Quantum Computation of Finite-Temperature Static and Dynamical Properties of Spin Systems Using Quantum Imaginary Time Evolution

Shi-Ning Sun, Mario Motta, Ruslan N. Tazhigulov, Adrian T. K. Tan, Garnet Kin-Lic Chan, and Austin J. Minnich

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA
IBM Quantum, IBM Research Almaden, San Jose, CA 95120, USA
Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

(Dated: September 9, 2020)

Developing scalable quantum algorithms to study finite-temperature physics of quantum many-body systems has attracted considerable interest due to recent advancements in quantum hardware. However, such algorithms in their present form require resources that exceed the capabilities of current quantum computers except for a limited range of system sizes and observables. Here, we report calculations of finite-temperature properties including energies, static and dynamical correlation functions, and excitation spectra of spin Hamiltonians with up to four sites on five-qubit IBM Quantum devices. These calculations are performed using the quantum imaginary time evolution (QITE) algorithm and made possible by several algorithmic improvements, including a method to exploit symmetries that reduces the quantum resources required by QITE, circuit optimization procedures to reduce circuit depth, and error mitigation techniques to improve the quality of raw hardware data. Our work demonstrates that the ansatz-independent QITE algorithm is capable of computing diverse finite-temperature observables on near-term quantum devices.

I. INTRODUCTION

Quantum computers have long been considered as a potential tool to simulate quantum many-body systems [1–3]. While near-term quantum devices have made rapid progress in simulating ground-state properties [4–10] and dynamics [11–15], the study of finite-temperature physics on quantum computers is less understood and established [16]. Early works on digital quantum simulation of finite-temperature physical systems involved thermalizing the quantum simulator by coupling to a bath comprised of ancilla qubits [17–19] or sampling thermal states using quantum versions of the Metropolis algorithm [20, 21]. These schemes require prohibitively large numbers of qubits and deep circuits and are hence out of reach for near-term quantum devices.

More practical variational algorithms have been proposed in recent years, such as protocols to construct thermal-field double states [22, 23] and machine learning-based methods to construct Gibbs states [24–28]. However, the accuracies of these variational schemes depend on the quality of the ansatz. While other non-variational alternatives exist, they are subject to various assumptions. For example, the minimal effective Gibbs ansatz [29] algorithm generates a minimal ensemble of pure states but presumes the eigenstate thermalization hypothesis.

Recently, the quantum imaginary time evolution (QITE) algorithm was introduced [30]. Compared to variational-based algorithms of imaginary time evolution on quantum computers [31–33], QITE is ansatz-independent. The QITE algorithm approximates imaginary time evolution with unitary operators over a domain of qubits and is able to reach the ground states of systems with a few sites. QITE can also be used to calculate finite-temperature quantities, for instance by combining with sampling techniques such as the minimal entangled typical thermal states (METTS) algorithm [34, 35], together denoted as the quantum METTS (QMETTS) algorithm. However, the original work on QITE [30] focused on the general formalism, while reduction and optimization of quantum resources were not thoroughly investigated. Subsequent development of QITE [36–39] proposed several variations of the original algorithm, but the practical evaluation of finite-temperature properties on existing quantum devices remains largely unaddressed.

Here, we report QITE-based calculations of finite-temperature static and dynamical properties of one-dimensional spin systems with up to four sites on five-qubit IBM Quantum devices. The computed observables include finite-temperature energies, static and dynamical correlation functions, and excitation spectra. These calculations are made possible by several algorithmic improvements. First, we exploit symmetries in the spin Hamiltonians to reduce Pauli strings in the QITE unitaries, thus reducing the overall required quantum resources. Second, circuit optimization procedures including gate decomposition and circuit recompilation are used to further reduce circuit depth. Third, error mitigation techniques, namely post-selection, read-out error mitigation and phase-and-scale correction, are used to improve the quality of raw hardware data. Our work demonstrates that with efficient use of quantum resources and effective error mitigation strategies, the ansatz-independent QITE algorithm is capable of com-

* garnetc@caltech.edu
† aminnich@caltech.edu
II. THEORY

A. Quantum imaginary time evolution (QITE)

We begin by reviewing the QITE algorithm in the context of a general Trotterization scheme of the imaginary time propagator. Consider imaginary time evolution on \( N \) qubits under a Hamiltonian \( \hat{H} = \sum_{m=1}^{M} \hat{h}[m] \), where each \( \hat{h}[m] \) acts on a local set of qubits. Since the local terms \( \hat{h}[m] \) are not commutative, we need to Trotterize the imaginary time propagator \( e^{-\hat{H}t} \) by grouping local terms \( \hat{h}[m] \) into Trotter terms \( \hat{H}[l] \) such that each \( \hat{H}[l] \) is a sum of local terms \( \hat{h}[m] \) and \( \hat{H} = \sum_{l=1}^{L} \hat{H}[l] \). For example, for a two-local Hamiltonian where each local term \( \hat{h}[m] \) acts on qubits \( m-1 \) and \( m \), setting \( L = 2, \hat{H}[1] = \sum_{m=1}^{[M/2]} \hat{h}[2m-1] \) and \( \hat{H}[2] = \sum_{m=1}^{[M/2]} \hat{h}[2m] \) corresponds to the even-odd Trotterization used in one-dimensional tensor network calculations of quantum many-body systems [40]. We consider first-order Trotterization [41] of the full imaginary time propagator \( e^{-\hat{H}t} \):

\[
e^{-\hat{H}t} = \left( \prod_{l=1}^{L} e^{-\Delta \tau \hat{H}[l]} \right)^{\tau / \beta} + O(\Delta \tau^2),
\]

where \( \beta \) is the number of imaginary time steps and \( \Delta \tau = \beta / \beta \).

The QITE algorithm approximates each imaginary time propagator \( e^{-\Delta \tau \hat{H}[l]} \) by a unitary operator

\[
e^{-i\Delta \tau \hat{G}[l]} = e^{-i\Delta \tau \sum_{\mu} x[l]_{\mu} \sigma_{\mu}},
\]

where \( x[l]_{\mu} \) are real coefficients and \( \sigma_{\mu} \) are Pauli strings. Here we use the notation \( \sigma_0 = I, \sigma_x = X, \sigma_y = Y, \sigma_z = Z \) to denote the identity and the Pauli matrices, so that each Pauli string can be written in the form \( \sigma_{\mu} = \bigotimes_{j=1}^{N-1} \sigma_{\mu_j} \), where \( \sigma_{\mu_j} \) acts on qubit \( j \) and \( \mu_j \in \{0, x, y, z\} \). The Pauli strings \( \sigma_{\mu} \) are chosen from the set

\[
P_{\hat{H}[l]} = \bigcup_{\hat{h}[m] \in \hat{H}[l]} P_{\hat{h}[m]},
\]

where \( P_{\hat{h}[m]} \) is the set all Pauli strings over a domain of \( D \) qubits larger than or equal to the support of \( \hat{h}[m] \). To apply the QITE unitaries, without an efficient decomposition scheme each unitary needs to be further Trotterized as

\[
e^{-i\Delta \tau \hat{G}[l]} = \prod_{\mu} e^{-i\Delta \tau x[l]_{\mu} \sigma_{\mu}} + O(\Delta \tau^2).
\]

The coefficient vector \( x[l] \) is found by minimizing the square of the difference between the unitarily evolved state \( e^{-i\Delta \tau \hat{G}[l]} \ket{\psi} \) and the imaginary-time-evolved state \( c[l]^{-1/2}e^{-\Delta \tau \hat{H}[l]} \ket{\psi} \), where \( c[l] = ||e^{-\Delta \tau \hat{H}[l]} \ket{\psi}||^2 \). This minimization results in a linear system

\[
A[l]x[l] = b[l],
\]

where

\[
A[l]_{\mu \nu} = \text{Re}(\langle \psi | \sigma_{\mu} \sigma_{\nu} | \psi \rangle),
\]

\[
b[l]_{\mu} = \frac{\text{Im}(\langle \psi | e^{-\Delta \tau \hat{H}[l]} \sigma_{\mu} | \psi \rangle)}{\Delta \tau c[l]^{1/2}}.
\]

