Variational quantum simulation of the critical Ising model with symmetry averaging

Troy J. Sewell [1,2,3,*] Ning Bao,[4] and Stephen P. Jordan[5,3]

[1]*Joint Center for Quantum Information and Computer Science, College Park, Maryland 20742, USA*
[2]*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*
[3]*University of Maryland Institute for Advanced Computer Studies, College Park, Maryland 20742, USA*
[4]*Computational Science Initiative, Brookhaven National Laboratory, Upton, New York 11973, USA*
[5]*Microsoft, Redmond, Washington 98052, USA*

Here we investigate the use of deep multiscale entanglement renormalization ansatz (DMERA) circuits as a variational ansatz. We use the exactly solvable one-dimensional critical transverse-field Ising model as a test bed. Numerically exact simulation of the quantum circuit ansatz can in this case be carried out to hundreds of qubits by exploiting efficient classical algorithms for simulating matchgate circuits. We find that, for this system, the DMERA strongly outperforms a standard quantum approximate optimization algorithm (QAOA)–style ansatz, and that a major source of systematic error in correlation functions approximated using the DMERA is the breaking of the translational and Kramers-Wannier symmetries of the transverse-field Ising model. We are able to reduce this error by up to four orders of magnitude by symmetry averaging, without incurring additional cost in qubits or circuit depth. We propose that this technique for mitigating systematic error could be applied to noisy intermediate-scale quantum (NISQ) simulations of physical systems with other symmetries.

I. INTRODUCTION

The multiscale entanglement renormalization ansatz (MERA) tensor networks represent quantum states on a d-dimensional lattice using a $(d + 1)$-dimensional representation. The extra dimension can be interpreted as scale, with slices of the network along this dimension corresponding to successively coarse-grained states. In one dimension, the treelike MERA tensor network can represent ground states of critical systems, reproducing polynomially decaying correlation functions and logarithmic scaling of subsystem entanglement entropy [1–3]. In the case where the local tensors are fermionic Gaussian unitaries, the networks can be viewed as a wavelet transform on fermion operators and can be rigorously shown to support good approximations of local free-fermion ground states [4–6].

The local unitary structure of a MERA tensor network may be interpreted as a quantum circuit which introduces further UV degrees of freedom scale by scale to entangle qubits of the target ground state. However, on a quantum computer it is more natural to allow increased circuit depth rather than local bond dimension as a means of increasing expressivity of an ansatz, giving rise to a class of quantum circuits known as a deep MERA (DMERA) [7], which could be used as a variational circuit for ground-state preparation. Of particular interest for near-term application is that local observables and correlation functions of ground states prepared by DMERA circuits feature an inherent resilience to local noise stemming from circuit topology [7]. It is also possible to prepare subregions of DMERA states on quantum computers much smaller than the total system size, which may help to leverage the capabilities of small quantum devices [8]. This has been demonstrated on an ion-trap quantum computer, where the critical ground-state subregion can be prepared as a robust fixed-point state of a local quantum channel derived from a DMERA quantum circuit [9].

Multiscale entanglement renormalization ansatz tensor networks can be contracted in polynomial time classically. However, the polynomial scaling with bond dimension is quite severe, e.g., $O(\chi^8)$ for one dimension and $O(\chi^{16})$ for two dimensions using the schemes proposed in [2]. Thus direct execution of the MERA on quantum computers can yield large polynomial speedups and can additionally serve as a method for preparing initial states in the context of a quantum algorithm for simulating quantum dynamics.

This work investigates the feasibility of multiscale circuits for variational ground-state preparation on quantum computers in the fashion of a variational quantum eigensolver [10,11] or more generally an ansatz circuit for some variational quantum algorithm [12], where some family of parametrized circuits is minimized according to some Hamiltonian energy in hopes that a good approximation of the ground state can be found. Another MERA-inspired quantum circuit ansatz for variational ground-state preparation was studied in Ref. [13]; however, the circuit ansatz studied in the present work more directly follows the constructions of Refs. [4,5,7].

Conformal field theories and the use of renormalization are of central importance to the study of quantum field theories, which may generally be viewed as deformations from critical fixed points of a renormalization-group flow. The prospect to simulate conformal field theories with critical lattice models and including the use of renormalization theory is seen as an important step in the ability to handle the more general

---
*tjsewell@umd.edu

study of field theory simulations on quantum computers [14]. Ground states of critical systems feature large correlation lengths which are challenging to reproduce using a more locally entangled ansatz such as matrix product states or short-depth quantum circuits. Scale invariance, however, makes these states somewhat simpler to describe using a multiscale ansatz which may incorporate this symmetry directly into the circuit parametrization. Gapped states may be better able to be represented by an ansatz with local correlations, but may also be prepared using a multiscale circuit with parameters that differ between scales and using fewer scales overall due to the lack of long-range correlations.

We benchmark the viability of the DMERA for variational state preparation by numerically optimizing for approximate ground states of the critical Ising spin chain in one dimension. We are able to find high-fidelity ground-state approximations using the relatively low circuit depth D of each scaling transformation. The ansatz states and local energy density converge to their exact values exponentially with D, with relative error in the energy density below 10^{-8} for $D = 6$. Key features of critical ground states in one dimension such as polynomially decaying correlation functions and logarithmic scaling subsystem entanglement entropy are found on average. (See Fig. 1 for a representation of the multiscale variational circuit with $D = 4$.)

Symmetry averaging is also used to improve the systematic error of local observable expectation values. This can be implemented in hybrid variational schemes by using classical postprocessing to average observables which are related by symmetries which are explicitly broken in the circuit ansatz. In the case of the critical Ising model, spatial translation and Kramers-Wannier symmetry are used, although we also note that these symmetries are still approximately reproduced in the ansatz states themselves, with smaller variance over these symmetry groups for increasing D. Similar ideas have been studied in the context of Refs. [15,16], where symmetry projection is used in postprocessing to improve broken symmetries of ansatz states.

