Letter

Avoiding barren plateaus via transferability of smooth solutions in a Hamiltonian variational ansatz

Antonio A. Mele,[1,2] Glen B. Mbeng,[3] Giuseppe E. Santoro,[2,4,5] Mario Collura,[2,6] and Pietro Torta[2]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*SISSA, Via Bonomea 265, I-34136 Trieste, Italy*
[3]*Universität Innsbruck, Technikerstraße 21 a, A-6020 Innsbruck, Austria*
[4]*International Centre for Theoretical Physics (ICTP), P.O. Box 586, I-34014 Trieste, Italy*
[5]*CNR-IOM, Consiglio Nazionale delle Ricerche - Istituto Officina dei Materiali, c/o SISSA Via Bonomea 265, 34136 Trieste, Italy*
[6]*INFN Sezione di Trieste, via Bonomea 265, 34136 Trieste, Italy*

(Received 6 July 2022; accepted 21 November 2022; published 19 December 2022)

A large ongoing research effort focuses on variational quantum algorithms (VQAs), representing leading candidates to achieve computational speed-ups on current quantum devices. The scalability of VQAs to a large number of qubits, beyond the simulation capabilities of classical computers, is still debated. Two major hurdles are the proliferation of low-quality variational local minima, and the exponential vanishing of gradients in the cost-function landscape, a phenomenon referred to as barren plateaus. In this work, we show that by employing iterative search schemes, one can effectively prepare the ground state of paradigmatic quantum many-body models, also circumventing the barren plateau phenomenon. This is accomplished by leveraging the transferability to larger system sizes of a class of iterative solutions, displaying an intrinsic smoothness of the variational parameters, a result that does not extend to other solutions found via random-start local optimization. Our scheme could be directly tested on near-term quantum devices, running a refinement optimization in a favorable local landscape with nonvanishing gradients.

DOI: 10.1103/PhysRevA.106.L060401

Introduction. Variational quantum algorithms (VQAs) [1,2] are among the main candidates for near-term practical applications of noisy intermediate-scale quantum (NISQ) devices [3]. VQAs are quantum-classical *hybrid* optimization schemes that have been successfully applied to quantum ground-state preparation [4–8] and classical optimization tasks [9], ranging from the solution of linear systems of equations [10] to quantum information [11]. In the standard VQA setting, one aims at minimizing the average energy of a problem Hamiltonian $\widehat{H}_{\rm targ}$ with respect to a variational state $|\psi(\boldsymbol{\gamma})\rangle$ prepared by a parameterized quantum circuit. This is accomplished by a feedback loop between a classical and a quantum machine: the quantum device is used to repeatedly prepare the ansatz state for a set of gate parameters $\boldsymbol{\gamma}$ and to estimate the cost function $E_{\rm var}(\boldsymbol{\gamma}) = \langle\psi(\boldsymbol{\gamma})|\widehat{H}_{\rm targ}|\psi(\boldsymbol{\gamma})\rangle$, while the optimization of the parameters is performed classically.

The optimization of the cost function $E_{\rm var}(\boldsymbol{\gamma})$ is known to be a difficult task [12]: only a careful choice of the ansatz is usually *expressive* enough to approximately find the ground state of $\widehat{H}_{\rm targ}$ and, at the same time, *trainable* enough for the optimization to succeed. In particular, the landscape of the cost function may not be easy to inspect for two reasons: the proliferation of low-quality local minima traps [13] and the exponential flattening of the landscape by increasing the number of qubits, a phenomenon dubbed *barren plateaus* [14], which can severely hinder the scalability of the VQA scheme beyond small system sizes amenable to classical simulations. Barren plateaus are linked to highly expressive parameterized quantum circuits [14–16], but they arise also in the context of less-expressive symmetry-preserving [17,18] or equivariant [19] *Ansätze*. A few recent studies have proposed different approaches to limit or avoid barren plateaus, by employing pretraining techniques [20], layerwise learning for classification tasks [21], identity-block initialization [22], or classical shadows [23,24].

Among the effective strategies to avoid low-quality local minima traps, we mention approaches [8,25–27] inspired by standard adiabatic quantum computation (AQC) [28,29], and iterative schemes [30,31], optimizing only a subset of gate parameters at each iteration and using this result as a warm-start guess for the next iterative step. These techniques proved particularly efficient for a class of variational states commonly dubbed Hamiltonian variational ansatz (HVA) [6,7,9,18,32–39], with an ansatz wave function of the form

$$|\psi(\boldsymbol{\gamma})\rangle = \prod_{m=1}^{P} e^{-i\gamma_{m,M}\widehat{H}_M}\cdots e^{-i\gamma_{m,1}\widehat{H}_1}|\psi_0\rangle. \tag{1}$$

Here, $m = 1\cdots P$ labels successive circuit layers, each in turn composed by $j = 1\cdots M$ alternating unitaries generated by Hamiltonian operators $\widehat{H}_j$, applied to a simple initial state $|\psi_0\rangle$. This family of ansatz wave functions draws inspiration from a digitized version of AQC (also known as digitized quantum annealing [40,41]). Indeed, the target Hamiltonian $\widehat{H}_{\rm targ}$ can be linearly decomposed in terms of the generators, so that each circuit layer resembles a Trotter split-up of the unitary evolution generated by $\widehat{H}_{\rm targ}$ for a small time step, with

the crucial difference that the angles $\gamma_{m,j}$ are not fixed by a Trotter split-up in this variational setting, but rather promoted to free variational parameters.