In our implementation, we expand the exponential \( e^{-\Delta \tau \hat{H}[l]} \) in \( b[l] \) and \( c[l] \) to second order in \( \Delta \tau \):

\[
b[l]_{\mu} = \frac{\text{Im}(\langle \psi | (-\hat{H}[l] + \Delta \tau \hat{H}[l]^2) \sigma_{\mu} | \psi \rangle)}{c[l]^{1/2}} + O(\Delta \tau^2),
\]

\[
c[l] = \langle \psi | 1 - 2\Delta \tau \hat{H}[l] + 2\Delta \tau^2 \hat{H}[l]^2 | \psi \rangle + O(\Delta \tau^3).
\]

To construct the linear systems, given the terms in Eqs. (6), (8) and (9) we measure operators of the form

\[
\sigma_{\mu} \sigma_{\nu}, \hat{H}[l] \sigma_{\mu}, \hat{H}[l]^2 \sigma_{\mu}, \hat{H}[l], \hat{H}[l]^2.
\]

The QITE algorithm is carried out by iterating the procedure of constructing the circuit from the QITE unitaries obtained at the previous imaginary time steps, measuring the operators in Eq. (10), constructing the linear system in Eq. (5), solving for \( x[l] \), and propagating the state with the new unitary \( e^{-i\Delta \tau \hat{G}[l]} \).

B. Finite-temperature dynamical correlation functions

Finite-temperature static observables have been previously computed on quantum hardware using the QMETTS algorithm by averaging over the observable evaluated from each METTS sample state [30]. In this work, we show that finite-temperature dynamical observables, in particular finite-temperature dynamical correlation functions, can be computed using a similar averaging procedure as for finite-temperature static observables.

On quantum computers, dynamical correlation functions can be calculated using the circuit reported in Refs. [42, 43]; this circuit has been recently used to compute neutron scattering cross-section [14] and magnon spectra [15] on quantum hardware. To obtain finite-temperature dynamical correlation functions, we insert
the exact expression of an observable ˆO on the QITE circuit into the dynamical correlation function circuit, resulting in the circuit shown in Fig. 1. The ancilla qubit is initialized in 0 and the system qubits are initialized in 0). Measuring X (Y) on the ancilla yields the real (imaginary) part of the dynamical correlation function on a binary part of the dynamical correlation function on a

The numerator trace and the denominator trace are either evaluated by full sampling over the entire Hilbert space, denoted as full trace evaluation, or by random sampling over a subspace of the Hilbert space, denoted as stochastic trace evaluation. If ˆO = U(t)V, Eq. (16) yields the finite-temperature dynamical correlation function ⟨U(t)V⟩β. ˆO can also be a static observable, in which case Eq. (16) yields the static observable at finite temperature.

III. METHODS

A. Pauli string reduction by Z2 symmetries

If we include all 4D Pauli strings over each domain consisting of D qubits, each QITE unitary e−i∆τG[1/2] applied as in Eq. (4) yields O(N4D) multi-qubit rotation gates by the standard rotation gate decomposition [44], which results in a circuit too deep on near-term quantum devices even for D = 2. Because of this prohibitive resource overhead, we describe a systematic method to reduce the number of Pauli strings in the QITE unitaries when the Hamiltonian and initial state have Z2 symmetries.

Z2 symmetries on qubit Hamiltonians have direct parallels with the stabilizer formalism in quantum error-correcting codes [45]. Suppose the Hamiltonian has d Z2 symmetries, i.e. ˆH commutes with elements of a group isomorphic to Z2d generated independently by d Pauli strings, and the initial state is in the +1 eigenspace of all d generators. If we regard the symmetry group Z2d as the stabilizer Ș, the symmetry sector of the initial state corresponds to the stabilizer subspace VS.

In stabilizer codes, the normalizer of the stabilizer N(Ș) includes all Pauli strings that commute with elements of the stabilizer Ș, and all valid operations on the code space are in the quotient group N(Ș)/Ș. Intuitively, to preserve Z2 symmetries, among all Pauli strings from P ˆH[1/2] the QITE unitaries should only include those from the quotient group N(Ș)/Ș. We now show that the original QITE algorithm subsumes the requirement that the Pauli strings should be chosen from N(Ș)/Ș because the action of the unitary e−i∆τG[1/2] with the Pauli strings from the unreduced set P ˆH[1/2] is the same as the action with Pauli strings from the reduced set P ˆH[1/2] ∩ N(Ș)/Ș. This result is stated as the following proposition, proved and discussed in Appendix A.

Proposition. Suppose QITE is applied to approximate the imaginary time propagator e−i∆τH[1/2] on the state |Ψ⟩. If there exists a stabilizer Ș such that every element of Ș commutes with ˆH[1/2] and |Ψ⟩ ∈ VS, then

(a) The action of e−i∆τG[1/2] on |Ψ⟩ with σµ ∈ P ˆH[1/2] is equivalent to the action with σµ ∈ P ˆH[1/2] ∩ N(Ș)/Ș;
(b) e−i∆τG[1/2] |Ψ⟩ ∈ VS.

Further reduction in the number of Pauli strings can be achieved by recalling from Ref. [30] that when the Hamiltonian and the initial state are real in the computational basis, the state after imaginary time evolution must be real. Thus, only Pauli strings with an odd number of Y need to be included in the QITE unitaries. Since Z2 symmetries and the conditions of a real Hamiltonian and initial state are independent, when both conditions are satisfied, the number of Pauli strings can be reduced using both conditions, in which case the reduced set of Pauli strings is modified to P ˆH[1/2] ∩ N(Ș)/Ș ∩ {σµ : ∑y δµy, y ≡ 1 (mod 2)}.
FIG. 2 Energy $E$ versus imaginary time $\beta$ simulated without noise or measurement sampling on a single initial state with and without reduction of the Pauli strings in the QITE unitaries by $Z_2$ symmetries. (a) Four-site TFIM with $J = h = 1$ and initial state $|0001\rangle$. The imaginary time step size is set to $\Delta \tau = 0.01$. The number of Pauli strings from three $D = 2$ domains is reduced from 16 to 6 by one $Z_2$ symmetry $Z_0Z_1Z_2Z_3$. (b) Four-site Heisenberg model with $J = \Delta = 1$ and initial state $(|0101\rangle + |1010\rangle)/\sqrt{2}$. The imaginary time step size is set to $\Delta \tau = 0.03$. The number of Pauli strings on the single $D = 4$ domain is reduced from 120 to 6 by two $Z_2$ symmetries $Z_0Z_1Z_2Z_3$ and $X_0X_1X_2X_3$. In both panels the energy trajectories using reduced numbers of Pauli strings match the energy trajectories without reduction, which also match the energy trajectories from exact imaginary time evolution.

In practice the Proposition is used inductively on the Trotter terms $\hat{H}[\ell]$, which implies the stabilizer need to be chosen such that every element of the stabilizer commutes with all $\hat{H}[\ell]$, or equivalently with $\hat{H}$. For spin Hamiltonians, the stabilizer generators are usually global $Z_2$ symmetries such as $Z^\otimes N$ and $X^\otimes N$. For general Hamiltonians, the $Z_2$ symmetries can be found by Gaussian elimination on the parity check matrix formed from the Hamiltonian terms [46].

To confirm our Pauli string reduction scheme, we compare the QITE energy trajectory as a function of imaginary time simulated without noise or measurement sampling on a single initial state with and without reduction of the Pauli strings in the QITE unitaries by $Z_2$ symmetries. The Hamiltonians we study include the transverse-anisotropic XXZ Hamiltonians, the stabilizer generators are usually global $Z_2$ symmetries such as $Z^\otimes N$ and $X^\otimes N$. For general Hamiltonians, the $Z_2$ symmetries can be found by Gaussian elimination on the parity check matrix formed from the Hamiltonian terms [46].

