

Universal variational quantum computation

Jacob Biamonte *

Skolkovo Institute of Science and Technology, 3 Nobel Street, Moscow 121205, Russia

 (Received 8 August 2020; revised 29 December 2020; accepted 12 January 2021; published 10 March 2021)

Variational quantum algorithms dominate contemporary gate-based quantum enhanced optimization, eigenvalue estimation, and machine learning. Here we establish the quantum computational universality of variational quantum computation by developing two objective functions which minimize to prepare outputs of arbitrary quantum circuits. The fleeting resource of variational quantum computation is the number of expected values which must be iteratively minimized using classical-to-quantum outer loop optimization. An efficient solution to this optimization problem is given by the quantum circuit being simulated itself. The first construction is efficient in the number of expected values for n -qubit circuits containing $O(\text{poly } \ln n)$ non-Clifford gates—the number of expected values has no dependence on Clifford gates appearing in the simulated circuit. The second approach yields $O(L^2)$ expected values whereas introducing not more than $O(\ln L)$ slack qubits for a quantum circuit partitioned into L gates. Hence, the utilitarian variational quantum programming procedure—based on the classical evaluation of objective functions and iterated feedback—is, in principle, as powerful as any other model of quantum computation. This result elevates the formal standing of the variational approach whereas establishing a universal model of quantum computation.

DOI: [10.1103/PhysRevA.103.L030401](https://doi.org/10.1103/PhysRevA.103.L030401)

Variational quantum algorithms reduce quantum state preparation requirements whereas necessitating measurements of individual qubits in the computational basis [1,2]. In the contemporary noisy intermediate-scale quantum (NISQ) enhanced technology setting [3], a sought reduction in coherence time is mediated through an iterative classical-to-quantum feedback and optimization process. Systematic errors which map to deterministic yet unknown control parameters—such as time variability in the application of specific Hamiltonians or poor pulse timing—can have less impact on variational algorithms as states are prepared iteratively and varied over to minimize objective function(s). These experimental advantages have made the variational approach to quantum computation the most widely studied gate-based approach today. See the reviews [4–6].

A variational quantum algorithm executes to prepare a state which minimizes an objective function. In the case of variational quantum approximate optimization (QAOA [1]), a state is prepared by alternating a Hamiltonian representing a penalty function (such as the NP-hard Ising embedding of 3-SAT) with a Hamiltonian representing local tunneling terms. The state is measured, and the resulting bit string serves as a candidate solution to minimize the penalty function. In the case of variational eigenvalue minimization [2], the state is repeatedly prepared and measured to obtain a set of expected values which is termwise calculated and collectively

minimized. In both approaches, the minimization process is iterated by updating Hamiltonian application times and/or gate angles. Several techniques in variational quantum computation are closely related to methods appearing in machine learning [7,8] in which a quantum circuit is tuned subject to a given training dataset.

Computational universality is a property of central study in both classical and quantum models of computation. Universal models, such as adiabatic quantum computation [9] both discrete and continuous quantum walks [10,11] and measurement-based quantum computation [12], have been proved computationally universal through constructions to emulate a universal set of quantum gates. This implies directly that the system has access to any polynomial time quantum algorithm on n qubits (the power of quantum algorithms in the class BQP).

An alternative notion of *universality* also exists in the literature. This strong notion is algebraic, wherein a system is called universal if its generating Lie algebra is proven to span $\mathfrak{su}(2^n)$ for n qubits. We call this *controllability*.

Evidently these two notions of universality can be inter-related: By proving that a controllable system can efficiently simulate a universal gate set, a controllable system becomes computationally universal. It is conversely anticipated by the strong Church-Turing-Deutsch principle [13] that a computationally universal system can be made to simulate any controllable system.

Herein we assume access to control sequences which can create quantum gates, such as Refs. [14–16]. Given a quantum circuit of L gates, preparing a state $|\psi\rangle = \prod_{l=1}^L U_l |0\rangle^{\otimes n}$ for unitary gates U_l , we construct a universal objective function that is minimized by $|\psi\rangle$. The objective function is engineered to have certain desirable properties. Importantly, we construct a gapped Hamiltonian where minimization past some fixed

*j.biamonte@skoltech.ru; <http://quantum.skoltech.ru>

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

tolerance ensures sufficient overlap with the desired output state $|\psi\rangle$.

Recent work of interest by Lloyd considered the controllability of QAOA sequences to create quantum gates [17]. The original goal of QAOA was to alternate target and driver Hamiltonians to evolve a system close to the target Hamiltonian’s ground state—thereby solving an optimization problem instance. Lloyd showed that alternating driver and target Hamiltonians can be programmed to give computationally universal dynamics [17]. Following Lloyd’s work [17], myself with two coauthors formulated Lloyd’s QAOA controllability result past one-dimensional lines of qubits [18].

