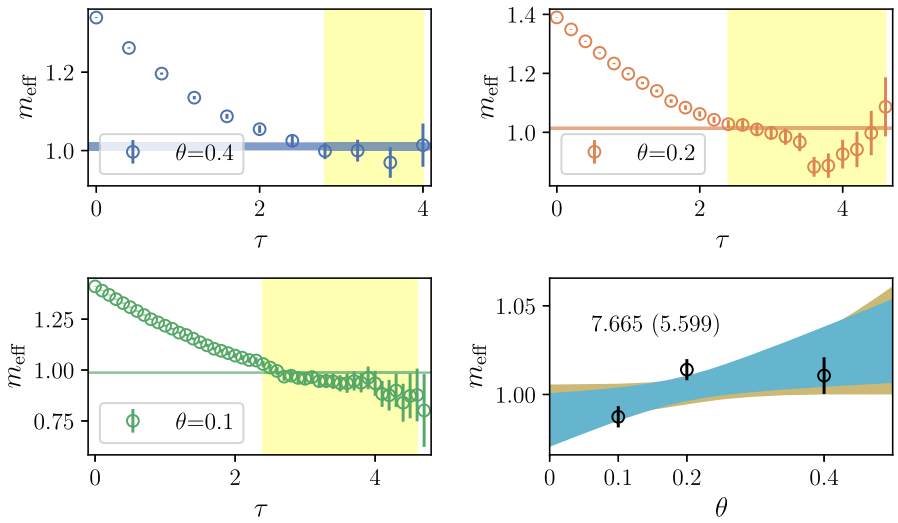


FIG. 30. Effective-mass plots corresponding to $(m^2, \lambda) = (0.2, 0.5)$. The bottom right plot shows the m_{eff} value from each plot averaged over the yellow window as a function of θ. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit which has a lower χ-squared.

Thus, the core state can be expressed as

$$
\begin{aligned}
|C\rangle &= \sum_{\substack{\boldsymbol{n} \\ \text{sum}(\boldsymbol{n}) \le R}} \frac{d_{\boldsymbol{n}}}{2^{\text{sum}(\boldsymbol{n})}} \bigotimes_{j=0}^{N-1} [(\hat{a}_j^\dagger)^{n_j} + A_{n_j,2}(\hat{a}_j^\dagger)^{n_j-2} + A_{n_j,4}(\hat{a}_j^\dagger)^{n_j-4} + \cdots + A_{n_j,n_j-2\lfloor \frac{n_j}{2} \rfloor}(\hat{a}_j^\dagger)^{n_j-2\lfloor \frac{n_j}{2} \rfloor}]|\mathbf{0}\rangle \\
&= \sum_{\substack{\boldsymbol{n} \\ \text{sum}(\boldsymbol{n}) \le R}} c'_{\boldsymbol{n}} (\hat{\boldsymbol{a}}^\dagger)^{\boldsymbol{n}} |\mathbf{0}\rangle \equiv C(\hat{\boldsymbol{a}}^\dagger)|\mathbf{0}\rangle. \qquad \text{(C3)}
\end{aligned}
$$

Define $\zeta = 0$ or 1. Comparing the coefficients of $(\hat{\boldsymbol{a}}^\dagger)^{\boldsymbol{n}}$ in the first and second lines, one gets that for $\boldsymbol{n}$ with $\text{sum}(\boldsymbol{n}) = R - \zeta$,

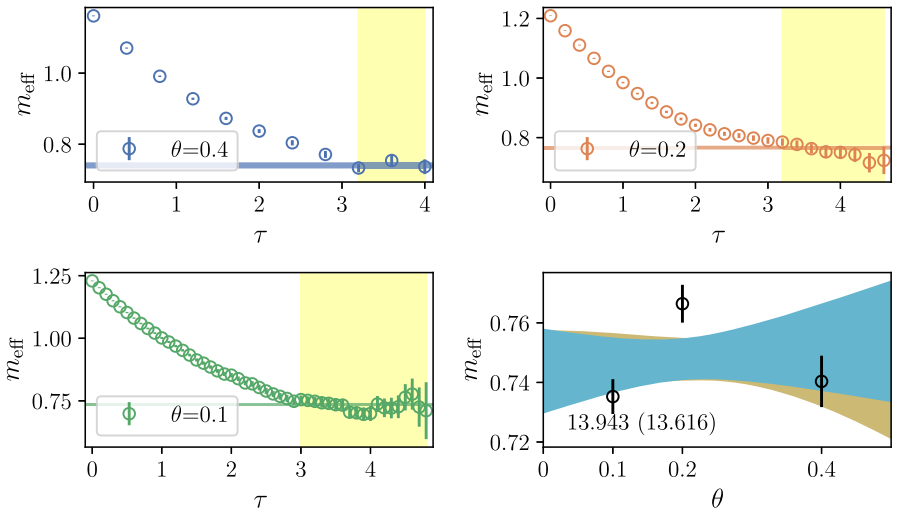


FIG. 31. Effective-mass plots corresponding to $(m^2, \lambda) = (0.1, 0.25)$. The bottom right plot shows the m_{eff} value from each plot averaged over the yellow window as a function of θ. The gold and blue bands are 68% confidence bands for quadratic $(a + b\theta^2)$ and linear $(a + b\theta)$ fitting functions, respectively. The value of the reduced χ-squared for the quadratic (linear) fit is listed without (with) the parenthesis. We take the final continuum value to be the one obtained with the fit which has a lower χ-squared.

$$c'_{\boldsymbol{n}} = \frac{d_{\boldsymbol{n}}}{2^{R-\zeta}}. \tag{C4}$$

Having solved for the coefficients $c'_{\boldsymbol{n}}$ accompanying the terms with the highest weights $\text{sum}(\boldsymbol{n}) = R - \zeta$, one can now work out the "descending coefficients." For instance, for $\boldsymbol{n}$ with $\text{sum}(\boldsymbol{n}) = R - \zeta - 2$,

$$c'_{\boldsymbol{n}} = \frac{d_{\boldsymbol{n}}}{2^{R-\zeta-2}} + \sum_{j=0}^{N-1} \frac{d_{\boldsymbol{n}+2\boldsymbol{e}_j}}{2^{R-\zeta}} A_{n_j+2,2}, \tag{C5}$$

where $\boldsymbol{e}_j$ is the unit vector in the direction j. For $\text{sum}(\boldsymbol{n}) = R - \zeta - 4$,

$$\begin{aligned} c'_{\boldsymbol{n}} = \frac{d_{\boldsymbol{n}}}{2^{R-\zeta-4}} &+ \sum_{j=0}^{N-1} \frac{d_{\boldsymbol{n}+2\boldsymbol{e}_j}}{2^{R-\zeta-2}} A_{n_j+2,2} + \sum_{j=0}^{N-1} \frac{d_{\boldsymbol{n}+4\boldsymbol{e}_j}}{2^{R-\zeta}} A_{n_j+4,4} \\ &+ \sum_{j\neq j'}^{N-1} \frac{d_{\boldsymbol{n}+2\boldsymbol{e}_j+2\boldsymbol{e}_{j'}}}{2^{R-\zeta}} A_{n_j+2,2} A_{n_{j'}+2,2}. \end{aligned} \tag{C6}$$

This procedure can thus be used to determine the coefficients of the standard core-state polynomial $C(\hat{\boldsymbol{a}}^\dagger)$ starting from the quadrature polynomial generator $C(\hat{\boldsymbol{\phi}})$. It is straightforward to invert the above equations. In other words, given any core-state polynomial $C(\hat{\boldsymbol{a}}^\dagger)$, one can obtain a unique polynomial $C^\phi(\hat{\boldsymbol{\phi}})$ which generates the same core state.

APPENDIX D: SPARSE QUANTUM-STATE PREPARATION

This appendix provides a brief overview of, and an illustrative example for, the sparse quantum-state preparation algorithm introduced in Ref. [110]. We proposed to use this method to prepare core states in Sec. V B.

Suppose the goal is to prepare an n-qubit state $|\psi\rangle = \sum_{x\in S} c_x |x\rangle$, where x are length-n bit strings taking values in some set $S \subset \{0,1\}^n$. Reference [110] provides a method for deducing the unitary transformation that converts $|\psi\rangle$ to the state $|\psi'\rangle = \sum_{x\in S'} c'_x |x\rangle$ with $|S'| < |S|$. Thus, by successively applying this algorithm, the state $|\psi\rangle$ can be mapped to a single computational basis state $|x\rangle$. This state can, in turn, be mapped to the state $|0\ldots0\rangle$ by the application of NOT (single-qubit Pauli-X) gates. Inverting the above series of transformations allows one to prepare the state $|\psi\rangle$ starting from the state $|0\ldots0\rangle$. This algorithm is especially efficient when the superposition is sparse, i.e., $|S| \ll 2^n$. As explained in the main text, this condition holds for the qubitzed (R, Q) core state.

As an example, consider applying the algorithm to convert the state $|\psi\rangle = c_0|0000\rangle + c_1|0101\rangle + c_2|0110\rangle + c_3|1011\rangle + c_4|1100\rangle + c_5|1101\rangle$ into a superposition of a smaller number of computational basis states. Repeated application of this subroutine then maps $|\psi\rangle$ to $|0000\rangle$. (For the notation and pseudocode, refer to Ref. [110], Algorithm 1.)

(1) *Initialization:* Initialize

$$T = S = [0000, 0101, 0110, 1011, 1100, 1101],$$
$$dif\_qubits = [], \qquad dif\_vals = [].$$

(2) *The first WHILE loop:* Scan over all qubits $b \in \{1, \ldots, n\}$ and look for the first qubit b such that the sizes of the sets $T_0 := \{x \in T | x[b] = 0\}$ and $T_1 := \{x \in T | x[b] = 1\}$ are as unequal as possible (with neither set being empty). This turns out to be $b = 1$; in this case, $T_0 = [0000, 1011]$ and

$T_1 = [0101, 0110, 1100, 1101]$, i.e., $|T_0| = 2$ and $|T_1| = 4$. Identify the bit value $v \in \{0, 1\}$ such that T_v is the smaller of these two sets (if the two sets have the same size, pick T_1). Now, append the qubit index b (i.e., $b = 1$ in this case) to dif_qubits, append the bit value v (i.e., $v = 0$ in this case) to dif_vals, and update T to the set of bit strings in T which take the value v at the identified qubit index b (i.e., $T \to T_0$ in this case). Thus, at this stage, one has

$$T = [0000, 1011], \quad dif\_qubits = [1], \quad dif\_vals = [0].$$

Since $|T| > 1$, once again scan over all qubits b to look for the first qubit that satisfies the above condition (now with the updated set T which has a reduced number of strings). This turns out to be $b = 0$; thus append this to dif_qubits. Since $|T_0| = |T_1|$ in this case, update T to $T = T_1$, and append 1 to dif_vals. This yields

$$T = [1011], \quad dif\_qubits = [1, 0], \quad dif\_vals = [0, 1].$$

Since $|T| = 1$, the first WHILE loop terminates.

(3) *Preparation for the second WHILE loop:* Remove the last value appended to dif_qubits, i.e., $b = 0$; store it as $dif = 0$. Also remove the last value of dif_vals. Finally, store the single element of T as x_1. This leaves

$$x_1 = 1011, \qquad dif = 0,$$
$$dif\_qubits = [1], \qquad dif\_vals = [0].$$

Now, define a new set T' which consists of all those strings $x \in S$ which take values in dif_vals on qubits in dif_qubits, i.e., $T' = [0000, 1011]$. Remove x_1 from T' to have the final assignments

$$T' = [0000], \qquad dif\_qubits = [1], \qquad dif\_vals = [0].$$

(4) *The second WHILE loop:* The second WHILE loop runs in exactly the same way as the first WHILE loop, except that the initial conditions now involve the set T', and the updated values of dif_qubits and dif_vals defined above. Since $|T'|$ is already 1, the loop does not run in this case.

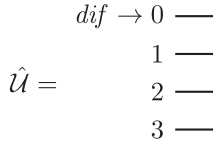
(5) *Preparation for circuit building:* Assign the single element in T' to the variable x_2. Thus, the relevant variable values at this stage are

$$x_1 = 1011, \qquad x_2 = 0000, \qquad dif = 0,$$
$$dif\_qubits = [1], \qquad dif\_vals = [0].$$

To summarize, the above steps identify two strings x_1 and x_2 which take equal values on the qubits in dif_qubits, and no other strings in S take the same values as x_1 and x_2 on dif_qubits. This property allows the strings x_1 and x_2 to be combined into a single computational basis state using the operations described in the next few steps, without introducing additional strings in the remaining set of bit strings in S.

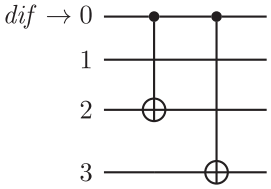
(6) *Circuit building:* Initialize $\hat{\mathcal{U}} = \hat{1}$.

(a) If $x_1[dif] \neq 1$, add a NOT gate to the line dif. In this case, $x_1[dif]$ is already 1.

$\hat{\mathcal{U}} =$ 

(D1)

(b) For all qubits b other than the dif qubit, if $x_1[b] \neq x_2[b]$, apply a CNOT gate on qubit b controlled on dif.