On a side note, this ansatz state can be regarded as a generalization of the quantum approximate optimization algorithm (QAOA) [9], originally devised for classical combinatorial optimization problems. Remarkably, by means of appropriate iterative schemes for constructing the layer parameters $\gamma_{m,j}$, it is often possible to efficiently single out optimal or nearly optimal variational parameters that are *smooth* functions [30,42–47] of the layer index m.

In this paper, we draw a connection between smooth optimal solutions—obtained by means of iterative methods—and barren plateaus, developing an efficient scheme to circumvent this issue. Our procedure leverages the *transferability* of an optimal *smooth* solution, obtained for small system size, to solve the same task with a larger number of qubits, where a direct optimization would fail due to barren plateaus. In a nutshell, the transferred smooth solution serves as an excellent warm start with low variational energy for the large system, and a subsequent refinement optimization is observed to be free of the barren plateau issue. Remarkably, even though other (nonsmooth) solutions for the small system can be obtained by standard random-start local optimization, they do not provide any useful warm start for larger systems and, crucially, a refinement optimization still suffers from barren plateaus in their neighborhood.

For definiteness, we focus on the ground-state preparation of the Heisenberg *XYZ* model [48] and of the antiferromagnetic longitudinal-transverse-field Ising model (LTFIM) [49], two ubiquitous models in quantum physics with rich phase diagrams, whose ground-state preparation with VQAs is affected by barren plateaus [17,19]. We select ansatz states in the form of Eq. (1), by choosing the generators $\widehat{H}_j$ in such a way to implement model symmetries into the variational wave functions. This leads to a restriction of the Hilbert space to the ground-state symmetry sector, boosting trainability, and a reduction in the number of independent Pauli correlators needed to compute the cost function.

Models and methods. The first class of models we consider is the spin-1/2 *XYZ* [48,50–52] Hamiltonian,

$$\widehat{H}_{\rm XYZ} = \sum_{j=1}^{N} \left(\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \Delta_{\rm Y} \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y + \Delta_{\rm Z} \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z\right), \tag{2}$$

restricting to the antiferromagnetic case $\Delta_{\rm Y}, \Delta_{\rm Z} > 0$. Second, we examine the antiferromagnetic LTFIM [49,53],

$$\widehat{H}_{\rm LTFIM} = \sum_{j=1}^{N} \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z - g_x \sum_{j=1}^{N} \hat{\sigma}_j^x - g_z \sum_{j=1}^{N} \hat{\sigma}_j^z, \tag{3}$$

with positive local fields $g_x, g_z > 0$. For both models, we examine even values of N and we assume periodic boundary conditions. In Supplemental Material (SM) [54], we briefly discuss the phase diagram of these models.

Our ansatz states are in the general form of Eq. (1) with $M = 2$ generating Hamiltonians only, defined to encode some symmetries of the model. To illustrate this idea for the *XYZ* case, let us split $\widehat{H}_{\rm XYZ}$ into two mutually noncommuting parts that refer to the *even*$(2j-1, 2j)$ and to the *odd* $(2j, 2j+1)$ bonds, $\widehat{H}_{\rm XYZ} = \widehat{H}_{\rm even} + \widehat{H}_{\rm odd}$, with

$$\widehat{H}_{\rm even} = \sum_{j=1}^{N/2} \left(\hat{\sigma}_{2j-1}^x \hat{\sigma}_{2j}^x + \Delta_{\rm Y} \hat{\sigma}_{2j-1}^y \hat{\sigma}_{2j}^y + \Delta_{\rm Z} \hat{\sigma}_{2j-1}^z \hat{\sigma}_{2j}^z\right), \tag{4}$$

and similarly for $\widehat{H}_{\rm odd}$. Next, in the spirit of AQC [28], imagine an *interpolating Hamiltonian* connecting $\widehat{H}_{\rm even}$ to the full $\widehat{H}_{\rm XYZ}$,

$$\widehat{H}(s) = s\widehat{H}_{\rm XYZ} + (1-s)\widehat{H}_{\rm even} = \widehat{H}_{\rm even} + s\widehat{H}_{\rm odd}, \tag{5}$$

with $s \in [0, 1]$. For $s = 0$, the ground state of $\widehat{H}(0) = \widehat{H}_{\rm even}$ is a valence-bond state of singlets on the even bonds, which is taken as the initial state.