We assume controllable state preparation followed by local Pauli measurements. These measurements are used to calculate an objective function (a Hamiltonian) which we prove minimizes to certify that the output of a target quantum circuit has been prepared. The objective function is expressed as a sum of terms in the Pauli basis. Each term corresponds to a simplistic measurement operator, and, hence, this approach avoids the long sequences of gates used to simulate Hamiltonian’s in the gate model. Whereas the objective function can be evaluated termwise, achieving tolerance $\sim\epsilon$ requires $\sim\epsilon^{-2}$ measurements—see Hoeffding’s inequality [19].

Structure. After introducing variational quantum computation as it applies to our setting, we construct an objective function (named a telescoping construction). The number of expected values has no dependence on Clifford gates appearing in the simulated circuit and is efficient for circuits with $O(\text{poly } \ln n)$ non-Clifford gates, making it amenable for near-term demonstrations. We then modify the Feynman-Kitaev clock construction and prove that universal variational quantum computation is possible by minimizing $O(L^2)$ expected values whereas introducing not more than $O(\ln L)$ slack qubits, for a quantum circuit partitioned into L Hermitian blocks.

We conclude by considering how the universal model of variational quantum computation can be utilized in practice. In particular, the given gate sequence prepares a state which will minimize the objective function. In practice, we think of this as providing a starting point for a classical optimizer. Given a T -gate sequence, we consider the first $L \leq T$ gates. This L -gate circuit represents an optimal control problem where the starting point is the control sequence to prepare the L gates. The goal is to modify this control sequence (shorten it) using a variational feedback loop. One would iterate this scenario, increasing L up to T .

I. VARIATIONAL QUANTUM COMPUTATION

We work in the standard setting of quantum computation using n qubits where we typically fix the computational basis $\mathcal{B}^{\otimes n}$ and use the standard qubit representation of the Pauli group algebra, satisfying the product identity $XY = iZ$, together with $X^2 = Y^2 = Z^2 = \mathbb{1}$. We will interchange the notation $\sigma_0 \equiv \mathbb{1}$, $\sigma_1 \equiv X$, $\sigma_2 \equiv Y$, $\sigma_3 \equiv Z$, and consider Hamiltonian’s $\mathcal{H} = \mathcal{H}^\dagger$ which act on the space of linear qubit maps $\mathcal{L}(\mathcal{B}^{\otimes n})$.

We wish to simulate the output of an L gate quantum circuit acting on the n -qubit product state $|0\rangle^{\otimes n}$. We have access to p appropriately bounded and tunable parameters to prepare

and vary over a family of quantum states. All coefficients herein are assumed to be accurate to not more than $\text{poly}(n)$ decimal places. We will define an objective function that when minimized will produce a state close to the desired quantum circuit output. We will provide a solution to the minimization problem. Establishing our results requires several definitions and supporting lemmas. Proofs not appearing after a lemma or theorem can be found in the Supplemental Material [20].

Definition 1 (Variational state space). The variational state space Γ of a p -parametrized n -qubit state preparation process is the union of $|\psi(\theta)\rangle$ over real assignments of θ ,

$$\Gamma \stackrel{\text{def}}{=} \bigcup_{\theta \in \mathbb{R}^{\times p}} \{|\psi(\theta)\rangle\} \subseteq \mathbb{C}_2^{\otimes n}. \quad (1)$$

Variational state space examples include preparing $|\psi(\theta)\rangle$ by a fixed quantum circuit (called an ansatz) with, e.g., $\theta \in [0, 2\pi)^{\times p}$ tunable parameters as

$$|\psi(\theta)\rangle = \prod_{l=1}^L U_l |0\rangle^{\otimes n}, \quad (2)$$

where typically U_l is adjusted by θ_l for $l = 1$ to p , although some gates might be fixed. There are various approaches to implement ansatz circuits, including the hardware efficient ansatz [21] which implements a daisy chain of two-body coupling gates or the brick layer (also known as the checkerboard) ansatz which is simply alternating stacks of nearest-neighboring coupling gates (see, for example, Ref. [22]).

Definition 2 (Variational sequence). A variational sequence is an assignment of θ^i to prepare a single state in a variational state space $|\psi(\theta^i)\rangle \in \Gamma$.

Examples of variational sequences include the general expression of a fixed quantum gate sequence acting on a product state, which can be expressed as (2).