$\hat{\mathcal{U}} =$ 

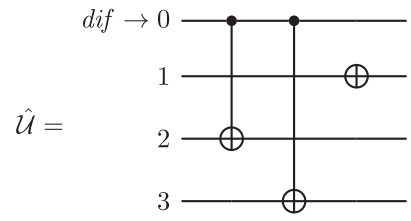
(D2)

$x_1[dif] = 1$ and $x_2[dif] = 0$ by design. Thus, the application of these CNOT gates leaves x_2 untouched while x_1 is transformed so that it agrees with x_2 on all qubits except the dif qubit. More concretely,

$$\begin{aligned}\hat{\mathcal{U}}|\psi\rangle &= \hat{\mathcal{U}}(c_0|\mathbf{0000}\rangle + c_1|0101\rangle + c_2|0110\rangle + c_3|\mathbf{1011}\rangle + c_4|1100\rangle + c_5|1101\rangle) \\ &= c_0|\mathbf{0000}\rangle + c_1|0101\rangle + c_2|0110\rangle + c_3|\mathbf{1000}\rangle + c_4|1111\rangle + c_5|1110\rangle. \end{aligned} \tag{D3}$$

Here and below, the strings x_1 and x_2, and their transformed versions, are highlighted.

(c) As mentioned earlier, x_1 and x_2 strings are the only strings that take the values dif_vals on the qubits dif_qubits. This step ensures that these values are all set to 1 so that the merging operation can be activated in the next step. In other words, for all qubits b in dif_qubits, if $x_1[b] = x_2[b] \neq 1$, a NOT gate is applied on line b.



(D4)

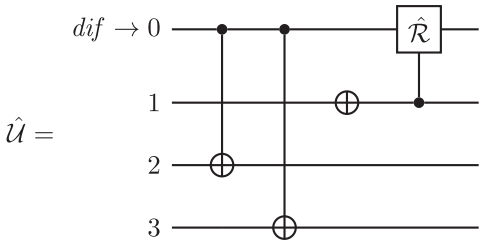
This gives

$$\begin{aligned}\hat{\mathcal{U}}|\psi\rangle &= c_0|\mathbf{0100}\rangle + c_1|0001\rangle + c_2|0010\rangle + c_3|\mathbf{1100}\rangle + c_4|1011\rangle + c_5|1010\rangle \\ &= (c_0|\mathbf{0}\rangle + c_3|\mathbf{1}\rangle) \otimes |\mathbf{100}\rangle + c_1|0001\rangle + c_2|0010\rangle + c_4|1011\rangle + c_5|1010\rangle. \end{aligned} \tag{D5}$$

(d) The highlighted strings can now be merged into one string by performing a suitable rotation $\hat{\mathcal{R}}$ on the qubit *dif* controlled on all qubits in *dif_qubits*. The rotation is defined so that

$$\hat{\mathcal{R}}(c_0|0\rangle + c_3|1\rangle) = \mathcal{N}_{c_0,c_3}|0\rangle,$$

where $\mathcal{N}_{c_0,c_3}$ is a complex amplitude.



(D6)

This is the final form for $\hat{\mathcal{U}}$. It gives the desired state $|\psi'\rangle$,

$$\begin{aligned}\hat{\mathcal{U}}|\psi\rangle &= \mathcal{N}_{c_0,c_3}|\mathbf{0100}\rangle + c_1|0001\rangle + c_2|0010\rangle \\ &\quad + c_4|1011\rangle + c_5|1010\rangle \equiv |\psi'\rangle, \end{aligned} \tag{D7}$$

which satisfies $|S'| = 5 < |S| = 6$.

In each iteration of the WHILE loops, the sizes of the sets T and T' are at least halved. Thus, the number of WHILE loops is $O(\lceil \log(|S|)\rceil)$. Each WHILE loop takes $O(|S|n)$ time as it involves looking at at most $|S|$ strings of length n. Thus, the total classical runtime of the above algorithm is $O(|S|n\log(|S|))$.

Since each WHILE loop adds an element to *dif_qubits*, $|dif\_qubits|$ is $O(\log(|S|))$. Step 6(a) involves at most one single-qubit gate while step 6(b) involves at most $n-1$ CNOT gates. In step 6(c), one adds $O(\log(|S|))$ one-qubit gates. In step 6(d), one must implement a rotation gate controlled on the $O(\log|S|)$ *dif_qubits*, which can be implemented using $O(\log(|S|))$ CNOT gates. Thus, the total number of CNOT gates is $O(n)$, and the total number of single-qubit gates is $O(\log(|S|))$.

Now, one instance of the above algorithm reduces the number of states in the superposition by one. Thus, this algorithm must be repeated $|S|$ times to deduce the unitary which maps the state $|\psi\rangle$ to a single computational basis state. Thus, the total classical runtime for determining the full quantum circuit turns out to be $O(|S|^2 n\log(|S|))$. The total number of CNOT gates is $O(|S|n)$, while the total number of one-qubit gates is $O(|S|\log(|S|)+n)$ (the addition of n arises because converting a single basis state to $|0\ldots0\rangle$ requires $O(n)$ single-qubit gates). In summary, if $|S| \ll 2^n$, this algorithm is efficient in converting the state $|\psi\rangle$ to $|0\ldots0\rangle$ and vice versa.

APPENDIX E: DIGITIZATION OF THE SQUEEZING OPERATOR

As discussed in Sec. V C of the main text, in order to make the single-mode squeezing operator, $\hat{S}(r)$, amenable to a discrete-variable implementation, it must first be mapped to a squeezing operator acting just on the truncated Hilbert space, i.e., $\hat{S}^{\Lambda}(r)$. Thereafter, $\hat{S}^{\Lambda}(r)$ is approximated by a first-order product expansion, $\hat{S}^{\Lambda,K}(r)$

$$\begin{aligned}\hat{S}(r) = e^{\frac{r}{2}(\hat{a}^{\dagger 2}-\hat{a}^2)} &\mapsto \hat{S}^{\Lambda}(r) \coloneqq e^{\frac{r}{2}[(\hat{a}^{\Lambda\dagger})^2-(\hat{a}^{\Lambda})^2]} \mapsto \hat{S}^{\Lambda,K}(r) \\ &\coloneqq (e^{-i\frac{r}{K}\hat{s}_0^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_2^{\Lambda}})^K(e^{-i\frac{r}{K}\hat{s}_1^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_3^{\Lambda}})^K. \end{aligned} \tag{E1}$$

To compare $\hat{S}^{\Lambda}(r)$ and $\hat{S}^{\Lambda,K}(r)$ (which act only on the truncated Hilbert space) with the squeezing operator $\hat{S}(r)$ (which acts on the full single-mode Hilbert space), we will embed the former in the untruncated single-mode Hilbert space by suitably padding the operator matrix with zero matrix elements to ensure the truncated operators do not act on Fock states beyond the cutoff Λ. In this appendix, we describe the errors arising from this truncation and Trotterization for both the single-mode and multimode cases.

For a single mode, the squeezing operator is applied to the rank-R core state to obtain the rank-R *Ansatz* $|\psi\rangle_R = \hat{S}(r)|C\rangle_R$. We demand a bound on the distance

$$\||\psi\rangle_R - |\psi^{\Lambda,K}\rangle_R\| \le \||\psi\rangle_R - |\psi^{\Lambda}\rangle_R\| + \||\psi^{\Lambda}\rangle_R - |\psi^{\Lambda,K}\rangle_R\| \le \epsilon^{(1)} \coloneqq \epsilon^{(1)}_{\text{trunc}} + \epsilon^{(1)}_{\text{trott}}, \tag{E2}$$

where $\||\psi\rangle\| \coloneqq \sqrt{\langle\psi|\psi\rangle}$ is the inner-product norm acting on the full single-mode Hilbert space, $|\psi^{\Lambda}\rangle_R \coloneqq \hat{S}^{\Lambda}(r)|C\rangle_R$, and $|\psi^{\Lambda,K}\rangle_R \coloneqq \hat{S}^{q,\Lambda,K}(r)|C\rangle_R$.

In the multimode case, the tensor product of single-mode squeezing operators is applied to the multimode rank-R, span-Q core state $|C\rangle_{R,Q}$ to obtain the (R,Q) *Ansatz* $|\psi\rangle_{R,Q} = \otimes_{j=0}^{N-1}\hat{S}_j(r)|C\rangle_{R,Q}$. In this case, we demand a bound on the distance

$$\||\psi\rangle_{R,Q} - |\psi^{\Lambda,K}\rangle_{R,Q}\| \le \||\psi\rangle_{R,Q} - |\psi^{\Lambda}\rangle_{R,Q}\| + \||\psi^{\Lambda}_{R,Q}\rangle - |\psi^{\Lambda,K}\rangle_{R,Q}\| \le \epsilon^{(N)} \coloneqq \epsilon^{(N)}_{\text{trunc}} + \epsilon^{(N)}_{\text{trott}}, \tag{E3}$$

where we use the same notation $\||\psi\rangle\|$ to denote the inner-product norm over the full multimode Hilbert space. Moreover, $|\psi^{\Lambda}\rangle_{R,Q} \coloneqq \otimes_{j=0}^{N-1}\hat{S}^{\Lambda}_j(r)|C\rangle_{R,Q}$, and $|\psi^{\Lambda,K}\rangle_{R,Q} \coloneqq \otimes_{j=0}^{N-1}\hat{S}^{q,\Lambda,K}_j(r)|C\rangle_{R,Q}$.

1. Truncation error

The goal is to compare the action of the full squeezing operator, $\hat{S}(r)$, and the approximate squeezing operator, $\hat{S}^{\Lambda}(r)$ on the single- and multimode core states. The single-mode rank-R core state is given by $|C\rangle_R = \bar{\sum}_n c_n|n\rangle$. Here, the sum $\bar{\sum}$ is restricted to a sum of $\frac{R}{2}+1$ terms, such that $|C\rangle$ is a symmetric rank-R core state as defined in Eq. (12). Introducing the notation $|n,r,\Lambda\rangle \coloneqq \hat{S}^{\Lambda}(r)|n\rangle$, the distance between the two states for the single-mode case is bounded by

$$\begin{aligned}\||\psi\rangle_R - |\psi^{\Lambda}\rangle_R\| &= \|(\hat{S}(r) - \hat{S}^{\Lambda}(r))|C\rangle_R\| \\ &\le \bar{\sum}_n |c_n|\||n,r\rangle - |n,r,\Lambda\rangle\| \\ &\le \sqrt{\frac{R}{2}+1}\, g(r,R,\Lambda),\end{aligned} \tag{E4}$$

where $g(r,R,\Lambda) \coloneqq \text{argmax}_{n\le R}\||n,r\rangle - |n,r,\Lambda\rangle\|$, and we have used the Cauchy-Schwartz inequality to bound $\bar{\sum}_n|c_n| = \bar{\sum}_n|c_n| \times 1 \le \sqrt{(\bar{\sum}_n|c_n|^2)(\bar{\sum}_n 1)}$. Similarly, the N-mode (R,Q) core state can be expressed as $|C\rangle_{R,Q} = \bar{\sum}_{\boldsymbol{n}} c_{\boldsymbol{n}}|\boldsymbol{n}\rangle$, where the sum $\bar{\sum}$ is restricted to the $N|c_{R,Q}|$ terms as defined in Eq. (26). It follows that,

$$\begin{aligned}\||\psi\rangle_{R,Q} - |\psi^{\Lambda}\rangle_{R,Q}\| &= \|(\otimes_{j=0}^{N-1}\hat{S}_j(r) - \otimes_{j=0}^{N-1}\hat{S}^{\Lambda}_j(r))|C\rangle_{R,Q}\| \\ &\le \bar{\sum}_{\boldsymbol{n}} |c_{\boldsymbol{n}}|\|(\otimes_{j=0}^{N-1}\hat{S}_j(r) - \otimes_{j=0}^{N-1}\hat{S}^{\Lambda}_j(r))|\boldsymbol{n}\rangle\| \\ &= \bar{\sum}_{\boldsymbol{n}} |c_{\boldsymbol{n}}|\|\otimes_{j=0}^{N-1}(|n_j,r\rangle) - \otimes_{j=0}^{N-1}(|n_j,r,\Lambda\rangle)\| \\ &\le \bar{\sum}_{\boldsymbol{n}} |c_{\boldsymbol{n}}|\sum_{j=0}^{N-1}\||n_j,r\rangle - |n_j,r,\Lambda\rangle\| \le N\sqrt{N|c_{R,Q}|}\, g(r,R,\Lambda),\end{aligned} \tag{E5}$$

where in the last line, we have used the fact that for normalized states $|\psi_0\rangle,\ldots,|\psi_{m-1}\rangle, |\phi_0\rangle,\ldots,|\phi_{m-1}\rangle$,

$$\begin{aligned}\|\otimes_{k=0}^{m-1}|\psi_k\rangle - \otimes_{k=0}^{m-1}|\phi_k\rangle\| &= \left\|\sum_{k=0}^{m-1}(|\psi_0\rangle\ldots|\psi_{k-1}\rangle)(|\psi_k\rangle - |\phi_k\rangle)(|\phi_{k+1}\rangle\ldots|\phi_{m-1}\rangle)\right\| \\ &\le \sum_{k=0}^{m-1}\||\psi_0\rangle\ldots|\psi_{k-1}\rangle\|\||\psi_k\rangle - |\phi_k\rangle\|\||\phi_{k+1}\rangle\ldots|\phi_{m-1}\rangle\| \\ &= \sum_{k=0}^{m-1}\||\psi_k\rangle - |\phi_k\rangle\|.\end{aligned} \tag{E6}$$