This suggests, in close analogy with QAOA, the following ansatz for the *XYZ* ground-state wave function:

$$|\psi(\boldsymbol{\beta}, \boldsymbol{\alpha})_P\rangle = \widehat{U}_P \cdots \widehat{U}_2 \, \widehat{U}_1 |\psi_0\rangle. \tag{6}$$

Here, $(\boldsymbol{\beta}, \boldsymbol{\alpha})_P = (\beta_1 \cdots \beta_P, \alpha_1 \cdots \alpha_P)$ are $2P$ variational parameters, and the unitary operators $\widehat{U}_m = \widehat{U}(\beta_m, \alpha_m)$, for $m = 1 \cdots P$, evolve the state according to $\widehat{H}_{\rm even}$ and $\widehat{H}_{\rm odd}$, in an alternating fashion,

$$\widehat{U}_m = \widehat{U}(\beta_m, \alpha_m) = e^{-i\beta_m \widehat{H}_{\rm even}} e^{-i\alpha_m \widehat{H}_{\rm odd}}. \tag{7}$$

As usual in the VQA framework, the goal is to minimize the variational energy,

$$E_N(\boldsymbol{\beta}, \boldsymbol{\alpha})_P = \langle \psi(\boldsymbol{\beta}, \boldsymbol{\alpha})_P | \widehat{H}_{\rm targ} | \psi(\boldsymbol{\beta}, \boldsymbol{\alpha})_P \rangle, \tag{8}$$

with $\widehat{H}_{\rm targ} = \widehat{H}_{\rm XYZ}$. The connection with AQC is restored in the $P \to \infty$ limit by setting specific values for $(\boldsymbol{\beta}, \boldsymbol{\alpha})_P$, as prescribed by a Trotter split-up [55] of the continuous-time AQC dynamics [31].

The ansatz state lies in the same symmetry subsector of the *XYZ* ground states for the following symmetries (see SM [54]): translations by *two* lattice spacings $\widehat{\rm T}^2$ (which maps $j \to j+2$), lattice inversion $\widehat{\rm I}$ (which maps $j \leftrightarrow N-j+1$), and parity $\widehat{\rm P}_b = \prod_j \hat{\sigma}_j^b$. Additionally, for the SU(2)-invariant Heisenberg model, this holds true for the total spin $\widehat{\rm S}^b_{\rm tot}$ ($b = x, y, z$) and $\widehat{\rm S}^2_{\rm tot}$, while for the U(1)-invariant *XXZ* model, only for $\widehat{\rm S}^z_{\rm tot}$. As a result, the cost function in Eq. (8) requires the evaluation of only six independent two-point correlators, which may be further reduced to four (two) by exploiting rotational symmetries in the *XXZ* (*XXX*) case.

The ansatz for the ground-state preparation of the LTFIM reads as in Eq. (6), with a single layer unitary given by

$$\widehat{U}_m = e^{i\beta_m \widehat{H}_{\rm X}} e^{-i\alpha_m \left(\widehat{H}_{\rm ZZ} - g_z \widehat{H}_{\rm Z}\right)}, \tag{9}$$

where we defined $\widehat{H}_{\rm ZZ}$, $\widehat{H}_{\rm Z}$, and $\widehat{H}_{\rm X}$ simply as the sum of nearest-neighbor interactions, Pauli-z and Pauli-x operators, respectively. In this setting, the initial state is simply the fully polarized state along x, $|\psi_0\rangle = |+\rangle^{\otimes N}$, once again bearing a direct connection with AQC state preparation for $P \to \infty$. The goal is to minimize the variational energy as in Eq. (8), now with $\widehat{H}_{\rm targ} = \widehat{H}_{\rm LTFIM}$.

Incidentally, we remark that similar HVA wave functions can be adopted for other models, encompassing, e.g., Hubbard models and quantum chemistry applications [32], where the ansatz form would still be given by a Trotter split-up of a digitized quantum annealing evolution.

The checkerboard structure in terms of local unitary gates of the ansatz in Eqs. (6) and (7) resembles the time-evolving

block decimation (TEBD) algorithm [56,57]. However, in our case, the unitary gates are not predetermined by the generator of the unitary dynamics, but they are globally optimized in a variational setting. Nonetheless, a natural interpretation in terms of light-cone spreading of quantum correlations emerges for both of our ansatz wave functions (see SM [54]). As a main consequence, the whole cost-function landscape, once rescaled by the system size N, becomes independent of N itself for $N > \tilde{N}_P$, where $\tilde{N}_P = 4P + 2$ and $\tilde{N}_P = 2P + 1$ for the *XYZ* and the LTFIM, respectively.

Results. In this work, we adopt an iterative interpolation scheme (INTERP) [30,31] which was originally formulated for standard QAOA applied to classical optimization problems. Here, we apply this heuristic to more general HVA wave functions as in Eq. (1), with the goal of quantum many-body ground-state preparation. Essentially, the idea is to perform a sequence of local optimizations for increasing values of P, each of them starting from an educated guess that is iteratively updated, by interpolating on the optimal parameters found at the previous step. Additional details on this algorithm are reported in the SM [54], where we also provide numerical evidence that both *XYZ* and LTFIM ground states can be efficiently prepared across their phase diagrams. The code for numerical simulations is written with QISKIT [58] (using as classical optimizer the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm [59]).