To verify our Pauli string reduction scheme, we compare the QITE energy trajectory as a function of imaginary time simulated without noise or measurement sampling on a single initial state with and without reduction of the Pauli strings in the QITE unitaries by $Z_2$ symmetries. The Hamiltonians we study include the transverse-field Ising model (TFIM) Hamiltonian

$$\hat{H}_{\text{TFIM}} = J \sum_{i=0}^{N-2} X_iX_{i+1} + h \sum_{i=0}^{N-1} Z_i$$

and the Heisenberg $XXZ$ Hamiltonian

$$\hat{H}_{\text{XXZ}} = J \sum_{i=0}^{N-2} (X_iX_{i+1} + Y_iY_{i+1} + \Delta Z_iZ_{i+1})$$

with open boundary conditions assumed for both.

In Fig. 2 we plot energy versus imaginary time calculated with QITE on a single initial state. The unreduced set of Pauli strings only includes Pauli strings with odd numbers of $Y$ because the Hamiltonian and initial state are real in the computational basis. We choose a sufficiently small imaginary time step size $\Delta \tau$ to ensure that the Trotter errors from expansion in Eq. (4) are negligible. Figure 2(a) plots the energy trajectory for the initial state $|0001\rangle$ in the four-site TFIM with $J = h = 1$. The Hamiltonian and the initial state have a $Z_2$ symmetry $Z_0Z_1Z_2Z_3$. By combining reduced Pauli strings from all three $D = 2$ domains, we obtain six Pauli strings in the QITE unitaries

$$X_0Y_1, Y_0X_1, X_1Y_2, Y_1X_2, X_2Y_3, Y_2X_3,$$  

compared to 16 Pauli strings without reduction by $Z_2$ symmetries. Figure 2(b) plots energy versus imaginary time of the initial state $(|0101\rangle + |1010\rangle)/\sqrt{2}$ on the four-site Heisenberg model with $J = \Delta = 1$. The Hamiltonian and the initial state have two $Z_2$ symmetries $Z_0Z_1Z_2Z_3$ and $X_0X_1X_2X_3$. The 120 Pauli strings in the QITE unitaries without reduction is reduced to the 6 Pauli strings

$$X_0Y_1Z_2, X_0Z_1Y_2, Y_0X_1Z_2, Y_0Z_1X_2, Z_0X_1Y_2, Z_0Y_1X_2,$$

In both panels of Fig. 2, the energy trajectories using reduced numbers of Pauli strings match the energy trajectories without reduction, which also match the energy trajectories from exact imaginary time evolution.

B. Circuit optimization

Even with reduction of Pauli strings in the QITE unitaries by $Z_2$ symmetries, applying the QITE unitaries as in Eq. (4) may still result in a circuit too deep to be implemented on current quantum hardware. In this section we describe circuit optimization techniques that further reduce circuit depth.

In two-site calculations, both the QITE circuit and the real time evolution circuit can be optimized to constant depth with a standard one- and two-qubit gate set, regardless of the number of imaginary and real time steps. For example, in two-site TFIM there is only one Pauli string $X_0Y_1$ in the QITE unitaries after reduction by the $Z_2$ symmetry $Z_0Z_1$. Suppose the unitary applied to the state at the $k$-th imaginary time step is $e^{-i\Delta \tau x_k X_0 Y_1}$. Then the unitaries at all imaginary time steps can be multiplied into a single two-qubit rotation gate $e^{-i\theta X_0 Y_1 / 2}$ where $\theta = 2\Delta \tau \sum_k x_k$. For real time evolution, the two-qubit operator $e^{-i\hat{H}t}$ is decomposed by the KAK decomposition [47–50] into six single-qubit gates and two CNOT gates.

In four-site calculations, neither the QITE circuit nor the real time evolution circuit is of constant depth. If we Trotterize the QITE unitaries as in Eq. (4) and similarly for the real time propagator, the circuit is too deep to be accurately implemented on existing quantum devices. Therefore, we recompile the circuit by fitting each QITE unitary $e^{-i\Delta \tau G[\ell]}$ or the real time propagator $e^{-i\hat{H}t}$ to a parametrized circuit [51–53] using an open-source tensor
network-based quantum simulation library [54]. In Fig. 3, we show the recompiled four-site QITE circuit, where the $U_3$ gate is a generic single-qubit gate defined as

$$U_3(\theta, \phi, \lambda) = \begin{pmatrix} \cos(\theta/2) & -e^{i\lambda} \sin(\theta/2) \\ e^{i\phi} \sin(\theta/2) & e^{i(\lambda+\phi)} \cos(\theta/2) \end{pmatrix}. \quad (21)$$

The four $U_3$ gates at the left constitute the base gate round. Each additional gate round consists of a layer of CNOT gates and a layer of single-qubit gates as shown in the dashed box. The additional gate rounds alternate between even-odd and odd-even pairs of qubits, so that the circuit shown consists of three gate rounds.

Post-selection is performed on $Z_2$ symmetries discussed in Sec. III A. When the Hamiltonian and the initial state have $Z_2$ symmetries, the final state after imaginary or real time evolution should have the same stabilizer parities as the initial state. However, during execution of the circuit, gate errors and qubit decoherence can induce nonzero overlap of the qubit state with the subspace of the wrong parity. Post-selection can mitigate these undesirable effects by discarding measurement outcomes with the wrong parity [56, 57].

We specifically consider the symmetry from a single stabilizer generator. If the operator to be measured is an ancilla operator, we can simply measure the stabilizer generator on all the system qubits and read off the parity without interfering with measurement of the ancilla. If the operator to be measured acts on system qubits, we need to simultaneously measure the operator and the stabilizer generator, which is possible because all operators in Eq. (10) commute with the stabilizer generator by our choice of Pauli strings in the QITE unitaries in Sec. III A. Specifically, each operator and the stabilizer generator can be simultaneously measured by using Clifford gates to transform the Pauli string components of the operator and the stabilizer generator until they are qubit-wise commuting, so that their expectation values can be read off on different qubits [58–61].

Figure 4 shows the circuit to simultaneously measure the stabilizer generator $Z_0Z_1Z_2Z_3$ and a Pauli string that commutes with it in a four-site QITE calculation. The sequence of CNOT gates after the QITE circuit in Fig. 4 transforms $Z_0Z_1Z_2Z_3$ to $Z_3$. Since $Z_3$ acts on a single qubit, it necessarily qubit-wise commutes with the transformed Pauli string. In practice, we stop applying CNOT gates when the transformed Pauli string becomes qubit-wise commuting with the transformed stabilizer generator, and therefore the number of CNOT gates added to the end of the circuit ranges from zero to three.

To account for noise in the final measurement, the built-in readout error mitigation routine in Qiskit [62] is applied to each measurement outcome. Because of the
small size of the systems we study, for $N$ qubits we carry out full calibration on all $2^N$ initial states. Application of the inverse of the calibration matrix to the raw measurement counts is performed by the default least-square fitting method.

We assess the effectiveness of applying post-selection and readout error mitigation at every imaginary step of QITE by simulating the finite-temperature energies of two-site and four-site TFIMs using full trace evaluation with measurement sampling and the noise model from ibmq_rome. In both panels of Fig. 5, QITE is applied with Pauli strings reduced and circuits optimized. In particular, four-site QITE unitaries are of domain size $D = 2$ and recompiled with three rounds of gates. From Fig. 5 we can see that both readout error mitigation and post-selection shift the raw data toward the exact data, confirming the effectiveness of both schemes in reducing the effect of noise. Furthermore, a combination of readout error mitigation and post-selection is observed to be most effective in mitigating the errors, which is not apparent on two sites in Fig. 5(a) presumably because of the small size of the system but clearly evident on four sites in Fig. 5(b).