Both the single-mode and multimode distances depend upon $g(r,R,\Lambda)$, which can be bounded above as follows:

$$\begin{aligned}g(r,R,\Lambda) &\coloneqq \text{argmax}_{n_1\le R}\||n_1,r\rangle - |n_1,r,\Lambda\rangle\| \\ &= \text{argmax}_{n_1\le R}\left[\sum_{n_2=0}^{\infty}|\langle n_2|\hat{S}(r) - \hat{S}^{\Lambda}(r)|n_1\rangle|^2\right]^{1/2} \\ &\le \underbrace{\text{argmax}_{n_1\le R}\left[\sum_{n_2=0}^{\Lambda}|\langle n_2|\hat{S}(r) - \hat{S}^{\Lambda}(r)|n_1\rangle|^2\right]^{1/2}}_{g^{\text{disc}}(r,R,\Lambda)} + \underbrace{\text{argmax}_{n_1\le R}\left[\sum_{n_2=\Lambda+1}^{\infty}|\langle n_2|\hat{S}(r)|n_1\rangle|^2\right]^{1/2}}_{g^{\text{leak}}(r,R,\Lambda)}.\end{aligned} \tag{E7}$$

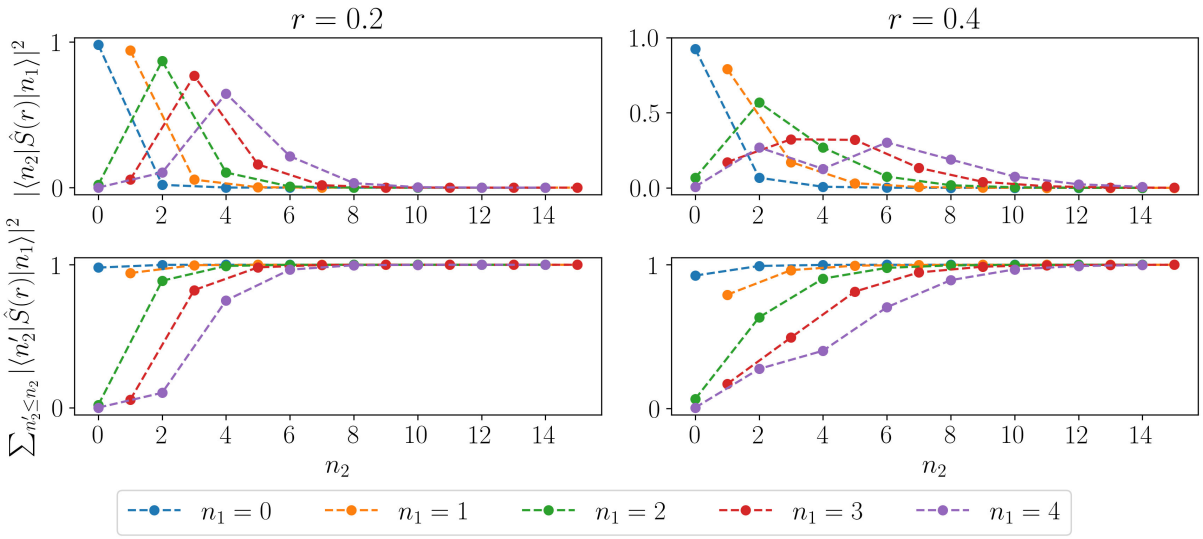


FIG. 32. Boson-number distribution for squeezed Fock states. Top panels show $|\langle n_2|\hat{S}(r)|n_1\rangle|^2$ as a function of n_2 for n_1 values up to 4 (which is the maximum rank of the core states considered in this work). Bottom panels show the cumulative distributions, $\sum_{n_2' \le n_2} |\langle n_2'|\hat{S}(r)|n_1\rangle|^2$, as a function of n_2. These plots demonstrate that the squeezing operator acted on a state with initial occupations $n_1 \le R$ has larger overlap with low-lying states n_2. In the top plots, for even (odd) values for n_1, only the nonvanishing matrix elements corresponding to even (odd) values of n_2 are plotted. In the bottom plots, the points corresponding to odd (even) n_2 values for even (odd) n_1 values are dropped as they equal their previous adjacent point. The dotted lines connecting the points are to guide the eye and do not represent a fit.

In the last line, we have used the fact that $\langle n_2|\hat{S}^{\Lambda}(r)|n_1\rangle = 0$ for $n_2 > \Lambda$. We will refer to the term $g^{\text{disc}}(r, R, \Lambda)$ as the matrix-element *discrepancy* and the second term $g^{\text{leak}}(r, R, \Lambda)$ as *leakage* [as it describes how much of the weight of initial state $|n_1\rangle$, with $n_1 \le R$, is leaked outside the truncated Hilbert space by $\hat{S}(r)$].

Using the analytical form for $|\langle n_2|\hat{S}(r)|n_1\rangle|^2$ (see Ref. [127]) and the unitarity of the squeezing operator, the value of $g^{\text{leak}}(r, R, \Lambda)$ can be calculated exactly. Figure 32 shows the boson-number distribution $|\langle n_2|\hat{S}(r)|n_1\rangle|^2$, together with the cumulative distribution $\sum_{n_2' \le n_2} |\langle n_2'|\hat{S}(r)|n_1\rangle|^2$ for $n_1 \le R = 4$, and for two characteristic values, $r = 0.2$, 0.4, encountered in this work. The plots illustrate that the distribution is peaked around small values of n_2, and the cumulative distribution quickly converges to one as n_2 is increased for these small r values.

We now upper bound $g^{\text{disc}}(r, R, \Lambda)$. Any matrix elements between even and odd Fock states vanish for both the full and truncated squeezing operators. Thus, consider the matrix elements with both n_1 and n_2 being odd or even. Upon Taylor expanding the exponential in $\hat{S}(r)$, the matrix element becomes

$$\begin{aligned}
\langle n_2|e^{\frac{r}{2}\sum_{n=0}^{\infty}\ell_n(|n+2\rangle\langle n|-|n\rangle\langle n+2|)}|n_1\rangle &= \sum_{p=0}^{\infty}\frac{r^p}{p!2^p}\langle n_2|\left[\sum_{n=0}^{\infty}\ell_n(|n+2\rangle\langle n|-|n\rangle\langle n+2|)\right]^p|n_1\rangle \\
&= \sum_{\Delta=0}^{\infty}\frac{r^{p_0+2\Delta}}{(p_0+2\Delta)!2^{p_0+2\Delta}}\underbrace{\langle n_2|\left[\sum_{n=0}^{\infty}\ell_n(|n+2\rangle\langle n|-|n\rangle\langle n+2|)\right]^{p_0+2\Delta}|n_1\rangle}_{f(\Delta)} \\
&\equiv \sum_{\Delta=0}^{\infty} t(\Delta). \qquad \text{(E8)}
\end{aligned}$$

In the second equality, the sum over integer-valued $p \ge 0$ is exchanged with a sum over Δ, where $p = p_0 + 2\Delta \ge 0$, and we have defined $p_0 := \frac{|n_2 - n_1|}{2}$. This latter sum, by construction, involves only nonvanishing matrix elements from the former sum. The term $f(\Delta)$, in particular, receives contributions from "paths" consisting of $p_0 + 2\Delta$ steps (of two units in Fock space) connecting n_1 and n_2. These steps can either be in the forward (corresponding to the $|n+2\rangle\langle n|$ term) or backward (corresponding to the $|n\rangle\langle n+2|$ term) direction. Each such path is composed of $p_0 + \Delta$ steps in the direction from n_1 to n_2, and Δ steps in the opposite direction. Thus, the total number of such

paths is $\binom{p_0+2\Delta}{\Delta}$, i.e., the number of ways in which one could intersperse the Δ steps in the opposite direction among the total number of steps. When $n_2 \geq n_1$, out of these paths, the highest numerical weight is carried by the one which goes all the way from n_1 to $n_2 + 2\Delta$, and then circles back to n_2. The contribution from this highest-weight path is

$$w(n_2 \geq n_1) \coloneqq (-1)^{\Delta} \underbrace{\ell_{n_1} \ldots \ell_{n_2-2} \ell_{n_2} \ldots \ell_{n_2+2\Delta-2}}_{n_1 \to n_2+2\Delta} \underbrace{\ell_{n_2+2\Delta-2} \ldots \ell_{n_2}}_{n_2+2\Delta \to n_2} = (-1)^{\Delta} \frac{(n_2+2\Delta)!}{\sqrt{n_1! n_2!}}. \tag{E9}$$

When $n_1 > n_2$, the highest-weight path goes from n_1 to $n_1 + 2\Delta$ and then circles back all the way to n_2. The contribution from this path is given by,

$$w(n_2 < n_1) \coloneqq (-1)^{\Delta} \underbrace{\ell_{n_1} \ldots \ell_{n_1+2\Delta-2}}_{n_1 \to n_1+2\Delta} \underbrace{\ell_{n_1+2\Delta-2} \ldots \ell_{n_1+2} \ell_{n_1} \ldots \ell_{n_2}}_{n_1+2\Delta \to n_2} = (-1)^{\Delta} \frac{(n_1+2\Delta)!}{\sqrt{n_1! n_2!}}. \tag{E10}$$

Thus,

$$|f(\Delta)| \leq \binom{p_0+2\Delta}{\Delta} \frac{(\max(n_1,n_2)+2\Delta)!}{\sqrt{n_1! n_2!}}, \tag{E11}$$

and the contribution at Δ can be bounded as

$$|t(\Delta)| \leq s(\Delta) \coloneqq \frac{r^{p_0}}{2^{p_0}\sqrt{n_1! n_2!}} \left(\frac{r}{2}\right)^{2\Delta} \frac{(\max(n_1,n_2)+2\Delta)!}{\Delta!(p_0+\Delta)!}. \tag{E12}$$

The corresponding matrix element for $\hat{S}^{\Lambda}(r)$ is obtained by dropping all contributions in the sum in Eq. (E8) from paths visiting states $|n\rangle$ with $n > \Lambda$. Thus, we obtain an analogous expansion,

$$\langle n_2|\hat{S}^{\Lambda}(r)|n_1\rangle = \sum_{\Delta=0}^{\infty} t^{\Lambda}(\Delta), \tag{E13}$$

where $|t^{\Lambda}(\Delta)| \leq s(\Delta)$. Furthermore, $t(\Delta) = t^{\Lambda}(\Delta)$ for $\Delta \leq \lfloor \frac{\Lambda - \max(n_1,n_2)+1}{2} \rfloor$ because no paths of these lengths can reach states beyond the cutoff. Thus, the difference between the relevant matrix elements of $\hat{S}(r)$ and $\hat{S}^{\Lambda}(r)$ can be bounded as

$$\begin{aligned} |\langle n_2|\hat{S}(r) - \hat{S}^{\Lambda}(r)|n_1\rangle| &= \left| \sum_{\Delta=\lfloor \frac{\Lambda-\max(n_1,n_2)+1}{2} \rfloor + 1}^{\infty} (t(\Delta) - t^{\Lambda}(\Delta)) \right| \\ &\leq 2 \sum_{\Delta=\lfloor \frac{\Lambda-\max(n_1,n_2)+1}{2} \rfloor + 1}^{\infty} s(\Delta), \end{aligned} \tag{E14}$$

for $n_1 \leq R$ and $n_2 \leq \Lambda$. Consider the ratio

$$b(\Delta) \coloneqq \frac{s(\Delta+1)}{s(\Delta)} = \left(\frac{r}{2}\right)^2 \frac{(\max(n_1,n_2)+2\Delta+1)(\max(n_1,n_2)+2\Delta+2)}{(\Delta+1)(p_0+\Delta+1)}. \tag{E15}$$

Observe that $b(\Delta) \to r^2 < 1$ as $\Delta \to \infty$. Furthermore, denoting $q_0 \coloneqq \max(n_1,n_2)$, we see that

$$b(\Delta+1) - b(\Delta) = \frac{a_2(q_0,p_0)\Delta^2 + a_1(q_0,p_0)\Delta + a_0(q_0,p_0)}{(\Delta+1)(p_0+\Delta+1)(\Delta+2)(p_0+\Delta+2)}, \tag{E16}$$

with

$$\begin{aligned} a_2(q_0,p_0) &\coloneqq 2 - 4q_0 + 4p_0, \\ a_1(q_0,p_0) &\coloneqq 6 - 10q_0 - 2q_0^2 + 12p_0, \\ a_0(q_0,p_0) &\coloneqq 4 - 5q_0 - 3q_0^2 + 8p_0 + q_0p_0 - q_0^2p_0. \end{aligned} \tag{E17}$$

Thus, when $n_1 = n_2 = 0$ (or equivalently when $q_0 = 0$ and $p_0 = 0$), the sequence of ratios $b(\Delta)$ increases monotonically with Δ. Since $b(\Delta)$ is upper bounded by r^2, in this case we can bound the above sequence by a geometric series as

$$|\langle 0|\hat{S}(r)-\hat{S}^{\Lambda}(r)|0\rangle| \le \tilde{G}_0(r,R,\Lambda) \coloneqq \frac{2s\left(\left\lfloor\frac{\Lambda-\max(n_1,n_2)+1}{2}\right\rfloor+1\right)}{1-r^2}. \tag{E18}$$

