Variational quantum eigensolver for the Heisenberg antiferromagnet on the kagome lattice

Joris Kattemölle [1,2,3] and Jasper van Wezel [1,2]

[1]*Institute for Theoretical Physics, University of Amsterdam, Science Park 904, Amsterdam, Netherlands*
[2]*QuSoft, CWI, Science Park 123, Amsterdam, Netherlands*
[3]*Department of Physics, University of Konstanz, D-78457 Konstanz, Germany*

Establishing the nature of the ground state of the Heisenberg antiferromagnet (HAFM) on the kagome lattice is well-known to be a prohibitively difficult problem for classical computers. Here, we give a detailed proposal for a variational quantum eigensolver (VQE) intending to solve this physical problem on a quantum computer. At the same time, this VQE constitutes an explicit experimental proposal for showing a useful quantum advantage on noisy intermediate-scale quantum devices because of its natural hardware compatibility. We classically emulate noiseless and noisy quantum computers with either 2D-grid or all-to-all connectivity and simulate patches of the kagome HAFM of up to 20 sites. In the noiseless case, the ground-state energy, as found by the VQE, approaches the true ground-state energy exponentially as a function of the circuit depth. Furthermore, VQEs for the HAFM on any graph can inherently perform their quantum computations in a decoherence-free subspace that protects against collective longitudinal and collective transversal noise, adding to the noise resilience of these algorithms. Nevertheless, the extent of the effects of other noise types suggests the need for error mitigation and performance targets alternative to high-fidelity ground-state preparation, even for essentially hardware-native VQEs.

I. INTRODUCTION

Despite decades of developments in numerical methods, the ground-state properties of the Heisenberg antiferromagnet (HAFM) on the kagome lattice (Fig. 1) remain elusive, owing to its geometrical frustration. The kagome HAFM is the prime candidate for exhibiting a unique phase of magnetism, the quantum spin liquid (QSL) [1–3], and forms a model for the magnetic properties of minerals like Herbertsmithite [4]. Approaches toward solving this ground state problem include exact diagonalization of finite-size patches [5] and the density matrix renormalization group (DMRG) method [2]. Next to the QSL, a valence bond crystal (VBC) has been proposed as the ground state of the kagome HAFM [6–10]. (See Ref. [5] and references therein for a more complete overview of the techniques and proposals.) All classical methods for finding the ground state of the kagome HAFM are ultimately limited, for example, by the inability to treat large patches (exact diagonalization) or the inability to describe highly entangled states (DMRG).

Quantum computation is a new player in this field that brings with it entirely novel possibilities. One method for finding ground states on a quantum computer is the variational quantum eigensolver (VQE) [11,12]. VQEs are especially suited for noisy intermediate-scale quantum (NISQ) [13] devices because of their relatively mild circuit depth requirements and inherent noise resilience [11,12,14,15]. A VQE is a variational method. What sets the VQE apart from classical variational methods is that the parametrized state is obtained by applying a parametrized quantum circuit to some easy-to-prepare reference state of the quantum computer's register.

The energy of the resulting state is obtained by performing measurements on many copies of that state. (Generally, the classical simulation of state preparation and measurement is intractable.) Parameter variation and optimization are still performed by a classical routine. So, a VQE can be seen as a classical variational method that uses a quantum computer as a subroutine for its function calls to the energy landscape.

Quantum computers can already outperform classical computers, and have hence obtained what is called quantum supremacy or a quantum advantage [16,17]. However, the tasks for which quantum computers can currently outperform classical computers have no known application; these tasks were designed purely for showing a quantum advantage. The
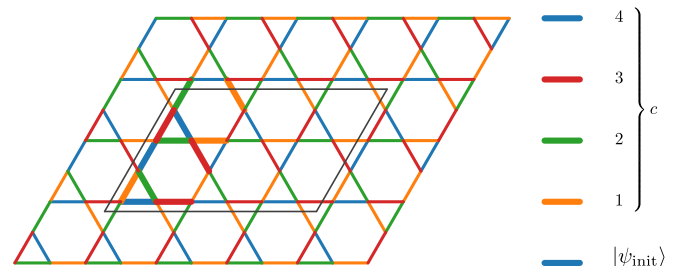

FIG. 1. The kagome lattice, with edges colored according to a minimal edge coloring with the smallest possible coloring unit cell (thicker lines). The ansatz of KVQE$_K$, with initial state $|\psi_{\text{init}}\rangle$ and circuit cycle c, is derived from this edge coloring by identifying every subset of edges having the same color with a layer of parametrized gates. The gray parallelogram delineates a periodic patch that is simulated in this paper.

milestone of a *useful* quantum advantage, where a quantum computer performs a useful task that cannot be performed on any classical computer, is still ahead [16,17].

In this paper, we design a VQE for the kagome HAFM as an explicit proposal for showing a useful quantum advantage on NISQ devices and for a method for finding the ground-state properties of the kagome HAFM. For NISQ VQE algorithms to give an advantage over purely classical methods, it has recently become increasingly clear that the structure of the problem must be close to the quantum hardware the VQE is run on [18]. For this reason, the central point of our paper is careful consideration of hardware compatibility and model choice. We explicitly consider the case where the quantum computer has the limited connectivity of a 2D grid. We refer to the resulting VQE as KVQE$_G$ for short. We refer to the VQE as KVQE$_K$ whenever we assume a quantum computer with at least kagome connectivity (this includes architectures with all-to-all connectivity). We tested KVQE$_{G/K}$ by noiseless and noisy emulation on a classical computer in a host of cases, which are detailed in Sec. II E. For both KVQE$_G$ and KVQE$_K$, the results of one case are reported in this paper. The remainder can be found in the Supplemental Material [19]. KVQE$_{G/K}$ only gives information about the ground-state properties of finite-size patches. As is standard practice in numerical methods, properties of the infinite system can be derived from a finite-size scaling approach, as in Ref. [20]. Such an approach is beyond the scope of the current paper, since there are not enough patches for which it is tractable to emulate KVQE classically, but would be fully possible on a quantum computer with on the order of 50 qubits.

KVQE$_{G/K}$ uses the Hamiltonian variational ansatz (HVA) [21]. In the HVA, to find the ground state of a Hamiltonian H, the ansatz state is obtained by first preparing some known, easy-to-prepare ground state of a Hamiltonian H_{init}. Thereafter, this state is evolved sequentially by terms from H and H_{init}. (Commuting terms may be evolved by simultaneously.) In the resulting circuit, every gate corresponds to time evolution along a term in H or H_{init}, where the parameter of that gate is set by the time duration of the evolution. The HVA itself does not specify H_{init} nor the sequence of terms the initial state is evolved by. Our choice of H_{init} and gate sequences for the cases presented in this paper are detailed in Figs. 1 and 2.

In addition, we propose to run a similar VQE for the HAFM on the periodic chain as an intermediate goal. We call this VQE CVQE for short. In contrast to KVQE$_{G/K}$, a system-size resource scaling, as presented in this paper, is possible for CVQE due to the many classically emulatable patches. Furthermore, CVQE is a suitable benchmark problem for quantum hardware because, in contrast to the kagome HAFM, the ground state of the HAFM on the chain can be computed efficiently classically through the Bethe ansatz [22–24]. This opens the possibility of comparing the optimal energy found by the VQE running on a quantum computer against the exact ground-state energy, even for chains with hundreds of sites. It is only for systems of up to approximately 50 qubits that similar benchmarks can be made for the kagome HAFM [5]. The explicit gate sequence we use for CVQE is depicted in Fig. 3. CVQE is similar to a VQE in Ref. [25]. Differences with the results in Ref. [25] are that we simulate a periodic
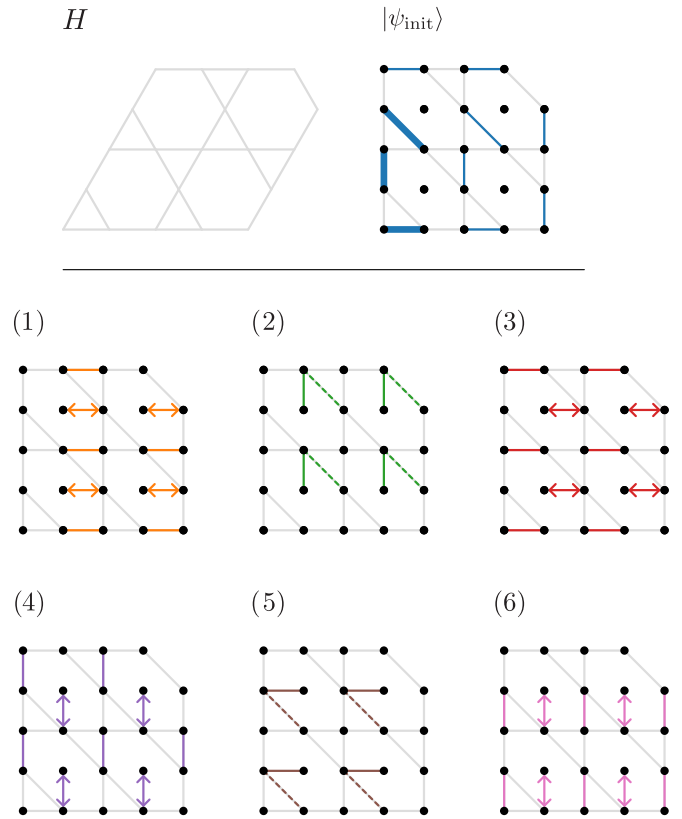

FIG. 2. Top left: The kagome lattice, of which only the 20-site open patch simulated by KVQE$_G$ is shown. For the kagome HAFM, spin-1/2 particles are placed on the vertices and the Heisenberg interaction is defined along the edges [Eq. (3)]. Top right: The initial state $|\psi_{init}\rangle$. Black dots represent qubits and solid blue lines represent singlets. A sheered kagome lattice is added in the background in gray as a guide to the eye. The bolder solid lines form one unit cell of the dimer covering, which can be used to extend the current ansatz to systems of arbitrary size. For open boundaries, some patching of the regular dimer covering is needed. Here, this consists of the two singlets at the far right. Bottom: The cycle c that is applied to $|\psi_{init}\rangle$ p times to obtain the ansatz state. Solid colored lines represent HEIS(α) gates. A HEIS(α) gate evolves two qubits according to the Heisenberg exchange interaction for a time given by α [Eq. (4)]. We use one parameter per HEIS gate, leading to $M = 30p$ parameters in total. Two-headed arrows indicate SWAP gates. Dashed, colored lines indicate along which bond of the kagome lattice the HEIS gates of that layer act effectively. The unit cell of the cycle equals that of the kagome lattice itself, and can hence be straightforwardly extended to larger system sizes.

chain instead of an open chain and that we use one parameter per gate (OPG) instead of one parameter per layer. We go beyond the results in Ref. [25] by the study of larger circuit depths and system sizes, a finite-size resource scaling, and a study of noise effects.

A. Hardware compatibility

KVQE$_G$ is exceptionally compatible with NISQ hardware for three reasons. First, because the Hamiltonian of the kagome HAFM is a spin Hamiltonian, it is directly a Hamiltonian defined on qubits, eliminating any overhead
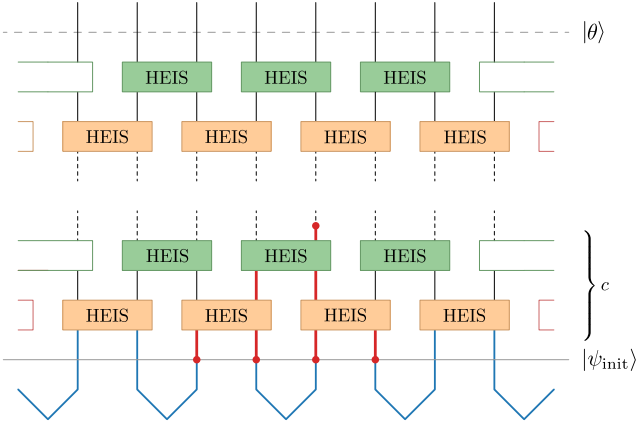
FIG. 3. The ansatz for CVQE, with time running from bottom to top. The initial state $|\psi_{\text{init}}\rangle$ consists of $n/2$ adjacent singlets, displayed in blue. The circuit that is applied to $|\psi_{\text{init}}\rangle$ consists of p repetitions of a cycle c, each time with new parameters, and with one parameter per gate. For one of the qubits, the past light cone (discussed in Sec. III A 2) that is due to a single cycle c (thus excluding the singlet generation) is displayed with thicker, red lines.

from fermion-to-spin maps. In contrast, many Hamiltonians for which VQEs are proposed are fermionic. Examples include those in quantum chemistry [12,14,26] and the Fermi-Hubbard model [15,21,27,28]. For a VQE to solve for the ground state of a fermionic Hamiltonian, it first needs to be mapped to a spin Hamiltonian, for example, by the Jordan-Wigner [29], Bravyi-Kitaev [30], or ternary-tree [31] transformations. Fermion-to-spin maps either increase the nonlocality of terms in the Hamiltonian or introduce additional qubits, in any case leading to an overhead in quantum resources.