In calculations of finite-temperature dynamical correlation functions, the ancilla qubit is in the state $|\pm\rangle$ before entangling with the system qubits. When there is a long sequence of gates in the real time propagator $e^{-i\hat{H}t}$, decoherence of the ancilla qubit such as amplitude damping to the qubit ground state $|0\rangle$ and depolarization will significantly affect the $X$ and $Y$ measurement results on the ancilla. To mitigate the effect of ancilla decoherence, we apply phase-and-scale correction [14, 15] as a single-step post-processing to the result at the end of the calculation. The only finite-temperature dynamical correlation function considered in this work is $\langle Z_0(t)Z_0 \rangle_\beta$, which is equal to 1 analytically at $t = 0$. Hence, we apply phase-and-scale correction by dividing the raw hardware $\langle Z_0(t)Z_0 \rangle_\beta$ at each $t$ by the raw hardware $\langle Z_0(t = 0)Z_0 \rangle_\beta$ to enforce the condition $\langle Z_0(t = 0)Z_0 \rangle_\beta = 1$.



IV. RESULTS

Experiments of computing finite-temperature observables were conducted on IBM Quantum devices ibmq_bogota and ibmq_rome [63], both of which consist of five qubits arranged on a chain with nearest-neighbor interactions and similar error rates. IBM’s open-source library Qiskit [64] was used to implement our algorithms on the devices. In each calculation, the $N$ system qubits $0, ..., N - 1$ are arranged adjacent to each other and the ancilla is closest to system qubit $0$.

The systems we study are sufficiently small that we apply QITE to approximate the full imaginary time propagator $e^{-\Delta \tau \hat{H}}$ at each imaginary time step, which is equivalent to setting $L = 1$ and $\hat{H}[1] = \hat{H}$ in Eq. (1). The QITE linear systems in Eq. (5) are solved by a conjugate gradient method. Because hardware noise and measurement sampling lead to ill-conditioned $A$ matrices in the QITE linear systems, we add a regularizer of 0.2 to the diagonal elements of each $A$ matrix in the four-site calculations.

Each calibration circuit used for readout error mitigation is repeated 1000 times; each Pauli string measurement circuit used to construct the QITE linear systems is repeated 8000 times. Error bars from full trace evaluation result only from measurement sampling and are the size of the markers in most figures; error bars from stochastic trace evaluation originate from both measurement sampling and initial state sampling. A detailed description of error bars in full and stochastic trace evaluation is given in Appendix B.

A. Two-site calculations

We study the two-site TFIM defined in Eq. (17) by setting $h = 1$ and varying $J$. The finite-temperature energy $\langle E \rangle_\beta$ and static correlation function $\langle X_0X_1 \rangle_\beta$ are calculated on the Hamiltonians with $J = \pm 1, \pm 3$, while the finite-temperature dynamical correlation function $\langle Z_0(t)Z_0 \rangle_\beta$ and excitation spectra are calculated on the Hamiltonian with $J = 3$. In all calculations, finite-temperature observables are calculated by full trace evaluation and the circuits are optimized according to the circuit optimization procedures in Sec. III B.

Figure 6 shows the finite-temperature energy $\langle E \rangle_\beta$ and static correlation function $\langle X_0X_1 \rangle_\beta$ of the two-site TFIM with $J = \pm 1, \pm 3$ from $\beta = 0$ to $\beta = 2$.

In both Fig. 6(a) and Fig. 6(b) the finite-temperature observables obtained on hardware are in good agreement with exact values. Further, if we regard each finite-temperature variable as a function of $J$, analyti-
FIG. 6 (a) Finite-temperature energy $\langle E \rangle_\beta$ and (b) static correlation function $\langle X_0 X_1 \rangle_\beta$ of the two-site TFIM with $J = \pm 1, \pm 3$ and $h = 1$ versus inverse temperature $\beta$. The imaginary time step size is set to $\Delta \tau = 0.1$. In each of the calculated observables, the mean absolute percentage error between hardware and exact results averaged over all $\beta$ ranges from 1% to 4%.

cally it can be shown that $\langle E \rangle_\beta (J) = \langle E \rangle_\beta (-J)$ and $\langle X_0 X_1 \rangle_\beta (J) = -\langle X_0 X_1 \rangle_\beta (-J)$. This relation is satisfied in the hardware data. In each of the observables we calculated, the mean absolute percentage error between hardware and exact results averaged over all $\beta$ ranges from 1% to 4%.

Next, finite-temperature dynamical properties were calculated on the two-site TFIM with $J = 3, h = 1$. The dynamical correlation function $\langle Z_0(t) Z_0 \rangle_\beta$ is evaluated from $\beta = 0$ to $\beta = 2$ and at real time from $t = 0$ to $t = 8\pi$. Figures 7(a) and 7(b) show the real and imaginary parts of $\langle Z_0(t) Z_0 \rangle_\beta$ at $\beta = 0.2$ and $\beta = 1.8$ up to $t = 4\pi$. From Figs. 7(a) and 7(b) we see that even without phase-and-scale correction, the real and imaginary parts of $\langle Z_0(t) Z_0 \rangle_\beta$ agree well with the exact results at both small and large $\beta$, presumably due to the constant and shallow depth of the real time evolution circuit.

The spectral density $S(\omega)$ is obtained by a discrete Fourier transform of the dynamical correlation function $\langle Z_0(t) Z_0 \rangle_\beta$. Specifically, at each $\beta$

$$S(\omega_k) = \frac{1}{n_t} \sum_{m=0}^{n_t-1} \langle Z_0(t_m) Z_0 \rangle_\beta e^{i\omega_k t_m}, \quad (23)$$

where $n_t$ is the total number of points in the time series, $t_m = m\Delta t$, and $\omega_k = 2\pi k/n_t\Delta t$. With this definition of Fourier transform, the peaks at positive (negative) frequencies correspond to emissions (absorptions) of excitations of the system.

In Fig. 7(c) we plot the excitation spectra of the two-site TFIM at $\beta = 0.2$. The exact excitation spectrum is obtained by a Fourier transform of the exact $\langle Z_0(t) Z_0 \rangle_\beta$ at the same points in real time as the $\langle Z_0(t) Z_0 \rangle_\beta$ obtained on hardware. From the plot, we can see that the hardware excitation spectrum agrees well with the exact

FIG. 7 Finite-temperature dynamical properties of the two-site TFIM with $J = 3, h = 1$. The imaginary time step size is set to $\Delta \tau = 0.1$. (a) Real and (b) imaginary parts of the finite-temperature dynamical correlation function $\langle Z_0(t) Z_0 \rangle_\beta$ at $\beta = 0.2$ and $\beta = 1.8$ versus real time $t$. (c) Finite-temperature excitation spectra $|S(\omega)|^2$ versus frequency $\omega$. Positive (negative) frequencies correspond to emissions (absorptions). (d) Amplitudes of the two emission peaks at $\omega = 7.18$ and $\omega = 5.94$. The hardware data capture finite-temperature dynamics of two-site TFIM across a wide range of temperatures.
excitation spectrum. The frequencies $\omega = 0, \pm 5.94, \pm 7.18$, at which the peaks in the excitation spectra are located, correspond to the excitation frequencies of $\omega = 0, \pm 6.00, \pm 7.21$ from exact diagonalization of the Hamiltonian. The deviation of the frequencies in the excitation spectra obtained from hardware $\langle Z_0(t)Z_0 \rangle_\beta$ compared to the frequencies obtained from exact diagonalization is due to the finite real time domain in our hardware calculations.

To analyze the evolution of the excitation spectra across different temperatures, we plot the amplitudes at the two emission frequencies versus $\beta$ in Fig. 7(d). Analytically, the amplitude of the transition from an initial state $|\Psi_i \rangle$ to a final state $|\Psi_f \rangle$ is $e^{-\beta E_f} |\langle \Psi_i |Z_0|\Psi_f \rangle|^2/Z$, where $E_f$ is the energy of the final state and $Z$ is the partition function. In the two-site TFIM, the only allowed transitions are between the two states in each of the two-dimensional eigenspaces of $Z_0Z_1$ with eigenvalues $\pm 1$. The frequency $\pm 7.18$ corresponds to a transition in the $+1$ eigenspace, where the ground state lies, and the frequency $\pm 5.94$ corresponds to a transition in the $-1$ eigenspace, where the first excited state lies. As the temperature decreases from infinite temperature ($\beta$ increases from 0), the populations in the two lowest states first increase until the ground state population dominates over that of the first excited state at around $\beta = 0.4$, a trend reproduced by the amplitudes obtained from hardware data in Fig. 7(d). Thus, Fig. 7 shows that quantum hardware accurately captures the finite-temperature dynamics of the two-site TFIM across a wide range of temperatures.