II. CIRCUIT CONSTRUCTION

We consider states which double the number of qubits with each scaling transformation, interleaving new qubits in the zero state between qubits from each of the previous layers. After ℓ scaling circuits we have a state on $L = 2^\ell$ qubits. Our goal is to prepare an approximation to the ground state of a corresponding critical Ising Hamiltonian. Each transformation is a quantum circuit $U(\theta)_\ell$ of depth D, parametrized in terms of angles θ,

$$|\psi_{\ell+1}\rangle = U_\ell(\theta)(|\psi_\ell\rangle \otimes |0\rangle^{\otimes 2^\ell}). \tag{1}$$

Approximate translation and scale invariance of the state are imposed by having each scale transformation be a periodic brickwork circuit invariant under translation by two sites and constraining each scaling transformation circuit at different layers to be made up of the same gates, albeit acting on different numbers of qubits. The approximate translation symmetry reduces the total number of variational parameters from scaling with the total number of gates $O(DL)$ to only with total circuit depth $O(D \log L)$. The approximate scale symmetry
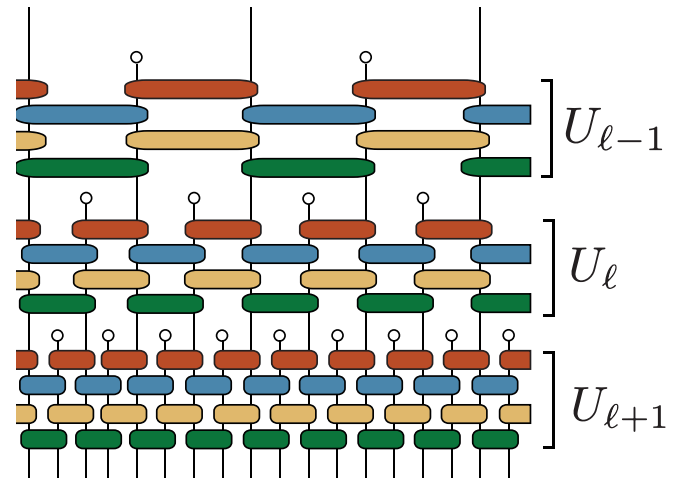
(a)
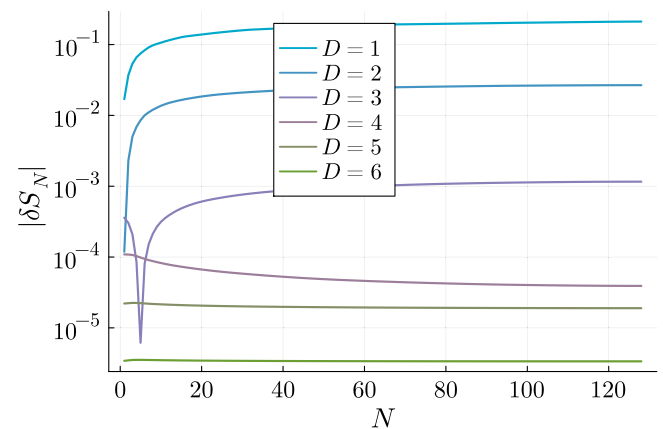
(b)

FIG. 1. (a) State preparation by successive application of $D = 4$ scale transformation circuits $U_\ell(\theta)$ with new qubits initialized in the state $|0\rangle$. Colors show the identity of local gate parameters imposed to replicate approximate translation and scaling symmetry. (b) Relative error in average entanglement entropy of N-qubit sybsystems for an $L = 256$ qubit state, which is exponentially small in D. The subsystem entropy scales logarithmically with N for critical ground states in one dimension, with DMERA circuits featuring an excess of entropy for $D < 3$ and a small entropy deficit for $D > 3$.

further reduces the number of circuit parameters down to only $O(D)$, the depth of a single scaling transformation circuit (see Fig. 1).

These symmetries cannot be imposed exactly due to the discrete nature of the gates and scaling transformations, so the exact symmetries are broken by the ansatz state yet retained as approximate symmetries. Deviations in local observables are exponentially small in D but constitute a major source of error for experimentally realistic values of D, e.g., $D = 1, 2, \ldots, 6$ as studied here. However, as shown in Fig. 4, the error in two-point correlation functions related by Kramers-Wannier symmetry are nearly out of phase with each other. Consequently, this source of error can be reduced by averaging expectation values over these known symmetries. This symmetry averaging lowers error in the correlation functions

by approximately two orders of magnitude, yielding relative error below 10^{-7} for correlation functions using the $D = 6$ state. This improvement is consistent across both of the example Hamiltonians considered in this work [Eqs. (3) and (6)], as shown in Figs. 4 and 5.

In [4–6], wavelet transformations on fermionic modes are used to construct approximations to free-fermion ground states. These constructions use knowledge of the underlying physics to analytically construct the desired wavelet and achieve accuracy exponential in the order of the wavelet (corresponding to circuit depth D). In this work we instead determine the variational parameters in our DMERA by numerical optimization. Inspired by the free-fermion MERA constructions of [4], we take

$$u(x, y) = \begin{bmatrix} \cos(x) & 0 & 0 & \sin(x) \\ 0 & \cos(y) & \sin(y) & 0 \\ 0 & -\sin(y) & \cos(y) & 0 \\ -\sin(x) & 0 & 0 & \cos(x) \end{bmatrix} \quad (2)$$

as the local gates in our circuit, where the two variational parameters specify the rotation on the odd- and even-parity subspaces of the pair of qubits. Because we use real-valued parity-conserving gates, the ansatz states will always be even-parity states with time-reflection symmetry. Similarly to [5], we find that variationally optimized parameters achieve substantially more accurate energy densities than analytically constructed parameters, even though the latter become exact in the limit of the DMERA of infinite depth or the MERA of infinite bond dimension.

III. MODEL AND VARIATIONAL OPTIMIZATION

We focus on the transverse-field Ising model at criticality, namely,

$$H_I = -\sum_{j=1}^{L} X_j X_{j+1} + Z_j. \quad (3)$$

This spin chain is known to be well described by a conformal field theory with $c = \frac{1}{2}$ at low energies, but is also integrable due to the Jordan-Wigner duality relating it to the free-fermion model

$$H = i \sum_{j=1}^{2L} \gamma_j \gamma_{j+1}, \quad (4)$$

where γ_j are Majorana operators. In this convention there are $2L$ Majorana operators for L spatial sites, with operators γ_{2j} and γ_{2j-1} together comprising the fermion at spatial site j. At the critical point the coupling between Majorana operators at the same spatial site and that at neighboring sites are equal, which is sometimes referred to as the half-shift symmetry, which allows for the simple description of the Hamiltonian in Eq. (4). For the spin Hamiltonian on an even number of sites with periodic boundary conditions, the ground state has even parity and is equivalent to the ground state of the Majorana Hamiltonian with antiperiodic boundary conditions, so $\gamma_{2L+1} = -\gamma_1$.

To achieve efficient numerical simulations, we exploit two properties. First, the gates of Eq. (2) are matchgates, which allows us to simulate their action on Gaussian fermion states efficiently as linear transformations on Majorana operators [17,18]. By constraining to matchgate tensors, tensor networks may be used to represent Gaussian fermion operations and states [19,20]. We use these techniques to compute two-point correlation functions out to large distances such as 512 lattice spacings. Second, the expectation values of local operators in MERA and DMERA states can be accelerated by exploiting the constant-width causal cones imposed by these circuits. Local reduced density matrices of the global pure state defined by the ansatz can be obtained as output from iterated quantum channels on a constant number of qubits. These local density matrices are all that is needed to compute the energy of the variational state under a local Hamiltonian. As the number of layers in the MERA is increased, the density matrix converges exponentially to the fixed point of this channel, which can be easily computed, as described in [9]. We tune variational parameters by minimizing the energy density of this fixed-point state.