In practice, one applies such iterative methods for a given system size N up to P layers, eventually finding a set of optimal variational parameters, $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N} = (\beta_1^* \cdots \beta_P^*, \alpha_1^* \cdots \alpha_P^*)$. Remarkably, the components of the two vectors $\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*$ are usually smooth functions of the layer index $m = 1 \cdots P$. For this reason, we dub them *smooth* solutions. This is consistently observed in all phases of our models, as shown in Fig. 1 at the critical point of *XXZ* (Heisenberg model) and close to the critical line of the LTFIM [49]. On top of that, we note that these smooth optimal curves are qualitatively similar for different system sizes. Inspired by this observation, we verify numerically that smooth optimal solutions $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G}$—obtained by applying INTERP to a small-size system with dimension N_G up to a certain value of P—can be transferred to solve the same task for a larger number of qubits, thus providing an effective educated guess. In the following, we will always indicate with N_G the "guess" size used to obtain the smooth optimal solution, which will be eventually transferred to a larger system with $N > N_G$ lattice sites. Unless otherwise stated, we set $N_G = 8$. In order to estimate the effectiveness of our transferability protocol, we define the *residual energy* as

$$\varepsilon_N(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G} = \frac{E_N(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G} - E_N^{\min}}{E_N^{\max} - E_N^{\min}}, \qquad (10)$$

where $E_N(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G}$ is the cost function in Eq. (8) for a system of size N, evaluated at fixed variational parameters $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G}$, while $E_N^{\min}$ ($E_N^{\max}$) is the ground-state (maximum) exact energy for such a size N. In Fig. 2, we plot this quantity for different points of the phase diagram of our models: strikingly, smooth optimal curves obtained for a small

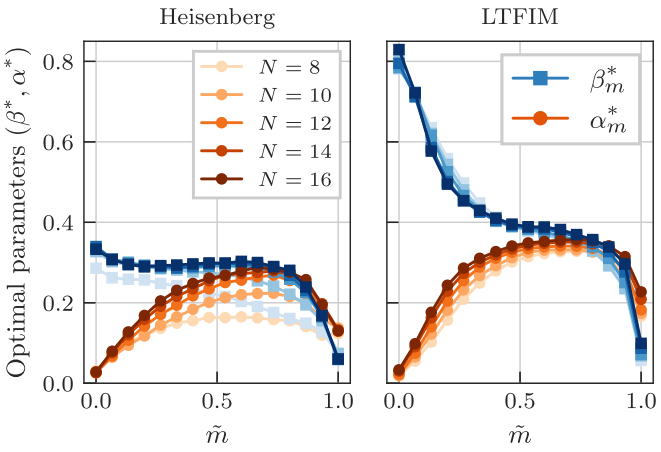

FIG. 1. Smooth optimal parameters $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N} = (\beta_1^* \cdots \beta_P^*, \alpha_1^* \cdots \alpha_P^*)$ obtained with INTERP (see main text), plotted vs the rescaled index $\tilde{m} \equiv (m-1)/(P-1)$ in the x-axis range [0,1]. Results are shown for the Heisenberg model ($\Delta_Y = 1$, $\Delta_Z = 1$) (left) and the LTFIM ($g_x = 1$, $g_z = 1$) (right) for $P = 16$, and they are qualitatively similar for different sizes N. These solutions are stable by further increasing the number of layers, P. Similar smooth solutions can be found for different points of the phase diagram.

system provide an excellent educated guess for the ground-state preparation up to $N = 24$ lattice sites.

A few comments are in order. The residual energy is usually a good proxy for the fidelity with the ground state. It may roughly evaluate at ≈0.5 when computed at a random point in the energy landscape, while its values obtained via transferability are remarkably low. A detailed study on the fidelity of transferred solutions with target ground states is carried out in the SM [54]. Second, the transferability of this class of smooth solutions found via INTERP holds true for larger values of P; in contrast, other equal-quality *nonsmooth* solutions for the small N_G-size system—obtained by means of random-start local optimization—do not provide any useful guess for the ground-state preparation of the same model with

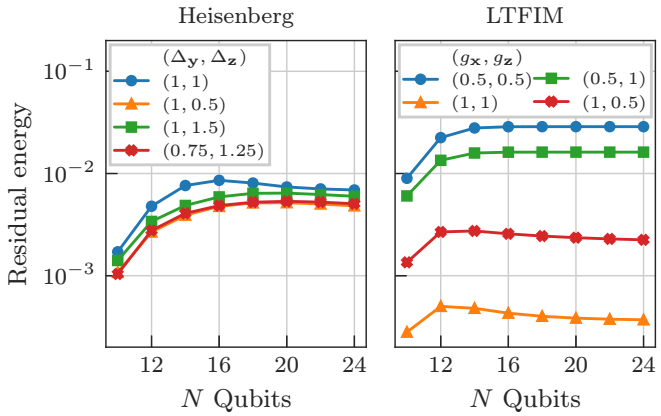

FIG. 2. Residual energies [Eq. (10)] for different system sizes using parameters from a small-size "guess system" ($N_G = 8$) computed for different flavors of the two models. We show the transferability of smooth optimal solutions $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G}$ with $P = 10$ for *XYZ* models (left) and the LTFIM (right).