In all other cases, $b(\Delta)$ decreases monotonically with Δ and is lower bounded by r^2. Thus, one can find the smallest value of $\Delta \ge \lfloor\frac{\Lambda-\max(n_1,n_2)+1}{2}\rfloor+1$ at which the ratio satisfies $\frac{s(\Delta+1)}{s(\Delta)} \le q$ for some $r^2 < q \le 1$. Suppose this value is denoted by Δ_q. Then, the above series can be bounded as,

$$\sum_{\Delta=\lfloor\frac{\Lambda-\max(n_1,n_2)+1}{2}\rfloor+1}^{\infty} s(\Delta) = \sum_{\Delta=\lfloor\frac{\Lambda-n_2+1}{2}\rfloor+1}^{\Delta_q-1} s(\Delta) + \sum_{\Delta=\Delta_q}^{\infty} s(\Delta) \le \left(\sum_{\Delta=\lfloor\frac{\Lambda-\max(n_1,n_2)+1}{2}\rfloor+1}^{\Delta_q-1} s(\Delta)\right) + \frac{s(\Delta_q)}{1-q}. \tag{E19}$$

Because $b(\Delta)$ can never drop below r^2 in this case, we set the threshold to be just 10% above this lower bound, i.e., $q = 1.1r^2$. This choice keeps the geometric-tail prefactor $\frac{1}{1-q}$ small and the resulting bound tight. With this choice of q, we have

$$|\langle n_2|\hat{S}(r)-\hat{S}^{\Lambda}(r)|n_1\rangle| \le \tilde{G}(r,R,\Lambda,n_1,n_2) \coloneqq 2\left(\sum_{\Delta=\lfloor\frac{\Lambda-\max(n_1,n_2)+1}{2}\rfloor+1}^{\Delta_q-1} s(\Delta)\right) + \frac{2s(\Delta_q)}{1-q}; \qquad n_1+n_2>0. \tag{E20}$$

Further, we define $\tilde{G}(r,R,\Lambda,0,0) = \tilde{G}_0(r,R,\Lambda)$. Therefore,

$$g^{\mathrm{disc}}_{r,R,\Lambda} \le G(r,R,\Lambda) \coloneqq \mathrm{argmax}_{n_1\le R}\left[\sum_{n_2=0}^{\Lambda}\tilde{G}(r,R,\Lambda,n_1,n_2)^2\right]^{1/2}. \tag{E21}$$

Thus, the final expressions for the truncation error are given by

$$\begin{aligned}
\|(\hat{S}(r)-\hat{S}^{\Lambda}(r))|C\rangle_R\| &\le \epsilon^{(1)}_{\mathrm{trunc}}(r,R,\Lambda) \coloneqq \sqrt{\frac{R}{2}+1}[G(r,R,\Lambda)+g^{\mathrm{leak}}(r,R,\Lambda)],\\
\|(\otimes_{j=0}^{N-1}\hat{S}_j(r)-\hat{\otimes}_{j=0}^{N-1}S^{\Lambda}_j(r))|C\rangle_{R,Q}\| &\le \epsilon^{(N)}_{\mathrm{trunc}}(r,R,\Lambda,Q) \coloneqq N\sqrt{N|c_{R,Q}|}[G(r,R,\Lambda)+g^{\mathrm{leak}}(r,R,\Lambda)].
\end{aligned} \tag{E22}$$

2. Trotter error

While we presented a state-dependent, i.e., (R,Q)-dependent, bound for the truncation error, it is common and often more straightforward to establish a state-independent bound for the Trotter error. Thus, we will bound,

$$\|\,|\psi^{\Lambda}\rangle_R - |\psi^{\Lambda,K}\rangle_R\| \le \|\hat{S}^{\Lambda,K}(r)-\hat{S}^{\Lambda}(r)\|, \tag{E23}$$

where $\|\hat{A}\|$ denotes the spectral norm of the operator which is induced by the inner-product norm $\||\psi\rangle\|$ on the full single-mode or multimode bosonic Hilbert space. The starting point is Eq. (53) of the main text, which is the first-order Trotter decomposition of $\hat{S}^{\Lambda,K}(r)$,

$$\hat{S}^{\Lambda,K}(r) = (e^{-i\frac{r}{K}\hat{s}^{\Lambda}_0}e^{-i\frac{r}{K}\hat{s}^{\Lambda}_2})^K(e^{-i\frac{r}{K}\hat{s}^{\Lambda}_1}e^{-i\frac{r}{K}\hat{s}^{\Lambda}_3})^K, \tag{E24}$$

with the $\hat{s}^{\Lambda}_{0,\ldots,3}$ operators defined in Eq. (52). Noting that for unitary operators $\hat{A}_0,\ldots,\hat{A}_{m-1},\hat{B}_0,\ldots,\hat{B}_{m-1}$,

$$\begin{aligned}
\|\hat{A}_0\ldots\hat{A}_{m-1}-\hat{B}_0\ldots\hat{B}_{m-1}\| &= \left\|\sum_{k=0}^{m-1}(\hat{A}_0\ldots\hat{A}_{k-1})(\hat{A}_k-\hat{B}_k)(\hat{B}_{k+1}\ldots\hat{B}_{m-1})\right\|\\
&\le \sum_{k=0}^{m-1}\|\hat{A}_1\ldots\hat{A}_{k-1}\|\|\hat{A}_k-\hat{B}_k\|\|\hat{B}_{k+1}\ldots\hat{B}_{m-1}\|\\
&= \sum_{k=0}^{m-1}\|\hat{A}_k-\hat{B}_k\|,
\end{aligned} \tag{E25}$$

one obtains,

$$\|\hat{S}^{\Lambda,K}(r)-\hat{S}^{\Lambda}(r)\| \le K(\|e^{-i\frac{r}{K}\hat{s}_0^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_2^{\Lambda}}-e^{-i\frac{r}{K}(\hat{s}_0^{\Lambda}+\hat{s}_2^{\Lambda})}\|+\|e^{-i\frac{r}{K}\hat{s}_1^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_3^{\Lambda}}-e^{-i\frac{r}{K}(\hat{s}_1^{\Lambda}+\hat{s}_3^{\Lambda})}\|). \tag{E26}$$

The Trotter error bound, following Ref. [117], is calculated as

$$\left\|e^{-i\frac{r}{K}\hat{s}_m^{\Lambda}}e^{-i\frac{r}{K}\hat{s}_{m+2}^{\Lambda}}-e^{-i\frac{r}{K}(\hat{s}_m^{\Lambda}+\hat{s}_{m+2}^{\Lambda})}\right\| \le \frac{r^2}{2K^2}\|[\hat{s}_m^{\Lambda},\hat{s}_{m+2}^{\Lambda}]\|, \tag{E27}$$

with $m=0,1$. This gives the following expression for the Trotter error bound for the single-mode squeezing operator

$$\|\hat{S}^{\Lambda,K}(r)-\hat{S}^{\Lambda}(r)\| \le \epsilon_{\mathrm{trott}}^{(1)}(r,\Lambda,K) := \frac{r^2}{2K}(\|[\hat{s}_0^{\Lambda},\hat{s}_2^{\Lambda}]\|+\|[\hat{s}_1^{\Lambda},\hat{s}_3^{\Lambda}]\|), \tag{E28}$$

and ultimately for the multimode product squeezing operator,

$$\left\|\bigotimes_{j=0}^{N-1}\hat{S}_j^{\Lambda,K}(r)-\bigotimes_{j=0}^{N-1}\hat{S}_j^{\Lambda}(r)\right\| \le \epsilon_{\mathrm{trott}}^{(N)}(r,\Lambda,K) := N\epsilon_{\mathrm{trott}}^{(1)}(r,\Lambda,K). \tag{E29}$$

While we will eventually compute the above commutator norms exactly to arrive at the bound, it is also useful to evaluate a looser but analytic bound. We first observe that the commutator evaluates to

$$\begin{aligned}[\hat{s}_m^{\Lambda},\hat{s}_{m+2}^{\Lambda}] &= \sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor}\sum_{n'=0}^{\lfloor\frac{\Lambda-4-m}{4}\rfloor}[\hat{T}_{4n+m},\hat{T}_{4n'+m+2}]\\ &= -\frac{1}{4}\sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor}\sum_{n'=0}^{\lfloor\frac{\Lambda-4-m}{4}\rfloor}\ell_{4n+m}\ell_{4n'+m+2}(|4n+m+2\rangle\langle 4n'+m+2|\delta_{n,n'+1}+|4n+m\rangle\langle 4n'+m+4|\delta_{n,n'})-\mathrm{H.c.}\\ &= -\frac{1}{4}\sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor}(\ell_{4n+m}\ell_{4n+m-2}|4n+m+2\rangle\langle 4n+m-2|+\ell_{4n+m}\ell_{4n+m+2}|4n+m\rangle\langle 4n+m+4|)-\mathrm{H.c.}\end{aligned} \tag{E30}$$

Thus, the norm of this commutator can be bounded as

$$\|[\hat{s}_m^{\Lambda},\hat{s}_{m+2}^{\Lambda}]\| \le \beta(\Lambda) := \frac{1}{2}\sum_{n=0}^{\lfloor\frac{\Lambda-2-m}{4}\rfloor}\ell_{4n+m}(\ell_{4n+m-2}+\ell_{4n+m+2}), \tag{E31}$$

which leads to

$$\left\|\bigotimes_{j=0}^{N-1}\hat{S}_j^{q,\Lambda,K}(r)-\bigotimes_{j=0}^{N-1}\hat{S}_j^{q,\Lambda}(r)\right\| \le \epsilon_{\mathrm{trott}}^{(N)}(r,\Lambda,K) \le \tilde{\epsilon}_{\mathrm{trott}}^{(N)}(r,\Lambda,K) := \frac{Nr^2\beta(\Lambda)}{2K}. \tag{E32}$$

Thus, $\epsilon_{\mathrm{trott}}^{(N)}(r,\Lambda,K)$ is $O(\frac{Nr^2\Lambda^3}{K})$.

3. Combined errors

The distance between the true state $|\psi\rangle_{R,Q}$ and the state obtained by the application of the truncated, Trotterized squeezing operator(s) satisfies

$$\||\psi\rangle_{R,Q}-|\psi^{\Lambda,K}\rangle_{R,Q}\|^2 = 2[1-\mathrm{Re}({}_{R,Q}\langle\psi|\psi^{\Lambda,K}\rangle_{R,Q})] \le (\epsilon^{(N)})^2. \tag{E33}$$

Since the core state, as well as the squeezing operator(s), are characterized by real coefficients, the overlap $\langle\psi|\psi^{\Lambda,K}\rangle$ is real. Further, we work in regimes where the net error $\epsilon^{(N)}<1$. This means that the fidelity $F=|\langle\psi|\psi^{\Lambda,K}\rangle|^2$ can be bounded as

$$F \ge \left[1-\frac{(\epsilon^{(N)})^2}{2}\right]^2. \tag{E34}$$

Similar results hold for the single-mode case with $|\psi\rangle_{R,Q}$ and $|\psi^{\Lambda,K}\rangle_{R,Q}$ replaced by $|\psi\rangle_R$ and $|\psi^{\Lambda,K}\rangle_R$, respectively, and $\epsilon^{(N)}$ replaced by $\epsilon^{(1)}$.

Assuming $\epsilon_{\mathrm{trunc}}^{(N)}=\epsilon_{\mathrm{trott}}^{(N)}$, we report in Table II of the main text, the smallest truncation Λ and the number of Trotter layers K that guarantee a fidelity $F\ge F_0$ based on the above computed bounds, for the system parameters considered in this work. The above estimates for the minimum value of cutoff and the number of Trotter layers result from

fairly conservative estimates of the errors, and thus, there is substantial scope for tightening these bounds. In particular, the Trotter error bound is state-independent, and a potentially tighter bound could be obtained by specializing the derivation to the highly constrained core state.

APPENDIX F: ON THE IMPLEMENTATION OF THE SQUEEZING OPERATOR

In Sec. V C, two implementation algorithms were introduced for implementing individual Trotter layers of the Trotterized squeezing operator in Eq. (53): the direct and SVD implementations. For a binary map of the bosonic register, direct implementation can introduce additional Trotter error (not accounted for by the analysis of the previous section), and is hence less favorable. An alternative method which works for both the unary and binary representations is based on the singular-value-decomposition (SVD) technique of Ref. [32]. We describe in this appendix how this algorithm can be applied to implement the squeezing operator. Additionally, the full squeezing operator can be implemented without Trotterization on a hybrid analog-digital device like that used in Ref [119]. We further describe this scheme in this appendix.

