Pre-optimizing variational quantum eigensolvers with tensor networks

Abid Khan,^{1,2,3} Bryan K. Clark,^{1,4} and Norm M. Tubman³

¹Department of Physics, University of Illinois Urbana-Champaign, Urbana, IL, United States 61801

²USRA Research Institute for Advanced Computer Science (RIACS), Mountain View, CA, 94043, USA

³NASA Ames Research Center, Moffett Field, CA, 94035, USA

⁴IQUIST and Institute for Condensed Matter Theory and NCSA Center for Artificial Intelligence Innovation,

University of Illinois at Urbana-Champaign, IL 61801, USA

(Dated: October 20, 2023)

The variational quantum eigensolver (VQE) is a promising algorithm for demonstrating quantum advantage in the noisy intermediate-scale quantum (NISQ) era. However, optimizing VQE from random initial starting parameters is challenging due to a variety of issues including barren plateaus, optimization in the presence of noise, and slow convergence. While simulating quantum circuits classically is generically difficult, classical computing methods have been developed extensively, and powerful tools now exist to approximately simulate quantum circuits. This opens up various strategies that limit the amount of optimization that needs to be performed on quantum hardware. Here we present and benchmark an approach where we find good starting parameters for parameterized quantum circuits by classically simulating VQE by approximating the parameterized quantum circuit (PQC) as a matrix product state (MPS) with a limited bond dimension. Calling this approach the variational tensor network eigensolver (VTNE), we apply it to the 1D and 2D Fermi-Hubbard model with system sizes that use up to 32 qubits. We find that in 1D, VTNE can find parameters for PQC whose energy error is within 0.5% relative to the ground state. In 2D, the parameters that VTNE finds have significantly lower energy than their starting configurations, and we show that starting VQE from these parameters requires non-trivially fewer operations to come down to a given energy. The higher the bond dimension we use in VTNE, the less work needs to be done in VQE. By generating classically optimized parameters as the initialization for the quantum circuit one can alleviate many of the challenges that plague VQE on quantum computers.

I. INTRODUCTION

The variational quantum eigensolver (VQE) is particularly well-suited for the noisy intermediate-scale quantum (NISQ) regime, where quantum computers are limited in size and coherence time. Some advantages of VQE are that its variational character can provide some degree of error mitigation in the parameterization of the gates [1–4] and that it features shallower circuits compared to more exact algorithms such as phase estimation and quantum approximate optimization algorithms (QAOA) [5–17]. The applications for VQE range over a number of different fields [18] including chemistry [3, 6, 8, 19–23], materials science [5, 24], and machine learning [25–27].

Despite the promise of parameterized quantum algorithms to provide advantages over classical methods [28, 29], several obstacles obstruct their realization. In particular, the parameterized quantum circuit (PQC) optimization landscape is plagued by the presence of barren plateaus [30–34], particularly starting from randomly parameterized quantum circuits, and local minima [35–37]. These problems have been explored in quantum chemistry applications, where circuits for computing molecular ground states can reach high-precision results using ini-

tializations based on mean-field Hartree-Fock or more sophisticated coupled-cluster-based solutions [6, 17, 38, 39]. There is active work on mitigating these challenges and finding ways to improve the performance of VQE [31, 32].

Another difficulty in demonstrating an advantage over classical algorithms using PQCs is the increasing sophistication of classical simulation algorithms. However, this also provides new possibilities for performing preoptimization on classical hardware. Several ideas of this type have been suggested recently for applications in quantum chemistry [40] and in related works these ideas were considered in detail with large-scale simulations up to 64 qubits [23, 39]. To perform simulations at these scales, interesting approximations for simulation of quantum circuits have to be employed [23, 39, 41]. While the above simulations have been performed at large scales, there are many alternative approaches in which these ideas can be further explored which include tensor network approaches [42, 43]. The ability of tensor networks to be deployed on powerful classical hardware accelerators, such as graphical and tensor processing units (GPUs and TPUs), raises the bar for quantum hardware to overcome.

Rather than perceiving this as an obstacle to quantum

advantage, one can instead view the success of sophisticated classical simulation techniques, including tensor network algorithms, as a path towards realizing quantum advantage: we can first classically simulate PQC optimization and then continue the work on quantum hardware. In this paper, we call this approach of classical optimization the variational tensor network eigensolver (VTNE). With VTNE, we bridge the gap between classical and quantum optimization by demonstrating how to find a good set of intermediate parameters for the VQE circuit by first approximately classically optimizing the VQE using tensor networks.

II. METHODS

We use MPSs as the tool for efficient quantum circuit simulation and optimization. To accomplish our goal of using MPSs as a pre-optimization tool, we start with a fixed gate structure on the quantum hardware, as might naturally be dictated by the physical device. For each set of parameters of the quantum circuit, we map it onto an approximate MPS. This MPS realization of the quantum circuit is then the starting point from which we compute an approximate energy and its derivative given the circuit parameters. The MPS is generated by starting with a tensor network where each unitary gate within the circuits is translated into a rank 2d tensor, where d represents the number of qubits the gate interfaces with, which is then contracted into an MPS.