We will define an objective function that can be efficiently calculated given access to a suitable quantum processor. Minimization of this objective function will be proven to produce a close two-norm approximation to the output of a given quantum circuit. Let us consider the most suitably general Hamiltonian acting on qubits (3).

Definition 3 (Objective function). We consider an objective function as the expected value of an operator expressed with real coefficients $\mathcal{J}_{\alpha_1 \alpha_2 \dots \alpha_n}^{a_1 a_2 \dots a_n}$ in the Pauli basis as

$$\mathcal{H} = \sum^{\text{poly}(n)} \mathcal{J}_{\alpha_1 \alpha_2 \dots \alpha_n}^{a_1 a_2 \dots a_n} \sigma_{\alpha_1}^{a_1} \sigma_{\alpha_2}^{a_2} \dots \sigma_{\alpha_n}^{a_n}, \quad (3)$$

where greek letters index Pauli matrices, roman letters index qubits, and the sum is over a $\text{poly}(n)$ bounded subset of the 4^n elements in the basis. The tensor (\otimes) is omitted in (3).

We are concerned primarily with Hamiltonians where $\mathcal{J}_{\alpha_1 \alpha_2 \dots \alpha_n}^{a_1 a_2 \dots a_n}$ is given and known to be nonvanishing for at most some $\text{poly}(n)$ terms. This wide class includes Hamiltonians representing electronic structure [21]. More generally, such Hamiltonians are of bounded cardinality.

Definition 4 (Cardinality). The number of terms in the Pauli basis $\{\mathbb{1}, X, Y, Z\}^{\otimes n}$ needed to express an objective function.

Definition 5 (Bounded objective function). An instance is called *bounded* when it is taken from a uniform family of objective functions of cardinality bounded by $\text{poly}(n)$.

We will now use the Cartesian tensor (\times) as states need not be proximally interacting and can be independently prepared. Hence, $|\phi\rangle^{\times O[\text{poly}(n)]}$ means we prepare polynomial many non-interacting copies of $|\phi\rangle$ to approximate each expected value.

Definition 6 (Polycomputable objective function). An objective function,

$$f: |\phi\rangle^{\times O[\text{poly}(n)]} \rightarrow \mathbb{R}_{\geq 0} \quad (4)$$

is called polycomputable provided $\text{poly}(n)$ independent physical copies of $|\phi\rangle$ can be efficiently prepared to evaluate a bounded objective function.

Efficiently computable objective function examples include calculating the expected value of $O(\ln n)$ products and sums over

$$\{\mathcal{H}, \langle \mathcal{H} \rangle, \cdot, +, \mathbb{R}\} \quad (5)$$

for bounded cardinality \mathcal{H} . Examples include

(2.i) Calculating the expected value of \mathcal{H} itself, which includes electronic structure Hamiltonians [2].

(2.ii) Calculating the dispersion $\text{var}(\mathcal{H}) = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2$ which vanishes if and only if the prepared state is an eigenstate of \mathcal{H} [23].

Acceptance (as follows) must be shown by providing a solution to the optimization problem defined by the objective function.

Definition 7 (Accepting a quantum state). An objective function f *accepts* $|\phi\rangle$ when given $O(\text{poly } n)$ copies of $|\phi\rangle$,

$$f(|\phi\rangle^{\times O(\text{poly}(n))}) = f(|\phi\rangle, |\phi\rangle, \dots, |\phi\rangle) < \Delta \quad (6)$$

evaluates strictly less than a chosen real parameter $\Delta > 0$.

The following Theorem 1 applies rather generally to variational quantum algorithms that minimize energy by adjusting a variational state to cause an objective function to accept. Herein acceptance will imply the preparation of a quantum state, which begs to establish the following.

Lemma 1 (Variational stability). Let non-negative $\mathcal{H} = \mathcal{H}^\dagger \in \mathcal{L}(\mathbb{C}_d)$ have spectral gap Δ and nondegenerate ground eigenvector $|\psi\rangle$ of eigenvalue 0. Consider then a unit vector $|\phi\rangle \in \mathbb{C}_d$ such that

$$\langle \phi | \mathcal{H} | \phi \rangle < \Delta, \quad (7)$$

it follows that

$$1 - \frac{\langle \phi | \mathcal{H} | \phi \rangle}{\Delta} \leq |\langle \phi | \psi \rangle|^2 \leq 1 - \frac{\langle \phi | \mathcal{H} | \phi \rangle}{\text{Tr}\{\mathcal{H}\}}. \quad (8)$$

A. Maximizing projection onto a circuit

We will now explicitly construct an elementary Hermitian penalty function that is non-negative with a nondegenerate lowest (0) eigenstate—so as to apply Lemma 1. Minimization of this penalty function prepares the output of a quantum circuit.