1. Implementation of the squeezing operator using a singular-value decomposition

Each operator $\hat{s}_m^{q,\Lambda}$ with $m \in \{0, \ldots, 3\}$ in Eq. (52) can be explicitly expressed as

$$\hat{s}_m^{q,\Lambda} = \hat{s}_m^{q,\Lambda,+} + \hat{s}_m^{q,\Lambda,-}, \tag{F1}$$

where

$$\begin{aligned} \hat{s}_m^{q,\Lambda,+} &:= \frac{i}{2} \sum_{n=0}^{\lfloor \frac{\Lambda-2-m}{4} \rfloor} \ell_{4n+m} |q^\Lambda(4n+m+2)\rangle\langle 4n+m|, \\ s_m^{q,\Lambda,-} &:= (s_m^{q,\Lambda,+})^\dagger. \end{aligned} \tag{F2}$$

Note importantly that $(s_m^{q,\Lambda,+})^2 = (s_m^{q,\Lambda,-})^2 = 0$. The above decomposition of $\hat{s}_m^{q,\Lambda}$ makes it amenable to an SVD implementation. Following Ref. [32], we introduce an ancilla qubit, and note that the operator $\hat{s}_m^{q,\Lambda} \otimes |0\rangle\langle 0|_{\text{anc}}$ can be decomposed as

$$\hat{s}_m^{q,\Lambda} |0\rangle\langle 0|_{\text{anc}} = \hat{\mathfrak{U}}^\dagger \hat{Z}_{\text{anc}} \hat{D}_m \hat{\mathfrak{U}}, \tag{F3}$$

where $\hat{Z}_{\text{anc}}$ is the Pauli-Z operator acting on the ancilla, $\hat{D}_m$ is a singular-value matrix which will be defined below, and the diagonalizing unitary is given by

$$\hat{\mathfrak{U}} = \hat{\mathrm{H}}_{\text{anc}} (|0\rangle\langle 0|_{\text{anc}} \hat{V}^\dagger + |1\rangle\langle 1|_{\text{anc}} \hat{W}^\dagger) \hat{P}. \tag{F4}$$

Here, $\hat{\mathrm{H}}_{\text{anc}}$ is the Hadamard gate acting on the ancilla qubit. The operators $\hat{V}$ and $\hat{W}$ feature in the SVD of $\hat{s}_a^{q,\Lambda,+}$

$$\begin{aligned} \hat{s}_m^{q,\Lambda,+} &= (\hat{I}^+)^2 \left(\frac{i}{2} \sum_{n=0}^{\lfloor \frac{\Lambda-2-m}{4} \rfloor} \ell_{4n+m} |q^\Lambda(4n+m)\rangle\langle q^\Lambda(4n+m)| \right) \hat{1} \\ &\equiv \hat{V} \hat{D}_m \hat{W}^\dagger, \end{aligned} \tag{F5}$$

with the singular-value matrix $\hat{D}_m$. Above, we have introduced the incrementor operator, $\hat{I}^+ |q^\Lambda(n)\rangle = (1 - \delta_{n,\Lambda}) |q^\Lambda(n+1)\rangle$ (with the decrementor operator being $\hat{I}^- |q^\Lambda(n)\rangle = (1 - \delta_{n,0}) |q^\Lambda(n-1)\rangle$). The operator $\hat{P}$ is defined by the action $\hat{P}|0\rangle_{\text{anc}} |\psi\rangle = |b\rangle_{\text{anc}} |\psi\rangle$, where the bit b is 0 when the state $|\psi\rangle \in \ker(\hat{s}^{q,\Lambda,+})$, and 1 otherwise. Thus, one finally arrives at

$$e^{-i\frac{r}{K}|0\rangle\langle 0|_{\text{anc}} \hat{s}_m^{q,\Lambda}} = \hat{\mathfrak{U}}^\dagger e^{-i\frac{r}{K}\hat{Z}_{\text{anc}} \hat{D}_m} \hat{\mathfrak{U}}, \tag{F6}$$

which implements the desired unitary operator acting on the bosonic qubit register upon initializing the ancilla qubit in state $|0\rangle_{\text{anc}}$.

The cost of implementing the diagonalizing transformation $\mathfrak{U}$ is dominated by the cost of the incrementor (controlled on the ancilla state). Using near-term strategies (which amount to introducing no or minimal additional ancilla qubits), this operation can be achieved with $O(n_q(\Lambda)^2)$ CNOT gates and $O(n_q(\Lambda)^2)$ single-qubit rotations [118]. A near-term strategy for implementing the diagonal unitary $\hat{D}_m$ is to decompose (using classical preprocessing) the exponent of the diagonal operator to a sum of products of Pauli Z operators acting on the qubits of the bosonic register (i.e., to find the Walsh-series expansion of the exponent [128]). In general, there are $2^{n_q(\Lambda)}$ such terms in the decomposition with up to $n_q(\Lambda)$ Pauli-Z operators in each term. Each of these terms can be exponentiated separately with no Trotter error, each costing $O(n_q(\Lambda))$ CNOT gates and single-qubit rotation gates. The total CNOT and single-qubit gate costs of implementing Eq. (F6) are, therefore, $O(n_q(\Lambda) 2^{n_q(\Lambda)})$ and $O(2^{n_q(\Lambda)})$, respectively. This exponential cost can be traded for a polynomial cost using far-term algorithms that employ a quantum phase-kickback routine and extra ancillary registers. Quantumly evaluating the coefficients ℓ_n can, in general, be costly [118]. Nonetheless, one can classically evaluate and store these values when acting on a given (R, Q) core state, and subsequently hardcode them in the quantum circuit.

Table III summarizes the cost associated with various (near-term) strategies for implementing each Trotter layer of the squeezing operator, i.e., $e^{-i\frac{r}{K}\hat{s}_m^{q,\Lambda}}$.

TABLE III. Summary of various qubit implementations of the single-mode squeezing operator. The above table details the complexity of a single Trotter layer. The Trotter error associated with all three methods is the same, and is discussed in the text and the previous appendix. When implementing a tensor product of such single-mode squeezing operators using the SVD method, one could reuse the ancilla.

Count	Direct (unary)	SVD (unary)	SVD (binary)
Qubits in bosonic register, i.e., $n_q(\Lambda)$	$\Lambda+1$	$\Lambda+1$	$\lceil \log_2(\Lambda+1)\rceil$
Ancilla qubits	0	1	1
CNOT gates per Trotter layer	$O(\Lambda)$	$O(\Lambda 2^{\Lambda})$	$O(\Lambda \log \Lambda)$
Single-qubit gates per Trotter layer	$O(\Lambda)$	$O(2^{\Lambda})$	$O(\Lambda)$

2. Implementation of the squeezing operator using a hybrid analog-digital simulator

Recall that according to Eq. (45) of the main text, the single-mode squeezing operator in the unary map can be expressed as

$$\hat{S}^{\mathrm{u},\Lambda}(r) = e^{\frac{r}{2}\sum_{n=0}^{\Lambda-2}\sqrt{(n+1)(n+2)}\left(|\mathrm{u}^{\Lambda}(n+2)\rangle\langle \mathrm{u}^{\Lambda}(n)|-|\mathrm{u}^{\Lambda}(n)\rangle\langle \mathrm{u}^{\Lambda}(n+2)|\right)}. \tag{F7}$$

Given the presentation of the unary-encoded bosonic state in Eq. (41), $\hat{S}^{\mathrm{u},\Lambda}$ can be regraded as evolution by a "squeezing Hamiltonian"

$$\hat{H}_{\mathrm{squeezing}} = \sum_{n=0}^{\Lambda-2}\sqrt{(n+1)(n+2)}(\hat{Y}_{n+2}\hat{X}_n - \hat{X}_{n+2}\hat{Y}_n) \tag{F8}$$

for time $r/4$.

In architectures such as that used in Ref. [119], tunable couplings between transmon qubits result in a nearest-neighbor XY Hamiltonian,

$$\hat{H}_{\mathrm{device}} = \sum_j \omega_j \hat{Z}_j + \sum_{\langle j,k\rangle}\frac{g_{j,k}}{2}(\hat{X}_j\hat{X}_k + \hat{Y}_j\hat{Y}_k) + \cdots, \tag{F9}$$

where the ellipsis denotes residual terms consisting of parametrically suppressed farther-neighbor $XX+YY$ terms. One can match the connectivity of the effective squeezing Hamiltonian to the connectivity of the device by mapping the unary encoding onto a ladder with legs 1 and 2. Next, one should match the Pauli content of the analog-device Hamiltonian by applying a site-dependent basis transformation.

The matching requires a new, modified unary representation, which we denote as ũ. In this modified unary representation, the mapping of states reads

$$\begin{aligned} |2j\rangle &\mapsto |\tilde{\mathrm{u}}^{\Lambda}(2j)\rangle = (-i)^j \hat{L}^{+}_{1,j}|0\rangle^{\otimes n_{\mathrm{u}}(\Lambda)}, \\ |2j+1\rangle &\mapsto |\tilde{\mathrm{u}}^{\Lambda}(2j+1)\rangle = (-i)^j \hat{L}^{+}_{2,j}|0\rangle^{\otimes n_{\mathrm{u}}(\Lambda)}, \end{aligned} \tag{F10}$$

where $\hat{L}^{+}_{1,j}$ and $\hat{L}^{+}_{2,j}$ are raising operators on leg 1 rung j and leg 2 rung j of the ladder, respectively. This modified representation is related to the previous unary representation u by an on site basis change. One can check that in this basis

$$\hat{X}_{1,j}\hat{X}_{1,j+1} + \hat{Y}_{1,j}\hat{Y}_{1,j+1} = 2i|\tilde{\mathrm{u}}^{\Lambda}(2j+2)\rangle\langle\tilde{\mathrm{u}}^{\Lambda}(2j)| + \mathrm{H.c.}, \tag{F11}$$

(and analogously for rung 2), as desired. Therefore, by setting $\omega_j = 0$ and tuning the couplings $g_{j,k}$ in experiment such that only nearest-neighbor transmons on the same leg couple, the squeezing Hamiltonian can be engineered. Consequently, evolving this system continuously for time t (with $g_{j,j+1} = \frac{r\ell_{2j}}{2t}$ on leg 1 and $g_{j,j+1} = \frac{r\ell_{2j+1}}{2t}$ on leg 2), implements the squeezing operator exactly, up to the errors associated with the residual terms.

In this work, the Gaussian part of the stellar *Ansatz* does not involve displacement operators due to symmetry constraints. However, as discussed in the next appendix, the implementation of the core state itself may involve displacements. Under the unary map, the displacement operator takes the form,

$$\hat{D}^{u,\Lambda}(\alpha) = e^{\sum_{n=0}^{\Lambda-1}\sqrt{n+1}(\alpha\hat{L}^{+,u,\Lambda}_{n+1}\hat{L}^{-,u,\Lambda}_{n} - \alpha^*\hat{L}^{+,u,\Lambda}_{n}\hat{L}^{-,u,\Lambda}_{n+1})}. \tag{F12}$$

To implement this operator via a unary mapping to the ladder geometry above, thus, requires not only couplings between legs at the same rung, but also diagonal couplings $2, j \to 1, j+1$. Those diagonal couplings between qubits would require either diagonal couplings between transmons, or (potentially) careful Floquet engineering of the residual second-neighbor terms in the device Hamiltonian $\hat{H}_{\mathrm{device}}$.

APPENDIX G: ON THE CONTINUOUS-VARIABLE APPROACH TO IMPLEMENTING THE (R,Q) *ANSATZ*

In this appendix, we discuss what ingredients are needed to implement our finite-rank R, span-Q *Ansatz* on a continuous-variable quantum platform. The advantage of continuous-variable platforms is that the scalar field can be faithfully encoded into the simulator's degrees of freedom without the need for any Hilbert-space truncation. Thus, one can avoid the truncation errors encountered in discrete-variable platforms, and the need to perform any extrapolation to the infinite truncation cutoff. Quantum simulations of quantum field theories with continuous-variable

platforms have gained significant attention [43,129–133]. Bosonic degrees of freedom occur in photonic quantum simulators, phononic modes in trapped ion systems, and electromagnetic fields in superconducting microwave cavities and circuit QED. Here, we assume a generic bosonic platform that can perform:

(1) *Single-mode Gaussian operations*: These include the squeezing $\hat{S}(\xi) = e^{\frac{1}{2}[\xi(\hat{a}^\dagger)^2 - \xi^* \hat{a}^2]}$ and displacement $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^* \hat{a}}$ operators with $\xi, \alpha \in \mathbb{C}$.
(2) *Multimode Gaussian operations*: These include passive rotations $R(\Phi) = e^{i\sum_{j,j'=0}^{N-1} \hat{a}_j^\dagger \Phi_{j,j'} \hat{a}_{j'}}$.
(3) *Single-boson additions* $\hat{a}_j^\dagger$.

Among the above Gaussian operations, single-mode displacements and passive rotations are often easier to perform than single-mode squeezing. For example, in a photonic device, the latter is limited to smaller squeezing-parameter magnitudes $|\xi| = r$. Roughly, this is because as r increases, the mean photon number in the squeezed mode grows as $\langle \hat{a}^\dagger \hat{a} \rangle = \sinh^2(r)$. States with such a large mean photon number are significantly more susceptible to decoherence from photon loss, and interactions with the environment. Gaussian operations form a continuous-variable analog of the Clifford group, and they can be efficiently simulated classically [134,135]. (In the field of quantum optics, squeezed states are often referred to as nonclassical because they exhibit reduced-noise quadratures. Yet, they possess Gaussian Wigner functions and *can* be simulated efficiently classically.) Thus, non-Gaussian operations are needed to achieve universal continuous-variable quantum computing. The non-Gaussian boson-addition primitive is considerably more demanding than any Gaussian gate because it typically relies on probabilistic schemes [136] or requires strong, precisely controlled interactions with ancillary systems (like qubits or other bosons), which are difficult to implement with high fidelity and efficiency [137]. In contrast, as demonstrated in Sec. V A, preparing our (R, Q) *Ansatz* core state on a discrete-variable quantum computer is rather straightforward.