A second reason that KVQE_G is close to NISQ hardware is that its gates are essentially native to multiple NISQ architectures. The HVA requires time evolution generated by terms in the Hamiltonian. For the HAFM, this amounts to turning on an exchange interaction between qubits, which is native to quantum dot architectures [32–35]. In Sec. IV, we show this interaction can also be realized on the superconducting hardware by Google AI Quantum [36] using a single native two-qubit gate and at most four single-qubit gates. (If a two-qubit gate is equal to a single native gate up to single-qubit rotations, we call the former gate essentially native.)

Next to the HVA, a well-known type of ansatz is the hardware-efficient ansatz (HEA) [12,37]. The HEA is hardware inspired; the circuit generating the ansatz state consists, by definition, of gates native to the hardware, avoiding the need to compile the ansatz into native gates. However, the HEA suffers from the barren plateau problem: the gradient of the energy cost function is exponentially small in the number of parameters [38]. The HVA, on the other hand, is problem inspired, and there is some evidence that it does not suffer from the barren plateau problem [39]. However, for execution on a quantum computer, gates in the HVA generally need to be compiled to gates native to that quantum computer. This increases the circuit depth, which is undesirable for NISQ devices. For KVQE_G, such compilation is not required on

quantum dot architectures, and only minimal compilation that does not increase the number of two-qubit gates is needed on Google's hardware. So, to summarize, KVQE_G on quantum dot and Google's superconducting hardware has the unique property that the HVA is essentially equal to the HEA.

Finally, KVQE_G is close to NISQ hardware because it runs on hardware with the connectivity of a 2D grid with minimal overhead. This is the connectivity that is also required for the surface code [40], and therefore much effort is put into designing platforms with grid connectivity [16,35,41–43].

Given a quantum computer with kagome connectivity, the hardware compatibility of KVQE_K goes even further because it would require no geometric mapping at all. Quantum architectures with kagome connectivity also form a proposal for fault-tolerant quantum computation [44], and superconducting quantum processors with said connectivity are under theoretical [45–47] and experimental [48,49] investigation. Unlike superconducting platforms, ion-trap quantum computers feature all-to-all connectivity [50–53], which naturally accommodates kagome connectivity. CVQE requires line connectivity, which is a limited connectivity available on most quantum computing platforms.

All circuits in our work can directly be used for the dynamic quantum simulation of the HAFM on the kagome lattice, offering an additional route toward a quantum advantage [54]. The dynamical simulation of the kagome HAFM has arguably less scientific relevance, and we therefore focus on ground-state simulation in the current paper.

B. Summary of numerical results

The task for KVQE_G reported here is the simulation of a 20-site open patch of the kagome HAFM. For this case, we assume a 24-qubit quantum computer with grid connectivity, the ability to natively implement the exchange interaction, SWAP, $\sqrt{Z}$, and X gates, and where every exchange interaction gate in the ansatz circuit has its own parameter. We emulate this quantum computer classically, as detailed in Sec. II E. The optimal state obtained by KVQE is compared to the exact ground state of the 20-site patch, which is obtained by exact diagonalization. In the noiseless case, we find that the optimal energy obtained by KVQE_G approaches the true ground-state energy exponentially as a function of the circuit depth. Also, the fidelity (overlap squared) between the optimal state and the true ground state approaches unity exponentially in the circuit depth. A fidelity of $>99.9\%$ is reached at a circuit depth of 99. More details are found in Fig. 4 and Sec. III A 1.

The task for KVQE_K reported here is the simulation of an 18-site *periodic* patch. For this case, we assume a noiseless 18-qubit quantum computer with all-to-all connectivity and the other settings as before [55]. Also the optimal energy obtained by KVQE_K approaches the true ground-state energy exponentially as a function of the circuit depth. The fidelity initially plateaus, but then continues exponentially toward unity. A fidelity of $>99.9\%$ is reached at a circuit of depth 151. We find the increment of the circuit depth required for obtaining a fidelity of $>99.9\%$ (as compared to the previous paragraph) not to be the result of the different ansatz but
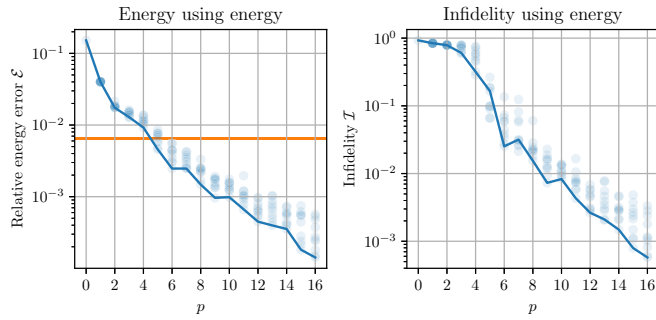
FIG. 4. Left: Semilog plot of the relative energy error [Eq. (5)] obtained by KVQE_G for 20 sites (Fig. 2), as a function of the number of cycles p. Translucent points represent the 10 local minima that were found by KVQE_G per p. At $p = 0$, only one point is present since no parameters need to be optimized. Every cycle is a circuit of depth 6 and uses 30 parameters. The solid line connects the lowest local minima $E(\theta^*)$. An orange horizontal line is drawn at the value of $\mathcal{E}$ corresponding to the energy of the first excited state (not restricting this state to any total spin sector). Right: Semilog plot of the infidelities of the states corresponding to the local minima in the left plot. The solid line connects the points that, for a given p, are lowest in *energy*. Although it occurs regularly, these points need not have the lowest infidelity.

rather to be the result of the difference in the systems that are simulated. More details are found in Fig. 5 and Sec. III A 1.

Under the same hardware assumptions, we let CVQE simulate a 20-site periodic chain. The fidelity and energy of the optimal state found by CVQE improve exponentially as a function of circuit depth, with a sudden improvement of performance after $p_{\mathrm{crit}} = 5$ due to the availability of systemwide entanglement after p_{crit}. The optimal state found by CVQE reaches a fidelity of >99.9% at a circuit of depth 19. More details are found in Fig. 6 and Sec. III A 2.

Considering the effects of noise, we find that without any modification, the HVA for the HAFM on any graph intrinsically performs its quantum computations inside a decoherence-free subspace (DFS) [56–60] that protects against noise that couples to the quantum registers' total spin operator in any direction (Sec. II B 3). This type of noise occurs in realistic scenarios, for example, when a single bath of long-wavelength modes of a bosonic bath (like the electromagnetic field) couples to the qubits collectively trans-
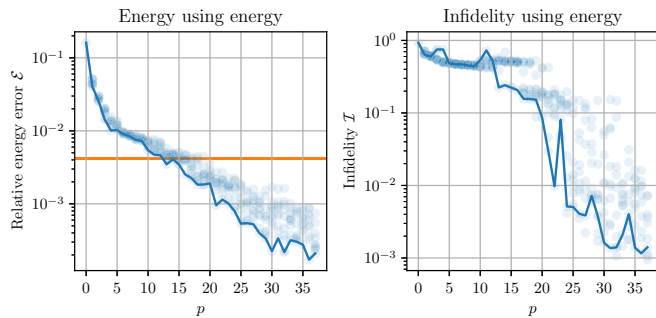

FIG. 5. Semilog plots similar to those in Fig. 4, but now for KVQE_K, simulating a periodic patch of 18 sites (Fig. 1). Per p, ten local minima are displayed. Every cycle is a circuit of depth 4 and uses 36 parameters.
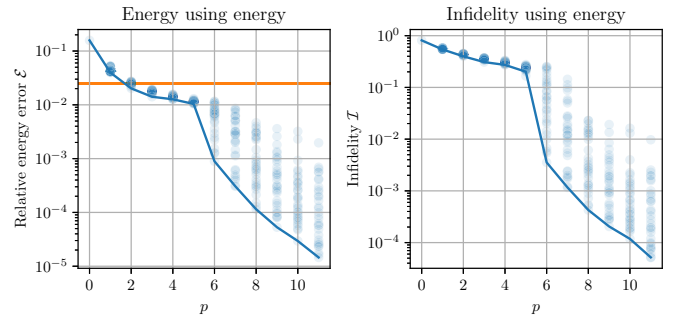

FIG. 6. Semilog plots similar to those in Fig. 4, but now for CVQE (Fig. 3), simulating the HAFM on a periodic chain of 20 sites. Per p, 32 local minima are displayed. Every cycle is a circuit of depth 2 and uses 20 parameters.

versely [61,62] and/or collectively longitudinally [53,56]. A VQE whose states occurring during the ansatz preparation remain in a DFS was treated before in Ref. [63], but differs from our setting in multiple essential ways (Appendix A). The fact that our proposed VQEs naturally perform their computations in a DFS has direct practical consequences, since it adds to the natural noise robustness of these VQEs against realistic noise sources.

We also study the degradation of the fidelity $\mathcal{F}$ between the true ground state and the output state of the VQE, as a function of the error probability p_e, for local depolarizing and bit-flip noise models (Sec. III B). We find that to prepare a ground state with fidelity $\mathcal{F}$, the error probability must be less than $p_e \approx (1 - \mathcal{F})/(nd)$, with n the number of sites and d the depth of the circuit. Using this relation, we estimate that to obtain the kagome HAFM ground state on a system of a hundred sites with a fidelity of 99.9%, error rates as low as $p_e = 10^{-7}$ may be needed. This is several orders of magnitude below known error thresholds for fault-tolerant quantum computation [64,65]. This makes the high-fidelity preparation of ground states of systems with on the order of a hundred sites unlikely on noisy hardware. Nevertheless, other classically nontrivial performance targets might be achieved by noisy hardware (Sec. V), while error mitigation may alleviate the stringent constraints on the physical error rates [66–74].

II. METHODS

A. VQE

In this subsection, we give a more detailed introduction to VQEs. It may be skipped by readers already familiar with VQEs. The problem of finding the ground-state energy of general k-local Hamiltonians is believed to be intractable even on quantum computers [75–77]. Nevertheless, there may be problem instance classes for which quantum computers could solve for the ground state efficiently. The VQE is a proposed general method for finding ground states on quantum computers.

Consider a quantum mechanical system with Hilbert space $\mathcal{H}$ of dimension N, Hamiltonian H with ground-state energy E_0, and a subset of parametrized states $\{|\theta\rangle\} \subseteq \mathcal{H}$, with $\theta \in \mathbb{R}^m$. To describe all of $\mathcal{H}$, it is necessary that $m = O(N)$. The

fact that

$$E(\theta) = \langle\theta|H|\theta\rangle \geqslant E_0$$

for all $|\theta\rangle \in \mathcal{H}$ is called the variational principle. For reasons of scalability, in variational methods, one generally employs a set of states described by $m = \mathrm{polylog}\, N$ parameters. Variational methods, like the VQE, seek to minimize $E(\theta)$ to hence establish an upper bound for the ground-state energy.

As input, a VQE receives a description of a Hamiltonian on n spin-$1/2$ particles,

$$H = \sum_{i=1}^{l} h_i H_i, \qquad (1)$$

with h_i real coefficients and H_i Hermitian operators. For k-local Hamiltonians, $l = \mathrm{poly}\, n$.

A VQE proceeds by first choosing an initial set of parameters θ. These initial parameters may be chosen at random or may be inspired by a classical approximate solution to the ground state, for example, by the Hartree-Fock ground state [78]. Then, a criterion is chosen, for example, that a maximum number of iterations has not been reached or that $E(\theta)$ has not reached a value below a given threshold. The VQE then proceeds as follows.

(1) While the criterion is true, repeat:

    (a) *Prepare the ansatz* $|\theta\rangle$.

    Prepare the initial state $|\psi_{\mathrm{init}}\rangle$. Apply a parametrized circuit $C(\theta)$ to obtain the state $|\theta\rangle = C(\theta)|\psi_{\mathrm{init}}\rangle$. The circuit $C(\theta)$ usually consists of gates on fixed positions, where every or some of the gates are parametrized.

    (b) *Measure and store* $E(\theta)$.