Theorem 1 (Telescoping construction). Consider $\prod_l U_l |0\rangle^{\otimes n}$ an L -gate quantum circuit preparing state $|\psi\rangle$ on n qubits and containing not more than $O[\text{poly}(\ln n)]$ non-Clifford gates. Then there exists a Hamiltonian $\mathcal{H} \geq 0$ on n qubits with $\text{poly}(L, n)$ cardinality, a (L, n) -independent gap Δ

and non-degenerate ground eigenvector $\in \text{span}\{\prod_l U_l |0\rangle^{\otimes n}\}$. In particular, a variational sequence exists causing the Hamiltonian to accept $|\phi\rangle$ viz., $0 \leq \langle \phi | \mathcal{H} | \phi \rangle < \Delta$ then Lemma 1 implies stability.

To prove Theorem 1 we first show existence of the penalty function. Construct Hermitian $\mathcal{H} \in \mathcal{L}(\mathbb{C}_2^{\otimes n})$ with $\mathcal{H} \geq 0$ such that there exists a nondegenerate $|\psi\rangle \in \mathbb{C}_2^{\otimes n}$ with the property that $\mathcal{H}|\psi\rangle = 0$. Define P_ϕ as a sum of projectors onto product states, i.e.,

$$P_\phi = \sum_{i=1}^n |1\rangle\langle 1|^{(i)} = \frac{n}{2} \left(\mathbb{1} - \frac{1}{n} \sum_{i=1}^n Z^{(i)} \right), \quad (9)$$

and consider (9) as the initial Hamiltonian, preparing state $|0\rangle^{\otimes n}$.

We will act on (9) with a sequence of gates $\prod_{l=1}^L U_l$ corresponding to the circuit being simulated as

$$h(k) = \left(\prod_{l=1}^{k \leq L} U_l \right) P_\phi \left(\prod_{l=1}^{k \leq L} U_l \right)^\dagger \geq 0, \quad (10)$$

which preserves the spectrum (i.e., $P_\phi|x\rangle = |x|_1|x\rangle$ for $x \in \{0, 1\}^n$ and $|\cdot|_1$ the Hamming weight). From the properties of P_ϕ it, hence, follows that $h(k)$ is non-negative and nondegenerate $\forall k \leq L$. We now consider the action of the gates (10) on (9).

At $k = 0$ from (9) there are n -expected values to be minimized plus a global energy shift that will play a multiplicative role as the circuit depth increases. To consider $k = 1$ we first expand a universal gate set expressed in the linear extension of the Pauli basis.

Interestingly, the coefficients $\mathcal{J}_{\alpha\beta\gamma}^{ab\dots c}$ of the gates will not serve as direct input(s) to the quantum hardware; these coefficients play a direct role in the classical step where the coefficients weight the sum to be minimized. Let us then consider single qubit gates, in general, form viz.,

$$e^{-i\mathbf{a}\cdot\sigma\theta} = \mathbb{1} \cos(\theta) - i\mathbf{a}\cdot\sigma \sin(\theta), \quad (11)$$

where \mathbf{a} is a unit vector and $\mathbf{a}\cdot\sigma = \sum_{i=1}^3 a_i \sigma_i$. So each single qubit gate increases the number of expected values by a factor of at most 4^2 . At first glance, this appears prohibitive, yet there are two factors to consider. The first is the following Lemma 2.

Lemma 2 (Clifford gate cardinality invariance). Let \mathcal{C} be the set of all Clifford circuits on n qubits, and let \mathcal{P} be the set of all elements of the Pauli group on n qubits. Let $C \in \mathcal{C}$ and $P \in \mathcal{P}$, then it can be shown that $CPC^\dagger \in \mathcal{P}$ or in other words $C(\sigma_\alpha^a \sigma_\beta^b \dots \sigma_\gamma^c)C^\dagger = \sigma_{\alpha'}^{a'} \sigma_{\beta'}^{b'} \dots \sigma_{\gamma'}^{c'}$, and so Clifford circuits act by conjugation on tensor products of Pauli operators to produce tensor products of Pauli operators.

For some U a Clifford gate, Lemma 2 shows that the cardinality of (10) is invariant. Non-Clifford gates increase the cardinality by factors $O(e^n)$ and so must be logarithmically bounded from above. Hence, telescopes bound the number of expected values by restricting to circuits with $k \sim O(\text{poly } \ln n)$ non-Clifford single qubit gates. Clifford gates do, however, modify the locality of terms appearing in the expected values—this is, i.e., prohibitive in adiabatic quantum computation yet arises here as local measurements.