We will now explore the feasibility of implementing the (R, Q) *Ansatz* with the above operations. Following Ref. [93], the single-mode finite-rank *Ansatz* can be decomposed into the above elementary operators as

$$\begin{aligned}|\psi\rangle_R &= \hat{S}(r)(c_0|0\rangle + c_2|2\rangle + \cdots + c_R|R\rangle) \\ &= \hat{S}(r)(c_0' + c_2'(\hat{a}^\dagger)^2 + \cdots + c_R'(\hat{a}^\dagger)^R)|0\rangle \\ &= \hat{S}(r)c_R' \prod_{n=0,2,\ldots,R} (\hat{a}^\dagger - \beta_n)|0\rangle \\ &= \hat{S}(r)c_R' \prod_{n=0,2,\ldots,R} \hat{D}(\beta_n^*)\hat{a}^\dagger \hat{D}^\dagger(\beta_n^*)|0\rangle, \end{aligned} \tag{G1}$$

where $\beta_n \in \mathbb{C}$ are the roots of the core-state polynomial $C(\hat{a}^\dagger) = c_0' + c_2'(\hat{a}^\dagger)^2 + \cdots + c_R'(\hat{a}^\dagger)^R$. Therefore, the single-mode core state can be prepared by performing a series of displaced single-boson additions on the vacuum state. The value of these displacements is given by the roots of the core-state polynomial. Thereafter, the full symmetric rank-R *Ansatz* can be prepared by applying the real squeezing operator to this core state.

Consider now the multimode case. Just like the single-mode case, a multimode core state is generated by a polynomial in the creation operators acting on the multimode vacuum, i.e., $|C\rangle = C(\hat{\boldsymbol{a}}^\dagger)|\mathbf{0}\rangle$. Since $C(\hat{\boldsymbol{a}}^\dagger) = C(a_0^\dagger, \ldots, \hat{a}_{N-1}^\dagger)$ is a multivariate polynomial, in general, it does not possess a discrete set of roots (unlike the single-mode core-state polynomial described above). Consequently, in general, an arbitrary multimode core state cannot be prepared by performing displaced single-boson additions to the vacuum. However, as shown in Ref. [138], the above single-mode state-preparation recipe can be generalized to prepare a smaller class of multimode finite-rank states. These states are prepared by interleaving single-boson additions with multimode Gaussian operations

$$\begin{aligned}|\psi\rangle_R^{\mathrm{IPAG}} &\coloneqq \hat{U}_G^{(R)} \hat{a}_0^\dagger \hat{U}_G^{(R-1)} \hat{a}_0^\dagger \cdots \hat{U}_G^{(1)} \hat{a}_0^\dagger \hat{U}_G^{(0)} |\mathbf{0}\rangle \\ &= (\hat{U}_G^{(R)} \cdots \hat{U}_G^{(0)})(\hat{U}_G^{(0)\dagger} \cdots \hat{U}_G^{(R-1)\dagger} \hat{a}_0^\dagger \hat{U}_G^{(R-1)} \cdots \hat{U}_G^{(0)}) \cdots (\hat{U}_G^{(0)\dagger} \hat{U}_G^{(1)\dagger} \hat{a}_0^\dagger \hat{U}_G^{(1)} \hat{U}_G^{(0)})(\hat{U}_G^{(0)\dagger} \hat{a}_0^\dagger \hat{U}_G^{(0)})|\mathbf{0}\rangle \\ &= \hat{U}_G \left[\alpha_0^{(R-1)} + \sum_{j=0}^{N-1} (u_{0,j}^{(R-1)} \hat{a}_j^\dagger + v_{0,j}^{(R-1)} \hat{a}_j)\right] \ldots \left[\alpha_0^{(0)} + \sum_{j=0}^{N-1} (u_{0,j}^{(0)} \hat{a}_j^\dagger + v_{0,j}^{(0)} \hat{a}_j)\right] |\mathbf{0}\rangle \\ &= \hat{U}_G |C\rangle_R^{\mathrm{IPAG}}, \end{aligned} \tag{G2}$$

where $\hat{U}_G^{(n)}$ with $n \in \{0, 1, \ldots, R\}$ are multimode Gaussian unitary operations, and the Gaussian $\hat{U}_G$ is defined as $\hat{U}_G \coloneqq \hat{U}_G^{(R)} \hat{U}_G^{(R-1)} \cdots \hat{U}_G^{(0)}$. In the third line, we have defined the action of the Gaussian operations $(\hat{U}_G^{(n)} \ldots \hat{U}_G^{(0)})^\dagger$, with $n \in \{0, \ldots, R\}$, by the Boguliubov transformation

$$\hat{A}_j^{(n)\dagger} \coloneqq \left(\hat{U}_G^{(n)} \cdots \hat{U}_G^{(0)}\right)^\dagger \hat{a}_j^\dagger \left(\hat{U}_G^{(n)} \cdots \hat{U}_G^{(0)}\right) = \alpha_j^{(n)} + \sum_{j'=0}^{N-1} \left(u_{j,j'}^{(n)} \hat{a}_{j'}^\dagger + v_{j,j'}^{(n)} \hat{a}_j\right). \tag{G3}$$

Here, $\alpha_j^{(n)} \in \mathbb{C}$ are complex displacement parameters while the matrices $U^{(n)} = [u_{j,j'}^{(n)}]$ and $V^{(n)} = [v_{j,j'}^{(n)}]$ with $u, v \in \mathbb{C}^{N\times N}$ satisfy $U^{(n)}U^{(n)\dagger} - V^{(n)}V^{(n)\dagger} = \mathbb{I}$ and $U^{(n)}V^{(n)T} = V^{(n)}U^{(n)T}$ to ensure the bosonic commutation relations $[A_j^{(n)}, A_{j'}^{(n)\dagger}] = \delta_{j,j'}$ and $[A_j^{(n)}, A_{j'}^{(n)}] = [A_j^{(n)\dagger}, A_{j'}^{(n)\dagger}] = 0$. States which exhibit the above decomposition are referred to as rank-R interleaved photon-added Gaussian (IPAG) states (photon-added since these were originally introduced in the context of quantum optics [138]). Similar to the single-mode case, the IPAG *Ansatz* can be prepared using a series of single-boson additions sandwiched between (multimode) Gaussian operations. The interleaved multimode Gaussian operations can be implemented by means of the Bloch-Messiah decomposition [96], which involves single-mode squeezings, single-mode displacements, and passive rotations.

As mentioned above, rank-R IPAG states constitute a strictly smaller subspace of rank-R states. This was shown in Ref. [138] by explicitly constructing a rank-2 state that cannot be written in the IPAG form. Thus, we must determine which (R, Q) states (if any) belong to the rank-R IPAG subspace. This amounts to checking if there exists a choice of Gaussian unitary transformations $\hat{U}_G^{(0)}, \ldots, \hat{U}_G^{(R)}$ in the IPAG state which makes it equal to the (R, Q) state of interest. While making general statements about the ability of an arbitrary (R, Q) state to be represented by an IPAG state is challenging, we provide indications for the simpler case of $R = 2$ with a simpler IPAG state with vanishing displacements, i.e., $\alpha_0^{(n)} = 0$. With these choices, the rank-2 IPAG core state has the form

$$\begin{aligned}
|C\rangle_2^{\mathrm{IPAG}} &= \left[\sum_{j=0}^{N-1}(u_{0,j}^{(1)}\hat{a}_j^\dagger + v_{0,j}^{(1)}\hat{a}_j)\right]\left[\sum_{j'=0}^{N-1}(u_{0,j'}^{(0)}\hat{a}_{j'}^\dagger + v_{0,j'}^{(0)}\hat{a}_{j'})\right]|\mathbf{0}\rangle \\
&= \left[\sum_{j=0}^{N-1} v_{0,j}^{(1)}u_{0,j}^{(0)} + \sum_{j=0}^{N-1} u_{0,j}^{(1)}u_{0,j}^{(0)}\hat{a}_j^{\dagger 2} + \sum_{j=0}^{N-1}\sum_{j'=0}^{j-1}\{u_{0,j}^{(1)}u_{0,j'}^{(0)} + u_{0,j'}^{(1)}u_{0,j}^{(0)}\}\hat{a}_j^\dagger a_{j'}^\dagger\right]|\mathbf{0}\rangle \\
&= \left[\sum_{j=0}^{N-1} v_{0,j}^{(1)}u_{0,j}^{(0)} + \sum_{j=0}^{N-1} u_{0,j}^{(1)}u_{0,j}^{(0)}\hat{a}_j^{\dagger 2} + \sum_{q=1}^{\frac{N}{2}}\sum_{j=0}^{N-1}\{u_{0,j}^{(1)}u_{0,j+q}^{(0)} + u_{0,j+q}^{(1)}u_{0,j}^{(0)}\}\hat{a}_j^\dagger a_{j+q}^\dagger\right]|\mathbf{0}\rangle,
\end{aligned} \tag{G4}$$

while the $(R, Q) = (2, Q)$ (symmetric) core state takes the form

$$|C\rangle_{2,Q} = \left[A + \sum_{q=0}^{Q} b_q \sum_{j=0}^{N-1}\hat{a}_j^\dagger\hat{a}_{j+q}^\dagger\right]|\mathbf{0}\rangle. \tag{G5}$$

To make the above states equal to each other, first, we must have $u_{0,j}^{(1)}u_{0,j}^{(0)} = b_0$ with $j \in \{0, \ldots, N-1\}$. Assuming $b_0 \neq 0$, it follows that $u_{0,j}^{(1)}, u_{0,j}^{(0)} \neq 0$. Now, consider the equivalence of the coefficients of the terms $\hat{a}_j^\dagger\hat{a}_{j+1}^\dagger$ in the $(2, Q)$ state,

$$u_{0,j}^{(1)}u_{0,j+1}^{(0)} + u_{0,j+1}^{(1)}u_{0,j}^{(0)} = \frac{b_0}{u_{0,j}^{(0)}}u_{0,j+1}^{(0)} + \frac{b_0}{u_{0,j+1}^{(0)}}u_{0,j}^{(0)} = b_1, \tag{G6}$$

which simplifies to

$$(u_{0,j+1}^{(0)})^2 + (u_{0,j}^{(0)})^2 = b_{1,0}u_{0,j}^{(0)}u_{0,j+1}^{(0)}, \tag{G7}$$

with $j \in \{0, \ldots, N-1\}$ and $b_{1,0} \coloneqq \frac{b_1}{b_0}$. Thus, the problem is reduced to the problem of N beads, $u_{0,0}^{(0)}, \ldots, u_{0,N-1}^{(0)}$, arranged on a closed necklace obeying an identical quadratic nearest-neighbor gluing rule $f_1(u_{0,j}^{(0)}, u_{0,j+1}^{(0)}) = 0$ with $f_1(x, y) \coloneqq x^2 + y^2 - b_{1,0}xy$. Thus, one may write $u_{0,j+1}^{(0)} = \lambda_{\sigma_j}u_{0,j}^{(0)}$, where $\lambda_\pm = \frac{b_{1,0}\pm\sqrt{b_{1,0}^2-4}}{2}$ and $\sigma_j \in \{\pm 1\}$. The closed ring condition is given by $\prod_{j=0}^{N-1}\lambda_{\sigma_j} = 1$. Since $\lambda_+\lambda_- = 1$, this condition is satisfied when exactly $\frac{N}{2}$ of the σ_j values are chosen to be $+1$. Once such a suitable choice of signs is picked, $u_{0,j}^{(0)} = u_{0,0}^{(0)}\lambda_{\sigma_0}\cdots\lambda_{\sigma_j}$ is a solution with arbitrary $u_{0,0}^{(0)} \in \mathbb{C}$.

The next step is to demand the equivalence of the $q = 2$ term, i.e., to consider the coefficients of the terms $\hat{a}_j^\dagger\hat{a}_{j+2}^\dagger$. This amounts to imposing the next-to-nearest-neighbor gluing constraint given by

$$(u_{0,j+2}^{(0)})^2 + (u_{0,j}^{(0)})^2 = b_{2,0}u_{0,j}^{(0)}u_{0,j+2}^{(0)}, \tag{G8}$$

where $b_{2,0} \coloneqq \frac{b_2}{b_0}$. Substituting the form of the solution obtained earlier, one gets

$$u_{0,0}^{(0)}(\lambda_{\sigma_{j+2}}^2 - b_{2,0}\lambda_{\sigma_{j+2}} + 1) = 0. \tag{G9}$$

Thus, a nontrivial solution only exists when the coefficients b_2, as well as b_1 and b_0 (which determine the λ_σ factors), of the (R, Q) *Ansatz* obey the above constraint. In other words, we cannot find a solution for a generic $(R = 2, Q)$ *Ansatz*.