For sufficiently shallow circuits, we find well-optimized variational parameters using gradient descent and the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm with restarts. For higher depth circuits we obtain variational parameters by starting from an optimized circuit for a given value of D with an additional row of gates inserted at the end of the circuit with parameters initialized near the identity and then reoptimizing to obtain a solution for $D + 1$. Similarly, we may alternatively insert two rows of near-identity gates anywhere within the scaling circuit to obtain a solution for $D + 2$. For the larger-depth circuits, the optimization by adding additional parameters to a smaller solution is better at finding good minima. Random initialization of all parameters is less likely to find an energy minimum any lower than found previously with shorter-depth circuits. We therefore rely on this technique of bootstrapping a better $D + 1$ minimum from a depth D minimum in order to find circuits which continue to show exponentially small error in the energy with circuit depth.

By using the total energy to estimate the average energy density, we are implicitly averaging over symmetries of the Hamiltonian, since the full symmetry group we consider for the critical Ising model relates each local term in the Hamiltonian to every other. By not preparing the whole state and only measuring some terms of the Hamiltonian, we run the risk of minimizing the energy in a nonvariational way, i.e., minimizing an estimate of the global energy below what is possible for the ground state, while inadvertently raising the energy for terms not measured because the global energy still cannot be below that of the ground state. While this could happen in principle, the spatial distribution of expectation values does not have a large variance due to the approximate translation symmetry still found in the circuit ansatz and so this likely would not happen in a catastrophic way in practice.

Imposing SWAP symmetry on the local gates would constrain the odd-parity rotation angle $y = 0$. For the isometry gates only the output legs are swapped, so instead this angle would be set to $y = -\pi/4$, resulting in an ansatz state which is symmetric with respect to spatial reflections. This choice of
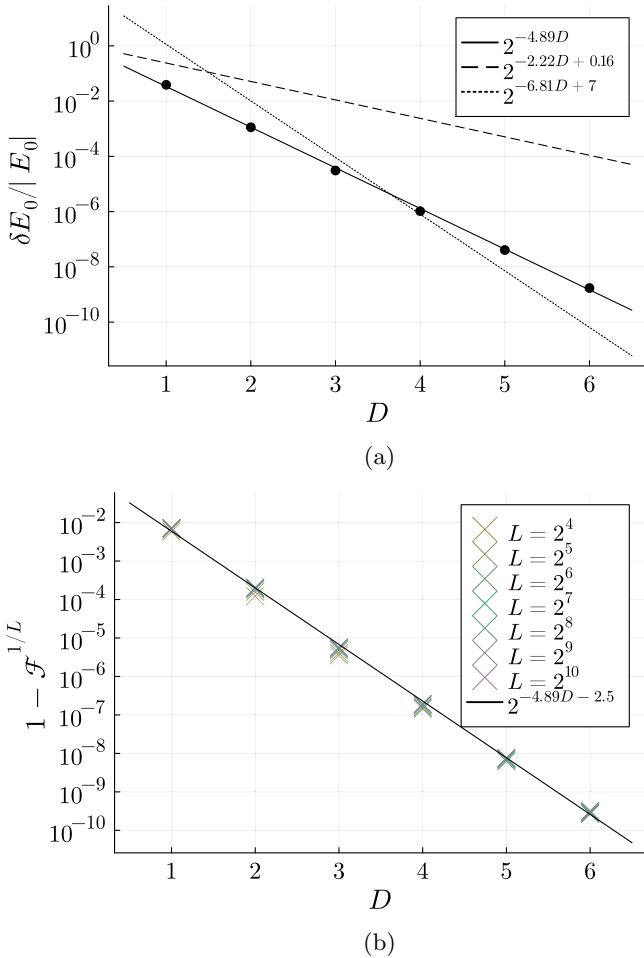
FIG. 2. (a) Relative error of variational fixed-point ground-state energy for channels of depth D, using the infinite volume energy density of $-4/\pi$. Trend lines show the exponential scaling of energy density error for our ansatz (solid line) as well as those from the analytic wavelet construction of [5] (dashed line) and the numerically optimized non-Gaussian MERA in [3] (dotted line). (b) Normalized state infidelity $1 - \mathcal{F}^{1/L}$ of states prepared by depth D scaling transformation for ground states of various system sizes L. Both quantities appear to decay exponentially with D with a coefficient of approximately -4.89.

parameters is optimal for the $D = 1$ case. However, for $D > 1$ we find that by relaxing this constraint, reflection-symmetry-broken states can be prepared using the same circuit depth which has lower energy and higher fidelity. By transforming the parameters $y \rightarrow -y$ for nonisometry gates and $y \rightarrow \pi/2 - y$, a reflected ansatz state with the same energy density and fidelity is prepared.

The circuit parameters we use throughout this paper for $D \leqslant 6$ can be found in the Appendix. The corresponding error in energy density appears to be exponentially small in D with a scaling coefficient of -4.89, seen in Fig. 2. We compare this fit line to two similar multiscale ground-state representations for the Ising model, the wavelet-based construction found in Ref. [5], which uses a very similar circuit ansatz but constructs the ideal parameters analytically rather than through energy minimization, and the more traditional MERA tensor network

states in Ref. [2] found through numerical optimization of the energy. In order to compare these different methods, we note that local gates of the circuit-based methods can be grouped together to form a MERA tensor network of bond dimension $\chi = 2^{D-1}$. So while the MERA from [2] does not have a circuit depth D, we will use this equivalence in order to compare ansatz states of similar complexity. Nevertheless, the circuit-based constructions of this work and [5] are much more constrained ansatz states compared to a generic MERA of bond dimension $\chi = 2^{D-1}$, due to the local circuit structure.

We find that while maintaining much of the simplicity of the wavelet construction of [5], our numerically optimized states are able to outperform those constructions at equal circuit depth and appear to scale with a larger scaling coefficient. Our ansatz states even outperform the general MERA states of [2] for short-depth circuits, but are beaten at larger bond dimensions, which should be expected due circuits being much more constrained than a MERA state of equivalent bond dimension.