    By linearity, $E(\theta) = \sum_i h_i\langle\theta|H_i|\theta\rangle$. Each expectation value $\langle\theta|H_i|\theta\rangle$ can be estimated by measuring the operator H_i repeatedly (each measurement requires a new preparation of $|\theta\rangle$) and taking the statistical average. See Ref. [11] for the expected number of measurements using this method, or Ref. [79], and references therein, for more efficient methods.

    (c) *Update* θ.

    Based on $E(\theta)$ and previous outcomes of $E(\theta)$, update θ according to some classical optimization algorithm.

(2) Return θ^*, which we define as the θ that achieved the lowest energy.

Physically relevant information, such as correlation functions, can now be extracted from $|\theta^*\rangle$ by repeatedly preparing and performing measurements on $|\theta^*\rangle$. Different VQEs differ in the way circuits are parametrized, how (an estimate for) $E(\theta)$ is obtained, and what specific optimization routine is used. These will be detailed in the subsequent sections.

B. Ansatz

1. Cyclic HVA

In the HVA [21], the initial state $|\psi_{\mathrm{init}}\rangle$ is the ground state of a Hamiltonian H_{init}. The Hamiltonian H_{init} is chosen in such a way that its ground state is known and easy to prepare. The ansatz state $|\theta\rangle$ is obtained by sequentially evolving along terms in H, according to some fixed sequence i,

$$|\theta\rangle = C(\theta)|\psi_{\mathrm{init}}\rangle,$$

with

$$C(\theta) = \exp(-\mathrm{i}\theta_M H_{i_M})\dots\exp(-\mathrm{i}\theta_2 H_{i_2})\exp(-\mathrm{i}\theta_1 H_{i_1}).$$

The M parameters are formed by the time duration of the M evolutions. Often, as for example in, but not limited to, Refs. [21,25,27,39], C consists of p cycles of a smaller circuit c, each time defined by the same sequence i of terms in the Hamiltonian. Every cycle gets its own set of m parameters. It is convenient to write θ as $\theta = (\theta_1, \dots, \theta_p)$, with $\theta_j = (\theta_{j_1}, \dots, \theta_{j_m})$. Then, a single cycle reads

$$c(\theta_j) = \exp(-\mathrm{i}\,\theta_{j_m}H_{i_m})\dots\exp(-\mathrm{i}\,\theta_{j_2}H_{i_2})\exp(-\mathrm{i}\,\theta_{j_1}H_{i_1}),$$

and so

$$|\theta\rangle = c(\theta_p)\dots c(\theta_1)|\psi_{\mathrm{init}}\rangle. \qquad (2)$$

We call this type of HVA the *cyclic* HVA. The cyclic HVA shows a close relation between static quantum simulation and dynamic quantum simulation; choosing $i = (1, \dots, l)$ (or a permutation thereof) and $\theta_j = (t/p, \dots, t/p)$ for all j, the cyclic HVA implements quantum time evolution for a target time t with p Trotter steps. In this way, the cyclic HVA can also mimic (but is more general than) adiabatic time evolution. In the case that no gap closes while adiabatically evolving from H_{init} to H, the HVA thus ensures that the ground state of H can be prepared with the ansatz (but without guaranties on the required circuit depth). This formed the initial motivation for the HVA in VQEs [21].

2. The HVA in this paper

We use the cyclic HVA, where the Hamiltonian (in units where $\hbar = 1$) is given by the HAFM Hamiltonian [cf. Eq. (1)]

$$H = \sum_{\langle i,j\rangle} \mathbf{S}^{(i)} \cdot \mathbf{S}^{(j)}, \qquad (3)$$

where $\mathbf{S}^{(i)} = (X_i, Y_i, Z_i)^T/2$, with X_i, Y_i, Z_i the Pauli matrices acting on spin i only, and where the sum runs over the edges $\langle i, j\rangle$ of some graph G. In this paper, G is either an open or a periodic patch of the kagome lattice or a periodic chain, and we take $H_{\mathrm{init}} = \sum_{\langle i,j\rangle'} \mathbf{S}^{(i)} \cdot \mathbf{S}^{(j)}$, with $\langle i, j\rangle'$ edges in a perfect matching of G. Since the ground state of a single term in H is the singlet state $|s\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, this means $|\psi_{\mathrm{init}}\rangle$ is a so-called dimer covering on the relevant physical system, where every dimer is a singlet state.

We define the gate HEIS (Heisenberg interaction) as time evolution along a single term in the HAFM Hamiltonian. In the computational basis of two qubits, it reads

$$\mathrm{HEIS}(\alpha) \equiv \mathrm{e}^{-\mathrm{i}\alpha/4}\mathrm{e}^{-\mathrm{i}\alpha\mathbf{S}^{(1)}\cdot\mathbf{S}^{(2)}}$$

$$= \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\alpha/2} & 0 & 0 & 0 \\ 0 & \cos(\alpha/2) & -\mathrm{i}\sin(\alpha/2) & 0 \\ 0 & -\mathrm{i}\sin(\alpha/2) & \cos(\alpha/2) & 0 \\ 0 & 0 & 0 & \mathrm{e}^{-\mathrm{i}\alpha/2} \end{pmatrix}. \qquad (4)$$

In a parametrized circuit [Eq. (2)], every instance of the HEIS gate gets its own parameter $\alpha = \theta_{j_k}$.

3. Symmetry

The initial states, Hamiltonian, and the ansatz circuits used in this paper possess an SU(2) rotational symmetry. In this section, we discuss how this leads to the following properties of our VQEs: (i) the variational manifold lays within the correct spin sector, (ii) they offer the opportunity of studying the physically relevant spin gap [2,5,20,80–82] of the kagome HAFM, and (iii) they intrinsically perform their computations in a DFS. More details are given in Appendix A.

First, (i) for the two-spin singlet state, the total spin quantum number S and the total magnetization quantum number S_z vanish, and therefore $S = S_z = 0$ for our initial states $|\psi_{\text{init}}\rangle$. By the SU(2) symmetry of our ansatz circuits, also $S = S_z = 0$ for the ansatz state $|\theta\rangle$ (Appendix A). The ground state of the HAFM on the kagome lattice is believed to have those same quantum numbers (Appendix A). Thus, the additional benefit of using the HVA over the HEA is that using the HVA, the variational manifold $\{|\theta\rangle\}$ automatically entirely lays within the correct spin sector.

The spin gap (ii) is defined as the energy difference between the lowest energy eigenstate in the $S = 0$ sector and the lowest energy eigenstate in the $S = 1$ sector [2,5,20,80–82]. Similar to the ideas in Ref. [83], KVQE$_{G/K}$ and CVQE can be extended to study the spin gap by running them with an $S = 1$ initial state, obtaining an optimal energy $E(\theta^*_{S=1})$, in addition to running them with a $S = 0$ dimer covering as the initial state, obtaining $E(\theta^*_{S=0})$. An estimate for the spin gap is then given by $E(\theta^*_{S=1}) - E(\theta^*_{S=0})$. $S = 1$ initial states can be obtained by changing one of the singlets in our initial state $|\psi_{\text{init}}\rangle$ into one of the three two-spin triplet states. Thus, we obtain a $(S = 1, S_z = 1)$, $(S = 1, S_z = 0)$, or $(S = 1, S_z = -1)$ dimer covering as the initial state, depending on which triplet state was chosen. In Appendix A, we show that the energy $E(\theta)$ is invariant under the choice of triplet state. Noise may brake the symmetry of the ansatz circuit. KVQE$_{G/K}$ and CVQE may abuse this noise to put amplitude on $S = 0$ states, thus obtaining unjustly low variational energies. This problem can be handled by symmetry verification [73] or by adding a term to the Hamiltonian that penalizes $S \neq 1$ states [11,84,85] (Appendix A).

For the performance of KVQE$_{G/K}$ and CVQE using $S = 1$ initial states, we expect to obtain results similar to those when using $S = 0$ initial states, as already presented in this paper. Correspondingly, the horizontal lines in Figs. 4–6 represent the first excited state above the ground state, where we put no restriction on the total spin sector of either the ground state or the first excited state. That is, the horizontal line need not represent the ground state in the $S = 1$ sector.

Finally, (iii), similar considerations lead to the conclusion that the VQEs in this paper naturally perform their computations in a DFS. If HEIS gates are native or can be compiled in terms of gates that preserve the total spin quantum number S at all times during gate execution, the state $|\psi\rangle$ occurring at any (nondiscretized) time during the ansatz circuit has total spin quantum number $S = 0$. It follows directly that $\mathcal{S}|\psi\rangle = 0$ for any $\mathcal{S} \in \mathcal{D} = \{\mathbf{S}^2, \mathbf{S}_x^{(\text{tot})}, \mathbf{S}_y^{(\text{tot})}, \mathbf{S}_z^{(\text{tot})}\}$. Here, $\mathbf{S}^2 \equiv \mathbf{S}^{(\text{tot})} \cdot \mathbf{S}^{(\text{tot})}$, with total spin operator $\mathbf{S}^{(\text{tot})} = \sum_{i=1}^{n} \mathbf{S}^{(i)}$ [cf. Eq. (3)]. Therefore, $|\psi\rangle$ is in a DFS that protects against continuous coupling to (any linear combination of) the operators in $\mathcal{D}$ (Appendix A).

4. Parameter multiplicity

In this paper, every Heisenberg gate in the ansatz gets its own parameter; we have one parameter per gate (OPG). Another possibility would be to have multiple HEIS gates per cycle share the same parameter. We call this one parameter per slice (OPS). We say the qubits sharing the same parameter are in the same slice.

A possible advantage of OPS is that, by choosing proper slices, we can ensure that the state produced by the circuit has the lattice symmetries expected to be present in the ground state. This would make the search space smaller by only restricting to states with the desired lattice symmetry.

Nevertheless, OPG has advantages over OPS. With OPS, we may overlook symmetry-broken ground states. For example, it is unknown whether the ground state of the kagome lattice has a spontaneously broken symmetry [5]. Second, even if the ground state does not break any symmetries, the depth of the OPG circuit for a given state may be lower than the depth of the OPS circuit that produces the same state. For NISQ devices, it is imperative to keep circuit depths as low as possible. Finally, the inherent noise resilience of VQEs may be compromised by choosing OPS over OPG. As an illustration, say we are given a noiseless quantum computer, a Hamiltonian, and a minimal-depth OPS circuit that produces the ground state of that Hamiltonian. Suppose that now a static but random over-rotation is added to each HEIS gate. Then it is very unlikely that the OPS circuit can still produce the correct ground state, no matter its parameters. When we lift the restriction that the gates in every slice share the same parameter, and hence go to an OPG circuit, the over-rotations can be absorbed into the parameters, and hence the ground state can still be produced without an increase in circuit depth.

C. Analysis

We assess the effectiveness of KVQE$_{G/K}$ and CVQE by running them for fixed system sizes but a varying number of cycles p. For every p, we plot the relative energy error $\mathcal{E}$ between the true ground-state energy, E_0, and the optimal energy found by the VQE, $E(\theta^*)$,

$$\mathcal{E} = \left| \frac{E(\theta^*) - E_0}{E_0} \right|. \tag{5}$$

Additionally, we plot the infidelity $\mathcal{I}$ between the true ground state $|E_0\rangle$ and the optimal state obtained by the VQE, $|\theta^*\rangle$,

$$\mathcal{I} = 1 - \mathcal{F} \equiv 1 - |\langle E_0|\theta^*\rangle|^2, \tag{6}$$

with $\mathcal{F}$ the fidelity between $|E_0\rangle$ and $|\theta^*\rangle$. Even in plots showing the infidelity, the corresponding VQE optimized the energy, not the infidelity. The infidelity is a useful figure of merit because it upper bounds the relative error in expectation value of any observable [86],

$$\frac{|\langle E_0|O|E_0\rangle - \langle\theta^*|O|\theta^*\rangle|}{\|O\|} \leqslant 4\sqrt{\mathcal{I}},$$

with $\| \cdot \|$ the operator norm, although this bound may be loose [87].

We obtain E_0 and $|E_0\rangle$, and thus $\mathcal{E}$ and $\mathcal{I}$, by exact diagonalization. For large system sizes, such as those needed

for obtaining a quantum advantage, this will no longer be possible.

D. Noise

On NISQ devices, noise will lead to errors during the preparation and measurement of the ansatz state. If the entire ansatz state preparation can be performed inside a DFS, these errors have no effect [56–59] (also see Appendix A). However, perfect DFSs only exists in idealized scenarios, and hence noise will generally degrade the performance of any VQE on a NISQ device [60].