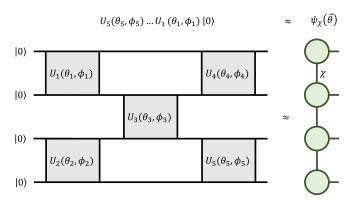


Figure 1. Contraction of a quantum circuit $|\psi_{PQC}\rangle$ into an MPS $|\psi_{\chi}\rangle$. With the contraction of each gate, the bond dimension is capped at χ .

The task of transforming a state obtained from a PQC into an MPS form presents challenges for highly entangled states due to the bond dimension of the MPS. To address this, we approximate the quantum circuit as an

MPS with a fixed bond dimension χ significantly smaller than the maximum bond dimension $\chi_{\rm max} = 2^{n_q/2}$, where n_q is the number of qubits. We then examine how the improved classical starting configuration depends on the bond dimension of the MPS which controls the computational complexity of the classical optimization. Starting with a good point, this approach may help alleviate the difficulties of executing VQE on quantum computers and set the stage for a more explicit demonstration of quantum advantage.

A. Model

In this work, we consider the one and two-dimensional Fermi-Hubbard model. This model is particularly interesting because its regular structure and relatively simple form suggest that it may be easier to implement on NISQ devices [44]. We anticipate the high-level approaches we introduce will also apply to other condensed-matter systems. The Hamiltonian for the Hubbard model is

$$H = -t \sum_{\langle i,j\rangle,\sigma\in\{\uparrow,\downarrow\}} (a_{i\sigma}^{\dagger} a_{j\sigma} + a_{j\sigma}^{\dagger} a_{i\sigma}) + U \sum_{i} n_{i\uparrow} n_{i\downarrow}, (1)$$

where $a_{i\sigma}^{\dagger}$, $a_{i\sigma}$ are fermionic creation and annihilation operators; $n_{i\sigma} = a_{i\sigma}^{\dagger} a_{i\sigma}$ and $\langle i,j \rangle$ correspond to nearest neighbors, t is the nearest neighbor hopping, and U is the on-site potential. Throughout the rest of this paper, we use t=1 and U=2 and work in the half-filled regime. We use the well-known Jordan-Wigner encoding of the fermionic Hamiltonian as a qubit Hamiltonian [45].

B. Ansatz

We consider the number-preserving ansatz used in ref. [5], which in the case of the Fermi-Hubbard model, is more general than its associated Hamiltonian variational ansatz (HVA) [38]. This ansatz consists of a parameterized number-preserving gate

$$NP(\theta,\phi) = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & \cos\theta & -i\sin\theta & 0\\ 0 & i\sin\theta & \cos\theta & 0\\ 0 & 0 & 0 & e^{i\phi} \end{pmatrix}$$
(2)

and a parameter-less fermionic swap gate

$$FSWAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3}$$

This circuit consists of qubits, labelled as (i, j, σ) , patterned on a 1D or 2D lattice where (i, j) specifies a lattice site position and $\sigma \in \{\uparrow, \downarrow\}$. (i, j, \uparrow) is always directly to the right of (i, j, \downarrow) . A layer of this ansatz starts with a set of two-qubit gates interacting between (i, j, \uparrow) and (i, j, \downarrow) . Following this, we have horizontal and vertical hopping gates between the four commuting sets of hopping terms $(i, j, \sigma) \leftrightarrow (i, j + 1, \sigma); (i, j, \sigma) \leftrightarrow (i + 1, j, \sigma);$ $(i, j+1, \sigma) \leftrightarrow (i, j+2, \sigma); (i+1, j, \sigma) \leftrightarrow (i+2, j, \sigma);$ where i and j are even. Between each set, a group of fermionic swaps is performed so that the corresponding sites being hopped between are consecutive. Illustrated in figure 2 for a 2×2 lattice, the number-preserving ansatz has shown success in capturing the ground state energy of the Fermi-Hubbard model for up to 24 qubits [5]. We prepend this ansatz in our simulations with $R_z(\theta)$ gates at each qubit. In testing, we found this layer helped to improve optimization.

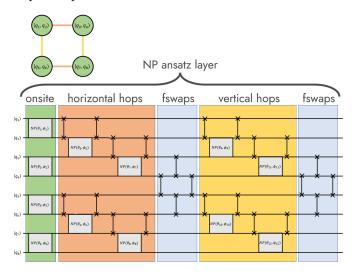


Figure 2. One layer of the number-preserving ansatz for the 2×2 lattice. Consisting of number-preserving gates defined in Eq. 2 and fermionic swap gates defined in Eq. 3, the layer starts with onsite interacting gates, followed by horizontal hopping gates, where fermionic swap operators are used to bring columns-wise nearest-neighbor same-spin wires together. Then, fermionic swaps are performed to bring rows-wise nearest neighbor same-spin wires together. Finally, fermionic swaps are used to return all the wires to their original positions.