This result should not be too surprising since the IPAG is a very restricted class of states. It may be more illuminating to consider a symmetric form of the IPAG state, or to directly use the IPAG state as the ground-state *Ansatz*. We did not consider this latter possibility in the present work given the greater classical cost of optimizing the IPAG states as opposed to the (R, Q) states. This direction, nonetheless, is worth exploring since the IPAG states greatly simplify circuit translation and implementation on bosonic quantum platforms.

[1] K. G. Wilson, Confinement of quarks, Phys. Rev. D **10**, 2445 (1974).
[2] M. Creutz, L. Jacobs, and C. Rebbi, Monte Carlo computations in lattice gauge theories, Phys. Rep. **95**, 201 (1983).
[3] I. Montvay and G. Münster, *Quantum Fields on a Lattice* (Cambridge University Press, Cambridge, England, 1994), 10.1017/CBO9780511470783.
[4] H. J. Rothe, *Lattice Gauge Theories: An Introduction* (World Scientific, Singapore, 2012), 10.1142/8229.
[5] C. Gattringer and C. Lang, *Quantum Chromodynamics on the Lattice: An Introductory Presentation*, Vol. 788 (Springer Science & Business Media, New York, 2009), 10.1007/978-3-642-01850-3.
[6] Y. Aoki, T. Blum, S. Collins, L. Del Debbio, M. Della Morte, P. Dimopoulos, X. Feng, M. Golterman, S. Gottlieb, R. Gupta *et al.*, FLAG review 2024, Phys. Rev. D **113**, 014508 (2026).
[7] U. Collaboration, A. Bazavov, F. Karsch, S. Mukherjee, and P. Petreczky, Hot-dense lattice QCD, Eur. Phys. J. A **55**, 1 (2019).
[8] Z. Davoudi, E. T. Neil, C. W. Bauer, T. Bhattacharya, T. Blum, P. Boyle, R. C. Brower, S. Catterall, N. H. Christ, V. Cirigliano *et al.*, Report of the Snowmass 2021 topical group on lattice gauge theory, arXiv:2209.10758.
[9] A. S. Kronfeld, T. Bhattacharya, T. Blum, N. H. Christ, C. DeTar, W. Detmold, R. Edwards, A. Hasenfratz, H.-W. Lin, S. Mukherjee *et al.*, Lattice QCD and particle physics, arXiv:2207.07641.
[10] M. Lüscher, Two-particle states on a torus and their relation to the scattering matrix, Nucl. Phys. **B354**, 531 (1991).
[11] L. Lellouch and M. Lüscher, Weak transition matrix elements from finite-volume correlation functions, Commun. Math. Phys. **219**, 31 (2001).
[12] R. A. Briceno, Z. Davoudi, and T. C. Luu, Nuclear reactions from lattice QCD, J. Phys. G **42**, 023101 (2015).
[13] R. A. Briceno, J. J. Dudek, and R. D. Young, Scattering processes and resonances from lattice QCD, Rev. Mod. Phys. **90**, 025001 (2018).
[14] M. T. Hansen and S. R. Sharpe, Lattice QCD and three-particle decays of resonances, Annu. Rev. Nucl. Part. Sci. **69**, 65 (2019).
[15] Z. Davoudi, W. Detmold, P. Shanahan, K. Orginos, A. Parreno, M. J. Savage, and M. L. Wagman, Nuclear matrix elements from lattice QCD for electroweak and beyond-standard-model processes, Phys. Rep. **900**, 1 (2021).
[16] C. Gattringer and K. Langfeld, Approaches to the sign problem in lattice field theory, Int. J. Mod. Phys. A **31**, 1643007 (2016).
[17] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, Taming the dynamical sign problem in real-time evolution of quantum many-body problems, Phys. Rev. Lett. **115**, 266802 (2015).
[18] M. Troyer and U.-J. Wiese, Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations, Phys. Rev. Lett. **94**, 170201 (2005).
[19] V. Goy, V. Bornyakov, D. Boyda, A. Molochkov, A. Nakamura, A. Nikolaev, and V. Zakharov, Sign problem in finite density lattice QCD, Prog. Theor. Exp. Phys. **2017**, 031D01 (2017).
[20] K. Nagata, Finite-density lattice QCD and sign problem: Current status and open problems, Prog. Part. Nucl. Phys. **127**, 103991 (2022).
[21] N. Klco, A. Roggero, and M. J. Savage, Standard Model physics and the digital quantum revolution: Thoughts about the interface, Rep. Prog. Phys. **85**, 064301 (2022).
[22] C. W. Bauer, Z. Davoudi, N. Klco, and M. J. Savage, Quantum simulation of fundamental particles and forces, Nat. Rev. Phys. **5**, 420 (2023).
[23] C. W. Bauer, Z. Davoudi, A. B. Balantekin, T. Bhattacharya, M. Carena, W. A. de Jong, P. Draper, A. El-Khadra, N. Gemelke, M. Hanada *et al.*, Quantum simulation for high-energy physics, PRX Quantum **4**, 027001 (2023).
[24] A. Di Meglio, K. Jansen, I. Tavernelli, C. Alexandrou, S. Arunachalam, C. W. Bauer, K. Borras, S. Carrazza, A. Crippa, V. Croft *et al.*, Quantum computing for high-energy physics: State of the art and challenges, PRX Quantum **5**, 037001 (2024).
[25] D. Beck, J. Carlson, Z. Davoudi, J. Formaggio, S. Quaglioni, M. Savage, J. Barata, T. Bhattacharya, M. Bishof, I. Cloet *et al.*, Quantum information science and technology for nuclear physics. input into us long-range planning, arXiv:2303.00113.
[26] T. Byrnes and Y. Yamamoto, Simulating lattice gauge theories on a quantum computer, Phys. Rev. A **73**, 022328 (2006).
[27] A. F. Shaw, P. Lougovski, J. R. Stryker, and N. Wiebe, Quantum algorithms for simulating the lattice Schwinger model, Quantum **4**, 306 (2020).
[28] A. Ciavarella, N. Klco, and M. J. Savage, Trailhead for quantum simulation of SU(3) Yang-Mills lattice gauge theory in the local multiplet basis, Phys. Rev. D **103**, 094501 (2021).

[29] A. Kan and Y. Nam, Lattice quantum chromodynamics and electrodynamics on a universal quantum computer, arXiv:2107.12769.
[30] H. Lamm, S. Lawrence, and Y. Yamauchi (NuQS Collaboration), General methods for digital quantum simulation of gauge theories, Phys. Rev. D **100**, 034518 (2019).
[31] J. F. Haase, L. Dellantonio, A. Celi, D. Paulson, A. Kan, K. Jansen, and C. A. Muschik, A resource efficient approach for quantum and classical simulations of gauge theories in particle physics, Quantum **5**, 393 (2021).
[32] Z. Davoudi, A. F. Shaw, and J. R. Stryker, General quantum algorithms for Hamiltonian simulation with applications to a non-Abelian lattice gauge theory, Quantum **7**, 1213 (2023).
[33] E. M. Murairi, M. J. Cervia, H. Kumar, P. F. Bedaque, and A. Alexandru, How many quantum gates do gauge theories require?, Phys. Rev. D **106**, 094504 (2022).
[34] M. L. Rhodes, M. Kreshchuk, and S. Pathak, Exponential improvements in the simulation of lattice gauge theories using near-optimal techniques, PRX Quantum **5**, 040347 (2024).
[35] H. Lamm, Y.-Y. Li, J. Shu, Y.-L. Wang, and B. Xu, Block encodings of discrete subgroups on a quantum computer, Phys. Rev. D **110**, 054505 (2024).
[36] P. Balaji, C. Conefrey-Shinozaki, P. Draper, J. K. Elhaderi, D. Gupta, L. Hidalgo, A. Lytle, and E. Rinaldi, Quantum circuits for SU(3) lattice gauge theory, Phys. Rev. D **112**, 054511 (2025).
[37] S. P. Jordan, K. S. Lee, and J. Preskill, Quantum computation of scattering in scalar quantum field theories, Quantum Inf. Comput. **14**, 1014 (2014).
[38] S. P. Jordan, K. S. Lee, and J. Preskill, Quantum algorithms for quantum field theories, Science **336**, 1130 (2012).
[39] S. P. Jordan, H. Krovi, K. S. Lee, and J. Preskill, BQP-completeness of scattering in scalar quantum field theory, Quantum **2**, 44 (2018).
[40] A. D. Bookatz, QMA-complete problems, Quantum Inf. Comput. **14**, 6 (2014); arXiv:1212.6312.
[41] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, SIAM J. Comput. **35**, 1070 (2006).
[42] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Soc., Providence, 2002), 10.1090/gsm/047.
[43] K. Marshall, R. Pooser, G. Siopsis, and C. Weedbrook, Quantum simulation of quantum field theory using continuous variables, Phys. Rev. A **92**, 063825 (2015).
[44] M. Bagherimehrab, Y. R. Sanders, D. W. Berry, G. K. Brennen, and B. C. Sanders, Nearly optimal quantum algorithm for generating the ground state of a free quantum field theory, PRX Quantum **3**, 020364 (2022).
[45] N. Klco and M. J. Savage, Minimally entangled state preparation of localized wave functions on quantum computers, Phys. Rev. A **102**, 012612 (2020).
[46] N. Klco and M. J. Savage, Systematically localizable operators for quantum simulations of quantum field theories, Phys. Rev. A **102**, 012619 (2020).
[47] N. Klco and M. J. Savage, Fixed-point quantum circuits for quantum field theories, Phys. Rev. A **102**, 052422 (2020).
[48] A. Kitaev and W. A. Webb, Wavefunction preparation and resampling using a quantum computer, arXiv:0801.0342.
[49] H. Nishi, T. Kosugi, and Y.-i. Matsushita, Implementation of quantum imaginary-time evolution method on nisq devices by introducing nonlocal approximation, npj Quantum Inf. **7**, 85 (2021).
[50] Y. Ge, J. Tura, and J. I. Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits, J. Math. Phys. (N.Y.) **60** (2019).
[51] Y. Dong, L. Lin, and Y. Tong, Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices, PRX Quantum **3**, 040305 (2022).
[52] L. Lin and Y. Tong, Near-optimal ground state preparation, Quantum **4**, 372 (2020).
[53] K. Choi, D. Lee, J. Bonitati, Z. Qian, and J. Watkins, Rodeo algorithm for quantum computing, Phys. Rev. Lett. **127**, 040505 (2021).
[54] T. D. Cohen and H. Oh, Efficient vacuum-state preparation for quantum simulation of strongly interacting local quantum field theories, Phys. Rev. A **109**, L020402 (2024).
[55] A. Luis, Quantum-state preparation and control via the Zeno effect, Phys. Rev. A **63**, 052112 (2001).
[56] Y. Zhan, Z. Ding, J. Huhn, J. Gray, J. Preskill, G. K. Chan, and L. Lin, Rapid quantum ground state preparation via dissipative dynamics, Phys. Rev. X **16**, 011004 (2026).
[57] Z. Ding, C.-F. Chen, and L. Lin, Single-ancilla ground state preparation via Lindbladians, Phys. Rev. Res. **6**, 033147 (2024).
[58] C. Feniou, O. Adjoua, B. Claudon, J. Zylberman, E. Giner, and J.-P. Piquemal, Sparse quantum state preparation for strongly correlated systems, J. Phys. Chem. Lett. **15**, 3197 (2024).
[59] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth *et al.*, The variational quantum eigensolver: A review of methods and best practices, Phys. Rep. **986**, 1 (2022).
[60] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).
[61] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley *et al.* (Google AI Quantum and Collaborators), Hartree-Fock on a superconducting qubit quantum computer, Science **369**, 1084 (2020).
[62] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, Quantum Sci. Technol. **4**, 014008 (2018).
[63] A. Anand, P. Schleich, S. Alperin-Lea, P. W. Jensen, S. Sim, M. Díaz-Tinoco, J. S. Kottmann, M. Degroote, A. F. Izmaylov, and A. Aspuru-Guzik, A quantum computing view on unitary coupled cluster theory, Chem. Soc. Rev. **51**, 1659 (2022).
[64] A. Romero, J. Engel, H. L. Tang, and S. E. Economou, Solving nuclear structure problems with the adaptive variational quantum algorithm, Phys. Rev. C **105**, 064317 (2022).