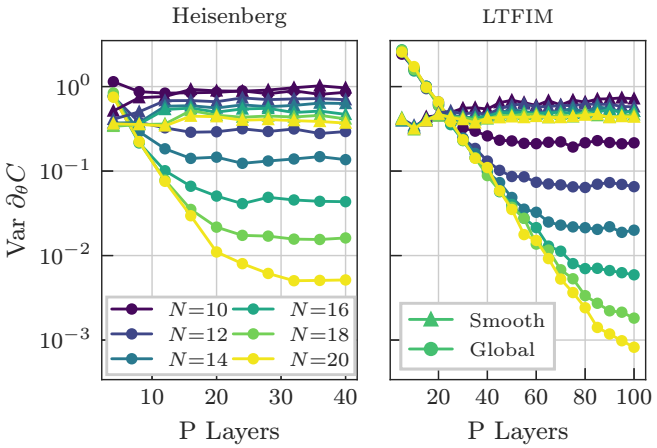

FIG. 3. Barren plateaus in the whole search space (data denoted as "Global"), contrasted with a qualitatively different trend in the ϵ-neighborhood of the transferred smooth solution $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G}$, obtained with INTERP for a small system size $N_G = 8$ (data denoted as "Smooth"). Here, we focus on a single partial derivative with respect to $\theta = \alpha_1$ [see Eqs. (7) and (9)] of the cost function in Eq. (8), rescaled in terms of the few independent correlators (see SM [54]) and dubbed C. We plot the sample variance of the partial derivative as a function of the number P of HVA layers in the circuit. We fix $\epsilon = 0.05$ and a batch of 1000 samples for each value of P and N.

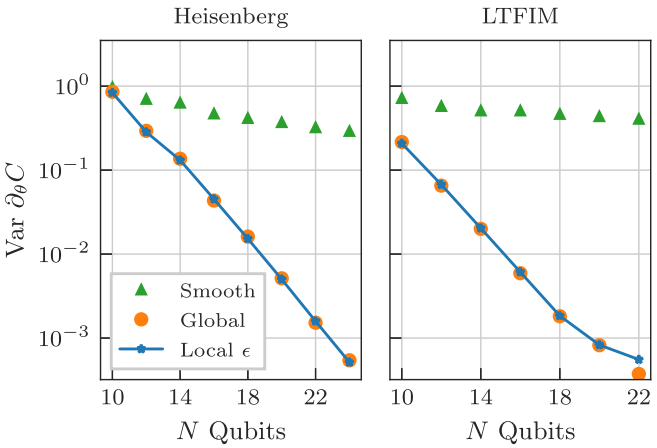

FIG. 4. Same quantity as in Fig. 3, here plotted vs the qubit number for a fixed circuit depth $P = 40$ ($P = 100$) for the Heisenberg model (LTFIM). The labels "Global" and "Smooth" refer to the same data of Fig. 3. We denote with "Local ϵ" a sampling performed in an ϵ-region centered around a random point. Data are averages over 20 random points (generated independently for each value of N) with $\epsilon = 0.05$. We use a constant batch size of 1000 samples, for each N and sampling region. A clear trend appears: the neighborhood of a random point (or that of a transferred nonsmooth solution; see SM [54]) exhibits the same exponential decay as the whole space, whereas this is not present in the neighborhood of the transferred smooth solution.

a larger number of qubits (see SM [54]). Finally, we tested the existence of smooth curves and their transferability to a larger number of qubits, also for the TFIM: our results are confirmed up to much larger sizes (see SM [54]) by leveraging a standard mapping to free fermions [60–63].

Despite the good educated guess provided by the transferability of smooth solutions, one may be tempted to refine the ground-state approximation for the N-size model, e.g., by aiming at a target value of fidelity such as 99.9%. However, for such large sizes, both the XYZ models and the LTFIM are affected by barren plateaus [17,19]. Therefore, any local optimization starting from a random point in the parameter space is doomed to fail on a realistic quantum device due to vanishingly small gradients requiring an exponential scaling of resources [14,64]. Remarkably—and this is the main result of our paper—we find that transferred smooth optimal solutions stand out in this respect: in their neighborhood, the landscape does not suffer from small gradients and a local optimization would succeed.