C. MPS from Quantum Circuit

Our goal now is to find good parameters for the variational circuit classically. To accomplish this, we will need to approximate the VQE circuit as an MPS, which is done as follows. Let

$$\left| \psi_{\text{PQC}} \left(\vec{\theta} \right) \right\rangle = U_n(\vec{\theta}_n) \dots U_2(\vec{\theta}_2) U_1(\vec{\theta}_1) \left| \psi_0 \right\rangle$$
 (4)

where $|\psi_0\rangle$ is an unparameterized starting state. We approximately represent this wave-function as a bond dimension χ MPS $|\psi_{\chi}(\vec{\theta})\rangle$ (See figure 1.) The approximation is performed via the time-evolving block decimation (TEBD) technique [46]. This construction provides us with a truncated representation of our PQC, so that we can classically optimize the energy function

$$E_{\chi}(\vec{\theta}) = \left\langle \psi_{\chi}(\vec{\theta}) \middle| H \middle| \psi_{\chi}(\vec{\theta}) \right\rangle, \tag{5}$$

where H is represented as a matrix product operator (MPO). Generally, k-local Hamiltonians have a simple MPO implementation [47, 48]. As we increase χ , the optimized energy E_{χ} gets closer to the exact energy $E_{\rm exact}$

$$E_{\text{exact}}(\vec{\theta}) = \left\langle \psi_{\text{PQC}}(\vec{\theta}) \middle| H \middle| \psi_{\text{PQC}}(\vec{\theta}) \right\rangle. \tag{6}$$

Note that $E_{\rm exact}(\vec{\theta}) = E_{\chi_{\rm max}}(\vec{\theta})$, where $\chi_{\rm max} = 2^{n_q/2}$. Throughout this paper, we use the ITensor package [49] to compute all our tensor network calculations.

D. Optimization

Given a bond dimension χ , the objective function that we optimize is Eq. 5. We begin our optimization by finding the ground state of the non-interacting (U=0)case. We optimize two non-interacting number-preserving ansatze; one in which spins occupy the even sites (for the spin-up determinant) and one where the spins occupy the odd sites (for the spin-down determinant) so that when we consider the full interacting system, the state starts with a checkerboard of up and down spin configurations. In this optimization, we need only half the qubits and can remove both the onsite gates, any swap gates required to cross over different flavored spins, and any hopping terms on the different flavored spins. This leaves less the half the number of parameters to optimize. We initialize the parameters using a Gaussian distribution $\mathcal{N}(0, 10^{-5})$, and we carry out the minimization using the Broyden-Fletcher-Goldfarb-Shannon (BFGS) algorithm [50, 51]. We terminate the optimization when the function tolerance reaches 10^{-7} or the gradient norm reaches 10^{-6} . Note that if $\chi < 2^{(n_x n_y/2)}$, this optimization will not necessarily give us parameters representing the exact ground state for the non-interacting case. After

performing the non-interacting optimization, we use those parameters to start the optimization for the interacting (U=2) case.

Classical techniques exist to efficiently compute the gradient of Eq. 5, which include automatic differentiation (AD) [52]. Here, we implement an approximate gradient scheme that does not require as much memory and only needs two circuit evaluations. To derive our approximation, we start with the gradient of the exact energy of the full PQC. For a unitary $U_k(\vec{\theta}_k)$ containing a parameter $\theta_{k,i} \in \vec{\theta}_k$, the derivative of the energy with respect to that parameter is

$$\frac{\partial E_{\text{exact}}(\vec{\theta})}{\partial \theta_{k,i}} = 2 \operatorname{Re} \left[\left\langle \psi_L^{(k)} \middle| \frac{\partial U_k^{\dagger}(\vec{\theta}_k)}{\partial \theta_{k,i}} \middle| \psi_R^{(k)} \right\rangle \right], \quad (7)$$

where

$$\left\langle \psi_L^{(k)} \right| = \left\langle 0 \middle| U_1^{\dagger}(\vec{\theta}_1) \dots U_{k-1}^{\dagger}(\vec{\theta}_{k-1}) \right. \tag{8}$$

$$\left|\psi_{R}^{(k)}\right\rangle = U_{k+1}^{\dagger}(\vec{\theta}_{k+1})\dots U_{n}^{\dagger}(\vec{\theta}_{n})H\left|\psi_{PQC}\right\rangle$$
 (9)

We can then iteratively compute the derivative with respect to each parameter by updating $\left\langle \psi_L^{(k)} \right|$ and $\left| \psi_R^{(k)} \right\rangle$ recursively:

$$\left\langle \psi_L^{(k-1)} \right| = \left\langle \psi_L^{(k)} \middle| U_{k-1}(\vec{\theta}_{k-1}) \right. \tag{10}$$

$$\left|\psi_{R}^{(k-1)}\right\rangle = U_{k}^{\dagger}(\vec{\theta}_{k})\left|\psi_{R}^{(k)}\right\rangle \tag{11}$$

Figure 3 depicts this process through tensor network diagrams. This gradient computation is exact in the limit of the maximal bond dimension. In our MPS approximation regime, we truncate $\left\langle \psi_L^{(k-1)} \right|$ and $\left| \psi_R^{(k-1)} \right\rangle$ in Eq. 11 to bond dimension χ after each iteration. This derivative approximates the derivative of Eq. 6, which is not the same as the derivative of Eq. 5.

III. RESULTS

We start by looking at the results for the 1D and 2D Hubbard models at various lattice sizes. For each lattice size (n_x, n_y) , the number of layers used per lattice configuration was chosen by using the results of [5] which provided depths that led exact-VQE to 0.99 fidelity with the true ground state for systems up to 24 qubits. We used those circuit depths, and for systems with more than 24 qubits, we linearly extrapolated from the depth vs qubit data. Table I shows the lattice configurations used along

n_x	n_y	qubits	layers
4	1	8	4
8	1	16	7
12	1	24	11
16	1	32	14
4	2	8	10
4	3	24	17
4	4	32	24

Table I. A table of lattice configurations with the number of layers used for the number-preserving ansatz.

with the number of layers used in the number-preserving ansatz.