Two common local noise models (against which an encoding into a DFS is generally impossible) are those involving depolarizing and bit-flip noise. Given an n–qubit density matrix ρ, the depolarizing channel acting on qubit q is given by the map

$$\rho \mapsto (1 - p_e)\rho + \frac{p_e}{3}(X_q \rho X_q + Y_q \rho Y_q + Z_q \rho Z_q). \quad (7)$$

Another common channel is the bit-flip channel:

$$\rho \mapsto (1 - p_e)\rho + p_e X_q \rho X_q. \quad (8)$$

These channels have a stochastic interpretation. For example, the latter channel can be interpreted as the following process: with probability $(1 - p_e)$, no error happens, and with probability p_e, a bit-flip error occurs.

Under stochastic noise of the type $\rho \mapsto \sum_i p_i A_i \rho A_i^\dagger$, with all A_i unitary, such as for the bit-flip and depolarizing channel, the ansatz circuit prepares a mixed state $\rho(\theta) = \sum_{\mathbf{i}} p_{\mathbf{i}}|\psi_{\mathbf{i}}(\theta)\rangle\langle\psi_{\mathbf{i}}(\theta)|$, with $p_{\mathbf{i}}$ the probability of noise realization $\mathbf{i}$ and $|\psi_{\mathbf{i}}(\theta)\rangle$ the pure state in case of known noise realization $\mathbf{i}$. Given a $\rho(\theta)$ of the aforementioned form, the energy expectation value becomes $E(\theta) = \text{tr}[H\rho(\theta)] = \sum_{\mathbf{i}} p_{\mathbf{i}}E^{(\mathbf{i})}(\theta)$, with $E^{(\mathbf{i})}(\theta) = \langle\psi_{\mathbf{i}}(\theta)|H|\psi_{\mathbf{i}}(\theta)\rangle$ the energy in case of noise realization $\mathbf{i}$. Likewise, the fidelity at given parameters θ can be written as $\mathcal{F}(\theta) = \sum_{\mathbf{i}} p_{\mathbf{i}}\mathcal{F}^{(\mathbf{i})}(\theta)$, with $\mathcal{F}^{(\mathbf{i})}(\theta) = |\langle E_0|\psi_{\mathbf{i}}(\theta)\rangle|^2$ the fidelity in case of noise realization $\mathbf{i}$. Because $\sum_{\mathbf{i}} p_{\mathbf{i}} = 1$, this can equivalently be phrased in terms of the infidelity; $\mathcal{I}(\theta) = \sum_{\mathbf{i}} p_{\mathbf{i}}\mathcal{I}^{(\mathbf{i})}(\theta)$, with $\mathcal{I}^{(\mathbf{i})}(\theta) = 1 - \mathcal{F}^{(\mathbf{i})}(\theta)$ the infidelity in case of noise realization $\mathbf{i}$.

For the remainder of this paper, we assume that after the preparation of the initial state, and after every layer of HEIS gates, the depolarizing channel acts separately on all qubits. We separately consider the case where instead the bit-flip channel acts at those spacetime locations. There are nd spacetime locations where an error may occur during a circuit run, with $d = |c|p + 1$ the depth of the circuit, where $|c|$ is the number of layers in the cycle c. Given θ^*, the fixed optimal parameters output by the VQE, and under the approximation that the occurrence of one or more errors during the execution of the ansatz circuit yields a state that is orthogonal to the true ground state, the noisy fidelity becomes

$$\mathcal{F}(\theta^*) \approx (1 - p_e)^{nd}\mathcal{F}_0(\theta^*), \quad (9)$$

with $\mathcal{F}_0 \equiv \mathcal{F}^{(0^L)}$ the fidelity in the noiseless case. Equivalently, under the same approximation, the infidelity becomes $\mathcal{I}(\theta^*) \approx 1 - (1 - p_e)^{nd}[1 - \mathcal{I}_0(\theta^*)]$, with $\mathcal{I} \equiv \mathcal{I}^{(0^L)}$.

E. Classical implementation

We emulate the quantum circuits in this paper using the homegrown, optionally GPU-accelerated, classical quantum emulator HEISENBERGVQE. Documentation, source code, and all generated data are available as Supplemental Material [19]. HEISENBERGVQE is tailored to running VQEs for the Heisenberg model on any graph. It is written in Python [88], with performance-critical code delegated to C via NumPy [89] if GPU acceleration is off, and CUDA via CuPy [90] if GPU acceleration is on.

We exploit the full access to the wave function, granted by classical emulation, for the computation of $E(\theta)$. We assume a noiseless quantum computer and use a gradient-based optimization method. Gradients are computed using backward-mode automatic differentiation [91], as implemented in Chainer [92].

For optimization of the cost function $E(\theta)$, we first choose initial parameters uniformly at random in the interval $[-10^{-3}, 10^{-3})$. There is some evidence that for the HVA, points close to the origin in parameter space are good starting points for local optimization [39]. We then use the BFGS algorithm, as implemented in SciPy [93], to find a local minimum. At every step of the BFGS routine, the energy and gradient of the energy are calculated. Here, we call these two steps together one *function call*. The steps of random parameter generation and local optimization (one round) are repeated a variable number of times. The parameters that achieve the lowest energy out of all local minimization rounds, θ^*, are stored together with $E(\theta^*)$. Thus, every data point in Figs. 4–6 corresponds to one local optimization, where the initial point of optimization is chosen independently and at random in the interval $[-10^{-3}, 10^{-3})$. Starting many rounds of local optimization from unrelated starting points has the benefit of straightforward parallelization. We note that the inner local classical optimization loop may be circumvented by using a full quantum eigensolver [94], based on linear combinations of unitary operations [95,96] at the cost of increased quantum resources and reduced success probability. We do not use this method here.

HEISENBERGVQE computes exact ground states using SciPy's wrapper of ARPACK, which implements the implicitly restarted Lanczos method [97]. Operator-vector multiplication is optionally GPU accelerated. The energy of the exact ground state and the exact ground state vector itself are used as a reference for the performance of the VQEs in this paper. Such reference will not be available for system sizes used in experiments showing a quantum advantage.

For the emulation of noisy quantum computation, we estimate $E(\theta)$ and $\mathcal{I}(\theta)$ by sampling $E^{(\mathbf{i})}$ and $\mathcal{I}^{(\mathbf{i})}$ according to the probability distribution p_i (Sec. II D) and computing the respective sample means. As the number of shots is increased, the sample means approach the true value of $E(\theta)$ and $\mathcal{I}(\theta)$, respectively. We call one stochastic execution of the ansatz circuit one *shot* of that circuit. In our implementation, every shot leads to one sample of $E(\theta)$ and $\mathcal{I}(\theta)$. For a finite number of shots, there is a statistical uncertainty in the sample means of $E(\theta)$ and $\mathcal{I}(\theta)$, which we characterize by 95% confidence intervals (CIs). These CIs are obtained by the default

implementation of the bootstrapping method [98] in SciPy [93]. At a fixed number of shots, reducing p_e or the number m of spacetime locations where an error may occur eventually leads to sets of shots where very few to no errors have occurred. This effect can lead to unjustly low CIs. (If no error occurred during any shot, the CI found by bootstrapping the data has size identically zero.) We exclude this possibility by increasing the number of shots per estimation of $E(\theta)$ and $\mathcal{I}(\theta)$ as a function of p_e or m in such a way that the expected number of shots where at least one error occurred equals 1024. Because of the significantly increased cost of noisy simulation, we only consider previously obtained, noiseless, locally optimal parameters for noisy simulation.

The Supplemental Material [19] also includes data and plots for systems and ansätze not reported in this paper, including simulations of the HAFM on the triangular lattice, other system sizes of the kagome lattice, simulations that use OPS (see Sec. II B), KVQE$_G$ simulating periodic patches, KVQE$_K$ simulating open patches, and runs where we use the infidelity as the cost function. Using the infidelity as a cost function is impractical on quantum computers (or even impossible if the ground state is not known) but may be used by classical computers to obtain further data on the expressibility of an ansatz. For all systems, data were stored in a human-readable format, and include the number of calls to the cost function by the BFGS routine, the wall-clock time of the classical emulation, the initial parameters, the parameters, energy, and infidelity of the local minima, as well as the noise type, error rate, and CIs in the case of noisy simulation.

III. RESULTS AND ANALYSIS

A. Noiseless results

1. Kagome

The explicit initial state and circuit used by KVQE$_G$ for the simulation of the HAFM on a 20-site patch of the kagome lattice is depicted in Fig. 2. For this patch, KVQE$_G$ uses 20 data qubits to represent the 20 sites of the patch, and an additional four qubits as swapping stations, used to realize kagome connectivity on grid architectures. The restriction of grid connectivity increases the circuit depth per cycle (assuming HEIS gates are native) from 4 to 6, and (in the thermodynamic limit) introduces one auxiliary qubit per three qubits. Because every gate gets its own parameter, the total number of parameters is $M = 30p$.

Noiseless results are displayed in Fig. 4. The relative energy error $\mathcal{E}$ decreases roughly exponentially in the range of all considered p. The number of function calls scales polynomially with p (data available as Supplemental Material [19]). KVQE$_G$ finds an energy lower than the energy of the first exited state for $p \geqslant 5$. There is no clear critical p after which $\mathcal{E}$ and/or $\mathcal{I}$ improve drastically. Nevertheless, $\mathcal{I}$ transitions to an improved exponential decay rate somewhere between $p = 3$ and $p = 5$, reaching a fidelity of >99.9% at $p \geqslant 16$. Under the hardware assumptions of Sec. I B, $p = 16$ amounts to $20/2 \times 3 = 30$ gates for the generation of the singlets, $30 \times 16 = 480$ HEIS gates, and $16 \times 16 = 256$ SWAP gates, giving a total of 766 gates and 480 parameters. The total depth

equals $3 + 6 \times 16 = 99$. To obtain the 10 local minima at $p = 16$, a total of 82 466 function calls were made.

The explicit initial state and circuit used by KVQE$_K$ for the simulation of an 18-site periodic patch are depicted in Fig. 1. For this patch, we assume a quantum computer with 18 qubits and all-to-all connectivity. Noiseless results are displayed in Fig. 5. Again, $\mathcal{E}$ decreases roughly exponentially in the range of all considered p, whereas the number of function calls scales polynomially with p [19]. KVQE$_K$ finds an energy lower than the energy of the first excited state for $p \geqslant 13$. After that same p, the performance of $\mathcal{I}$ improves significantly.

We note that for p such that the ansatz is not expressive enough to attain energies below the first excited state (in the present case, this is for $p < 13$), there need not be a relation between the energy of an ansatz state and its overlap with the ground state; in this regime, the ansatz may, for example, prepare states with large overlap with the first excited state and zero overlap with the ground state. It is only for p such that the ansatz can reach energies below the first excited state that a decrease in energy must be met with a decrease in infidelity. Data obtained using the infidelity as the cost function do show uniform exponential decay of $\mathcal{I}(p)$ [19].

In the absence of noise, KVQE$_K$ (using the energy as the cost function) reaches a fidelity of >99.9% at $p \geqslant 37$. Under the hardware assumptions of Sec. I B, $p = 37$ amounts to $18/2 \times 3 = 27$ gates for the generation of the singlets and $36 \times 37 = 1332$ HEIS gates, giving a total of 1359 gates and 1332 parameters. The total depth equals $3 + 4 \times 37 = 151$. To obtain the 10 local minima at $p = 37$, a total of 191 582 function calls were made.

One striking feature of these results is that the rate of exponential decay of the optimal energy and infidelity as a function of the number of cycles p is significantly smaller (in absolute value) than those same rates obtained by KVQE$_G$. This difference in performance is not because of the difference in the structure of the ansatz but because of the difference in the systems being simulated; the performance of KVQE$_G$ in simulating the periodic patch of Fig. 1 (in this case, the ansatz for KVQE$_G$ is obtained by extending the ansatz of Fig. 2 in such a way that its effective HEIS gates cover the periodic patch of Fig. 1) does not differ significantly from the performance of KVQE$_K$ on that same patch [19]. Vice versa, the performance of KVQE$_K$ (in this case, the ansatz for KVQE$_K$ is obtained by extending the ansatz of Fig. 1 in such a way that its effective HEIS gates cover the open patch of Fig. 2) in simulating the 20-site open patch of Fig. 2 (top left) does not differ significantly from the performance of KVQE$_G$ on that same patch [19].