Figures 3 and 4 illustrate this important point. For conciseness, we show data for $(\Delta_Y = 1, \Delta_Z = 1)$ and $(g_x = 1, g_z = 1)$, but our results extend to other points of the phase diagrams. Specifically, in Fig. 3, we plot the variance of a representative gradient component of the cost function in Eq. (8), as customary in studies on barren plateaus [14,15,18], which is sampled at random in the whole landscape. As expected, its exponential decay with the system size N confirms the presence of barren plateaus. However, if we sample the same gradient component only in a neighborhood of radius ϵ of the transferred smooth solution, its magnitude does not show any appreciable exponential decay. This result is clearly observed for both classes of models under exam and it is further evidenced in Fig. 4, showing data for a fixed value of P: the exponential decrease of the gradients in the whole search space is equivalent to that in a neighborhood of radius ϵ of any given set of variational parameters, with the exception of the smooth transferred curve $(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*)|_{P,N_G}$. Once more, also this local-landscape property does not extend to the neighborhood of other transferred nonsmooth solutions, which neither provide a useful educated guess for the large system nor solve the barren plateau issue for a local optimization. This is shown in the SM [54], along with data supporting the effectiveness of a refinement optimization, performed classically in the neighborhood of the transferred smooth curve.

Incidentally, for each value of N, the sample variance in the whole search space saturates after a certain circuit depth P, as argued in [14,64] and clearly shown in Fig. 3. This fact is usually linked to the ansatz parameterized quantum circuit approaching an approximate 2-design [65–67] on its symmetry subspace [17].

Finally, our findings pave the way to an improved scheme to prepare the ground state of this class of many-body quantum systems with a large number of qubits: the smooth optimal curves can be found *classically* for a small system and then transferred to solve the same task for larger N, beyond the reach of classical simulations. The *quantum* device would only be needed for a refinement optimization, in the absence of barren plateaus.

Conclusions. We tackled many-body ground-state preparation via problem-inspired VQAs and provided extensive numerical evidence on the transferability of a class of optimal *smooth* solutions—obtained by means of iterative schemes for a small number of qubits—to solve the same task for larger system sizes. Remarkably, these solutions

provide an excellent educated guess for the ground-state wave function, as opposed to other solutions that can be easily obtained for small systems without appropriate iterative schemes. These results are confirmed up to larger sizes for the TFIM.

Our procedure overcomes the well-known (and not yet fully addressed) difficulties related to the highly nontrivial structure of the variational energy landscape. On top of avoiding low-quality local minima traps daunting random-start local optimization [13,30], we provided evidence of a remarkable feature of this class of solutions: the cost-function landscape is observed to become free of barren plateaus in their neighborhood, potentially allowing for further effective refinement optimizations with a quantum device on a classically obtained smooth guess.

This work paves the way to a plethora of novel exciting research directions. Our effective way of approaching ground-state preparation for larger many-body systems may allow one to deal with two-dimensional lattice models, ranging from spin systems to Hubbard-like systems, with or without disorder. Furthermore, an interesting proposal would be to test our scheme in the presence of noise and for the mitigation of noise-induced barren plateaus [68].

On a theoretical side, it might be interesting to analytically prove the transferability and landscape properties of smooth solutions found via INTERP. Previous numerical and analytical results on reusable optimal variational parameters exist, either among typical instances of a problem or across different system sizes [43,69–75]. These "parameter concentration" results are usually limited to shallow circuits, while here we focus on smooth optimal solutions for large values of P, providing a link between solution transferability and the local absence of barren plateaus.

Interestingly, the existence of smooth QAOA schedules has also been described in the context of a theoretical characterization of optimal QAOA protocols [76,77] and in a generalized variational setting inspired by counterdiabatic (CD) evolution [45]. We conjecture that the transferability property of smooth optimal curves might be related to nonlinear adiabatic schedules, similar, e.g., to the one obtained analytically in the simpler setting of a Grover search [78]. We speculate that a VQA, endowed with a Trotter-inspired ansatz as in Eq. (1) and optimized by means of iterative techniques, could be able to single out these solutions, possibly showing a mild dependence on the system size.

Finally, our scheme could be directly tested with near-term technology on real quantum devices, beyond the size limits of classical computation.

Acknowledgments. We thank J. J. Meyer, S. Khatri, R. Suzuki, Y. Quek, J. Denzler, J. S. Kottman, and J. Eisert for useful discussion. The research was partly supported by EU Horizon 2020 under ERC-ULTRADISS, Grant Agreement No. 834402. A.A.M. acknowledges support from BMBF (FermiQP and Hybrid). G.B.M. acknowledges support from the Austrian Science Fund through the SFB BeyondC Project No. F7108-N38. G.E.S. acknowledges that his research has been conducted within the framework of the Trieste Institute for Theoretical Quantum Technologies (TQT).

[1] M. Cerezo, A. Arrasmith, R. Babbush, S. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. Coles, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

[2] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

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

[4] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev, VQE method: A short survey and recent developments, arXiv:2103.08505.

[5] 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).

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

[7] D. Wierichs, C. Gogolin, and M. Kastoryano, Avoiding local minima in variational quantum eigensolvers with the natural gradient optimizer, Phys. Rev. Res. **2**, 043246 (2020).

[8] L. Lumia, P. Torta, G. B. Mbeng, G. E. Santoro, E. Ercolessi, M. Burrello, and M. M. Wauters, Two-dimensional $\mathbb{Z}_2$ lattice gauge theory on a near-term quantum simulator: Variational quantum optimization, confinement, and topological order, PRX Quantum **3**, 020320 (2022).