Figure 4 shows the relative energy error of the ground state as we increase the layer depth for 1×8 and 2×4 system sizes at different bond dimensions. When the VQE is performed using the full bond dimension ($\chi = 256$), increasing the number of layers leads to more accurate ground state energies. Figure 4 also tells us for a given depth, how well we optimize at a certain bond dimension. In the case of the 1D (8x1) system, we find that at one layer, optimizing at bond dimension 16 is enough to completely optimize our ansatz. This is evident by observing that optimizing at bond dimensions 32 and 256 yields the same relative energy error as with bond dimension 16. When we add another layer, a bond dimension of 16 is no longer enough to represent the state exactly, but optimizing at a bond dimension of 32 is. Once we are at 4 layers the bond dimension 32 optimization is no longer able to exactly represent the PQC. Overall, we see that as we add layers to the ansatz, we require more bond dimension to more accurately represent our PQC and drive our optimization down. Turning to the 2D (4x2) system, we find that neither bond dimensions 16 nor 32 are enough to fully represent the PQC even for one layer.

VQE optimizations are performed by capping the MPS bond dimension χ_b , resulting in optimized parameters θ_b^* . We can then compute the energy given by contracting a PQC with these parameters into an MPS of bond dimension χ_a as

$$E_{\chi_a}(\vec{\theta_b^*}) = \left\langle \psi_{\chi_a}(\vec{\theta_b^*}) \middle| H \middle| \psi_{\chi_a}(\vec{\theta_b^*}) \right\rangle, \tag{12}$$

Note that when $\chi_a = 2^{n_x n_y}$, this quantity yields the exact energy given by the PQC state with the parameters θ_h^* .

Figure 5 shows the relative percent energy errors for the Fermi-Hubbard model VQE for different system sizes. Shown on the plots are $E_{\chi_b}(\vec{\theta_b^*})$, $E_{\min(2^{n_x n_y}, 512)}(\vec{\theta_b^*})$, and DMRG energies as a function of the optimization bond dimension χ_b . For larger system sizes, the exact energies

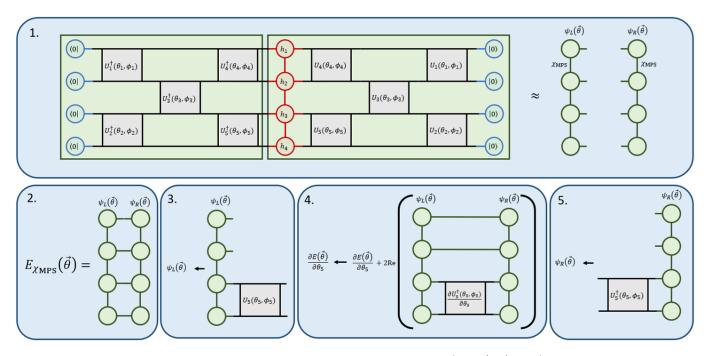


Figure 3. Procedure for computing the approximate energy (1-2) $E_{\chi_{\text{MPS}}}(\vec{\theta}) = \langle \psi_{\chi}(\vec{\theta}) | H | \psi_{\chi}(\vec{\theta}) \rangle$ and (3-5) its gradient. (1) We first construct two MPSs $\psi_L \approx \langle 0 | U_1^{\dagger}(\vec{\theta}_1) U_2^{\dagger}(\vec{\theta}_2) \dots U_n^{\dagger}(\vec{\theta}_n)$ and $\psi_R \approx H U_n(\vec{\theta}_n) \dots U_2(\vec{\theta}_2) U_1(\vec{\theta}_1) | 0 \rangle$. (2) The approximate energy is then the contraction of these two MPSs $E_{\text{MPS}} = \langle \psi_L | \psi_R \rangle$. (3) We pop out the last gate from ψ_L by applying the inverse to it. (4) We then compute gradient terms with respect to the parameters in the gate by performing the contraction $\langle \psi_L | \vec{\nabla} U_n(\vec{\theta}_n) | \psi_R \rangle$. (5) We move the popped gate from ψ_L to ψ_R . We repeat (3-5) until ψ_L is parameter-less.

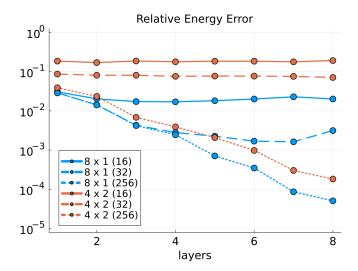


Figure 4. Relative energy error vs layer depth for a 8×1 (blue) and 4×2 (red) Hubbard model for bond dimensions $\chi = 16$ (solid lines), $\chi = 32$ (dashed lines), and the full bond dimension $\chi_{\rm exact} = 256$ (dotted lines).

that the optimized energies are being compared to are extrapolated from lower bond dimension DMRG energies.