A final argument supporting the utility of telescopes is that the initial state is restricted primarily by the initial Hamiltonian having only a polynomial number of nonvanishing coefficients in the Pauli basis. In practice—using today’s hardware—it should be possible to prepare an ϵ -close two-norm approximation to any product state $\bigotimes_{k=1}^n \cos \theta_k |0\rangle + e^{i\phi_k} \sin \theta_k |1\rangle$ which is realized by modifying the projectors in (9) with a product of single qubit maps $\bigotimes_{k=1}^n U_k$. Other more complicated states would also be possible.

To finish the proof of Lemma 1, the variational sequence is given by the description of the gate sequence itself. That is,

$$h(k) \left(\prod_{l=1}^{k \leq L} U_l \right) |x\rangle = |x|_1 \left(\prod_{l=1}^{k \leq L} U_l \right) |x\rangle \quad (12)$$

is minimized for x the string of zeros in $\{0, 1\}^n$. Hence, a state can be prepared causing the Hamiltonian to accept and stability applies (Lemma 1).

To explore telescopes in practice, let us then explicitly consider the quantum algorithm for state overlap (also known as a *swap test* see, e.g., Ref. [24]). This algorithm has an analogous structure to phase estimation, a universal quantum primitive of error-corrected quantum algorithms.

Example 1. We are given two d -qubit states $|\rho\rangle$ and $|\tau\rangle$ which will be nondegenerate and minimal eigenvalue states of some initial Hamiltonian(s) on $n + 1$ qubits,

$$h(0)|+, \rho, \tau\rangle = 0 \quad (13)$$

corresponding to the minimization of $\text{poly}(n/2) + 1$ expected values where the first qubit (superscript 1 below) adds one term and is measured in the X basis. The controlled swap gate takes the form

$$[U_{\text{swap}}]_m^1 = \frac{1}{2}(\mathbb{1}^1 + Z^1) \otimes \mathbb{1}^m + \frac{1}{2}(\mathbb{1}^1 - Z^1) \otimes S^m, \quad (14)$$

where $m = (i, j)$ indexes a qubit pair and the exchange operator of a pair of qubit states is $\mathcal{S} = \mathbb{1} + \sigma \cdot \sigma$. For the case of $d = 1$ we arrive at the simplest (three-qubit) experimental demonstration. At the minimum ($= 0$), the expected value of the first qubit being in logical zero is $\frac{1}{2} + \frac{1}{2}|\langle \rho | \tau \rangle|^2$. The final Hadamard gate on the control qubit is considered in the measurement step.

Telescopes provide some handle on what we can do without adding additional slack qubits yet fail to directly prove universality in their own right. The crux lies in the fact that we are only allowed some polynomial in $\ln n$ non-Clifford gates (which opens an avenue for classical simulation, see Refs. [25,26]). Interestingly however, we considered the initial Hamiltonian in (9) as a specific sum over projectors. We instead could bound the cardinality by some polynomial in n . Such a construction will now be established: requires the addition of slack qubits. The universal construction then follows.

B. Maximizing projection onto the history state

We will now prove the following Theorem 2 which establishes universality of the variational model of quantum computation.

Theorem 2. Consider a quantum circuit of L gates on n -qubits producing state $\prod_l U_l |0\rangle^{\otimes n}$. Then there exists an ob-

jective function (Hamiltonian \mathcal{H}) with nondegenerate ground state, cardinality $O(L^2)$, and spectral gap $\Delta \geq O(L^{-2})$ acting on $n + O(\ln L)$ qubits such that acceptance implies efficient preparation of the state $\prod_l U_l |0\rangle^{\otimes n}$. Moreover, a variational sequence exists causing the objective function to accept.

To construct an objective function satisfying Theorem 2, we modify the Feynman-Kitaev clock construction [27,28]. Coincidentally (and tangential to our objectives here), this construction is also used in certain definitions of the complexity class quantum-Merlin-Arthur (QMA), the quantum analog of NP, through the QMA-complete problem k -LOCAL HAMILTONIAN [28].