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

[10] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles, Variational quantum linear solver, arXiv:1909.05820.

[11] K. C. Tan and T. Volkoff, Variational quantum algorithms to estimate rank, quantum entropies, fidelity, and Fisher information via purity minimization, Phys. Rev. Res. **3**, 033251 (2021).

[12] L. Bittel and M. Kliesch, Training Variational Quantum Algorithms Is NP-Hard, Phys. Rev. Lett. **127**, 120502 (2021).

[13] E. R. Anschuetz and B. T. Kiani, Quantum variational algorithms are swamped with traps, Nat. Commun. **13**, 7760 (2022).

[14] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 4812 (2018).

[15] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, PRX Quantum **3**, 010313 (2022).

[16] C. O. Marrero, M. Kieferová, and N. Wiebe, Entanglement induced barren plateaus, arXiv:2010.15968.

[17] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, Diagnosing barren plateaus with tools from quantum optimal control, Quantum **6**, 824 (2022).

[18] R. Wiersema, C. Zhou, Y. de Sereville, J. F. Carrasquilla, Y. B. Kim, and H. Yuen, Exploring entanglement and optimization within the Hamiltonian variational ansatz, PRX Quantum **1**, 020319 (2020).

[19] J. J. Meyer, M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert, Exploiting symmetry in variational quantum machine learning, arXiv:2205.06217.

[20] J. Dborin, F. Barratt, V. Wimalaweera, L. Wright, and A. G. Green, Matrix product state pre-training for quantum machine learning, arXiv:2106.05742.

[21] A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib, Layerwise learning for quantum neural networks, Quantum Machine Intell. **3**, 5 (2021).

[22] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, Quantum **3**, 214 (2019).

[23] M. Kliesch and I. Roth, Theory of quantum system certification, PRX Quantum **2**, 010201 (2021).

[24] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn, Avoiding barren plateaus using classical shadows, PRX Quantum **3**, 020365 (2022).

[25] S. H. Sack and M. Serbyn, Quantum annealing initialization of the quantum approximate optimization algorithm, Quantum **5**, 491 (2021).

[26] P. Torta, G. B. Mbeng, C. Baldassi, R. Zecchina, and G. E. Santoro, Quantum approximate optimization algorithm applied to the binary perceptron, arXiv:2112.10219.

[27] C. Lyu, V. Montenegro, and A. Bayat, Accelerated variational algorithms for digital quantum simulation of many-body ground states, Quantum **4**, 324 (2020).

[28] T. Albash and D. A. Lidar, Adiabatic quantum computation, Rev. Mod. Phys. **90**, 015002 (2018).

[29] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, Science **292**, 472 (2001).

[30] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices, Phys. Rev. X **10**, 021067 (2020).

[31] G. B. Mbeng, R. Fazio, and G. Santoro, Quantum annealing: A journey through digitalization, control, and hybrid quantum variational schemes, arXiv:1906.08948.

[32] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).

[33] W. W. Ho, C. Jonay, and T. H. Hsieh, Ultrafast variational simulation of nontrivial quantum states with long-range interactions, Phys. Rev. A **99**, 052332 (2019).

[34] M. M. Wauters, G. B. Mbeng, and G. E. Santoro, Polynomial scaling of the quantum approximate optimization algorithm for ground-state preparation of the fully connected p-spin ferromagnet in a transverse field, Phys. Rev. A **102**, 062404 (2020).

[35] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos *et al.*, Self-verifying variational quantum simulation of lattice models, Nature (London) **569**, 355 (2019).

[36] G. Matos, S. Johri, and Z. Papić, Quantifying the efficiency of state preparation via quantum variational eigensolvers, PRX Quantum **2**, 010309 (2021).

[37] C.-Y. Park, Efficient ground state preparation in variational quantum eigensolver with symmetry breaking layers, arXiv:2106.02509.

[38] N. Astrakhantsev, G. Mazzola, I. Tavernelli, and G. Carleo, Algorithmic phases in variational quantum ground-state preparation, arXiv:2205.06278.

[39] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong, Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families, arXiv:1907.13623.

[40] R. Barends, A. Shabani, L. Lamata, J. Kelly, A. Mezzacapo, U. L. Heras, R. Babbush, A. G. Fowler, B. Campbell, Y. Chen *et al.*, Digitized adiabatic quantum computing with a superconducting circuit, Nature (London) **534**, 222 (2016).

[41] G. B. Mbeng, R. Fazio, and G. E. Santoro, Optimal quantum control with digitized quantum annealing, arXiv:1911.12259.

[42] G. B. Mbeng, Quantum annealing and digital quantum ground state preparation algorithms, Ph.D. thesis, SISSA, Trieste, 2019, available at https://iris.sissa.it.

[43] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin *et al.*, Quantum approximate optimization of the long-range Ising model with a trapped-ion quantum simulator, Proc. Natl. Acad. Sci. **117**, 25396 (2020).

[44] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size, Quantum **6**, 759 (2022).

[45] J. Wurtz and P. J. Love, Counterdiabaticity and the quantum approximate optimization algorithm, Quantum **6**, 635 (2022).