At first, the 18-site periodic patch of Fig. 1 might seem to be easier to simulate than the 20-site open patch of Fig. 2 (top left) because of the difference in the number of sites. However, the graph defining the periodic patch has 36 edges, as opposed to 30 edges for the open patch. Therefore, the optimization problem corresponding to the periodic patch has more constraints, and can therefore naturally be more challenging. Additionally, boundary effects, which are absent in the periodic patch, may lead to a ground state that is easier to prepare. We leave an investigation of the latter issue for further work.
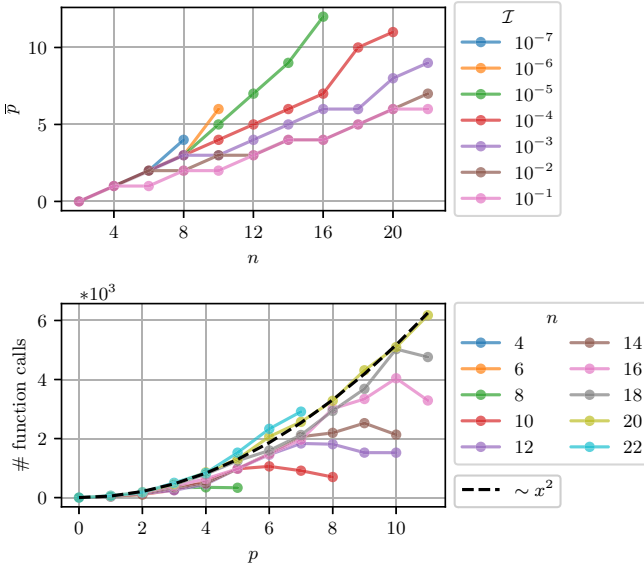
FIG. 7. Top: The number of cycles $\overline{p}$ required to prepare the approximate ground state using CVQE as a function of the number of sites n, for various target infidelities $\mathcal{I}$. Results were obtained as in Fig. 6, using the energy as the cost function during classical optimization of the variational parameters. Bottom: The number of function calls needed to reach a single local optimum, averaged over all 32 local optima that were reached per tuple (n, p). For some n, data were not obtained for all $p = 0, 1, \ldots, 11$, either because a negligible infidelity was reached ($n \leqslant 14$) or because of increasing computational cost ($n = 22$). A quadratic fit to the data for $n = 20$ (black, dashed) is added as a guide to the eye. We consider one function call as one computation of the energy $E(\theta)$ and one computation of the gradient $\nabla E(\theta)$ (Sec. II E).

2. Chain

The explicit initial state and circuit used by CVQE for the simulation of a 20-site periodic chain are depicted in Fig. 3. Any quantum processing unit with grid, kagome, or all-to-all connectivity naturally embeds subsets of qubits with (at least) the connectivity of a periodic chain. Because every gate gets its own parameter, the total number of parameters is $M = 20p$.

Noiseless results are displayed in Fig. 6. Both the relative energy error $\mathcal{E}$ and the infidelity $\mathcal{I}$ (Sec. II C) initially decrease exponentially as a function of p, reaching an energy that is below the first excited state for $p \geqslant 2$. Both functions show a sudden improvement after $p_{\mathrm{crit}} = 5$. From p_{crit} to $p_{\mathrm{crit}} + 1$, $\mathcal{E}$ drops by an order of magnitude and $\mathcal{I}$ drops by two orders of magnitude. For $p > 5$, both functions again decrease roughly exponentially with a rate that is greater in magnitude than before. At the same time, the number of function calls (as defined in Sec. II E), grows polynomially with p (Fig. 7, bottom). A fidelity of $>99.9\%$ is reached for $p \geqslant 8$. Assuming HEIS gates are native and that singlets can be created with a circuit of depth 3 (Sec. IV), $p = 8$ amounts to a circuit with $20/2 \times 3 = 30$ gates for the preparation of singlets and $20 \times 8 = 160$ HEIS gates, giving a total of 190 gates and 160 parameters. The depth of the circuit is $3 + 8 \times 2 = 19$. The optimization routine for finding the 32 local minima at $p = 8$ used 104 890 function calls. A fidelity of $>99.99\%$ is reached at $p = 11$ cycles, using a total of 197 685 function calls.

A plausible explanation of the sudden improvement of CVQE beyond $p_{\mathrm{crit}} = 5$ is in terms of the past light cone. The past light cone of a qubit q after a circuit C consists of all qubits q' for which there exists a past-directed path through C that connects q to q'. It is only when q' is in the past light cone of q that C can build up entanglement between q and q'. Also see Fig. 3. The ground state of the HAFM on the chain is known to possess long-range entanglement [99]. The sudden improvement of performance is a clear sign of a ground state with long-range entanglement. At p_{crit}, there is no qubit whose past light cone at the end of C covers the entire chain. After p_{crit}, the past light cone of *every* qubit at the end of C covers the entire chain.

For KVQE_G not all qubits' past light cones cover the entire system for the first time at an identical number of cycles [100]. For KVQE_K, the past light cone of every qubit covers the entire patch for the first time at $p = 3$. However, unlike CVQE, the states that can be reached by KVQE_K at this depth are still very far from the true ground state, and no sudden improvement of performance is observed.

3. Sistem-size scaling

The noiseless simulations carried out in this paper display an exponential decay of the infidelity [Eq. (6)] as a function of the number of cycles p. However, for the scalability of VQEs, it is essential that this rate of decay does not decrease too strongly (e.g., exponentially) with the system size. Alternatively, we may study the number of cycles $\overline{p}$ that is required to reach a given target infidelity $\mathcal{I}$. For $\mathrm{KVQE}_{G/K}$, obtaining a meaningful estimate for $\overline{p}$ is likely infeasible with classical emulation. This is because there are too few classically tractable and structurally similar patches for a systematic system-size scaling.

For CVQE, however, rings of size $n = 2, 6, \ldots, 22$ give us sufficient classically tractable and structurally similar patches. Results showing $\overline{p}$, together with the average number of function calls per local optimum, are displayed in Fig. 7.

The data in Fig. 7 indicate that $\overline{p}$ is upper bounded by a polynomial of low degree. Additionally, the data indicate that the number of function calls to the energy landscape is upper bounded by a function quadratic in p. Therefore, the total run time, excluding any sampling overhead, appears to scale as a polynomial of a polynomial, itself a polynomial.

B. Noisy results

Using the method outlined in Sec. II E, we compute the energy and infidelity obtained under the bit-flip and depolarizing noise models (Sec. II D) at all *locally* optimal noiseless parameters of Figs. 4–6 (corresponding to all translucent plot points in these plots), for $p_e \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$. Results for the depolarizing channel, showing the noisy values of the *optimal* noiseless data points, are displayed in Figs. 8–10.

We observed no significant difference between the effect of the depolarizing and the bit-flip channel. Therefore, only the results regarding the former are shown. The similarity of effect can be understood as follows. The effect of the depolarizing channel on a state on the Bloch sphere is isotropic, while that of the bit-flip channel is not [101]. For example, eigenstates of the Pauli-X operator are affected by the depolarizing
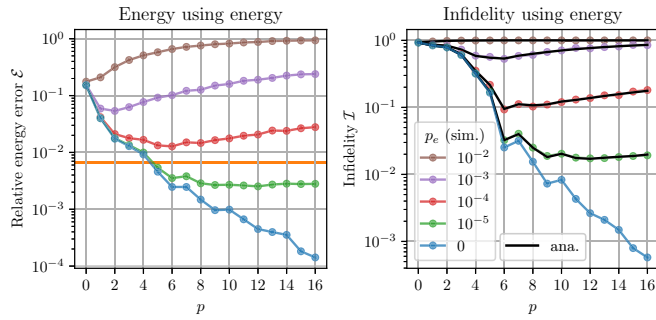
FIG. 8. Left: As Fig. 4, now including the effects of depolarizing noise with error probability p_e (legend shown in the right plot). At all p_e, data were obtained by computing the energy at $\theta^*(p)$, the noiseless optimal variational parameters as returned by the VQE. The blue line shows the noiseless energy $E(\theta^*)$ and is identical to the blue line in Fig. 4. Error bars marking the 95% confidence intervals are smaller than any of the plot points and are therefore omitted. Right: As the left plot, but now showing the infidelity between the true ground state and the states corresponding to the plot points of the left plot. Additionally, the analytical estimate for the infidelity [Eq. (9)], which is based on the noiseless infidelity $\mathcal{I}(\theta^*)$, is shown in black.

channel but are invariant under the bit-flip channel. Nevertheless, given that the input states of the channels have no preferred direction, the detrimental effect of the depolarizing and bit-flip channels are equal on average. (Technically, the twirl of a bit-flip channel with noise parameter p_e over SU(2) yields the depolarizing channel with the same noise parameter, which can be verified straightforwardly using the results of Ref. [102].) Therefore, if the states occurring during the ansatz state preparation or our VQEs are sufficiently isotropic, one expects roughly equal effects of the depolarizing and bit-flip channels.

The locally optimal angles that did not attain the lowest energy in the absence of noise may do so in the presence of noise. We found this effect to make no qualitative difference. More generally, Refs. [103,104] showed that VQEs can possess *noise adaptivity*; when a noisy ansatz circuit is run, the optimal parameters obtained in the presence of that noise (the noise-aware variational parameters) may attain a lower energy than the optimal parameters obtained in the absence of that noise (the noise-unaware parameters). We have only
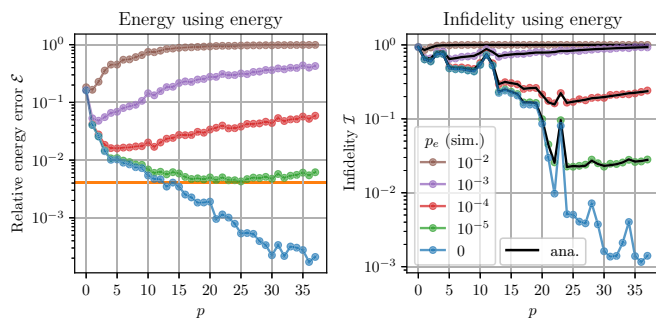

FIG. 9. As Fig. 5, now including the effects of depolarizing noise with error probability p_e. Definitions for the noisy data are as in Fig. 8.
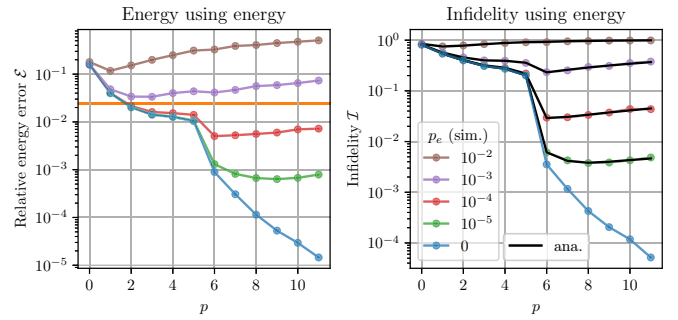

FIG. 10. As Fig. 6, now including the effects of depolarizing noise with error probability p_e. Definitions for the noisy data are as in Fig. 8.

optimized the energy landscape in the absence of noise, but a constrained form of noise adaptivity my still arise due to the freedom to choose the best noiseless locally optimal parameters in the noisy case. We indeed observed (with numerical significance, as set by a clear separation of the respective CIs) that the locally optimal angles that did not attain the lowest energy in the absence of noise do incidentally obtain the lowest energy in the presence of noise. However, the effect is insignificant on the scale of the fluctuations of the lines in Figs. 8–10. The latter variations arise because, per p, the VQEs did not always find the global minimum in the noiseless case. In that case, by chance, a lower minimum may be found at p compared to $p + 1$, even though the globally minimum energy at $p + 1$ is upper bounded by the minimum energy at p. Because the overall trend of the energy as a function of p is downward, this must lead to fluctuations in the energy as a function of p.

The noisy data for KVQE_G (periodic KVQE_K) on 20 sites (18 sites) suggest that, already for systems too small to obtain a quantum advantage, error rates between 10^{-4} and 10^{-5} are necessary to achieve an energy that is lower than the first excited state. This makes it highly unlikely that these energies will be obtained on NISQ hardware without error mitigation. Furthermore, the data show no observable discrepancy between the infidelity obtained by numerical simulation and the estimate of Eq. (9). Under the reasonable assumption that a similar agreement continues to hold for larger systems, we may use Eq. (9) to obtain estimates for the required error rates for systems too large for classical emulation. Thus, even if a VQE is able to prepare a ground state perfectly, $\mathcal{I} \approx 1 - (1 - p_e)^{nd}$ for depolarizing and bit-flip noise models. Inverting this relation, we have

$$p_e \approx \frac{\mathcal{I}}{nd}$$

for $p_e \ll 1$. If an infidelity of 10^{-3} is demanded for a system of a hundred qubits (as in Refs. [72,105]), and assuming the ground state can be prepared perfectly with a depth of $d = n = 100$, already an error rate of $p_e \approx 10^{-7}$ is required. This is an error rate several orders of magnitude lower than what can currently be achieved and also several orders of magnitude below known error thresholds for fault-tolerant quantum computation [64,65].
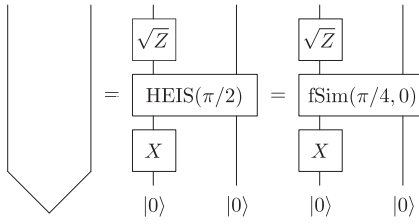
FIG. 11. Circuits for preparing the singlet state up to a global phase (left), using gates native to quantum dots (middle) and Google's superconducting hardware (right).