IV. NUMERICAL BENCHMARKING

Having minimized the energy density for circuits with $D \leqslant 6$, we can compare our ansatz states with the true ground state for a variety of other measures. We compare the state fidelity and subsystem fidelities for different (sub)system sizes and circuit depths as well as logarithmic scaling of subsystem entropy. Finally, we look at the two-point correlation functions and show how averaging local observables over symmetry groups broken by the circuit ansatz can greatly improve the estimation accuracy of these observables.

We prepare finite-size ground states using the fixed-point optimized circuit parameters and applying some number of scaling transformation layers $\mathcal{L}$ to prepare the state on $L = 2^{\mathcal{L}}$ qubits. Interestingly, the infidelity of our ansatz states seem to be exponentially small in D, with the same scaling coefficient of -4.89 as the energy density (see Fig. 2). We compare different-size systems by normalizing the state infidelity as $1 - \mathcal{F}^{1/L}$, which we find clusters together ansatz states of the same D but different system sizes L.

In Fig. 3 we compare the infidelity of N-qubit subsystems of an $L = 256$ qubit system, normalized by subsystem size. This behaves similarly to that of global state fidelity with identical exponential scaling in D and clustering of subsystem infidelity from the normalization. The smallest subsystems, however, have an even smaller normalized infidelity than that of the global state and so local observables should be even more accurate than global state properties.

The relative error in average subsystem entropy is shown in Fig. 1, showing exponential accuracy with D. Critical ground states in one dimension feature more long-range entanglement than area-law states, which would have constant entropy in one dimension. Instead, subsystem entropy scales logarithmically with the subsystem size, which becomes a major challenge of strictly local circuits which can only entangle qubits within a finite range. Multiscale circuits introduce entanglement at all scales and therefore can prepare states with this scaling of subsystem entropy. In fact, the ansatz states
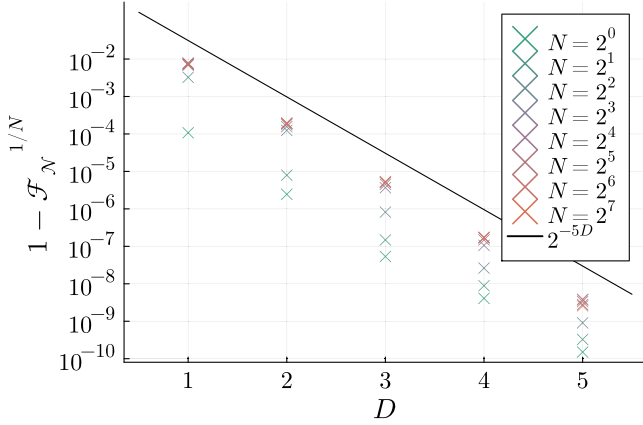
FIG. 3. Normalized infidelity of N-qubit subsystems of an $L = 256$ qubit state, averaged over different subsystem locations.

for $D \leqslant 3$ have on average an excess of entropy rather than a deficit.

V. AVERAGING BROKEN SYMMETRIES

Although we impose a quasitranslation invariance in the local gates of our ansatz, the actual circuit breaks this translation invariance within the circuit layer, preparing a state which is only exactly invariant under translations by $L/2$ sites. Despite this, an approximate symmetry is still maintained, so local observables related by translations have a small variance which decreases exponentially with D.

In the Majorana representation there are two operators γ_{2i-1} and γ_{2i} at site i, so a translation by n sites amounts to shifting the index of all Majorana operators by $2n$, i.e., $\gamma_j \rightarrow \gamma_{j+2n}$. At criticality, the Ising model possesses an additional symmetry which constrains the expectation value of local Z and XX operators to be identical [21]. The Kramers-Wannier transformation maps between these sets of operators as

$$Z_i \rightarrow X_i X_{i+1}, \quad X_i X_{i+1} \rightarrow Z_{i+1}, \tag{5}$$

which is a symmetry of the Hamiltonian. This symmetry becomes more apparent in the Majorana representation, since the Jordan-Wigner transformation maps Z_i to $\gamma_{2i-1}\gamma_{2i}$ and $X_i X_{i+1}$ to $\gamma_{2i}\gamma_{2i+1}$, so the Kramers-Wannier transformation amounts to shifting the Majorana index by one and is sometimes known as a half-shift symmetry of the Majorana operators since two of these transformations amount to a spatial translation by one site. This transformation is clearly a symmetry of the Hamiltonian in Eq. (4), but is broken away from criticality when the Z and XX operators have different coefficients.

In Fig. 4(a) we look at the relative error in the expectation values of the two sets of quadratic Majorana operators $\gamma_{2i-1}\gamma_{2j}$ and $\gamma_{2i}\gamma_{2j+1}$. Translation symmetry dictates that the ground-state expectation values are the same for equal $i - j$, so we can average over the group of symmetries, the cyclic group on L elements, and we plot the relative error after translation averaging in Fig. 4(a). These two sets of observables decay inverse polynomially in the distance between the two Majorana operators as $|i - j|^{-1}$ and are related to each other by the Kramers-Wannier transformation.
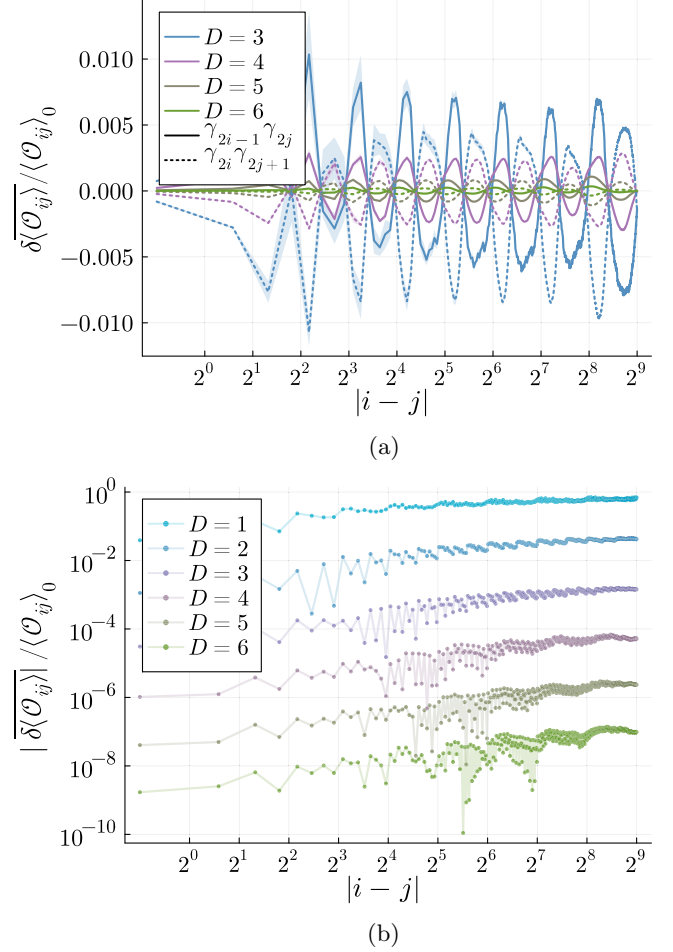

(a)

(b)

FIG. 4. (a) Relative error in the translation-averaged expectation value of the quadratic Majorana operators $\gamma_{2i-1}\gamma_{2j}$ (solid lines) and $\gamma_{2i}\gamma_{2j+1}$ (dashed lines) at fixed distance $|i - j|$. These two cases are related by a Kramers-Wannier transformation and have systematic errors which appear oscillatory on a logarithmic scale and nearly out of phase with each other. (b) Further averaging over the Kramers-Wannier symmetry cancels much of the error of quadratic Majorana observables, reducing the error in the expectation value estimate by orders of magnitude.