Our VQE simulations using MPSs display a convergence pattern similar to that of DMRG, albeit with some overhead. We observe that the relative energy errors decrease as the bond dimension increases, highlighting the importance of the bond dimension in obtaining accurate ground state energies. Furthermore, our results indicate that representing the PQC with a larger MPS bond dimension, $\chi_a > \chi_b$, while retaining the optimization parameters, yields energies that are comparable or even better than those obtained with an MPS bond dimension χ_b . In other words, for $\chi_a > \chi_b$, we have $E_{\chi_a}(\vec{\theta_b^*}) \lesssim E_{\chi_b}(\vec{\theta_b^*})$. This feature tells us that when we only have classical access to a bond dimension of χ_b , and therefore energy $E_{\chi_b}(\vec{\theta_b^*})$, the true energy of the ansatz $E_{\chi_a}(\vec{\theta_b^*})$, where $\chi_a = 2^{n_x n_y}$, when ran on quantum hardware, will have equal or lower energies.

For most 1D Hamiltonians of interest, the convergence of DMRG is very rapid and in practice, large enough bond dimension and system size are accessible numerically [53]. In our 1D VQE simulations, we find that, just

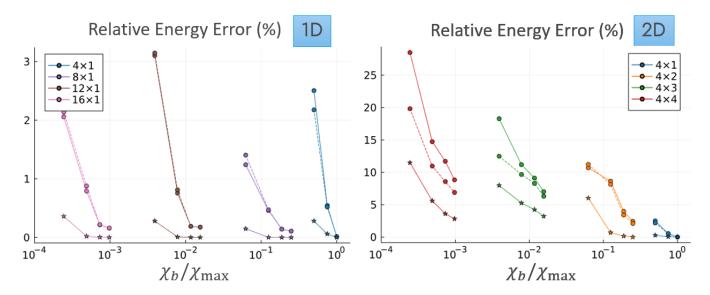


Figure 5. Relative percent energy errors for the Fermi-Hubbard model VQE for different system sizes. The x-axis represents the bond dimension χ_b that the circuit ansatz is represented as normalized to the maximum bond dimension necessary to exactly represent the circuit for an $n_x \times n_y$ sized system ($\chi_{\text{exact}} = 2^{n_x n_y}$). The left plot shows a comparison of 1D Hubbard models, and the right plot shows a comparison of 2D systems. The solid lines represent the VQE relative energy when the circuit is approximated using bond dimension χ_b , and the dashed lines represent the VQE relative energy when the circuit is approximated by an MPS of bond dimension $\chi = \min(512, 2^{n_x n_y})$. DMRG energies (starred points) at these optimization bond dimensions are also shown.

as in DMRG, we only need a relatively small bond dimension to get close to the ground state. As a result, in figure 5 we see a very weak dependence between the relative energy errors and the system size. For all 1D systems (up to 32 qubits), a bond dimension of $\chi_b = 16$ gets us a relative energy error of $\lesssim 3\%$.

Focusing on 2D systems, as shown in figure 5, we observe that the relative energy errors tend to increase with larger lattice sizes, indicating that achieving an accurate ground state energy becomes more challenging as the system size grows. This is consistent with our expectations, as larger systems exhibit a more complex entanglement structure, which necessitates a higher bond dimension for accurate representation. Nevertheless, the classical VQE demonstrates a similar trend as DMRG as the bond dimension is tuned. To further assess the performance of our VQE simulations, we show in figure 6 the fidelity errors for selected system sizes which corroborates the effectiveness of our approach in approximating the ground state of the Fermi-Hubbard model.

Most importantly, we study whether doing VQE on top of the parameters that were classically found for the PQC converges faster than starting with random parameters and running VQE without any classical pre-optimization.

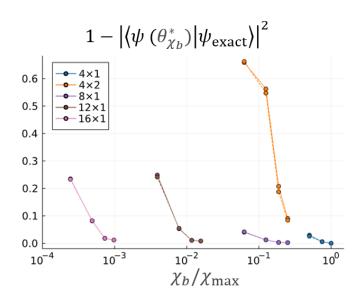


Figure 6. Infidelity measurements between VQE optimized ground states and the exact DMRG obtained ground state.

To explore this, we compare parameters found via classi-

cal optimization with random initial parameters. Looking at the 4×2 lattice (where $\chi_{\text{max}} = 256$), we sample 10 parameter sets $\{\vec{\theta}_0^{(i)}\}_{i=1}^{10}$, with each parameter randomly picked from a Gaussian distribution $\mathcal{N}(0, 10^{-3})$. We compare these parameters with the classically optimized parameters $\vec{\theta}_{\chi}$ obtained above. We run a full VQE simulation with the Adam optimizer ($\alpha = 0.001$) using parameter sets $\{\vec{\theta}_0^{(i)}\}_{i=1}^{10}$ and $\vec{\theta}_{\chi_b}$. Figure 7a shows the average relative energy error vs the optimization step (or number of gradient calls) for $\vec{\theta}_0$, $\vec{\theta}_{16}$, $\vec{\theta}_{32}$, and $\vec{\theta}_{64}$. Overall we find that optimizing classically beforehand with an MPS backend significantly saves the number of gradient evaluations needed to reach lower energies. For example, we save about 400 gradient calls by performing an approximate VQE with a bond dimension 16 MPS and about 1,000 gradient calls using a bond dimension 32 MPS. We also test the robustness of this method by removing the step involving optimizing the non-interacting case instead. That is, we use VTNE to obtain optimized parameters $\vec{\theta}_{\chi_b}^{(i)}$ for each $\vec{\theta}_0^{(i)}$. We find similar advantages shown in figure 7. In general, for ansatz where we do not have a good starting point, a large number of computations are needed to match the classically optimized starting configuration. This suggests that classical optimization before employing quantum hardware plays an instrumental role in guiding the quantum algorithm toward the global minima more efficiently, which is especially crucial given the current limited quantum resources.