The data for the performance of CVQE on 20 sites under noise (Fig. 10) indicate that an energy below the first excited state may be reached on quantum computers with linear connectivity and error rates between 10^{-3} and 10^{-4}. Current rates fall between 10^{-2} and 10^{-3} [106,107] [108]. Thus, if technological improvements are able to bring down the error rates by one order of magnitude, proof-of-principle implementations of CVQE with on the order of 20 qubits can already be carried out. Feasibility is further improved by error mitigation.

IV. QUANTUM HARDWARE IMPLEMENTATION

The HEIS gate is directly native to quantum dot architectures. This also allows native implementation of the SWAP gate on these devices since $\mathrm{SWAP} = i\,\mathrm{HEIS}(\pi)$. Assuming also X and $\sqrt{Z}$ gates are native, singlets can be created with a circuit of depth three (see Fig. 11). Thus, on these devices, essentially no compilation is needed for $\mathrm{KVQE}_{G/K}$ and CVQE.

Similarly, on Google Quantum AI's superconducting hardware, compilation of $\mathrm{KVQE}_{G/K}$ and CVQE's gates induces no overhead in the number of two-qubit gates. This is because Google's native parameterized two-qubit gate, the fermionic simulation or fSim gate [36], is related to the HEIS gate by

$$\mathrm{HEIS}(\alpha) = \mathrm{RZ}_0(\alpha/2)\,\mathrm{RZ}_1(\alpha/2)\,\mathrm{fSim}(\alpha/2, \alpha). \quad (10)$$

Here, $\mathrm{RZ}_0(\theta) = \mathrm{RZ}(\theta) \otimes \mathbb{1}$ and $\mathrm{RZ}_1(\theta) = \mathbb{1} \otimes \mathrm{RZ}(\theta)$, with $\mathrm{RZ}(\theta) = e^{-i\theta Z/2}$, and Z the Pauli-Z operator. See Appendix B for details on the resolution of a technical subtlety regarding Eq. (10). The SWAP gate is related to the fSim gate by

$$\mathrm{SWAP} = \sqrt{Z_0}\sqrt{Z_1}\mathrm{fSim}(\pi/2, \pi).$$

Hence, a SWAP gate can be implemented by using one layer of RZ rotations and a single fSim gate. As for the quantum dot architectures, singlets can be created with a circuit of depth three (Fig. 11).

By adding the angles of subsequent RZ rotations, ℓ layers of HEIS and/or SWAP gates can be implemented by at most $\ell + 1$ layers of single-qubit RZ rotations and ℓ layers of fSim gates. Depending on the specific circuit, further reductions may be possible by using that $\mathrm{RZ}_0(\beta)\mathrm{RZ}_1(\beta)$ commutes with $\mathrm{fSim}(\theta, \phi)$ and addition of RZ rotation angles. For an example, see Fig. 12.

Current hardware does not yet simultaneously have grid connectivity (for more than four qubits) and the ability to essentially natively implement the exchange interaction for all parameter values. Quantum dot architectures can natively
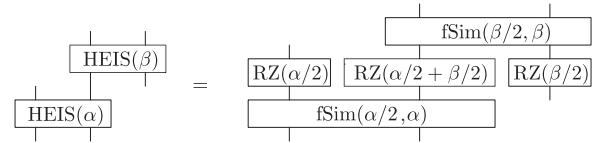

FIG. 12. Example of the compilation of HEIS gates into fSim gates and single-qubit RZ rotations for $0 \leqslant \alpha/2 \leqslant \pi$.

implement the exchange interaction, but are not yet available with grid connectivity. However, this connectivity may become available in the future, as detailed proposals already exist [32,33,35,109]. Google AI Quantum can implement the exchange interaction essentially natively for all parameter values and can demonstrate grid connectivity [16], but is not yet able to combine these two features in a single processor. This has, however, been expressed as a future goal. (See Sec. C of the Supplemental Material of Ref. [16].) They are already able to implement any two-qubit gate (so including the HEIS gate for all parameter values) using three native two-qubit gates [110]. Current hardware is already capable of efficiently performing CVQE for small problem sizes of open [36] or closed chains [35]. In such experiments, the observation of a critical circuit depth could form an early goal and would indicate the ability to generate and find ground states with systemwide entanglement.

V. DISCUSSION AND OUTLOOK

In this paper, we have introduced and studied VQEs for the HAFM on the kagome lattice ($\mathrm{KVQE}_{G/K}$) as exceptionally hardware-native algorithms with the potential of showing a useful quantum advantage on NISQ devices. In the noiseless case, the energies found by these VQEs decay exponentially with circuit depth and the VQEs appear to be scalable, as indicated by the system-size scaling carried out for CVQE. Furthermore, we showed that our VQEs naturally perform their quantum computations in a DFS that protects against collective longitudinal and transversal noise. However, other common noise types put extraordinary demands on the allowed error rates if an energy is to be reached that is below the first excited state or if the ground state is to be prepared with high fidelity. The following factors alleviate these demands.

(1) *Performance target*. The requirement of reaching an energy below the first excited state, or an infidelity of 10^{-3}, may be needlessly demanding. First, arguably a useful quantum advantage can be claimed once a VQE for a system of on the order of a hundred sites finds variational energies lower than those that can be found by state-of-the-art classical variational techniques [2,111,112]. Second, there are indications that performance targets on physically relevant observables other than the energy, such as spin-spin or dimer-dimer correlation functions, may be much more benign to VQEs [87]. Finally, with minimal adjustments to the classical optimization routine, the VQEs in this paper can be turned into variational quantum thermalizer algorithms [113,114]. In this setting, a natural performance target for a useful quantum advantage would be to prepare a thermal state with a target temperature inaccessible to classical numerical methods. For systems with on the order of a hundred qubits, the typical
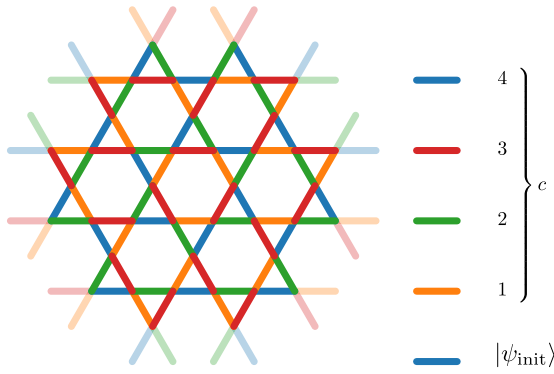
FIG. 13. One unit cell of a four coloring of the kagome lattice (opaque edges), derived from the 36-site VBC (blue lines) of Refs. [6–10]. To four color the entire lattice using this unit cell, a hexagonal tiling of these unit cells needs to be made. This four coloring directly translates to an ansatz for KVQE_K. On quantum computers with grid connectivity, we propose to use the same initial state, but define the cycle c similarly to the cycle defined for KVQE_G defined in Fig. 2.

energy associated with such a target temperature may be well above that of the first excited state.

(2) *Noise robustness.* We have computed the energy and infidelity of the output state of a noisy quantum computation at the previously found noiseless locally optimal parameters. This leaves little room for noise robustness. If parameters are optimized in the presence of noise, as on real quantum devices, lower energies may be found due to the noise adaptivity of the parameters [103,104,115].

(3) *Error mitigation.* By combining multiple runs of multiple quantum circuits with classical pre- and postprocessing, the error in estimates of expectation values of observables [such as the error in $E(\theta)$] can be reduced [66–70], and can in principle even be made arbitrarily small [71], even in the absence of quantum error correction. Such error-mitigation techniques typically lead to a sampling overhead that increases exponentially with the number of qubits, circuit depth, and error rate. However, the base of the exponent may be brought close to unity and sampling is highly (quantum) parallelizable [72]. Various error-mitigation techniques that exist specifically for VQEs may be used in conjunction [73,74].

Whether a combination of these factors allows for a useful quantum advantage on pre-error-corrected devices can likely only be answered by the quantum computing community as a whole. Technologically, there is a need for the further reduction of physical error rates. Furthermore, for the implementation of error mitigation and additional techniques such as circuit knitting [116–118], the tight and highly parallelized integration of quantum and classical resources, also known as quantum-centric supercomputing [72], is required. Additionally, further theoretical advances in the performance characterization and performance targets of classical quantum algorithms is needed to be able to claim a useful quantum advantage on near-term quantum devices.

In Refs. [6–10], the proposed ground state of the kagome HAFM is a 36-site VBC. We propose to use this VBC as the initial VQE state on quantum computers with at least that same number of data qubits (Fig. 13). For large patches, this

has the potential to answer whether this VBC is the ground state with very low circuit depths.

Note added. Recently, an independent but similar work appeared by Bosse and Montanaro [87]. At a high level, Bosse and Montanaro's work shares the same motivation, methods, and results as the current paper. Differences include the exact mapping from the kagome lattice to a grid, the numerical implementation, and the simulated patches of the kagome lattice. In addition, Bosse and Montanaro report data on the extraction of observables and an investigation of the barren-plateau problem [38], where we include an ansatz for quantum computers with kagome connectivity and focus on the experimental realization by offering explicit compilation into native gates and a study of the effects of noise.

APPENDIX A: SYMMETRY, THE SPIN GAP, AND DECOHERENCE-FREE SUBSPACES

In this Appendix, we give some background on symmetry and the total spin operators, and add detail to the discussion in Sec. II B 3.

The total spin operator is defined by $\mathbf{S}^2 \equiv \mathbf{S}^{(\mathrm{tot})} \cdot \mathbf{S}^{(\mathrm{tot})}$, with $\mathbf{S}^{(\mathrm{tot})} = \sum_{i=1}^{n} \mathbf{S}^{(i)}$ [cf. Eq. (3)], and is related to the total spin quantum number S by $\mathbf{S}^2|\psi\rangle = S(S+1)|\psi\rangle$ for eigenstates $|\psi\rangle$ of $\mathbf{S}^2$. The total magnetization operator in the i direction is defined as $\mathbf{S}_i^{(\mathrm{tot})} = \sum_{j=1}^{n}[\mathbf{S}^{(j)}]_i$ and is related to the magnetization quantum number S_z by $\mathbf{S}_z^{(\mathrm{tot})}|\psi\rangle = S_z|\psi\rangle$ for eigenstates $|\psi\rangle$ of $\mathbf{S}_z$. The operators $\mathbf{S}^2$ and $\mathbf{S}_z^{(\mathrm{tot})}$ commute and are hence simultaneously diagonalizable. Acting with the ladder operator $\mathbf{S}_{\pm}^{(\mathrm{tot})} = \mathbf{S}_x^{(\mathrm{tot})} \pm \mathrm{i}\,\mathbf{S}_y^{(\mathrm{tot})}$ on a state with quantum number S_z raises (lowers) the S_z quantum number of that state with unity, given that the new value of S_z lays between (or is equal to) $-S$ and S. It annihilates that state otherwise. The Hamiltonian H, the ansatz circuits C, and initial states $|\psi_{\mathrm{init}}\rangle\langle\psi_{\mathrm{init}}|$ in this paper commute with $\mathbf{S}_x^{(\mathrm{tot})}$, $\mathbf{S}_y^{(\mathrm{tot})}$, and $\mathbf{S}_z^{(\mathrm{tot})}$ (and hence with $\mathbf{S}^2$ and $\mathbf{S}_{\pm}^{(\mathrm{tot})}$), and therefore H, C, and $|\psi_{\mathrm{init}}\rangle\langle\psi_{\mathrm{init}}|$ possess a SU(2) symmetry. This means, among other things, that the ansatz circuits C conserve the quantum numbers S and S_z. See, e.g., Ref. [119] for more background on total spin operators and symmetry.