Feynman developed a time-independent Hamiltonian that induces unitary dynamics to simulate a sequence of gates [27]. Consider Feynman’s Hamiltonians,

$$\begin{aligned} \tilde{\mathcal{H}}_t &= U_t \otimes |t\rangle\langle t-1| + U_t^\dagger \otimes |t-1\rangle\langle t|, \\ \tilde{\mathcal{H}}_{\text{prop}} &= \sum_{t=1}^L \tilde{\mathcal{H}}_t, \end{aligned} \quad (15)$$

where the Hamiltonian (15) acts on a clock register (right of \otimes) with orthogonal clock states 0 to L and an initial state $|\xi\rangle$ (left). Observation of the clock in state $|L\rangle$ after some time $s = s_*$ produces

$$\mathbb{1} \otimes \langle L| e^{-i \cdot s \cdot \mathcal{H}_{\text{prop}}} |\xi\rangle \otimes |0\rangle = U_L \cdots U_1 |\xi\rangle. \quad (16)$$

The Hamiltonian $\mathcal{H}_{\text{prop}}$ in (15) can be modified as (17) so as to have the history state (18) as its ground state,

$$\begin{aligned} -U_t \otimes |t\rangle\langle t-1| - U_t^\dagger \otimes |t-1\rangle\langle t| + |t\rangle\langle t| + |t-1\rangle\langle t-1| \\ = 2\mathcal{H}_t \geq 0, \end{aligned} \quad (17)$$

where \mathcal{H}_t is a projector. Then $\mathcal{H}_{\text{prop}} = \sum_{t=1}^L \mathcal{H}_t$ has the history state,

$$|\psi_{\text{hist}}\rangle = \frac{1}{\sqrt{L+1}} \sum_{t=0}^L U_t \cdots U_1 |\xi\rangle \otimes |t\rangle, \quad (18)$$

as its ground state as for any input state $|\xi\rangle$ where $0 = \langle \psi_{\text{hist}} | \mathcal{H}_{\text{prop}} | \psi_{\text{hist}} \rangle$. This forms the building blocks of our objective function. We will, hence, establish Theorem 2 by a series of lemmas. We let $P_0 \stackrel{\text{def}}{=} |0\rangle\langle 0|$.

Lemma 3 (Degeneracy lifting). Adding the tensor product of a projector on the first clock qubit with a telescope,

$$\mathcal{H}_{\text{in}} = V \left(\sum_{i=1}^n P_1^{(i)} \right) V^\dagger \otimes P_0 \quad (19)$$

lifts the degeneracy of the ground space of $\mathcal{H}_{\text{prop}}$ and the history state with fixed input as

$$\frac{1}{\sqrt{L+1}} \sum_{t=0}^L \prod_{l=1}^t U_l (V|0\rangle^{\otimes n}) \otimes |t\rangle \quad (20)$$

is the nondegenerate ground state of $J\mathcal{H}_{\text{in}} + K\mathcal{H}_{\text{prop}}$ for real $J, K > 0$.

Lemma 4 (Gap existence). For appropriate non-negative J and K , the operator $J\mathcal{H}_{\text{in}} + K\mathcal{H}_{\text{prop}}$ is gapped with a

nondegenerate ground state and, hence, Lemma 1 applies with

$$\Delta \geq \max \left\{ J, \frac{K\pi^2}{2(L+1)^2} \right\}. \quad (21)$$

Lemma 5 (Log space embedding $\mathcal{H}_{\text{prop}}$). The clock space of $\mathcal{H}_{\text{prop}}$ embeds into $O(\ln L)$ slack qubits, leaving the ground space of $J\mathcal{H}_{\text{in}} + K\mathcal{H}_{\text{prop}}$ and the gap invariant.

Lemma 6 (Existence and acceptance). The objective function $J\mathcal{H}_{\text{in}} + K\mathcal{H}_{\text{prop}}$ satisfies Theorem 2. The gate sequence $\prod_l U_l |0\rangle^{\otimes n}$ is accepted by the objective function from Lemma 6 thereby satisfying Theorem 2.

We will add K -identity gates to boost the probability of the desired circuit output state $|\phi\rangle = \prod_{l=1}^L U_l |0\rangle^{\otimes n}$. From Lemma 1, we have that

$$1 - \frac{\langle \phi | \mathcal{H} | \phi \rangle}{\Delta} \leq |\langle \phi | \psi_{\text{hist}} \rangle|^2 = \frac{1}{1 + \frac{L+1}{K}}, \quad (22)$$

whenever $\langle \phi | \mathcal{H} | \phi \rangle < \max \{ J, \frac{K\pi^2}{2(L+1)^2} \}$. For large enough $K > L$, the right-hand side of (22) approaches unity, satisfying the theorem.