[65] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Cloud quantum computing of an atomic nucleus, Phys. Rev. Lett. **120**, 210501 (2018).
[66] O. Kiss, M. Grossi, P. Lougovski, F. Sanchez, S. Vallecorsa, and T. Papenbrock, Quantum computing of the ^{6}Li nucleus via ordered unitary coupled clusters, Phys. Rev. C **106**, 034325 (2022).
[67] S. Rethinasamy, E. Guo, A. Wei, M. M. Wilde, and K. D. Launey, Neutron-nucleus dynamics simulations for quantum computers, Quantum Sci. Technol. **11**, 015052 (2026).
[68] N. Klco, E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage, Quantum-classical computation of Schwinger model dynamics using quantum computers, Phys. Rev. A **98**, 032331 (2018).
[69] Y. Y. Atas, J. Zhang, R. Lewis, A. Jahanpour, J. F. Haase, and C. A. Muschik, SU(2) hadrons on a quantum computer via a variational approach, Nat. Commun. **12**, 6499 (2021).
[70] R. C. Farrell, M. Illa, A. N. Ciavarella, and M. J. Savage, Scalable circuits for preparing ground states on digital quantum computers: The Schwinger model vacuum on 100 qubits, PRX Quantum **5**, 020315 (2024).
[71] A. Crippa, K. Jansen, and E. Rinaldi, Analysis of the confinement string in $(2+1)$-dimensional quantum electrodynamics with a trapped-ion quantum computer, arXiv:2411.05628.
[72] M. Fromm, O. Philipsen, M. Spannowsky, and C. Winterowd, Simulating Z_2 lattice gauge theory with the variational quantum thermalizer, Eur. Phys. J. Quantum Technol. **11**, 1 (2024).
[73] R. C. Farrell, M. Illa, A. N. Ciavarella, and M. J. Savage, Quantum simulations of hadron dynamics in the Schwinger model using 112 qubits, Phys. Rev. D **109**, 114510 (2024).
[74] Z. Davoudi, C.-C. Hsieh, and S. V. Kadam, Scattering wave packets of hadrons in gauge theories: Preparation on a quantum computer, Quantum **8**, 1520 (2024).
[75] Z. Davoudi, C.-C. Hsieh, and S. V. Kadam, Quantum computation of hadron scattering in a lattice gauge theory, arXiv:2505.20408.
[76] A. T. Than, Y. Y. Atas, A. Chakraborty, J. Zhang, M. T. Diaz, K. Wen, X. Liu, R. Lewis, A. M. Green, C. A. Muschik *et al.*, The phase diagram of quantum chromodynamics in one dimension on a quantum computer, Nat. Commun. **16**, 10288 (2025).
[77] X.-D. Xie, X. Guo, H. Xing, Z.-Y. Xue, D.-B. Zhang, and S.-L. Zhu (QuNu Collaboration), Variational thermal quantum simulation of the lattice Schwinger model, Phys. Rev. D **106**, 054509 (2022).
[78] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, A review of barren plateaus in variational quantum computing, Nat. Rev. Phys. **7**, 174 (2025).
[79] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio *et al.*, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).
[80] E. R. Anschuetz and B. T. Kiani, Quantum variational algorithms are swamped with traps, Nat. Commun. **13**, 7760 (2022).
[81] N. Klco and M. J. Savage, Digitization of scalar fields for quantum computing, Phys. Rev. A **99**, 052335 (2019).
[82] J. Liu, Z. Li, H. Zheng, X. Yuan, and J. Sun, Towards a variational Jordan-Lee-Preskill quantum algorithm, Mach. Learn. Sci. Technol. **3**, 045030 (2022).
[83] J. Barata, N. Mueller, A. Tarasov, and R. Venugopalan, Single-particle digitization strategy for quantum computation of a ϕ^4 scalar field theory, Phys. Rev. A **103**, 042410 (2021).
[84] A. Hardy, P. Mukhopadhyay, M. S. Alam, R. Konik, L. Hormozi, E. Rieffel, S. Hadfield, J. Barata, R. Venugopalan, D. E. Kharzeev *et al.*, Scattering processes from quantum simulation algorithms for scalar field theories, PRX Quantum **7**, 010343 (2026).
[85] K. Yeter-Aydeniz, E. F. Dumitrescu, A. J. McCaskey, R. S. Bennink, R. C. Pooser, and G. Siopsis, Scalar quantum field theories as a benchmark for near-term quantum computers, Phys. Rev. A **99**, 032306 (2019).
[86] S. Thompson and G. Siopsis, Quantum computation of phase transition in interacting scalar quantum field theory, Quantum Inf. Process. **22**, 396 (2023).
[87] A. C. Li, A. Macridin, S. Mrenna, and P. Spentzouris, Simulating scalar field theories on quantum computers with limited resources, Phys. Rev. A **107**, 032603 (2023).
[88] M. Turco, G. Quinta, J. Seixas, and Y. Omar, Quantum simulation of bound state scattering, PRX Quantum **5**, 020311 (2024).
[89] N. A. Zemlevskiy, Scalable quantum simulations of scattering in scalar field theory on 120 qubits, Phys. Rev. D **112**, 034502 (2025).
[90] J. Ingoldby, M. Spannowsky, T. Sypchenko, S. Williams, and M. Wingate, Real-time scattering on quantum computers via Hamiltonian truncation, arXiv:2505.03878.
[91] J. D. Watson, J. Bringewatt, A. F. Shaw, A. M. hilds, A. V. Gorshkov, and Z. Davoudi, Quantum algorithms for simulating nuclear effective field theories, arXiv:2312.05344.
[92] W. Loinaz and R. Willey, Monte Carlo simulation calculation of the critical coupling constant for two-dimensional continuum φ^4 theory, Phys. Rev. D **58**, 076003 (1998).
[93] U. Chabaud, D. Markham, and F. Grosshans, Stellar representation of non-Gaussian quantum states, Phys. Rev. Lett. **124**, 063605 (2020).
[94] U. Chabaud and S. Mehraban, Holomorphic representation of quantum computations, Quantum **6**, 831 (2022).
[95] Š. Blahnik and S. Shandera, Natural disorder distributions from measurement, arXiv:2405.02214.
[96] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods* (CRC Press, Boca Raton, 2023).
[97] X. Han, S. A. Hartnoll, and J. Kruthoff, Bootstrapping matrix quantum mechanics, Phys. Rev. Lett. **125**, 041601 (2020).
[98] D. Berenstein and G. Hulsey, Bootstrapping simple QM systems, arXiv:2108.08757.
[99] Z. Ozzello and Y. Meurice, Bootstrap methods for digitized scalar field theory, Proc. Sci., LATTICE2023 (**2024**) 163 [arXiv:2312.10559].

[100] P. Stevenson, Gaussian effective potential: Quantum mechanics, Phys. Rev. D **30**, 1712 (1984).
[101] P. Stevenson, Gaussian effective potential. II. $\lambda\phi_4$ field theory, Phys. Rev. D **32**, 1389 (1985).
[102] R. D. Somma, Quantum simulations of one dimensional quantum systems, arXiv:1503.06319.
[103] A. Macridin, A. Li, S. Mrenna, and P. Spentzouris, Bosonic field digitization for quantum computers, Phys. Rev. A **105**, 052405 (2022).
[104] A. Macridin, P. Spentzouris, J. Amundson, and R. Harnik, Digital quantum computation of fermion-boson interacting systems, Phys. Rev. A **98**, 042312 (2018).
[105] G. K. Brennen, P. Rohde, B. C. Sanders, and S. Singh, Multiscale quantum simulation of quantum field theory using wavelets, Phys. Rev. A **92**, 032315 (2015).
[106] S. Rychkov and L. G. Vitale, Hamiltonian truncation study of the φ^4 theory in two dimensions, Phys. Rev. D **91**, 085011 (2015).
[107] R. D. Somma, Quantum computation, complexity, and many-body physics, arXiv:quant-ph/0512209.
[108] P. C. Encinar, A. Agustí, and C. Sabín, Digital quantum simulation of beam splitters and squeezing with IBM quantum computers, Phys. Rev. A **104**, 052609 (2021).
[109] N. M. Tubman, C. Mejuto-Zaera, J. M. Epstein, D. Hait, D. S. Levine, W. Huggins, Z. Jiang, J. R. McClean, R. Babbush, M. Head-Gordon *et al.*, Postponing the orthogonality catastrophe: Efficient state preparation for electronic structure simulations on quantum devices, arXiv:1809.05523.
[110] N. Gleinig and T. Hoefler, An efficient algorithm for sparse quantum state preparation, in *2021 58th ACM/IEEE Design Automation Conference (DAC)* (2021), 10.1109/DAC18074.2021.9586240.
[111] T. M. de Veras, L. D. da Silva, and A. J. da Silva, Double sparse quantum state preparation, Quantum Inf. Process. **21**, 204 (2022).
[112] X.-M. Zhang, T. Li, and X. Yuan, Quantum state preparation with optimal circuit depth: Implementations and applications, Phys. Rev. Lett. **129**, 230504 (2022).
[113] S. Fomichev, K. Hejazi, M. S. Zini, M. Kiser, J. Fraxanet, P. A. M. Casares, A. Delgado, J. Huh, A.-C. Voigt, J. E. Mueller *et al.*, Initial state preparation for quantum chemistry on quantum computers, PRX Quantum **5**, 040339 (2024).
[114] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Simulating physical phenomena by quantum networks, Phys. Rev. A **65**, 042323 (2002).
[115] Y. Tong, V. V. Albert, J. R. McClean, J. Preskill, and Y. Su, Provably accurate simulation of gauge theories and bosonic systems, Quantum **6**, 816 (2022).
[116] B. Peng, Y. Su, D. Claudino, K. Kowalski, G. H. Low, and M. Roetteler, Quantum simulation of boson-related Hamiltonians: Techniques, effective Hamiltonian construction, and error analysis, Quantum Sci. Technol. **10**, 023002 (2025).
[117] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter error with commutator scaling, Phys. Rev. X **11**, 011020 (2021).
[118] Z. Davoudi, A. F. Shaw, and J. R. Stryker, General quantum algorithms for Hamiltonian simulation with applications to a non-Abelian lattice gauge theory, Quantum **7**, 1213 (2023).
[119] T. I. Andersen, N. Astrakhantsev, A. H. Karamlou, J. Berndtsson, J. Motruk, A. Szasz, J. A. Gross, A. Schuckert, T. Westerhout, Y. Zhang *et al.*, Thermalization and criticality on an analogue–digital quantum simulator, Nature (London) **638**, 79 (2025).
[120] S. Harmalkar, H. Lamm, and S. Lawrence, Quantum simulation of field theories without state preparation, arXiv:2001.11490.
[121] E. J. Gustafson and H. Lamm, Toward quantum simulations of $\mathbb{Z}_2$ gauge theory without state preparation, Phys. Rev. D **103**, 054507 (2021).
[122] P. Stornati, A. Acin, U. Chabaud, A. Dauphin, V. Parigi, and F. Centrone, Variational quantum simulation using non-Gaussian continuous-variable systems, Phys. Rev. Res. **6**, 043212 (2024).
[123] A. D. Linde, A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems, Phys. Lett. B **108B**, 389 (1982).
[124] A. Albrecht and P. J. Steinhardt, Cosmology for grand unified theories with radiatively induced symmetry breaking, Phys. Rev. Lett. **48**, 1220 (1982).
[125] N. Gupta, C. D. White, and Z. Davoudi, Data for: Euclidean-Monte-Carlo-informed ground-state preparation for quantum simulation of scalar field theory, 10.5281/zenodo.18463645 (2026).
[126] J. Gong and P. Aravind, Expansion coefficients of a squeezed coherent state in the number state basis, Am. J. Phys. **58**, 1003 (1990).
[127] F. De Oliveira, M. S. Kim, P. L. Knight, and V. Bužek, Properties of displaced number states, Phys. Rev. A **41**, 2645 (1990).
[128] B. Golubov, A. Efimov, and V. Skvortsov, *Walsh Series and Transforms: Theory and Applications* (Springer Science & Business Media, New York, 2012), Vol. 64, 10.1007/978-94-011-3288-6.
[129] Z. Davoudi, N. M. Linke, and G. Pagano, Toward simulating quantum field theories with controlled phonon-ion dynamics: A hybrid analog-digital approach, Phys. Rev. Res. **3**, 043072 (2021).
[130] R. G. Jha, F. Ringer, G. Siopsis, and S. Thompson, Toward quantum computations of the O(3) model using qumodes, Proc. Sci., LATTICE2023 (**2024**) 230 [arXiv:2308.06946].
[131] V. Ale, N. M. Bauer, R. G. Jha, F. Ringer, and G. Siopsis, Quantum computation of SU(2) lattice gauge theory with continuous variables, J. High Energy Phys. 06 (2025) 084.
[132] R. A. Briceño, R. G. Edwards, M. Eaton, C. González-Arciniegas, O. Pfister, and G. Siopsis, Toward coherent quantum computation of scattering amplitudes with a measurement-based photonic quantum processor, Phys. Rev. Res. **6**, 043065 (2024).
[133] E. Crane, K. C. Smith, T. Tomesh, A. Eickbusch, J. M. Martyn, S. Kühn, L. Funcke, M. A. DeMarco, I. L. Chuang, N. Wiebe *et al.*, Hybrid oscillator-qubit quantum processors: Simulating fermions, bosons, and gauge fields, arXiv:2409.03747.
[134] S. D. Bartlett, B. C. Sanders, S. L. Braunstein, and K. Nemoto, Efficient classical simulation of continuous

variable quantum information processes, Phys. Rev. Lett. **88**, 097904 (2002).

[135] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[136] B. Marco and Z. Alessandro, Manipulating light states by single-photon addition and subtraction, in *Progress in Optics*, Vol. 55 (Elsevier, New York, 2010), pp. 41–83, 10.1016/B978-0-444-53705-8.00002-3.

[137] M. Hofheinz, E. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. O'connell, H. Wang, J. M. Martinis, and A. Cleland, Generation of Fock states in a superconducting quantum circuit, Nature (London) **454**, 310 (2008).

[138] U. Chabaud, G. Ferrini, F. Grosshans, and D. Markham, Classical simulation of Gaussian quantum circuits with non-Gaussian input states, Phys. Rev. Res. **3**, 033018 (2021).