For the HAFM on the chain, there is formal proof that the ground state lays within the $S = 0$ sector, for example, via the Bethe ansatz [23] or through the more general result by Lieb and Mattis [120]. For the (tripartite) kagome lattice, such formal proof is unknown. There is, however, substantial evidence that $S = S_z = 0$ for the ground state of the HAFM on (patches of) the kagome lattice [2,5,20,80–82].

We now prove the invariance of the variational energy $E(\theta)$ under the choice of two-qubit triplet state. The two-qubit

triplet states are

$$|t_1\rangle = |00\rangle, \quad |t_0\rangle = \mathbf{S}_-^{(1,2)}|t_1\rangle/\sqrt{2} = (|01\rangle + |10\rangle)/\sqrt{2},$$

$$|t_{-1}\rangle = \mathbf{S}_-^{(1,2)}\mathbf{S}_-^{(1,2)}|t_1\rangle/2 = |11\rangle,$$

with $\mathbf{S}_\mp^{(1,2)}$ the lowering (raising) operator on two qubits. Without loss of generality, change the first singlet of the dimer covering $|\psi_{\text{init}}\rangle$ into any of the triplet states, and denote the remainder of the initial state, which is a $(S = 0, S_z = 0)$ dimer covering on $n - 2$ qubits, by $|\text{dim}\rangle$. Thus, we obtain the initial state $2^{-m/2}[\mathbf{S}_-^{(1,2)}]^m|t_1\rangle|\text{dim}\rangle$, with $m \in \{0, 1, 2\}$ specifying the specific triplet state. Note $\mathbf{S}_+^{(1,2)\dagger} = \mathbf{S}_\pm^{(1,2)}$ and $\mathbf{S}_\mp^{(\text{tot})\dagger} = \mathbf{S}_\pm^{(\text{tot})}$. Then,

$$
\begin{aligned}
E(\theta) &= 2^{-m}\langle t_1|[\mathbf{S}_+^{(1,2)}]^m\langle\text{dim}|C^\dagger HC[\mathbf{S}_-^{(1,2)}]^m|t_1\rangle|\text{dim}\rangle \\
&= 2^{-m}\langle t_1|\langle\text{dim}|[\mathbf{S}_+^{(\text{tot})}]^m C^\dagger HC[\mathbf{S}_-^{(\text{tot})}]^m|t_1\rangle|\text{dim}\rangle \\
&= \langle t_1|\langle\text{dim}|C^\dagger HC|t_1\rangle|\text{dim}\rangle,
\end{aligned}
$$

which does not depend on $m \in \{0, 1, 2\}$.

Under the influence of noise, the ansatz circuit $C(\theta)$ will generally not produce a pure state $|\theta\rangle$, but rather a mixed state ρ_θ. If the noise breaks the SU(2) symmetry of $C(\theta)$, $\text{KVQE}_{G/K}$ and CVQE (running with $S = 1$ initial states) may abuse this noise to put amplitude on $S = 0$ states, thus obtaining unjustly low variational energies. This may be dealt with by symmetry verification [73]. Alternatively, a penalty term $A_{P_1}(\mathbf{S}^2 - 2\mathbb{1})^2$, with $A_{P_1} > 0$, may be added to H [11,84,85]. In the Pauli basis, this penalty term has $O(n^4)$ terms, and may hence be costly to compute in every iteration of $\text{KVQE}_{G/K}$ and CVQE.

We now give a concise introduction to DFSs [56–59] and a previous work on DFSs in the context of VQEs. To introduce DFSs [121], consider a general system-bath interaction $\tilde{H} = H_S + H_B + H_{SB}$, where $H_{S(B)}$ acts nontrivially on the system (bath) only. Without loss of generally, $H_{SB} = \sum_\alpha \mathcal{S}_\alpha \mathcal{B}_\alpha$, where $\mathcal{S}_\alpha(\mathcal{B}_\alpha)$ acts nontrivially on the system (bath) only. A DFS is a system subspace containing those states $|\psi\rangle_S$ for which

$$\mathcal{S}_\alpha|\psi\rangle_S = \lambda_\alpha|\psi\rangle_S \tag{A1}$$

for all α, and with λ_α independent of $|\psi\rangle_S$. States in a DFS are not affected by the system-bath coupling at all; for an initial product state $|\Psi\rangle = |\psi\rangle_S|\varphi\rangle_B$, it follows $\text{tr}_B(e^{-it\tilde{H}}|\Psi\rangle\langle\Psi|e^{it\tilde{H}}) = e^{-itH_S}|\psi\rangle_S\langle\psi|_S e^{itH_S}$. If, furthermore, H_S does not take $|\psi\rangle_S$ outside the DFS, states in a DFS remain in the DFS and evolve unitarily despite the nontrivial system-bath coupling. This is generalized straightforwardly to cases where $\tilde{H}$ is time dependent (with known time dependence of H_S).

In Ref. [63], a programmable analog quantum simulator is used to simulate the lattice Schwinger model, where the states occurring during their ansatz state preparation remain an approximate eigenstate of $\mathbf{S}_z^{(\text{tot})}$ and a charge conjugation and spatial reflection (CP) operation after modifications to their variational ansatz.

APPENDIX B: THE fSim GATE

The native, parametrized two-body gate of Google AI Quantum, called the fermionic simulation gate, reads

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

and has been demonstrated experimentally for all $\theta \in [0, \pi/2]$ and $\phi \in [-\pi, \pi]$ [36]. Equation (10) shows how the HEIS gate is implemented using the fSim gate. This relation shows that during the unconstrained optimization of the $\text{KVQE}_{G/K}$ or CVQE variational parameters, fSim gates may occur with a parameter θ that falls outside the range for which implementation has been demonstrated. Nevertheless, fSim(θ, ϕ) gates for general $\theta \in [-\pi, \pi]$ can be implemented by a single fSim gate after remapping θ and inserting appropriate single-qubit Z gates, as is shown by the identity

$$
\text{fSim}(\theta, \phi) = \begin{cases} Z_0 Z_1 \text{fSim}(\theta - \pi, \phi) & : -\pi \leqslant \theta < -\pi/2 \\ Z_0 \text{fSim}(-\theta, \phi) Z_0 & : -\pi/2 \leqslant \theta < 0 \\ \text{fSim}(\theta, \phi) & : 0 \leqslant \theta < \pi/2 \\ Z_0 \text{fSim}(-\theta + \pi, \phi) Z_1 & : \pi/2 \leqslant \theta < \pi, \end{cases}
$$

with $Z_0 = Z \otimes \mathbb{1}$ and $Z_1 = \mathbb{1} \otimes Z$.

[1] H.-C. Jiang, Z.-Y. Weng, and D. N. Sheng, Density Matrix Renormalization Group Numerical Study of the Kagome Antiferromagnet, Phys. Rev. Lett. **101**, 117203 (2008).

[2] S. Yan, D. A. Huse, and S. R. White, Spin-liquid ground state of the $S = 1/2$ kagome Heisenberg antiferromagnet, Science **332**, 1173 (2011).

[3] P. W. Anderson, Resonating valence bonds: A new kind of insulator? Mater. Res. Bull. **8**, 153 (1973).

[4] M. R. Norman, Colloquium: Herbertsmithite and the search for the quantum spin liquid, Rev. Mod. Phys. **88**, 041002 (2016).

[5] A. M. Läuchli, J. Sudan, and R. Moessner, $S = \frac{1}{2}$ kagome Heisenberg antiferromagnet revisited, Phys. Rev. B **100**, 155142 (2019).

[6] J. Marston and C. Zeng, Spin-Peierls and spin-liquid phases of kagomé quantum antiferromagnets, J. Appl. Phys. **69**, 5962 (1991).

[7] P. Nikolic and T. Senthil, Physics of low-energy singlet states of the kagome lattice quantum Heisenberg antiferromagnet, Phys. Rev. B **68**, 214415 (2003).

[8] R. R. P. Singh and D. A. Huse, Ground state of the spin-1/2 kagome-lattice Heisenberg antiferromagnet, Phys. Rev. B **76**, 180407(R) (2007).

[9] R. R. P. Singh and D. A. Huse, Triplet and singlet excitations in the valence bond crystal phase of the kagome lattice Heisenberg model, Phys. Rev. B **77**, 144415 (2008).

[10] G. Evenbly and G. Vidal, Frustrated Antiferromagnets with Entanglement Renormalization: Ground State of the spin-$\frac{1}{2}$

Heisenberg Model on a Kagome Lattice, Phys. Rev. Lett. **104**, 187203 (2010).

[11] 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).

[12] 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).

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

[14] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Scalable Quantum Simulation of Molecular Energies, Phys. Rev. X **6**, 031007 (2016).

[15] J.-M. Reiner, F. Wilhelm-Mauch, G. Schön, and M. Marthaler, Finding the ground state of the Hubbard model by variational methods on a quantum computer with gate errors, Quantum Sci. Technol. **4**, 035005 (2019).

[16] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, David A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[17] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[18] D. S. França and R. García-Patrón, Limitations of optimization algorithms on noisy quantum devices, Nat. Phys. **17**, 1221 (2021).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.106.214429 for all data reported in this paper, the code that generated it, and additional data.

[20] A. M. Läuchli, J. Sudan, and E. S. Sørensen, Ground-state energy and spin gap of spin-$\frac{1}{2}$ kagomé-Heisenberg antiferromagnetic clusters: Large-scale exact diagonalization results, Phys. Rev. B **83**, 212401 (2011).

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

[22] H. Bethe, Zur theorie der metalle, Z. Phys. **71**, 205 (1931).

[23] F. Franchini, *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems* (Springer International Publishing, Cham, Switzerland, 2017)

[24] J. Caux, Correlation functions of integrable models: A description of the ABACUS algorithm, J. Math. Phys. **50**, 095214 (2009).

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

[26] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).

[27] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, Strategies for solving the Fermi-Hubbard model on near-term quantum computers, Phys. Rev. B **102**, 235122 (2020).

[28] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, A. Bengtsson, S. Boixo, M. Broughton, B. B. Buckley *et al.*, Observation of separated dynamics of charge and spin in the fermi-hubbard model, arXiv:2010.07965.

[29] M. Nielsen, The fermionic canonical commutation relations and the Jordan-Wigner transform, School of Physical Sciences, The University of Queensland **59** (2005).

[30] S. B. Bravyi and A. Yu Kitaev, Fermionic quantum computation, Ann. Phys. **298**, 210 (2002).

[31] Z. Jiang, A. Kalev, W. Mruczkiewicz, and H. Neven, Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning, Quantum **4**, 276 (2020).

[32] C. J. van Diepen, T.-K. Hsiao, U. Mukhopadhyay, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen, Quantum Simulation of Antiferromagnetic Heisenberg Chain with Gate-Defined Quantum Dots, Phys. Rev. X **11**, 041025 (2021).

[33] P. Barthelemy and L. M. K. Vandersypen, Quantum dot systems: A versatile platform for quantum simulations, Annalen der Physik **525**, 808 (2013).

[34] D. Loss and D. P. DiVincenzo, Quantum computation with quantum dots, Phys. Rev. A **57**, 120 (1998).

[35] N. W. Hendrickx, W. I. L. Lawrie, M. Russ, F. van Riggelen, S. L. de Snoo, R. N. Schouten, A. Sammak, G. Scappucci, and M. Veldhorst, A four-qubit germanium quantum processor, Nature (London) **591**, 580 (2021).

[36] B. Foxen, C. Neill, A. Dunsworth, P. Roushan, B. Chiaro, A. Megrant, J. Kelly, Z. Chen, K. Satzinger, R. Barends *et al.*, Demonstrating a Continuous Set of Two-Qubit Gates for Near-Term Quantum Algorithms, Phys. Rev. Lett. **125**, 120504 (2020).

[37] 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).

[38] 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).

[39] 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).

[40] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[41] C. D. Hill, E. Peretz, S. J. Hile, M. G. House, M. Fuechsle, S. Rogge, M. Y. Simmons, and L. C. L. Hollenberg, A surface code quantum computer in silicon, Sci. Adv. **1**, e1500707 (2015).

[42] C. K. Andersen, A. Remm, S. Lazar, S. Krinner, N. Lacroix, G. J. Norris, M. Gabureac, C. Eichler, and A. Wallraff, Repeated quantum error detection in a surface code, Nat. Phys. **16**, 875 (2020).

[43] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, and L. DiCarlo, Scalable Quantum Circuit and Control for a Superconducting Surface Code, Phys. Rev. Appl. **8**, 034021 (2017).

[44] A. Hutter, J. R. Wootton, and D. Loss, Parafermions in a Kagome Lattice of Qubits for Topological Quantum Computation, Phys. Rev. X **5**, 041040 (2015).