Interestingly, the relative errors in these two sets of correlation functions appear oscillatory when plotted on a logarithmic scale in the distance $|i - j|$, which we attribute to the breaking of continuous scaling symmetry down to an approximate symmetry due to the discrete scaling transformations. The amplitudes of these deviations decay with D, but also appear to be nearly out of phase with each other for the two sets of operators related by the Kramers-Wannier transformation. Averaging over this additional $\mathbb{Z}_2$ symmetry, we can cancel much of this remaining systematic error due to the respective deviations being nearly out of phase, resulting in a much more accurate estimation of the observables after averaging of the Kramers-Wannier symmetry, seen in Fig. 4(b).

We can assess how well this cancellation by Kramers-Wannier symmetry works by comparing the average amount of error of correlators at a fixed distance, denoted by $\overline{|\delta\langle\mathcal{O}_{ij}\rangle|}$, with the amount of error in the symmetry-averaged
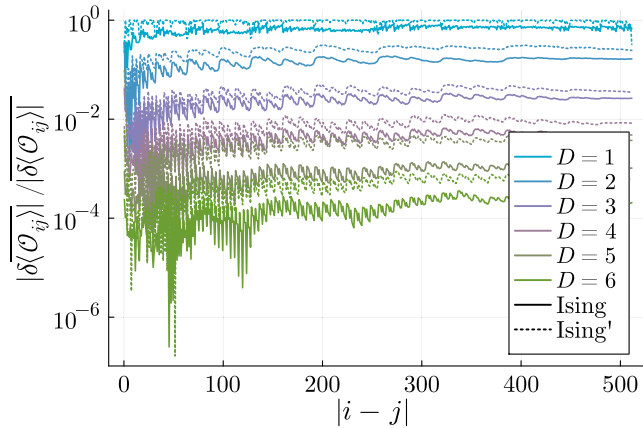
FIG. 5. Improvement in accuracy from symmetry averaging also increases with D. The ratio is plotted between the error magnitude in the symmetry-averaged correlation function, denoted by $|\delta\overline{\langle\mathcal{O}_{ij}\rangle}|$, and the average error in particular expectation values with fixed $|i-j|$, denoted by $\overline{|\delta\langle\mathcal{O}_{ij}\rangle|}$. Also included are the ratios for approximate ground states of the modified Ising model $H_{I'}$, which show similar improvement from symmetry averaging.

expectation value, denoted by $|\delta\overline{\langle\mathcal{O}_{ij}\rangle}|$. In Fig. 5 we see that this cancellation works remarkably well, with the symmetry-averaged mean showing accuracy multiple orders of magnitude better than the typical local observable. This shows that states of larger circuit depth feature both a lower amount of variance in their oscillatory expectation values and a higher degree of error cancellation, which compound to produce highly accurate estimators of local observables after symmetry averaging.

In Fig. 5 we also include values from circuits approximating the ground state of an alternate Ising model

$$H_{I'} = \sum_{i=1}^{L} -X_i X_{i+1} + X_{i-1} Z_i X_{i+1}. \qquad (6)$$

This Hamiltonian is related to the Ising model by a constant depth unitary $\prod_j \frac{1}{\sqrt{2}}(-I + iX_j X_{j+1})$, which smears the single Pauli Z term into a three-body operator, but is still equivalent to a quadratic fermion model by Jordan-Wigner duality. This other set of parameters for the $H_{I'}$ ground state feature similar degrees of error cancellation.

While these two Hamiltonians share much of the same physics due to their equivalence up to a short-depth local transformation, the variational circuits we find to approximate their respective ground states are not made equivalent by any short-depth local transformation. Both sets of circuits preparing ground-state approximations to H_I or $H_{I'}$ are scale-invariant variational circuits with different sets of parameters. It would be possible to prepare approximate ground states for H_I by using the circuits for $H_{I'}$ and then applying the short-depth transformation that relates the two Hamiltonians, or vice versa, but the resulting circuits are not scale invariant due to the final circuit layer differing from all the previous ones.

This breaking of scale symmetry is explicitly used in some applications of the MERA, where an initial and a final layer

of tensors, which differ from the scale-invariant layers, are used and can be referred to as transition layers. These may be beneficial to include explicitly because the finite volume and lattice spacing of the model already break scale invariance at the largest and smallest scales and so a circuit with complete scale symmetry may not be the most ideal representation of the ground state. We restrict our attention to circuits with identical variational parameters for all scaling transformations within the quantum circuit in this work and to this extent we consider the circuits found for ground states of H_I and $H_{I'}$ to give independent confirmation to the efficacy of these circuits.

Subsystem entropy also varies with location due to the lack of exact translation invariance. However, entropy is a much more coarse-grained property of a state and does not display a great degree of error canceling from translation averaging. That is, the subsystem entropies are typically biased in a particular direction for all subsystems, so the average entropy is not much more accurate than that of a typical subsystem.

Here we specifically consider the translational invariance and Kramers-Wannier symmetries of two-point functions on a one-dimensional N-site translationally invariant lattice with periodic boundary conditions. This symmetry averaging procedure can be directly generalized to arbitrary groups of symmetries, e.g., S_n symmetries on permutation-symmetric systems, the translation and rotation symmetries in a fixed number of spatial dimensions, or internal global or gauge symmetries, which could be either discrete or continuous.

Averaging over translations by a symmetry group G when measuring an observable comes at no cost in qubit count or circuit depth. Rather, the requirement is for repeated executions of the circuit and measurement. Such repeated measurements are already necessary, as $1/\epsilon^2$ measurements are needed to measure an observable with statistical error of order ϵ. Instead of repeating $1/\epsilon^2$ measurements on an individual choice of G translation of a given observable, one can instead choose a uniformly random G translation of the observable for each measurement. In the absence of any *a priori* information to suggest some particular G translation as preferable, this G averaging will in the worst case yield equivalent accuracy to a randomly chosen but fixed G translation and in the best case can result in greatly improved accuracy due to cancellation of errors of opposite sign. The latter situation is observed in the two examples H_I and $H_{I'}$ considered here. It remains for future work to investigate which models exhibit such a strong benefit from cancellation and which do not.