IV. DISCUSSION

It is worth distinguishing our approach from alternative approaches that work directly with the best bond-dimension MPS [54–56]. In our approach, we find parameters for a given class of parameterized circuits which can be chosen to be shallow or commensurate with the hardware. These parameters can then be utilized to initialize quantum states on the device. Approaches that work directly with the MPS and then add parameterized gates on top of them are inherently forced to work with deeper circuits as quantum MPS scales quadratically with the bond dimension [57–59] and linearly with the system size. To put this in context, an arbitrary MPS of bond dimension χ consists of $\lceil \log_2 2\chi \rceil$ -qubit unitary operators. Such unitaries require

CNOTs
$$\leq \frac{23}{16} \Delta_{\chi}^2 \chi^2 - 3\Delta_{\chi} \chi + \frac{4}{3}$$
 (13)
 $\sim \mathcal{O}(\chi^2)$

CNOTS, where $\Delta_{\chi} = 2^{\lceil \log_2 2\chi \rceil - \log_2 2\chi}$ [60]. Put an-

other way, to construct a circuit with a constraint of depth D, then the bond dimension of the MPS must be $\chi \lesssim \mathcal{O}(\sqrt{D/N})$, where N is the system size. Considering Eq. 13, a $\chi=64$ MPS could require up to 7,660 CNOT gates per site. In addition to this depth dependency, there is also significantly less freedom in choosing the type of circuit architecture being optimized.

With regards to both the approach that maps DMRG MPS states to circuits and VTNE, increasing bond dimension leads better performance. However, when it comes to implementing the circuits obtained from these algorithms on quantum hardware, other factors also become relevant. VTNE focuses on generic circuits that will often be shallow or tailored to various hardware constraints, while circuits from DMRG-optimized MPSs have different costs with regards to their implementation. This allows DMRG to achieve states with lower energies for a given bond dimension but at the cost of deeper circuits when the MPS is implemented on quantum hardware. Crucially, if we fix the circuit depth, DMRG-generated MPSs are restricted to bond dimensions that depend on the system size $(\sqrt{D/N})$. This constraint results in higher energy states when compared to VTNE as we increase the system size. Thus, when circuit depth is a limiting factor, VTNE followed by VQE is a competitive and efficient approach for achieving low energy states.

V. CONCLUSIONS

We have demonstrated that VTNE, which classically pre-optimizes circuits by approximately simulating VOE using MPS significantly aids VQE optimization, with the bond dimension of the MPS tuning how much work (or the number of gradient evaluations) is saved on quantum hardware. The work here explores ideas that have been discussed in the recent literature [40] and provides a complementary approach to other approximate quantum circuit simulations [23, 39]. For 1D Hamiltonians, the convergence of our method is rapid, with the difference between the energy of the classically optimized parameters and the true ground state energy only depending weakly on system size. In this case, most, if not all of the optimization can be performed classically, and in this regime, this method becomes an algorithm for state-preparation on quantum hardware. In contrast, 2D systems exhibit increased relative energy errors with larger lattice sizes due to more complex entanglement structures, emphasizing the necessity of higher bond dimensions for accurate representation. For 2D systems, our algorithm serves as pre-optimization for generating an initial set of parameters for VQE.

VQE Simulations with various initial parameters for the 4x2 Hubbard Model

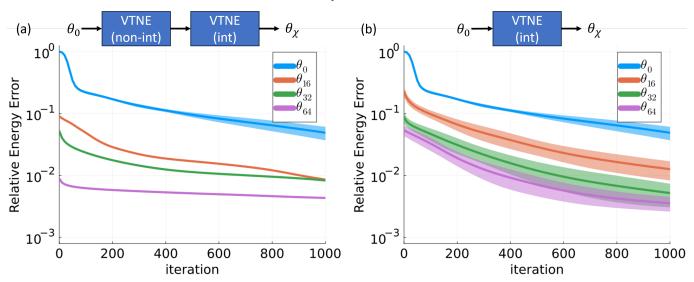


Figure 7. Relative energy error vs optimization step for different parameter initialization methods for the 4×2 Hubbard Model. Shown are the mean relative energy errors at each step, averaged over 10 sampled parameter sets. The blue line represents initial parameters $\vec{\theta}_0$, randomly sampled from a Gaussian distribution $\mathcal{N}(0, 10^{-3})$. The red, green, and purple lines correspond to optimizations initialized with parameters $\vec{\theta}_{16}$, $\vec{\theta}_{32}$, and $\vec{\theta}_{64}$, respectively. In (a), these parameters are found by first optimizing the non-interacting Hubbard Model first, then the interacting case. In (b), the parameters are obtained by classically optimizing an MPS-assisted VQE simulation at the specified bond dimensions, with the classical optimizer initialized using the parameter set $\vec{\theta}_0$.

Our work illustrates the effectiveness of using an approximate tensor network backend for VQE, facilitating accurate ground state energy estimation and efficient circuit initialization for large system sizes. This approach stands to enhance the scalability and feasibility of VQE on near-term quantum hardware and extends its applicability to a wide range of quantum many-body and chemistry problems.