B. Four-site calculations

We next proceed to four-site spin systems. We study the four-site TFIM defined in Eq. (17) with $J = 3, h = 1$. Full trace evaluation is employed unless otherwise specified.

First, let us consider the gate count in the four-site circuits. For the four-site TFIM with $D = 2$, after reduction by the $Z_2$ symmetry $Z_0 Z_1 Z_2 Z_3$ there are six weight-two Pauli strings, which are the ones given in Eq. (19). If we Trotterize the QITE unitaries as in Eq. (4), each unitary requires 12 CNOT gates by the standard rotation gate decomposition [44], which becomes unfeasible on near-term quantum hardware after the first few imaginary time steps. When $D = 4$, even after reduction by one $Z_2$ symmetry there are still 28 Pauli strings in each QITE unitary. After Trotterization and rotation gate decomposition, each QITE unitary requires more than 50 CNOT gates for a single imaginary time step. Hence, when $D = 2$, we compute finite-temperature observables both by Trotterizing and by recompiling the QITE unitaries with three gate rounds; when $D = 4$, we only recompile the QITE unitaries with three gate rounds. To obtain dynamical correlation functions, we additionally recompile the real time propagator $e^{-iHt}$ with five gate rounds. The number of gate rounds is chosen such that the fidelity in Eq. (22) is at least 0.999 on average in each calculation.

Figure 8 shows the finite-temperature energy $\langle E \rangle_\beta$ and static correlation functions $\langle X_0 X_1 \rangle_\beta, \langle X_0 X_2 \rangle_\beta, \langle X_0 X_3 \rangle_\beta$ of the four-site TFIM. From the figure, we can see that the finite-temperature observables calculated with Trotterized $D = 2$ QITE unitaries deviate from those calculated with $D = 2$ recompiled QITE unitaries or $D = 4$ recompiled QITE unitaries after $\beta = 0.1$. This deviation is due to the deep circuit resulting from 12 layers of CNOT gates per imaginary time step, compared to 3 layers of CNOT gates per imaginary time step in the recompiled circuit. Moreover, even the recompiled QITE unitaries are not able to track the exact finite-temperature observables after the first few $\beta$. In particular, the slope is reversed compared to the exact result after $\beta = 0.5$. QITE up to $\beta = 0.5$ corresponds to 5 imaginary time steps and hence 15 layers of CNOT gates, which is almost at the limit of circuit depth on these quantum devices.

We examine more closely the calculations with recompiled QITE unitaries and focus on the data at $\beta \leq 0.5$. For both $D = 2$ and $D = 4$, the recompiled QITE unitaries apply the same number of layers of gates at each

\[ \text{Exact} \quad D = 2, \text{recompiled} \quad D = 2, \text{Trotterized} \quad D = 4, \text{recompiled} \]

\[ \beta \quad E \quad X_0 X_1 \beta \quad X_0 X_2 \beta \quad X_0 X_3 \beta \]

FIG. 8 (a) Finite-temperature energy $\langle E \rangle_\beta$ and static correlation functions (b) $\langle X_0 X_1 \rangle_\beta$ (c) $\langle X_0 X_2 \rangle_\beta$ (d) $\langle X_0 X_3 \rangle_\beta$ of the four-site TFIM with $J = 3, h = 1$ versus inverse temperature $\beta$ with different QITE unitaries. The imaginary time step size is set to $\Delta \tau = 0.05$. The $D = 2$ QITE unitaries are either Trotterized as in Eq. (4) or recompiled, while all $D = 4$ QITE unitaries are recompiled. The results with recompiled QITE unitaries are closer to exact results than the results with Trotterized QITE unitaries due to circuit depth. Between the calculations with recompiled unitaries, $D = 4$ is not necessarily closer to exact results than $D = 2$ for all observables possibly due to the increased influence of hardware noise in the larger linear systems.
imaginary time step, so the difference in calculated observables is formally caused by the difference in domain sizes. Since the domain size should grow with correlation length and hence with imaginary time [30], we should expect the hardware results to be closer to the exact results with $D = 4$ than with $D = 2$ especially at large $\beta$. However, this hypothesis only holds for $\langle X_0 X_2 \rangle$ and $\langle X_0 X_3 \rangle$ but not for $\langle E \rangle$ and $\langle X_0 X_1 \rangle$. Failure of this hypothesis is likely due to that fact that the larger 28-dimensional linear system arising from $D = 4$ incorporates more errors from hardware noise compared to the smaller 6-dimensional linear system arising from $D = 2$.

To explore the scalability of our method, we compare stochastic trace evaluation with full trace evaluation in calculating the finite-temperature energy of the four-site TFIM. Stochastic trace evaluation is performed by uniformly selecting initial states in the full trace evaluation result with recompiled $D = 2$ QITE unitaries. In Fig. 9, we plot the stochastic trace evaluation results with 10 and 20 samples along with the full trace evaluation and exact results; the inset shows the running average of $\langle E \rangle$ versus number of samples $n_{\text{samples}}$. As can be seen from the figure, random sampling with 10 samples already reproduced the results from full sampling on all 16 initial states, indicating that using scalable sampling schemes is a promising approach to studying larger systems.

Finally, in Fig. 10 we show the dynamical properties of the four-site TFIM with $J = 3, h = 1$ at $\beta = 0.2$. The calculation is implemented by recompiling $D = 2$ QITE unitaries with three gate rounds and real time propagation with five gate rounds. Figure 10(a) shows the real and imaginary parts of $\langle Z_0(t) Z_0 \rangle$ after phase-and-scale correction. With this correction, both the real and the imaginary parts show good agreement with the exact result. Figure 10(b) shows the excitation spectra from hardware data accurately reproduces not only the frequencies $\omega = 0, \pm 4.90, \pm 6.37, 7.84$ but also the peak amplitudes.

The favorable agreement of the hardware $\langle Z_0(t) Z_0 \rangle$ with the exact result is in contrast with the deviation of finite-temperature static observables from the exact results in Fig. 8. In fact, the raw hardware $\langle Z_0(t) Z_0 \rangle$ at $t = 0$ is $0.821 + 0.397i$, which is far from the exact value 1, indicating that phase-and-scale correction has a significant effect in correcting raw hardware data. Even though phase does not enter the static observables we computed, lack of a scale correction scheme for the static observables may explain their large deviation from the exact values compared to dynamical observables. More-
V. CONCLUSION AND OUTLOOK

Our work demonstrates that finite-temperature physics of quantum many-body systems is accessible with near-term quantum hardware. With methods to reduce required quantum sources and mitigate errors in raw hardware data, QITE enables the practical calculation of finite-temperature energies, static and dynamical correlation functions, and spectral densities of excitations.

On two sites, static and dynamical observables for a wide range of temperatures are accurately captured by quantum hardware. An important factor underlying this accuracy is the constant depth of the circuit in both QITE and real time evolution. Constant depth in QITE allowed us to extend QITE-based finite-temperature calculations from a single site \cite{30}; constant depth in real time evolution allowed us to reproduce exact finite-temperature dynamical correlation functions on quantum hardware without phase-and-scale correction as compared to previous studies \cite{14, 15}.

On four sites, finite-temperature static observables calculated on quantum hardware with circuit recompilation are in reasonable agreement with exact results at $\beta \leq 0.5$. We were also able to accurately reproduce the finite-temperature dynamical correlation function using phase-and-scale correction at a high temperature $\beta \sim 0.2$. However, accurate determination of observables at lower temperatures still appears challenging using the current recompilation scheme where the QITE unitaries are recompiled separately at each imaginary time step. Therefore, treating larger systems and lower temperatures will require additional reduction of circuit depth such as recompilation with merged imaginary time steps \cite{38} or lower error rates on quantum devices either from efficient error mitigation for imaginary time or from improvements in hardware.