Finally are faced with considering self-inverse gates. Such gates (U) have a spectrum $\text{Spec}(U) \subseteq \{\pm 1\}$, are bijective to idempotent projectors ($P^2 = P = P^\dagger$), viz. $U = \mathbb{1} - 2P$, and if V is a self-inverse quantum gate, so is the unitary conjugate $\tilde{V} = GV G^\dagger$ under arbitrary G . Shi showed that a set comprising the controlled not gate (also known as the Feynman gate) plus any one-qubit gate whose square does not preserve the computational basis is universal [16]. Consider Hermitian,

$$R(\theta) = X \sin(\theta) + Z \cos(\theta), \quad (23)$$

then,

$$e^{i\theta Y} = R(\pi/2)R(\theta). \quad (24)$$

Hence, a unitary Y rotation is recovered by a product of two Hermitian operators. A unitary X rotation is likewise recovered by the composition (24) when considering Hermitian $Y \sin(\theta) - Z \cos(\theta)$. The universality of self-inverse gates is then established with constant overhead. Hence, and to conclude, the method introduces not more than $O(L^2)$ expected values whereas requiring not more than $O(\ln L)$ slack qubits for an L gate quantum circuit.

C. Ansatz states and a combinatorial quantum circuit area law

Our construction of universal variational quantum computation has not considered whether a restricted form of ansatz is capable of universal quantum computation at some arbitrary depth as Lloyd [17] and others [18] have. Instead, the objective function to be minimized is defined in terms of the unitary gates arising in the target circuit to be simulated. What ansatz states are then required to simulate a given target circuit?

This question appears to be difficult, and not much is currently known. In the case of QAOA it was recently shown by myself and coauthors that the ability of an ansatz to approximate the ground-state energy of a satisfiability instance worsens with increasing problem density (the ratio of constraints to variables) [29]. These related results, however, do not immediately apply to our interests here.

Towards our goals, we show that reasonable depth circuits might saturate bipartite entanglement—the depth of these circuits scales with the number of qubits and depends on the interaction geometry present in a given quantum processor. Consider the following.

An ebit is a unit of entanglement contained in a maximally entangled two-qubit (Bell) state. A quantum state with q ebits of entanglement (quantified by any entanglement measure) contains the same amount of entanglement (in that measure) as q -Bell states.

Lemma 7. Let c be the depth of two-qubit controlled rotation gates in the n -qubit hardware-efficient ansatz. Then the maximum possible number of ebits across any bipartition is

$$E_b = \min\{\lfloor n/2 \rfloor, c\}.$$

In a low-depth circuit, the underlying geometry of the processor heavily dictates c above. For example, for a line of qubits and for a ring, the minimal c required to possibly maximize E_b is $\sim n/2$ and $\sim n/4$ respectfully. However, in the case of a grid, the minimal depth scales as $\sim \sqrt{2}/2$.

Hence, if we wish to simulate a quantum algorithm described by a low-depth circuit, having access to a grid architecture could provide an intrinsic advantage. Specifically, our combinatorial quantum circuit area law establishes that an objective circuit generating $k < \lfloor n/2 \rfloor$ ebits across every bipartition, must be simulated by an ansatz of, at least, minimal required circuit depth $\sim \sqrt{k}$ on a grid.

Although this does establish a preliminary relationship, the general case remains unclear at the time of writing. For example, given a quantum circuit with application time t^* which outputs $|\psi\rangle$, what is the minimal $t(\epsilon) \leq t^*$ for a control sequence to provide an ϵ close two-norm approximation to $|\psi\rangle$?

II. DISCUSSION

We have established that variational quantum computation admits a universal model. The gate sequence being simulated serves as an upper bound showing that a control sequence exists to minimize the expected values. Expected values are then the fleeting resource of the variational model.

Although error correction is assumed in our universality proofs, the techniques we develop should augment possibilities in the NISQ setting, particularly, with the advent of error suppression techniques [30,31]. Importantly, variational quantum computation forms a universal model in its own right and is not (in principle) limited in application scope.

An interesting feature of variational quantum computation is how many-body Hamiltonian terms are realized as part of the measurement process. This is in contrast with leading alternative models of universal quantum computation.

In the gate model, many-body interactions must be simulated by sequences of two-body gates. The adiabatic model applies perturbative gadgets to approximate many-body interactions with two-body interactions [9,32]. The variational model simulates many-body interactions by local measurements. Moreover the coefficients weighting many-body terms need not be implemented by the quantum hardware directly; this weight is compensated for in the classical process. Finally,

as many quantum states can cause a considered objective function to accept, the presented model is, therefore, partially agnostic to how states are prepared.

Variational counterparts to an increasing number of celebrated quantum algorithms have been recently developed, including the solution to linear [33] (and nonlinear [34]) systems. Indeed, computational universality implies that such variational incarnations generally exist. Yet the present results certainly do not rule out significant overhead reductions of task tailored variational quantum algorithms.

ACKNOWLEDGMENTS

The author acknowledges support from the project, Leading Research Center on Quantum Computing (Agreement No. 014/20).