Transformational Tools and Technology (TTT) Project. A.K. acknowledges support from USRA NASA Academic Mission Services under contract No. NNA16BD14C through participation in the Feynman Quantum Academy internship program. BKC acknowledges support from the Department of Energy grant DOE DESC0020165.

ACKNOWLEDGEMENTS

We are grateful for support from NASA Ames Research Center. We acknowledge funding from the NASA ARMD

S. McArdle, X. Yuan, and S. Benjamin, Physical Review Letters 122, 180501 (2019), publisher: American Physical Society.

^[2] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Phys. Rev. A 95, 042308 (2017).

^[3] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A.

- de Jong, and I. Siddiqi, Phys. Rev. X 8, 011021 (2018).
- [4] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Physical Review A 98, 062339 (2018), publisher: American Physical Society.
- [5] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, Phys. Rev. B 102, 235122 (2020).
- [6] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz (2017).
- [7] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, npj Quantum Information 6, 1 (2020), number: 1 Publisher: Nature Publishing Group.
- [8] P. O'Malley, R. Babbush, I. Kivlichan, J. Romero, J. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. White, P. Coveney, P. Love, H. Neven, A. Aspuru-Guzik, and J. Martinis, Physical Review X 6, 031007 (2016), publisher: American Physical Society.
- [9] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, S. Filipp, and I. Tavernelli, Phys. Rev. A 98, 022322 (2018).
- [10] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, Algorithms 12, 10.3390/a12020034 (2019).
- [11] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Phys. Rev. A 97, 022304 (2018).
- [12] S. H. Sureshbabu, D. Herman, R. Shaydulin, J. Basso, S. Chakrabarti, Y. Sun, and M. Pistoia, Parameter setting in quantum approximate optimization of weighted problems (2023), arXiv:2305.15201 [quant-ph].
- [13] P. C. Lotshaw, K. D. Battles, B. Gard, G. Buchs, T. S. Humble, and C. D. Herold, Phys. Rev. A 107, 062406 (2023).
- [14] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm (2014), arXiv:1411.4028 [quant-ph].
- [15] V. Kremenetski, A. Apte, T. Hogg, S. Hadfield, and N. M. Tubman, Quantum alternating operator ansatz (qaoa) beyond low depth with gradually changing unitaries (2023), arXiv:2305.04455 [quant-ph].
- [16] V. Kremenetski, T. Hogg, S. Hadfield, S. J. Cotton, and N. M. Tubman, Quantum alternating operator ansatz (qaoa) phase diagrams and applications for quantum chemistry (2021), arXiv:2108.13056 [quant-ph].
- [17] 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, and K. B. Whaley, Postponing the orthogonality catastrophe: efficient state preparation for electronic structure simulations on quantum devices (2018), arXiv:1809.05523 [quant-ph].
- [18] N. P. Sawaya, D. Marti-Dafcik, Y. Ho, D. P. Tabor, D. Bernal, A. B. Magann, S. Premaratne, P. Dubey,

- A. Matsuura, N. Bishop, W. A. de Jong, S. Benjamin, O. D. Parekh, N. Tubman, K. Klymko, and D. Camps, Hamlib: A library of hamiltonians for benchmarking quantum algorithms and hardware (2023), arXiv:2306.13126 [quant-ph].
- [19] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature 549, 242 (2017), number: 7671 Publisher: Nature Publishing Group.
- [20] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Phys. Rev. A 95, 020501 (2017).
- [21] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Phys. Rev. X 8, 031022 (2018).
- [22] Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, K. M. Beck, A. Blinov, V. Chaplin, M. Chmielewski, C. Collins, S. Debnath, K. M. Hudek, A. M. Ducore, M. Keesan, S. M. Kreikemeier, J. Mizrahi, P. Solomon, M. Williams, J. D. Wong-Campos, D. Moehring, C. Monroe, and J. Kim, npj Quantum Information 6, 1 (2020), number: 1 Publisher: Nature Publishing Group.
- [23] J. W. Mullinax and N. M. Tubman, Large-scale sparse wavefunction circuit simulator for applications with the variational quantum eigensolver (2023), arXiv:2301.05726 [quant-ph].
- [24] A. Jahin, A. C. Y. Li, T. Iadecola, P. P. Orth, G. N. Perdue, A. Macridin, M. S. Alam, and N. M. Tubman, Phys. Rev. A 106, 022434 (2022).
- [25] A. Cervera-Lierta, J. S. Kottmann, and A. Aspuru-Guzik, PRX Quantum 2, 020329 (2021).
- [26] J. Rogers, G. Bhattacharyya, M. S. Frank, T. Jiang, O. Christiansen, Y.-X. Yao, and N. Lanatà, Error mitigation in variational quantum eigensolvers using probabilistic machine learning (2021).
- [27] M. Ostaszewski, L. M. Trenkwalder, W. Masarczyk, E. Scerri, and V. Dunjko, Reinforcement learning for optimization of variational quantum circuit architectures (2021).
- [28] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev, Materials Theory 6, 2 (2022).
- [29] M. S. Jattana, F. Jin, H. De Raedt, and K. Michielsen, Phys. Rev. Appl. 19, 024047 (2023).
- [30] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nature Communications 9, 4812 (2018), number: 1 Publisher: Nature Publishing Group.
- [31] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Nature Communications 12, 1791 (2021), number: 1 Publisher: Nature Publishing Group.
- [32] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn, PRX Quantum 3, 020365 (2022).
- [33] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, PRX Quantum 3, 010313 (2022).
- [34] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone,
 L. Cincio, and P. J. Coles, Nature Communications 12,
 6961 (2021), number: 1 Publisher: Nature Publishing