Scalable methods need to be employed for systems of larger size. For this reason, we examined how stochastic trace evaluation performs in calculating finite-temperature observables compared to full trace evaluation. We found that on four sites stochastic trace evaluation reproduced full trace evaluation results accurately in the temperature regime we studied. Compared to the previously proposed QMETTS algorithm, stochastic trace evaluation has zero autocorrelation time. A detailed comparison of QITE-based computation of finite-temperature observables with different sampling schemes is certainly a topic worth exploring. Furthermore, with the availability of more qubits \cite{65, 66}, trading increased computational time due to sampling for an increased number of qubits via constructing density matrix purification states \cite{22, 23} may be another feasible direction for studying finite-temperature physics on near-term quantum hardware.

ACKNOWLEDGEMENTS

The authors thank Yi Wang, Xiuqi Ma, Sarah Sheldon and Tanvi Gujarati for helpful discussions. S.S., A.T.K.T., and A.J.M. are supported by NSF Grant Number 1839204. R.N.T. and G.K.C. are supported by the US Department of Energy, Office of Science, Grant Number 19374. S.S. acknowledges Jody Burks and Gavin Jones for helping with access to IBM Quantum devices.

APPENDIX A: PROOF OF PAULI STRING REDUCTION BY $Z_2$ SYMMETRIES

In Sec. III A we introduced a scheme to reduce Pauli strings in the QITE unitaries by $Z_2$ symmetries. We mentioned that rather than impose $Z_2$ symmetries in choosing the Pauli strings in the QITE unitaries, the original QITE algorithm subsumes the preservation of $Z_2$ symmetries. We now restate the proposition and present a proof that derives directly from the QITE linear systems in Eq. (5).

Proposition. Suppose QITE is applied to approximate the imaginary time propagator $e^{-\Delta t \hat{H}[l]}$ on the state $|\Psi\rangle$. If there exists a stabilizer $S$ such that every element of $S$ commutes with $\hat{H}[l]$ and $|\Psi\rangle \in V_S$, then

(a) The action of $e^{-i \Delta t \hat{G}[l]}$ on $|\Psi\rangle$ with $\sigma_{\mu} \in P_{\hat{H}[l]}$ is equivalent to the action with $\sigma_{\mu} \in P_{\hat{H}[l]} \cap \mathbb{N}(S)/S$,

(b) $e^{-i \Delta t \hat{G}[l]} |\Psi\rangle \in V_S$.

Proof. Pick $\sigma_{\mu} \notin \mathbb{N}(S)$. Since $e^{-\Delta t \hat{H}[l]}$ commutes with elements of $S$ and $|\Psi\rangle \in V_S$, for any $s \in S$ we have $\langle \Psi | e^{-\Delta t \hat{H}[l]} \sigma_{\mu} s | \Psi \rangle = - \langle \Psi | s e^{-\Delta t \hat{H}[l]} \sigma_{\mu} | \Psi \rangle$, which implies $\langle \Psi | e^{-\Delta t \hat{H}[l]} \sigma_{\mu} | \Psi \rangle = 0$. Hence

\[ b[l]_{\mu} = \frac{\text{Im} \langle \Psi | e^{-\Delta t \hat{H}[l]} \sigma_{\mu} | \Psi \rangle}{\Delta \tau e[l/2]} = 0 \]  

Now fix the column index $s$ such that $\sigma_{\nu} \notin \mathbb{N}(S)$, then for any $s \in S$, $\langle \Psi | \sigma_{\mu} \sigma_{\nu} s | \Psi \rangle = - \langle \Psi | s \sigma_{\mu} \sigma_{\nu} | \Psi \rangle$, which implies $\langle \Psi | \sigma_{\mu} \sigma_{\nu} | \Psi \rangle = 0$. Hence

\[ A_{\mu \nu} = \text{Re} \langle \Psi | \sigma_{\mu} \sigma_{\nu} | \Psi \rangle = 0 \]

Since $A$ is Hermitian and real, $A_{\mu \nu} = A_{\nu \mu} = A_{\mu \nu} = 0$. Thus the linear system has the block-diagonal form

\[ \begin{pmatrix} A[l] & 0 \\ 0 & A[l]^{\prime \prime} \end{pmatrix} \begin{pmatrix} x[l] \\ x[l]^{\prime \prime} \end{pmatrix} = \begin{pmatrix} b[l] \\ 0 \end{pmatrix}, \]  

(A3)
where the quantities with single primes are indexed by $\mu$ such that $\sigma_\mu \in \mathcal{N}(S)$ and those with double primes are indexed by $\mu'$ such that $\sigma_{\mu'} \notin \mathcal{N}(S)$. By setting $x[l][\nu'][\sigma_{\nu'}]$ to 0, the linear system is reduced to $A[l]'x[l]' = b[l]'$.

To show that the set of $\sigma_\mu$ can be reduced from $\mathcal{N}(S)$ to $\mathcal{N}(S)/S$, suppose $\sigma_\mu$ and $\sigma_{\mu'}$ belong to the same coset in $\mathcal{N}(S)/S$, then $\sigma_{\mu'} = \pm \sigma_\mu s$ for some $s \in S$. In the QITE unitary $e^{-i\Delta \tau \tilde{G}[l]} = \sum_{k=0}^{\infty} (-i \Delta \tau)^k \sum_{\mu} x[l][\mu] \sigma_{\mu}^k$, each term in the sum is a power of $-i \Delta \tau$ times a product of the form $\prod_{\nu}(x[l][\nu] \sigma_{\nu})$. If a product term contains $x[l][\mu] \sigma_{\mu'}$, the action of this term on $|\Psi \rangle$ is proportional to

$$
\left( \prod_{\nu'} x[l][\nu'] \sigma_{\nu'} \right) (x[l][\mu] \sigma_{\mu'}) \left( \prod_{\nu'} x[l][\nu'] \sigma_{\nu'} \right) |\Psi \rangle \tag{A4}
$$

In the product over $\nu'$, each $\sigma_{\nu'} \in \mathcal{N}(S)$, so $\prod_{\nu'}(x[l][\nu] \sigma_{\nu}) |\Psi \rangle \in V_S$. Then Eq. (A4) is equivalent to

$$
\left( \prod_{\nu'} x[l][\nu'] \sigma_{\nu'} \right) (\pm x[l][\mu] \sigma_{\mu}) \left( \prod_{\nu'} x[l][\nu'] \sigma_{\nu'} \right) |\Psi \rangle \tag{A5}
$$

Since this applies to every pair of Pauli strings in the same coset, $\tilde{G}[l]$ can be written as

$$
\tilde{G}[l] = \sum_{\mu} x[l][\mu] \sigma_{\mu} \tag{A6}
$$

where $\mu$ is chosen such that $\sigma_{\mu} \in \mathcal{P}_{\mathcal{N}(S)/S}$, $x[l][\mu] = \sum_{\mu'} \eta_{\mu'} x[l][\mu', \eta_{\mu'} = \pm 1$ and $\mu'$ is chosen such that $\sigma_{\mu'} \in \mathcal{P}_{\mathcal{N}(S)/S}$.

Since all Pauli strings on the exponent of $e^{-i \Delta \tau \tilde{G}[l]}$ commute with elements of $\mathcal{S}$, $e^{-i \Delta \tau \tilde{G}[l]}$ commutes with elements of $\mathcal{S}$ and hence $e^{-i \Delta \tau \tilde{G}[l]} |\Psi \rangle \in V_S$. 

Our Pauli string reduction scheme is related to the qubit encoding scheme that removes redundant qubits by exploiting $\mathbb{Z}_2$ symmetries reported in Ref. [46]. In the qubit encoding scheme, a Hamiltonian over some number of qubits is transformed to another Hamiltonian over a smaller number of qubits by a series of Clifford gates. Our Pauli string reduction scheme coincides with the qubit encoding scheme when the domain size $D$ equals the total number of qubits $N$, in the sense that the reduced set of Pauli strings in our scheme exactly corresponds to all Pauli strings in the encoded Hamiltonian with redundant qubits removed in the qubit encoding scheme.