The only cost for symmetry averaging is potentially having to sample multiple circuits in order to prepare different subregions for measuring the collection of observables to be averaged. This makes it a viable tool for implementation on noisy intermediate-scale quantum (NISQ) hardware and less resource intensive than doing a full projection onto the trivial representation of the symmetry group as in Ref. [15]. Since Pauli strings only have eigenvalues ± 1, their variance is $1 - \langle\mathcal{O}\rangle^2$ for an observable $\mathcal{O}$. The symmetry-averaged expectation value will still only have measurement outcomes ± 1, so aggregating the measurements from different observables will result in a sample with variance $1 - \langle\overline{\mathcal{O}}\rangle^2$.

This technique is complementary to the widely used Richardson extrapolation and randomized resampling meth-

ods introduced in [22], as it mitigates systematic error induced by the ansatz rather than random error introduced by the hardware. These observations also suggest that imposing strict symmetries could be overly restrictive for a circuit ansatz. Instead, it could be more advantageous to consider ansatz states which can break these symmetries while retaining accurate symmetry-averaged expectation values. With more quantum resources available, one could consider a coherent analog of this error mitigation scheme in which states with approximate symmetries are projected onto the proper symmetric subspace after preparation to achieve higher fidelity states. Specifically, one can project onto the G-invariant subspace preparing a control register in the uniform superposition, executing a controlled G translation on the target state, and then measuring the control register in the Hadamard basis.

## VI. COMPARISON TO a QUANTUM ALTERNATING OPERATOR ANSATZ

Another natural approach to variational state preparation is using a strictly local circuit by evolving commuting terms in the Hamiltonian in sequence to form a quantum alternating operator ansatz [23,24]. Using a short-depth ansatz makes it feasible for NISQ devices and these circuits may be constrained to conserve symmetries of the Hamiltonian, such as preparing only translation invariant states. In this case, they do not benefit from symmetry averaging, unlike the DMERA. This approach was used to find exact variational circuits for ground states of the critical Ising model on $2p$ spins using p rounds of a quantum approximate optimization algorithm (QAOA) ansatz [25] and has been implemented experimentally on an ion-trap device [26].

Exact variational circuits are not possible with fewer than $L/2$ rounds, as this is the minimum circuit depth needed for the ansatz to "see the whole graph" [27]. That is, the past causal cone of some subregion of a local circuit ansatz contains the entire initial state only when the circuit depth is L or greater, making global entanglement possible. Before a local circuit can see the whole system it also is unaware of system size. So, even though the ground state of a finite-size Ising chain with periodic boundary conditions has energy density lower than the infinite volume value of $-4/\pi$, a local circuit with depth less than L could not prepare a state with energy below this threshold. This also means the energy density for a local circuit ansatz will be identical for all system sizes greater than the circuit depth.

The energy density and normalized ground-state infidelity for a local QAOA circuit ansatz are shown in Fig. 6. The circuits are initialized using parameters provided in [25] for preparing exact ground states for system sizes $L = 2p$ and then the energy is optimized for systems with $L > 2p$ so that the circuit does not see the whole system and therefore cannot prepare the exact ground state. The resulting states give us an idea of how short-depth local circuits can perform when preparing approximate ground states. Both errors in energy density and state fidelity appear consistent with polynomial decay in the circuit depth. This contrasts with the DMERA, which achieves exponentially improving accuracy with increasing D.
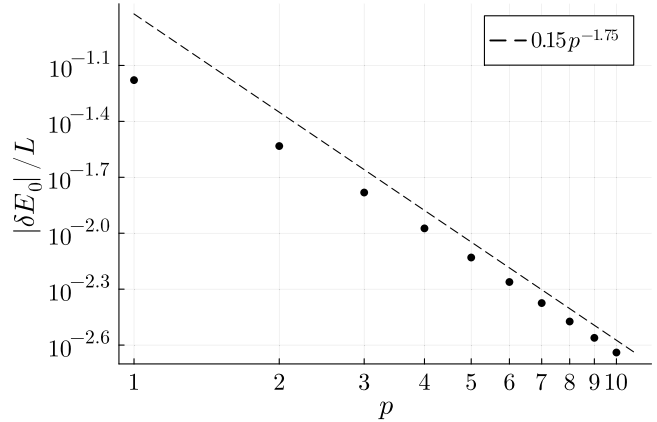
(a)
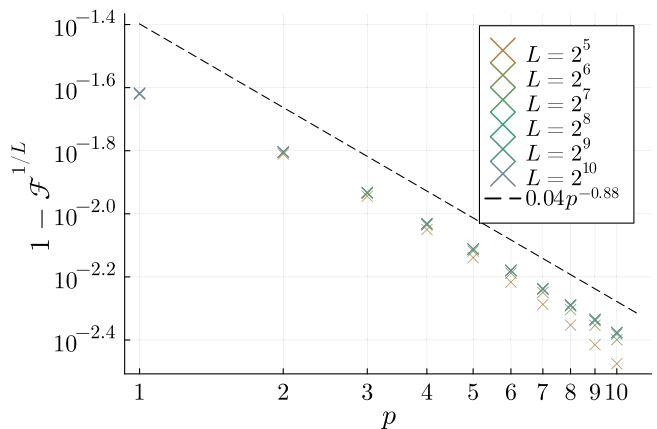
(b)

FIG. 6. (a) Error of variational energy density from the infinite volume value of $-4/\pi$ against total circuit depth or number of rounds p for systems $L > 2p$. (b) Normalized pure state infidelity versus number of rounds p for various system sizes $L > 2p$. Errors in both the energy density and normalized infidelity for short depth ansatz states appear consistent with polynomial decay in p for large system sizes.

VII. DISCUSSION

The critical free-fermion model is special in many ways, but we expect similar circuit architectures to be successful more broadly for states of local Hamiltonians. In these situations, locality and separation of length scales have been invaluable tools in physical analysis, which can also be leveraged in designing good quantum circuits for preparing physically relevant states. Adapting to systems which are nonintegrable could be done by allowing a more general set of local gates which goes beyond Gaussian fermion operations (i.e., universal rather than matchgate circuits). Additionally, states with finite correlation length ξ could be prepared using $O(\log \xi)$ scaling layers optimized independently, which may be interpreted as a nontrivial renormalization-group flow which disentangles the state at scales beyond the correlation length. The variational ansatz states we find are able to outperform similar analytic constructions while maintaining much of their simplicity compared to a more general

variational ansatz such as a large bond dimension MERA state. Finding a way to more precisely formalize variational multiscale circuits as genuine scaling transformations beyond state preparation may also find use in algorithms for simulating conformal field theories and quantum field theories using closely ideas related from renormalization and wavelet theory [14,28–31]. For NISQ applications, mitigation of error under tight hardware constraints is a key consideration. The symmetry averaging methods described here for mitigating ansatz-induced error complement prior work showing that the DMERA's renormalization structure provides intrinsic resilience against hardware-induced error [7,9].