The author declares no competing interests.

The data that support the findings of this Letter are available within the article. Data Availability. No data sets were generated. Materials supporting the conclusions drawn in this Letter are available in the Supplemental Material [20].

-
- [1] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [2] 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).
- [3] J. Preskill, Quantum Computing in the NISQ era, and beyond, *Quantum* **2**, 79 (2018).
- [4] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, Parameterized quantum circuits as machine learning models, *Quantum Sci. Technol.* **4**, 043001 (2019).
- [5] 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).
- [6] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, Quantum optimization using variational algorithms on near-term quantum devices, *Quantum Sci. Technol.* **3**, 030503 (2018).
- [7] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, *Nature (London)* **549**, 195 (2017).
- [8] G. Verdon, M. Broughton, and J. Biamonte, A quantum algorithm to train neural networks using low-depth circuits, [arXiv:1712.05304](https://arxiv.org/abs/1712.05304).
- [9] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, Adiabatic quantum computation is equivalent to standard quantum computation, in *45th Annual IEEE Symposium on Foundations of Computer Science, Rome, 2004* (IEEE, Piscataway, NJ, 2004), pp. 42–51.
- [10] A. M. Childs, Universal Computation by Quantum Walk, *Phys. Rev. Lett.* **102**, 180501 (2009).
- [11] N. B. Lovett, S. Cooper, M. Everitt, M. Trevers, and V. Kendon, Universal quantum computation using the discrete-time quantum walk, *Phys. Rev. A* **81**, 042330 (2010).
- [12] M. Van den Nest, A. Miyake, W. Dür, and H. J. Briegel, Universal Resources for Measurement-Based Quantum Computation, *Phys. Rev. Lett.* **97**, 150504 (2006).
- [13] D. Deutsch, Quantum theory, the Church-Turing principle, and the universal quantum computer, *Proc. R. Soc. London A, Ser. A* **400**, 97 (1985).
- [14] S. C. Benjamin and S. Bose, Quantum Computing with An Always-On Heisenberg Interaction, *Phys. Rev. Lett.* **90**, 247901 (2003).
- [15] S. C. Benjamin and S. Bose, Quantum computing in arrays coupled by “always-on” interactions, *Phys. Rev. A* **70**, 032314 (2004).
- [16] Y. Shi, Both toffoli, and controlled-not need little help to do universal quantum computing, *Quantum Info. Comput.* **3**, 84 (2003).
- [17] S. Lloyd, Quantum approximate optimization is computationally universal, [arXiv:1812.11075](https://arxiv.org/abs/1812.11075).
- [18] M. E. S. Morales, J. D. Biamonte, and Z. Zimborás, On the universality of the quantum approximate optimization algorithm, *Quant. Info. Proc.* **19**, 291 (2020).
- [19] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Am. Stat. Assoc.* **58**, 13 (1963).
- [20] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevA.103.L030401> for proofs not appearing after a lemma or theorem.
- [21] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules, and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [22] A. V. Uvarov, A. S. Kardashin, and J. D. Biamonte, Machine learning phase transitions with a quantum processor, *Phys. Rev. A* **102**, 012415 (2020).
- [23] A. Kardashin, A. Uvarov, D. Yudin, and J. Biamonte, Certified variational quantum algorithms for eigenstate preparation, *Phys. Rev. A* **102**, 052610 (2020).
- [24] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, Learning the quantum algorithm for state overlap, *New J. Phys.* **20**, 113022 (2018).
- [25] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, *Quantum* **3**, 181 (2019).
- [26] S. Bravyi and D. Gosset, Improved Classical Simulation of Quantum Circuits Dominated by Clifford Gates, *Phys. Rev. Lett.* **116**, 250501 (2016).
- [27] R. P. Feynman, Quantum mechanical computers, *Opt. News* **11**, 11 (1985).
- [28] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical, and Quantum Computation* (American Mathematical Society, Boston, 2002).
- [29] V. Akshay, H. Philathong, M. E. S. Morales, and J. D. Biamonte, Reachability Deficits in Quantum Approximate Optimization, *Phys. Rev. Lett.* **124**, 090504 (2020).
- [30] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, *Phys. Rev. X* **7**, 021050 (2017).

- [31] 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).
- [32] J. D. Biamonte and P. J. Love, Realizable Hamiltonians for universal adiabatic quantum computers, *Phys. Rev. A* **78**, 012352 (2008).
- [33] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles, Variational quantum linear solver, [arXiv:1909.05820](https://arxiv.org/abs/1909.05820).
- [34] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, *Phys. Rev. A* **101**, 010301(R) (2020).