However, because the weight of a Pauli string can change during the Clifford transformation, the two schemes differ when $D < N$. On the one hand, some Pauli strings can decrease in weight after encoding. If we include all Pauli strings with domain size $D$ in the encoded Hamiltonian, these Pauli strings might include those with domain size $D' > D$ in the original Hamiltonian, thus increasing the total number of Pauli strings. On the other hand, some Pauli strings can increase in weight after encoding and result in an increased cost of the QITE algorithm. As an example, consider performing QITE on a Hamiltonian with periodic boundary condition and the $\mathbb{Z}_2$ symmetry $Z_0 Z_1 Z_2 Z_3$. One of the $D = 2$ Pauli strings is $X_0 Y_3$. In the qubit encoding scheme, the symmetry operator $Z_0 Z_1 Z_2 Z_3$ is transformed to $Z_3$ so that qubit 3 can be eliminated, but the weight-two Pauli string $X_0 Y_3$ is transformed to the higher-weight Pauli string $X_0 X_1 Y_2$, thus requiring a larger QITE domain and increasing the overall cost of the algorithm. Therefore, in the present work, we use $\mathbb{Z}_2$ symmetries to reduce the number of Pauli strings in the QITE unitaries rather than eliminate redundant qubits.

**APPENDIX B: ERROR BARS IN TRACE EVALUATION**

We describe calculation of error bars of finite-temperature observables in full and stochastic trace evaluation. Here we use $E(Q)$ and $\text{Var}(Q)$ to denote the mean and variance of a quantity $Q$. The error in $Q$ is the square root of its variance.

A finite-temperature observable $\langle \hat{O} \rangle_\beta \equiv O$ has the expression $O = \sum_i P_i O_i / \sum_i P_i$, where $P_i = || | \Psi_i(\beta/2) \rangle ||^2$ is the (unnormalized) probability and $O_i = \langle \Psi_i(\beta/2) | \hat{O} | \Psi_i(\beta/2) \rangle$ is the expectation value of the observable after imaginary time evolution on the ith basis state $| \Psi_i \rangle$. On quantum computers, each probability $P_i$ is built up from the energy expectation value at each imaginary time step.

$$
P_i = \prod_{k=0}^{n_{\beta/2}-1} e^{-2 \Delta \tau E_i,k}, \tag{B1}
$$

where $E_i,k = \langle \Phi_i(k\Delta \tau/2) | \hat{H} | \Phi_i(k\Delta \tau/2) \rangle$ and $n_{\beta/2} = \beta/2 \Delta \tau$; each $O_i$ is the expectation value of the observable on the QITE-evolved state $| \Psi_i(\beta/2) \rangle$. Note that here both the exact imaginary-time-evolved state $| \Phi_i(\beta/2) \rangle$ and the QITE-evolved state $| \Psi_i(\beta/2) \rangle$ are consistent with the definitions in Sec. II.B.

In full trace evaluation $O$ is regarded a function of the $P_i$ and $O_i$, which are random variables because of measurement sampling on quantum computers. $\text{Var}(O)$ can be evaluated by expanding $O$ to first order in all $P_i$ and $O_i$ and assuming all $P_i$ and $O_i$ are independent, which then gives

$$
\text{Var}(O) = \sum_{i=1}^{2^N} [E(P_i)^2 \text{Var}(O_i) + (E(O_i) - E(O))^2 \text{Var}(P_i)] \left( \sum_{i=1}^{2^N} E(P_i) \right)^2 \tag{B2}
$$

In stochastic trace evaluation, we need to consider initial state sampling on top of measurement sampling. Define the numerator $N = n_{\text{samples}}^{-1} \sum_{i=1}^{n_{\text{samples}}} P_i O_i$ and the denominator $D = n_{\text{samples}}^{-1} \sum_{i=1}^{n_{\text{samples}}} P_i$ so that $O = N/D$. By first-order expansion of $O$ in $N$ and $D$ and assuming...
\( N \) and \( D \) are independent,

\[
\text{Var}(O) = \frac{E(N)^2 \text{Var}(D) + E(D)^2 \text{Var}(N)}{E(D)^4}.
\]  

(B3)

Now expanding \( N \) and \( D \) to first order in all variables and assuming all variables are independent, we have

\[
\text{Var}(N) = \frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} \left[ (E(P_i)E(O_i) - E(N))^2 + E(P_i)^2 \text{Var}(O_i) + E(O_i)^2 \text{Var}(P_i) \right],
\]  

(B4)

\[
\text{Var}(D) = \frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} \left[ (E(P_i) - E(D))^2 + \text{Var}(P_i) \right].
\]  

(B5)