APPENDIX: FREE FERMIONS

Some basics of the representation of Gaussian fermionic states and operations are outlined here following [32]. See also the Grassmann representation in [18].

A Gaussian fermion state ρ can be fully characterized by the covariance matrix

$$\Gamma_{j,k} = \frac{i}{2}\text{tr}(\rho[\gamma_j, \gamma_k]). \tag{A1}$$

For n local fermionic modes a_i we can define $2n$ Majorana modes

$$\gamma_{2i-1} = \frac{a_i + a_i^\dagger}{2}, \tag{A2}$$

$$\gamma_{2i} = \frac{a_i - a_i^\dagger}{2}. \tag{A3}$$

These operators are Hermitian and obey the usual fermionic anticommutation relations $\gamma_i, \gamma_j = 2I$, but also $\gamma_i^2 = \frac{I}{2}$.

A Gaussian fermion state satisfies $\Gamma^2 \leqslant -I$, with equality holding for pure states. The expectation value of quadratic fermion observables can be directly read from the covariance matrix, while Wick's theorem allows higher-order observables to be calculated in terms of the quadratic expectation values.

Subsystems of a Gaussian state are defined by restricting the covariance matrix to a subset of the Majorana operators, eliminating the rows and columns associated with the discarded operators. Any mixed Gaussian state can be considered a subsystem of some pure Gaussian state. The covariance matrix for a mixed state can be purified by adding additional Majorana modes, introducing the new rows and columns with entries so that the new covariance matrix satisfies $\Gamma^2 = -I$.

Although we can relate the fermionic system to a system of qubits using a Jordan-Wigner duality, this operation of taking fermionic subsystems is a different restriction of the Hilbert space from the tracing out of qubit degrees of freedom. However, all even-parity operators on the reduced fermionic

and qubit subsystems will agree. Odd-parity operators have zero expectation value in Gaussian states.

A Gaussian unitary transformation always exists which transforms a Gaussian state into a collection of independent fermion modes, each with excitation probability p_j, so every Gaussian state has a product spectrum. Gaussian unitary operations are equivalent to linear transformations on the fermion modes, which act on the covariance matrix via an orthogonal transformation $\Gamma' = O\Gamma O^T$. The Gaussian unitary transformation that decouples the Gaussian fermion state block diagonalizes the covariance matrix into 2×2 blocks with off-diagonal entries $\pm i\lambda_j$.

Because all Gaussian states have a product spectrum, the entropy of any mixed Gaussian state is the sum of entropies S_j from the states eigenmodes, each of which can be written in terms of the respective eigenvalue of the covariance matrix

$$S_j = -\frac{1+\lambda_j}{2}\ln\frac{1+\lambda_j}{2} - \frac{1-\lambda_j}{2}\ln\frac{1-\lambda_j}{2}$$

$$= \ln 2 - \ln\sqrt{1-\lambda_j^2} - \lambda_j\ln\sqrt{\frac{1+\lambda_j}{1-\lambda_j}}. \tag{A4}$$

The fidelity between two mixed Gaussian states can also be computed in terms of their covariance matrices [33],

$$\mathcal{F}_{\rho,\sigma} = 2^{-n/2}\det(I - \Gamma_\rho\Gamma_\sigma)^{1/4}\det\big(I - \sqrt{I + \Gamma_{\rho,\sigma}^2}\big)^{1/4}, \tag{A5}$$

with $\Gamma_{\rho,\sigma} = \frac{\Gamma_\rho + \Gamma_\sigma}{I - \Gamma_\rho\Gamma_\sigma}$. If at least one of these states is pure then the simpler expression can be used, $\mathcal{F}_{\rho,\sigma} = |\det[(\Gamma_\rho + \Gamma_\sigma)/2]|^{1/4}$.

The quadratic fermionic Hamiltonian we work with, which is dual to the critical transverse-field Ising model, is

$$H = i\sum_{j=1}^{2L}\gamma_j\gamma_{j+1}. \tag{A6}$$

For a finite-size spin chain with periodic boundary conditions, the dual quadratic fermion Hamiltonian must have antiperiodic boundary conditions for the even-parity states and periodic boundary conditions for odd-parity states. Thus, with periodic boundary conditions, the full Ising Hamiltonian is not strictly dual to a single quadratic fermion Hamiltonian, but all of its eigenstates are dual to a Gaussian fermion states.

To avoid these issues we will focus just on the Majorana fermion model with antiperiodic boundary conditions for our numerics and in this case our circuits may be interpreted as local operations on fermionic degrees of freedom. In fact, because we are working with a free-fermion model, we will also restrict our gates to be unitary Gaussian fermion operations, equivalent to Bogoliubov transformations.

Because local matchgate circuits are equivalent to linear transformations on the fermionic modes, this subgroup of quantum states and operations and can be simulated by classical polynomial time algorithms [17] and may implemented numerically using covariance matrix methods or via fermionic linear optics [18]. Under Jordan-Wigner duality, local two-qubit gates become orthogonal transformations on four consecutive Majorana operators. These techniques can be extended to tensor networks as matchgate tensor networks [19,20].

The two-qubit gates $u(x, y)$ can be written implemented as an orthogonal transformations $\tilde{u}(x', y')$ on four Majorana operators, using the parameters $x' = x + y$ and $y' = x - y$, as follows:

$$\tilde{u}(x', y') \equiv \begin{bmatrix} \cos(x') & 0 & -\sin(x') & 0 \\ 0 & \cos(y') & 0 & \sin(y') \\ \sin(x') & 0 & \cos(x') & 0 \\ 0 & -\sin(y') & 0 & \cos(y') \end{bmatrix}. \tag{A7}$$

The variational parameters found and used in this paper are reported as θ_D for a depth D ansatz with $2D$ parameters [see Eq. (A8)]. Here the $(2i - 1)$th parameter is the x' parameter for the ith layer of gates within the scaling transformation and the $(2i - 1)$th parameter is the y' parameter for the same layer of gates. We also found alternate sets of parameters for the modified Ising model in Eq. (6), listed in (A9). These circuits have similar features and the symmetry-averaging improvement is displayed in Fig. 5 along with those of the original Ising model states:

$\theta_1 = [0.431\,88, -1.138\,91],$

$\theta_2 = [0.1379, -0.563\,74, -0.534\,56, 0.180\,71],$

$\theta_3 = [-1.797\,16, -1.518\,91, 0.644\,86, 2.0904, 0.109\,94, -0.314\,94],$

$\theta_4 = [-1.665\,28, -1.551\,01, 1.051\,14, 1.829\,04, 0.434\,26, -0.749\,51, 0.073\,49, -0.207\,64],$

$\theta_5 = [-1.610\,46, -1.563\,58, 1.291\,07, 1.694\,99, 0.801\,32, -1.066\,96, 0.308\,63, -0.545\,39, 0.051\,58, -0.146\,51],$

$\theta_6 = [-1.586\,82, -1.568\,31, 1.426\,66, 1.6278, 1.084\,03, -1.281\,63, 0.620\,53, -0.854\,36, 0.2323,$
$\qquad - 0.4140, 0.038\,62, -0.110\,17];$ \hfill (A8)

$\tilde{\theta}_1 = [-0.221\,07 - 1.791\,87],$

$\tilde{\theta}_2 = [0.379\,21, -1.586\,72, -0.235\,88, -0.8906],$

$\tilde{\theta}_3 = [0.097\,44, -1.623\,52, -0.862\,47, -0.294\,72, -0.186\,93, 0.512\,59],$

$\tilde{\theta}_4 = [-0.0058, -0.210\,63, 1.6268, 0.783\,19, -1.087\,15, 0.1295, 2.778\,86, -0.087\,55],$

$\tilde{\theta}_5 = [-0.083\,57, -1.736\,96, -0.092\,69, -0.140\,14, 0.141\,62, 1.328\,26, -1.058\,85, 0.113\,38, -0.421\,56, -0.406\,56],$

$\tilde{\theta}_6 = [0.036\,14, -1.562\,55, -1.114\,74, -0.281\,92, 0.554\,76, 2.599\,16, -0.088\,26, 0.103\,61,$
$\qquad - 1.058\,98, -1.746\,65, -0.050\,94, 0.254\,21].$ \hfill (A9)

[1] G. Vidal, Class of Quantum Many-Body States That Can Be Efficiently Simulated, Phys. Rev. Lett. **101**, 110501 (2008).

[2] G. Evenbly and G. Vidal, Algorithms for entanglement renormalization, Phys. Rev. B **79**, 144108 (2009).

[3] G. Evenbly and G. Vidal, Quantum criticality with the multi-scale entanglement renormalization ansatz, in *Strongly Correlated Systems*, edited by A. Avella and F. Mancini, Springer Series in Solid-State Sciences Vol. 176 (Springer, Berlin, Heidelberg, 2013), pp. 99–130.

[4] G. Evenbly and S. R. White, Entanglement Renormalization and Wavelets, Phys. Rev. Lett. **116**, 140403 (2016).

[5] J. Haegeman, B. Swingle, M. Walter, J. Cotler, G. Evenbly, and V. B. Scholz, Rigorous Free-Fermion Entanglement Renormalization from Wavelet Theory, Phys. Rev. X **8**, 011003 (2018).

[6] F. Witteveen, V. Scholz, B. Swingle, and M. Walter, Quantum circuit approximations and entanglement renormalization for the Dirac field in 1+1 dimensions, Commun. Math. Phys. **389**, 75 (2022).

[7] I. H. Kim and B. Swingle, Robust entanglement renormalization on a noisy quantum computer, arXiv:1711.07500.

[8] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[9] T. J. Sewell and S. P. Jordan, Preparing renormalization group fixed points on NISQ hardware, arXiv:2109.09787.

[10] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[11] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods and best practices, Phys. Rep. **986**, 1 (2022).

[12] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New J. Phys. **18**, 023023 (2016).

[13] Q. Miao and T. Barthel, A quantum-classical eigensolver using multiscale entanglement renormalization, arXiv:2108.13401.

[14] T. J. Osborne and A. Stottmeister, Quantum simulation of conformal field theory, arXiv:2109.14214.

[15] K. Seki, T. Shirakawa, and S. Yunoki, Symmetry-adapted variational quantum eigensolver, Phys. Rev. A **101**, 052340 (2020).

[16] T. Tsuchimochi, M. Taii, T. Nishimaki, and S. L. Ten-no, Adaptive construction of shallower quantum circuits with quantum spin projection for fermionic systems, arXiv:2205.07097.

[17] R. Jozsa and A. Miyake, Matchgates and classical simulation of quantum circuits, Proc. R. Soc. A **464**, 3089 (2008).

[18] S. Bravyi, Lagrangian representation for linear optics, Quantum Inf. Comput. **5**, 216 (2005).

[19] N. Schuch and B. Bauer, Matrix product state algorithms for gaussian fermionic states, Phys. Rev. B **100**, 245121 (2019).

[20] A. Jahn, M. Gluza, F. Pastawski, and J. Eisert, Holography and criticality in matchgate tensor networks, Sci. Adv. **5**, eaaw0092 (2019).

[21] D. Radicevic, Spin structures and exact dualities in low dimensions, arXiv:1809.07757.

[22] K. Temme, S. Bravyi, and J. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[23] S. Hadfield, Z. Wang, B. O'Gorman, E. Rieffel, D. Venturelli, and R. Biswas, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, Algorithms **12**, 34 (2019).

[24] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[25] W. W. Ho and T. H. Hsieh, Efficient variational simulation of non-trivial quantum states, SciPost Phys. **6**, 029 (2019).

[26] D. Zhu, S. Johri, N. M. Linke, K. A. Landsman, C. H. Alderete, N. H. Nguyen, A. Y. Matsuura, T. H. Hsieh, and C. Monroe, Generation of thermofield double states and critical ground states with a quantum computer, in *Proceedings of the National Academy of Sciences* (2020), Vol. 117.

[27] E. Farhi, D. Gamarnik, and S. Gutmann, The quantum approximate optimization algorithm needs to see the whole graph: Worst case examples, arXiv:2004.09002.

[28] 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).

[29] 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).

[30] F. Witteveen and M. Walter, Bosonic entanglement renormalization circuits from wavelet theory, SciPost Phys. **10**, 143 (2021).

[31] T. J. Osborne and A. Stottmeister, Conformal field theory from lattice fermions, Commun. Math. Phys. **398**, 219 (2023).

[32] S. Bravyi and D. Gosset, Complexity of quantum impurity problems, Commun. Math. Phys. **356**, 451 (2017).

[33] B. G. Swingle and Y. Wang, Recovery map for fermionic Gaussian channels, J. Math. Phys. **60**, 072202 (2019).