[46] G. E. Crooks, Performance of the quantum approximate optimization algorithm on the maximum cut problem, arXiv:1811.08419.

[47] L. T. Brady, L. Kocia, P. Bienias, A. Bapat, Y. Kharkov, and A. V. Gorshkov, Behavior of analog quantum algorithms, arXiv:2107.01218.

[48] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, London, 1982).

[49] A. A. Ovchinnikov, D. V. Dmitriev, V. Y. Krivnov, and V. O. Cheranovskii, Antiferromagnetic Ising chain in a mixed transverse and longitudinal magnetic field, Phys. Rev. B **68**, 214406 (2003).

[50] R. J. Baxter, One-dimensional anisotropic Heisenberg chain, Ann. Phys. **70**, 323 (1972).

[51] M. P. M. den Nijs, Derivation of extended scaling relations between critical exponents in two-dimensional models from the one-dimensional Luttinger model, Phys. Rev. B **23**, 6111 (1981).

[52] E. Ercolessi, S. Evangelisti, F. Franchini, and F. Ravanini, Essential singularity in the Renyi entanglement entropy of the one-dimensional XYZ spin-$\frac{1}{2}$ chain, Phys. Rev. B **83**, 012402 (2011).

[53] P. Sen, Quantum-fluctuation-induced spatial stochastic resonance at zero temperature, Phys. Rev. E **63**, 040101(R) (2001).

[54] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.106.L060401 for a description of phase diagrams, an in-depth analysis of the HVA and its symmetries, the light-cone interpretation, and details on the INTERP algorithm and its performance. Also, we provide additional numerical evidence on the transferability of smooth solutions and on their effectiveness in avoiding barren plateaus, contrasted with random-start (nonsmooth) solutions. Finally, we report our results for large-scale simulations for the TFIM.

[55] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[56] G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, Phys. Rev. Lett. **91**, 147902 (2003).

[57] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, Phys. Rev. Lett. **93**, 040502 (2004).

[58] A. Gadi, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen *et al.*, Qiskit: An open-source framework for quantum computing, version 0.7.2, Zenodo (2019), doi:10.5281/zenodo.2562111.

[59] C. Zhu, R. H. Byrd, and J. Nocedal, L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization, ACM Trans. Math. Softw. **23**, 550 (1997).

[60] P. Jordan and E. Wigner, On the Pauli exclusion principle, Z. Phys. **47**, 631 (1928).

[61] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, Ann. Phys. **16**, 407 (1961).

[62] G. B. Mbeng, R. Fazio, and G. E. Santoro, Optimal quantum control with digitized quantum annealing, arXiv:1911.12259.

[63] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Quantum approximate optimization algorithm for MAXCUT: A fermionic view, Phys. Rev. A **97**, 022304 (2018).

[64] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nat. Commun. **12**, 1791 (2021).

[65] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, J. Math. Phys. **48**, 052104 (2007).

[66] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).

[67] J. Haferkamp, Random quantum circuits are approximate unitary t-designs in depth $o(nt^{5+o(1)})$, Quantum **6**, 795 (2022).

[68] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, Nat. Commun. **12**, 6961 (2021).

[69] G. Verdon, M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni, Learning to learn with quantum neural networks via classical neural networks, arXiv:1907.05415.

[70] V. Akshay, D. Rabinovich, E. Campos, and J. Biamonte, Parameter concentrations in quantum approximate optimization, Phys. Rev. A **104**, L010401 (2021).

[71] A. Galda, X. Liu, D. Lykov, Y. Alexeev, and I. Safro, Transferability of optimal QAOA parameters between random graphs, arXiv:2106.07531.

[72] F. G. S. L. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, For fixed control parameters the quantum approximate optimization algorithm's objective function value concentrates for typical instances, arXiv:1812.04170.

[73] M. Streif and M. Leib, Training the quantum approximate optimization algorithm without access to a quantum processing unit, Quantum Sci. Technol. **5**, 034008 (2020).

[74] F. Sauvage, S. Sim, A. A. Kunitsa, W. A. Simon, M. Mauri, and A. Perdomo-Ortiz, FLIP: A flexible initializer for arbitrarily-sized parametrized quantum circuits, arXiv:2103.08572 (2021).

[75] A. Cervera-Lierta, J. S. Kottmann, and A. Aspuru-Guzik, Meta-variational quantum eigensolver: Learning energy profiles of parameterized Hamiltonians for quantum simulation, PRX Quantum **2**, 020329 (2021).

[76] G. B. Mbeng, L. Arceci, and G. E. Santoro, Optimal working point in digitized quantum annealing, Phys. Rev. B **100**, 224201 (2019).

[77] L. T. Brady, L. Kocia, P. Bienias, A. Bapat, Y. Kharkov, and A. V. Gorshkov, Behavior of analog quantum algorithms, arXiv:2107.01218.

[78] J. Roland and N. J. Cerf, Quantum search by local adiabatic evolution, Phys. Rev. A **65**, 042308 (2002).