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. 21, 467 (1982).
[2] S. Lloyd, Universal Quantum Simulators, Science 273, 1073 (1996).
[3] I. M. Georgescu, S. Ashhab, and F. Nori, Quantum simulation, Rev. Mod. Phys. 86, 153 (2014).
[4] 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).
[5] 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).
[6] 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 549, 242 (2017).
[7] J. Colless, V. Ramasesh, D. Dahlen, M. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. de Jong, and I. Siddiqi, Computation of Molecular Spectra on a Quantum Processor with an Error-Resilient Algorithm, Phys. Rev. X 8, 011021 (2018).
[8] A. Kandala, K.Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature 567, 491 (2019).
[9] H. Ma, M. Govoni, and G. Galli, Quantum simulations of materials on near-term quantum computers, npj Quantum Information 5, 106 (2019).
[10] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. DEmura, A. Dunsworth, D. Eppens, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Muczikerwicz, J. Mutus, O. Namaan, M. Neeley, C. Neill, H. Neven, M. Y. Niu, T. E. O’Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh, and A. Zalcman, Hartree-Fock on a superconducting qubit quantum computer, Science 369, 1084 (2020).
[11] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks, and C. Monroe, Emergence and Frustration of Magnetism with Variable-Range Interactions in a Quantum Simulator, Science 340, 583 (2013).
[12] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature 551, 601 (2017).
[13] A. Smith, M. S. Kim, F. Pollmann, and J. Knolle, Simulating quantum many-body dynamics on a current digital quantum computer, npj Quantum Inf. 5, 106 (2019).
[14] A. Chiesa, F. Tacchino, M. Grossi, P. Santini, I. Tavagnacco, G. Vincetti, D. Gerace, and S. Carretta, Quantum hardware simulating four-dimensional inelastic neutron scattering, Nat. Phys. 15, 455 (2019).
[15] A. Francis, J. K. Freericks, and A. F. Kemper, Quantum computation of magnon spectra, Phys. Rev. B 101, 014411 (2020).
[16] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science, arXiv:2001.03685.
[17] B. M. Terhal and D. P. DiVincenzo, Problem of equilibration and the computation of correlation functions on a quantum computer, Phys. Rev. A 61, 022301 (2000).
[18] D. Poulin and P. Wocjan, Sampling from the Thermal Quantum Gibbs State and Evaluating Partition Functions with a Quantum Computer, Phys. Rev. Lett. 103, 220502 (2009).
[19] A. Riera, C. Gogolin, and J. Eisert, Thermalization in nature and on a quantum computer, Phys. Rev. Lett. 108, 080402 (2012).
[21] M.-H. Yung and A. Aspuru-Guzik, A quantum–quantum Metropolis algorithm, Proc. Natl. Acad. Sci. 109, 754 (2012).
[22] J. Martyn and B. Swingle, Product spectrum ansatz and the simplicity of thermal states, Phys. Rev. A 100, 032107 (2019).
[23] J. Wu and T. H. Hsieh, Variational Thermal Quantum Simulation via Thermofield Double States, Phys. Rev. Lett. 123, 220502 (2019).
[24] Y. Wang, G. Li, and X. Wang, Variational ansatz-based quantum simulation, Phys. Rev. A 100, 062323 (2020).
[25] G. Verdon, J. Marks, S. Nanda, S. Leichenauer, and J. Hidary, Quantum Hamiltonian-Based Models and the Variational Quantum Thermalizer Algorithm, arXiv:1910.02071.
[26] A. N. Chowdhury, G. H. Low, and N. Wiebe, A Variational Quantum Algorithm for Preparing Quantum Gibbs States, arXiv:2002.00055.
[27] C. Zoufal, A. Lucchi, and S. Woerner, Variational Quantum Boltzmann Machines (2020), arXiv:2006.06004.
[28] Y. Wang, G. Li, and X. Wang, Variational quantum Gibbs state preparation with a truncated Taylor series (2020), arXiv:2005.08797.
[29] J. Cohn, F. Yang, K. Najafi, B. Jones, and J. K. Freericks, Minimal effective Gibbs ansatz: A simple protocol for extracting an accurate thermal representation for quantum simulation, Phys. Rev. A 102, 022622 (2020).
[30] M. Motta, C. Sun, A. T. K. Tan, M. J. O’Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. 10.1038/s41567-019-0704-4 (2019).
[31] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution, npj Quantum Inf. 5, 75 (2019).
[32] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum 3, 191 (2019).
[33] 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).
[34] S. R. White, Minimally Entangled Typical Quantum States at Finite Temperature, Phys. Rev. Lett. 102, 190601 (2009).
[35] E. M. Stoudenmire and S. R. White, Minimally entangled typical thermal state algorithms, New J. Phys. 12, 055026 (2010).
[36] K. Yeter-Aydeniz, R. C. Pooser, and G. Siopsis, Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and Lanczos algorithms, npj Quantum Inf. 6, 63 (2020).
[37] H. Nishi, T. Kosugi, and Y. ichiro Matsushita, Implementation of quantum imaginary-time evolution method on nisq devices: Nonlocal approximation, arXiv:2005.12715.
[38] N. Gomes, F. Zhang, N. F. Berthelsen, C.-Z. Wang, K.-M. Ho, P. P. Orth, and Y. Yao, Efficient step-merged quantum imaginary time evolution algorithm for quantum chemistry, arXiv:2006.15371.
[39] K. Yeter-Aydeniz, G. Siopsis, and R. C. Pooser, Scattering in the Ising Model Using Quantum Lanczos Algorithm, arXiv:2008.08763.
[40] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, Phys. Rev. Lett. 93, 040502 (2004).
[41] H. F. Trotter, On the Product of Semi-Groups of Operators, Proc. Am. Math. Soc. 10, 545 (1959).
[42] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Quantum algorithms for fermionic simulations, Phys. Rev. A 64, 022319 (2001).
[43] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Simulating physical phenomena by quantum networks, Phys. Rev. A 65, 042323 (2002).
[44] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information (Cambridge University Press, 2019).
[45] D. Gottesman, Theory of fault-tolerant quantum computation, Phys. Rev. A 57, 127 (1998).
[46] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, Tapering off qubits to simulate fermionic hamiltonians, arXiv:1701.08213.
[47] N. Khaneja, R. Brockett, and S. J. Glaser, Time optimal control in spin systems, Phys. Rev. A 63, 032308 (2001).
[48] B. Kraus and J. I. Cirac, Optimal creation of entanglement using a two-qubit gate, Phys. Rev. A 63, 062309 (2001).
[49] F. Vatan and C. Williams, Optimal quantum circuits for general two-qubit gates, Phys. Rev. A 69, 032315 (2004).
[50] G. Vidal and C. M. Dawson, Universal quantum circuit for two-qubit transformations with three controlled-NOT gates, Phys. Rev. A 69, 010301 (2004).
[51] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles, Quantum-assisted quantum compiling, Quantum 3, 140 (2019).
[52] T. Jones and S. C. Benjamin, Quantum compilation and circuit optimisation via energy dissipation, arXiv:1811.03147.
[53] K. Heya, Y. Suzuki, Y. Nakamura, and K. Fujii, Variational Quantum Gate Optimization, arXiv:1810.12745.
[54] J. Gray, quimb: A python package for quantum information and many-body calculations, J. Open Source Softw. 3, 819 (2018).
[55] R. Jozsa, Fidelity for Mixed Quantum States, J. Mod. Opt. 41, 2315 (1994).
[56] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. Campbell, Efficient quantum state preparations in variational quantum eigensolver by partitioning into commuting families, arXiv:1907.13623.
[57] G. Vidal and C. M. Dawson, Universal quantum circuit for two-qubit transformations with three controlled-NOT gates, Phys. Rev. A 69, 010301 (2004).
[58] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles, Quantum-assisted quantum compiling, Quantum 3, 140 (2019).
[59] T. Jones and S. C. Benjamin, Quantum compilation and circuit optimisation via energy dissipation, arXiv:1811.03147.
[62] A. Asfaw, L. Bello, Y. Ben-Haim, S. Bravyi, N. Bronn, L. Capelluto, A. C. Vazquez, J. Ceroni, R. Chen, A. Frisch, J. Gambetta, S. Garion, L. Gil, S. D. L. P. Gonzalez, F. Harkins, T. Imamichi, D. McKay, A. Mezzacapo, Z. Minev, R. Movassagh, G. Nannicini, P. Nation, A. Phan, M. Pistoia, A. Rattew, J. Schaefer, J. Shabani, J. Smolin, K. Temme, M. Tod, S. Wood, and J. Wootton., Learn Quantum Computation using Qiskit, http://community.qiskit.org/textbook.

[63] J. Gambetta and D. McClure, Hitting a Quantum Volume Chord: IBM Quantum adds six new systems with Quantum Volume 32, https://www.ibm.com/blogs/research/2020/07/qv32-performance/.

[64] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernandez, J. Carballo-Franquis, A. Chen, C.-F. Chen, J. M. Chow, A. D. Crocles-Gonzales, A. J. Cross, A. Cross, J. Cruz-Benito, C. Culver, S. D. L. P. Gonzalez, E. D. L. Torre, D. Ding, E. Dumitrescu, I. Duran, P. Eendebak, M. Everitt, I. F. Sertage, A. Frisch, A. Fuhrer, J. Gambetta, B. G. Gago, J. Gomez-Mosquera, D. Greenberg, I. Hamamura, V. Havlicek, J. Hellmers, uakasz Herok, H. Horii, S. Hu, T. Imamichi, T. Itoko, A. Javadi-Abhari, N. Kanazawa, A. Karazeev, K. Kruslich, P. Liu, Y. Luh, Y. Maeng, M. Marques, F. J. Martin-Fernandez, D. T. McClure, D. McKay, S. Meehala, A. Mezzacapo, N. Moll, D. M. Rodriguez, G. Nannicini, P. Nation, P. Ollitrault, L. J. O’Riordan, H. Paik, J. Prez, A. Phan, M. Pistoia, V. Pruteyanov, M. Reuter, J. Rice, A. R. Davila, R. H. P. Rudy, M. Ryu, N. Sathaye, C. Schnabel, E. Schoute, K. Setia, Y. Shi, A. Silva, Y. Siriaichi, S. Sivarajah, J. A. Smolin, M. Soeken, H. Takahashi, I. Tavernelli, C. Taylor, P. Taylour, K. Trabing, M. Treinish, W. Turner, D. Vogt-Lee, C. Vuillot, J. A. Wildstrom, J. Wilson, E. Winston, C. Wood, S. Wood, S. Winer, I. Y. Akhalwaya, and C. Zoufal, Qiskit: An Open-source Framework for Quantum Computing (2019), language: en.

[65] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. Gnilk, T. Itoko, N. Kanazawa, A. Kandala, G. A. Keefe, K. Kruslich, W. Landers, E. P. Lewandowski, D. T. McClure, G. Nannicini, A. Narasgond, H. M. Nayfeh, E. Pritchett, M. B. Rothwell, S. Srinivasan, N. Sundaresan, C. Wang, K. X. Wei, C. J. Wood, J.-B. Yau, E. J. Zhang, O. E. Dial, J. M. Chow, and J. M. Gambetta, Demonstration of quantum volume 64 on a superconducting quantum computing system, arXiv:2008.08571.

[66] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysz, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandr, J. R. McClean, M. McEwen, A. Megraint, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Quantum supremacy using a programmable superconducting processor, Nature 574, 505 (2019).