[45] S. Schmidt and J. Koch, Circuit QED lattices: towards quantum simulation with superconducting circuits, Ann. Phys. **525**, 395 (2013).

[46] M. D. Kim and J. Kim, Scalable quantum computing model in the circuit-QED lattice with circulator function, Quant. Info. Proc. **16**, 192 (2017).

[47] Mun Dae Kim, Quantum simulation scheme of two-dimensional xy-model Hamiltonian with controllable coupling, Quant. Info. Proc. **18**, 1 (2019).

[48] D. L. Underwood, W. E. Shanks, J. Koch, and A. A. Houck, Low-disorder microwave cavity lattices for quantum simulation with photons, Phys. Rev. A **86**, 023837 (2012).

[49] A. J. Kollár, M. Fitzpatrick, and A. A. Houck, Hyperbolic lattices in circuit quantum electrodynamics, Nature (London) **571**, 45 (2019).

[50] J. I. Cirac and P. Zoller, Quantum Computations with Cold Trapped Ions, Phys. Rev. Lett. **74**, 4091 (1995).

[51] K. Mølmer and A. Sørensen, Multiparticle Entanglement of Hot Trapped Ions, Phys. Rev. Lett. **82**, 1835 (1999).

[52] D. Kielpinski, C. Monroe, and D. J. Wineland, Architecture for a large-scale ion-trap quantum computer, Nature (London) **417**, 709 (2002).

[53] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, 14-Qubit Entanglement: Creation and Coherence, Phys. Rev. Lett. **106**, 130506 (2011).

[54] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Toward the first quantum simulation with quantum speedup, Proc. Natl. Acad. Sci. **115**, 9456 (2018).

[55] For the sake of comparison of the gate count, we assume the same native gates as before. To the best of our knowledge, quantum computers that support both the essentially native implementation of the exchange interaction and kagome connectivity do not exist yet, so on current devices with all-to-all connectivity there will be some transpilation overhead.

[56] G. M. Palma, K.-A. Suominen, and A. K. Ekert, Quantum computers and dissipation, Proc. R. Soc. London A **452**, 567 (1996).

[57] P. Zanardi and M. Rasetti, Error avoiding quantum codes, Mod. Phys. Lett. B **11**, 1085 (1997).

[58] L.-M. Duan and G.-C. Guo, Reducing decoherence in quantum-computer memory with all quantum bits coupling to the same environment, Phys. Rev. A **57**, 737 (1998).

[59] D. A. Lidar, I. L. Chuang, and K. B. Whaley, Decoherence-Free Subspaces for Quantum Computation, Phys. Rev. Lett. **81**, 2594 (1998).

[60] J. Kattemölle and J. van Wezel, Dynamical fidelity susceptibility of decoherence-free subspaces, Phys. Rev. A **99**, 062340 (2019).

[61] R. H. Dicke, Coherence in spontaneous radiation processes, Phys. Rev. **93**, 99 (1954).

[62] P. Kirton, M. M. Roses, J. Keeling, and E. G. D. Torre, Introduction to the Dicke model: From equilibrium to nonequilibrium, and vice versa, Adv. Quantum Technol. **2**, 1800043 (2019).

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

[64] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Surface code quantum computing with error rates over 1%, Phys. Rev. A **83**, 020302(R) (2011).

[65] K. Dalton, C. K. Long, Y. S. Yordanov, C. G. Smith, C. H. W. Barnes, N. Mertig, and D. R. M. Arvidsson-Shukur, Variational quantum chemistry requires gate-error probabilities below the fault-tolerance threshold, arXiv:2211.04505.

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

[67] B. Koczor, Exponential Error Suppression for Near-Term Quantum Devices, Phys. Rev. X **11**, 031057 (2021).

[68] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, Virtual Distillation for Quantum Error Mitigation, Phys. Rev. X **11**, 041036 (2021).

[69] C. Song, J. Cui, H. Wang, J. Hao, H. Feng, and Y. Li, Quantum computation with universal error mitigation on a superconducting quantum processor, Sci. Adv. **5**, eaaw5686 (2019).

[70] D. Bultrini, M. H. Gordon, P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Unifying and benchmarking state-of-the-art quantum error mitigation techniques, (2021), arXiv:2107.13470 [quant-ph].

[71] E. van den Berg, Z. K. Minev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors, arXiv:2201.09866.

[72] S. Bravyi, O. Dial, J. M. Gambetta, D. Gil, and Z. Nazario, The future of quantum computing with superconducting qubits, J. Appl. Phys. **132**, 160902 (2022).

[73] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).

[74] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo, Experimental error mitigation via symmetry verification in a variational quantum eigensolver, Phys. Rev. A **100**, 010302(R) (2019).

[75] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, SIAM J. Comput. **35**, 1070 (2006).

[76] S. Piddock and A. Montanaro, The complexity of antiferromagnetic interactions and 2D lattices, Quantum Inf. Comput. **17**, 636 (2017).

[77] D. Aharonov and T. Naveh, Quantum NP—a survey, arXiv:quant-ph/0210077.

[78] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry Introduction to Advanced Electronic Structure Theory* (Dover publications Inc., Mineola, New York, 2018).

[79] H.-Y. Huang, R. Kueng, and J. Preskill, efficient Estimation of Pauli Observables by Derandomization, Phys. Rev. Lett. **127**, 030503 (2021).

[80] P. Lechminant, B. Bernu, C. Lhuillier, L. Pierre, and P. Sindzingre, Order versus disorder in the quantum Heisenberg antiferromagnet on the kagomé lattice using exact spectra analysis, Phys. Rev. B **56**, 2521 (1997).

[81] P. Sindzingre and C. Lhuillier, Low-energy excitations of the kagomé antiferromagnet and the spin-gap issue, Europhys. Lett. **88**, 27009 (2009).

[82] H. Nakano and T. Sakai, Numerical-diagonalization study of spin gap issue of the kagome lattice Heisenberg antiferromagnet, J. Phys. Soc. Jpn. **80**, 053704 (2011).

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

[84] I. G. Ryabinkin, S. N. Genin, and A. F. Izmaylov, Constrained variational quantum eigensolver: Quantum computer search engine in the Fock space, J. Chem. Theory Comput. **15**, 249 (2019).

[85] K. Kuroiwa and Y. O. Nakagawa, Penalty methods for a variational quantum eigensolver, Phys. Rev. Res. **3**, 013197 (2021).

[86] M. J. S. Beach, R. G. Melko, T. Grover, and T. H. Hsieh, Making trotters sprint: A variational imaginary time ansatz for quantum many-body systems, Phys. Rev. B **100**, 094434 (2019).

[87] J. L. Bosse and A. Montanaro, Probing ground state properties of the kagome antiferromagnetic Heisenberg model using the variational quantum eigensolver, arXiv:2108.08086.

[88] Python Software Foundation, Python language reference, version 2.8, https://docs.python.org/3.8/ (2019).

[89] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith *et al.*, Array programming with NumPy, Nature (London) **585**, 357 (2020).

[90] R. Okuta, Y. Unno, D. Nishino, S. Hido, and C. Loomis, CuPy: A NumPy-Compatible Library for NVIDIA GPU Calculations, in *Proceedings of Workshop on Machine Learning Systems (LearningSys) in the Thirty-First Annual Conference on Neural Information Processing Systems* (NIPS, Long Beach, CA, 2017), http://learningsys.org/nips17/assets/papers/paper_16.pdf.

[91] M. A. Nielsen, *Neural Networks and Deep Learning* (Determination Press, 2015), Vol. 25.

[92] S. Tokui, K. Oono, S. Hido, and J. Clayton, Chainer: A next-generation open source framework for deep learning, in *Proceedings of Workshop on Machine Larning Systems (LearningSys) in the Twenty-Ninth Annual Conference on Neural Information Processing Systems*, Vol. 5 (NIPS, 2015), pp. 1, http://learningsys.org/papers/LearningSys_2015_paper_33.pdf.

[93] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Meth. **17**, 261 (2020).

[94] S. Wei, H. Li, and G. Long, A full quantum eigensolver for quantum chemistry simulations, Research **2020**, 1 (2020).

[95] Gui-Lu Long, General quantum interference principle and duality computer, Commun. Theor. Phys. **45**, 825 (2006).

[96] Andrew M. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, Quantum Inf. Comput. **12**, 901 (2012).

[97] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users' Guide* (Society for Industrial and Applied Mathematics, Philadelphia, 1998).

[98] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall/CRC, Boca Raton, FL, 1994).

[99] J. I. Latorre, E. Rico, and G. Vidal, Ground state entanglement in quantum spin chains, Quantum Inf. Comput. **4**, 48 (2004).

[100] Let us focus on KVQE$_G$, and lay out a coordinate system over the 24 qubits used in Fig. 2 (top right). We put the origin (0,0) at the bottom left qubit, the qubit directly above at (0,1), and the qubit directly to the right of the origin at (1,0). At $p = 2$, there is no qubit whose past light cone in C covers the entire system. At $p = 3$, there are qubits, such as the bottom right (4,0), bottom left (0,0), top left (0,4) and middle (2,2) qubits, whose past light cone in C covers the entire system. There are, however, still some qubits for which this is not the case, such as the qubits at (3,4), (4,3) and the top middle (2,0). At $p = 4$, the past light cone of the latter qubits covers the entire system, except for the qubit at (3,4). It is only after $p = 5$ cycles that its past light cone covers the entire system.

[101] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

[102] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, J. Opt. B: Quantum Semiclass. Opt. **7**, S347 (2005).

[103] J. Kattemölle and G. Burkard, Effects of correlated errors on the quantum approximate optimization algorithm, arXiv:2207.10622.

[104] E. Fontana, N. Fitzpatrick, D. M. Ramo, R. Duncan, and I. Rungger, Evaluating the noise resilience of variational quantum algorithms, Phys. Rev. A **104**, 022403 (2021).

[105] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter Error with Commutator Scaling, Phys. Rev. X **11**, 011020 (2021).

[106] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, and D. M. Lucas, High-Fidelity Quantum Logic Gates using Trapped-Ion Hyperfine Qubits, Phys. Rev. Lett. **117**, 060504 (2016).

[107] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[108] The gate infidelities typically reported in the literature do not correspond directly to the error rates as defined in Eqs. (7) and (8), but for the purpose of the order-of-magnitude estimates here these discrepancies are insignificant.

[109] R. Li, L. Petit, D. P. Franke, J. P. Dehollain, J. Helsen, M. Steudtner, N. K. Thomas, Z. R. Yoscovits, K. J. Singh, S. Wehner *et al.*, A crossbar network for silicon quantum dot qubits, Sci. Adv. **4**, eaar3960 (2018).

[110] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, Nat. Phys. **17**, 332 (2021).

[111] S. Depenbrock, I. P. McCulloch, and U. Schollwöck, Nature of the Spin-Liquid Ground State of the $s = 1/2$ Heisenberg Model on the Kagome Lattice, Phys. Rev. Lett. **109**, 067201 (2012).

[112] Y.-C. He, M. P. Zaletel, M. Oshikawa, and F. Pollmann, Signatures of Dirac Cones in a DMRG Study of the Kagome Heisenberg Model, Phys. Rev. X **7**, 031020 (2017).

[113] G. Verdon, J. Marks, S. Nanda, S. Leichenauer, and J. Hidary, Quantum Hamiltonian-Based Models and the Variational Quantum Thermalizer Algorithm (2019), arXiv:1910.02071.

[114] J. Foldager, A. Pesah, and L. K. Hansen, Noise-assisted variational quantum thermalization, Sci. Rep. **12**, 3862 (2022).

[115] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Phys. Rev. A **95**, 042308 (2017).

[116] S. Bravyi, G. Smith, and J. A. Smolin, Trading Classical and Quantum Computational Resources, Phys. Rev. X **6**, 021043 (2016).

[117] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, Simulating Large Quantum Circuits on a Small Quantum Computer, Phys. Rev. Lett. **125**, 150504 (2020).

[118] J. Sun, S. Endo, H. Lin, P. Hayden, V. Vedral, and X. Yuan, Perturbative Quantum Simulation, Phys. Rev. Lett. **129**, 120505 (2022).

[119] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics* (Cambridge University Press, Cambridge, 2020).

[120] E. Lieb and D. Mattis, Ordering energy levels of interacting spin systems, J. Math. Phys. **3**, 749 (1962).

[121] J. Kempe, D. Bacon, D. A. Lidar, and K. B. Whaley, Theory of decoherence-free fault-tolerant universal quantum computation, Phys. Rev. A **63**, 042307 (2001).