- Group.
- [35] E. R. Anschuetz, Critical points in quantum generative models (2021).
- [36] E. R. Anschuetz and B. T. Kiani, Nature Communications 13, 7760 (2022), number: 1 Publisher: Nature Publishing Group.
- [37] A. Arrasmith, Z. Holmes, M. Cerezo, and P. J. Coles, Quantum Science and Technology 7, 045015 (2022), publisher: IOP Publishing.
- [38] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A 92, 042303 (2015).
- [39] M. R. Hirsbrunner, D. Chamaki, J. W. Mullinax, and N. M. Tubman, Beyond mp2 initialization for unitary coupled cluster quantum circuits (2023), arXiv:2301.05666 [quant-ph].
- [40] U. Baek, D. Hait, J. Shee, O. Leimkuhler, W. J. Huggins, T. F. Stetina, M. Head-Gordon, and K. B. Whaley, Say no to optimization: A non-orthogonal quantum eigensolver (2022), arXiv:2205.09039 [quant-ph].
- [41] J. Chen, H.-P. Cheng, and J. K. Freericks, Journal of Chemical Theory and Computation 17, 841 (2021), pMID: 33503376, https://doi.org/10.1021/acs.jctc.0c01052.
- [42] Y. Zhou, E. M. Stoudenmire, and X. Waintal, Phys. Rev. X 10, 041038 (2020).
- [43] T. Ayral, T. Louvet, Y. Zhou, C. Lambert, E. M. Stoudenmire, and X. Waintal, A density-matrix renormalization group algorithm for simulating quantum circuits with a finite fidelity (2022).
- [44] F. Arute, K. Arva, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, A. Bengtsson, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, Y.-A. Chen, B. Chiaro, R. Collins, S. J. Cotton, W. Courtney, S. Demura, A. Derk, A. Dunsworth, D. Eppens, T. Eckl, C. Erickson, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, J. A. Gross, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. Huggins, L. B. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. N. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, M. Marthaler, O. Martin, J. M. Martinis, A. Marusczyk, S. McArdle, J. R. McClean, T. McCourt, M. McEwen, A. Megrant, C. Mejuto-Zaera, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Newman, M. Y. Niu, T. E. O'Brien, E. Ostby, B. Pató, A. Petukhov, H. Putterman, C. Quintana, J.-M. Reiner, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, P. Schmitteckert, M. Szalay, N. M. Tubman, A. Vainsencher, T. White,

- N. Vogt, Z. J. Yao, P. Yeh, A. Zalcman, and S. Zanker, Observation of separated dynamics of charge and spin in the fermi-hubbard model (2020), arXiv:2010.07965 [quant-ph].
- [45] S. E. Barnes and S. Maekawa, Journal of Physics: Condensed Matter 14, L19 (2001).
- [46] G. Vidal, Phys. Rev. Lett. 93, 040502 (2004).
- [47] R. Orús, Annals of Physics 349, 117 (2014).
- [48] J. C. Bridgeman and C. T. Chubb, Journal of Physics A: Mathematical and Theoretical **50**, 223001 (2017).
- [49] M. Fishman, S. R. White, and E. M. Stoudenmire, SciPost Phys. Codebases , 4 (2022).
- [50] C. G. BROYDEN, IMA Journal of Applied Mathematics 6, 76 (1970), https://academic.oup.com/imamat/articlepdf/6/1/76/2233756/6-1-76.pdf.
- [51] R. Fletcher, The Computer Journal 13, 317 (1970), https://academic.oup.com/comjnl/articlepdf/13/3/317/988678/130317.pdf.
- [52] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, Phys. Rev. X 9, 031041 (2019).
- [53] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull (Simons Collaboration on the Many-Electron Problem), Phys. Rev. X 5, 041041 (2015).
- [54] M. S. Rudolph, J. Miller, J. Chen, A. Acharya, and A. Perdomo-Ortiz, Synergy between quantum circuits and tensor networks: Short-cutting the race to practical quantum advantage (2022), arXiv:2208.13673 [quant-ph].
- [55] M. Foss-Feig, D. Hayes, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, J. M. Pino, and A. C. Potter, Phys. Rev. Res. 3, 033002 (2021).
- [56] S.-J. Ran, Phys. Rev. A **101**, 032310 (2020).
- [57] E. J. Gustafson, A. C. Y. Li, A. Khan, J. Kim, D. M. Kurkcuoglu, M. S. Alam, P. P. Orth, A. Rahmani, and T. Iadecola, Preparing quantum many-body scar states on quantum computers (2023), arXiv:2301.08226 [quant-ph].
- [58] W. Huggins, P. Patil, B. Mitchell, K. B. Whaley, and E. M. Stoudenmire, Quantum Science and Technology 4, 024001 (2019).
- [59] M. Mottonen and J. J. Vartiainen, Decompositions of general quantum gates (2005), arXiv:quant-ph/0504100 [quant-ph].
- [60] V. Shende, S. Bullock, and I. Markov, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 25, 1000 (2006).