


Quantum Power Method by a Superposition of Time-Evolved States

Kazuhiro Seki^{1,*} and Seiji Yunoki^{1,2,3}

¹Computational Quantum Matter Research Team, RIKEN Center for Emergent Matter Science (CEMS), Saitama 351-0198, Japan

²Computational Materials Science Research Team, RIKEN Center for Computational Science (R-CCS), Hyogo 650-0047, Japan

³Computational Condensed Matter Physics Laboratory, RIKEN Cluster for Pioneering Research (CPR), Saitama 351-0198, Japan

 (Received 11 August 2020; revised 19 November 2020; accepted 7 January 2021; published 26 February 2021)

We propose a quantum-classical hybrid algorithm of the power method, here dubbed as the quantum power method, to evaluate $\hat{\mathcal{H}}^n|\psi\rangle$ with quantum computers, where n is a non-negative integer, $\hat{\mathcal{H}}$ is a time-independent Hamiltonian of interest, and $|\psi\rangle$ is a quantum state. We show that the number of gates required for approximating $\hat{\mathcal{H}}^n$ scales linearly in the power and the number of qubits, making it a promising application for near-term quantum computers. Using numerical simulation, we show that the power method can control systematic errors in approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ for n as large as 100. As an application, we combine our method with a multireference Krylov-subspace-diagonalization scheme to show how one can improve the estimation of ground-state energies and the ground-state fidelities found using a variational-quantum-eigensolver scheme. Finally, we outline other applications of the quantum power method, including several moment-based methods. We numerically demonstrate the connected-moment expansion for the imaginary-time evolution and compare the results with the multireference Krylov-subspace diagonalization.

DOI: [10.1103/PRXQuantum.2.010333](https://doi.org/10.1103/PRXQuantum.2.010333)

I. INTRODUCTION

Numerically solving quantum many-body systems is one of the most useful approaches for yet challenging issues in condensed-matter physics and quantum chemistry [1–4]. With classical computers, a repeated multiplication of a Hamiltonian $\hat{\mathcal{H}}$ of interest to a properly chosen state, i.e., the power iteration, is an essential element of various practical and advanced numerical techniques such as Krylov-subspace methods [5], including the Lanczos method [6–11], and polynomial-expansion methods [12]. Such methods allow for calculating not only ground states but also dynamics [13–17] of quantum many-body systems. A major obstacle in these methods is, however, the exponential growth of the dimension of the Hilbert space with its system size N . The Lanczos method has been implemented also with the variational Monte Carlo technique to systematically improve variational states towards

the exact ground state [18]. While the variational Monte Carlo method allows for substantially larger N than the full-Hilbert space approaches, an affordable number of the Lanczos iterations is practically limited to a few due to the $O(N^n)$ number of terms constituting $\hat{\mathcal{H}}^n$.

Recently, simulating quantum many-body systems with quantum computers [19–23] attracts great interest due to experimental realizations of and advances on quantum devices [24–33]. Quantum computers will allow for a rather more direct access to quantum states defined in a Hilbert space of potentially huge dimensions that cannot be treated with classical computers. At present, quantum computers are prone to noises and computations have to be accomplished with a small number of gates. In this regard, the variational-quantum-eigensolver (VQE) scheme has been proposed to simulate quantum many-body systems using noisy intermediate-scale quantum devices [34] and classical computers in a hybrid manner [35–41]. VQE calculations with noisy quantum devices are now becoming affordable for fairly larger systems [42] than in the earlier studies. While the majority of VQE schemes is devoted for gate reduction at the expense of the increased number of measurements, a measure-and-reuse technique has been proposed for reducing the number of qubits [43]. Such a qubit-reuse technique has been demonstrated by

*kazuhiro.seki@riken.jp

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

evaluating the ground-state energy of the one-dimensional Heisenberg model accurately only with a few trapped-ion qubits [44].

Moreover, to bypass variational parameter optimization and ansatz-state cultivation inherent in the VQE scheme, several versions of Krylov-subspace methods have been proposed. The quantum Lanczos (QLanczos) method [45] generates a Krylov subspace by evolving a reference state with an approximate imaginary-time evolution [46–49]. The multireference-selected quantum Krylov (MRSQK) algorithm generates a set of states spanning a Krylov subspace by evolving selected reference states in real time [50]. As a related method, a quantum version of the filter-diagonalization (QFD) method has been developed [51]. A version of the inverse-iteration method suitable for quantum computers [52] makes use of an integral representation of the inverse of the Hamiltonian [53]. Recently, an implementation of the exact imaginary-time evolution with the help of ancillary qubits and Grover’s search algorithm has been proposed [54]. Subspace-diagonalization schemes, with subspaces not restricted to a Krylov subspace but intended to approximate a particular set of eigenspaces of the Hamiltonian of interest, have been implemented for calculating not only the ground state but also excited states of correlated quantum-chemistry systems [55–60].

In this paper, we propose a quantum power method, a version of the power method suitable for quantum-classical hybrid computing of quantum many-body systems. The method is based on a time-discretized form of the higher-order derivative $\hat{\mathcal{H}}^n = i^n d^n \hat{U}(t)/dt^n|_{t=0}$ of the time-evolution operator $\hat{U}(t) = e^{-i\hat{H}t}$, by which the Hamiltonian power $\hat{\mathcal{H}}^n$ is represented as a linear combination of $\hat{U}(t)$ at different time (t) variables close to $t = 0$. The approximated Hamiltonian power retains its Hermiticity by engaging the time-discretized formalism with a central-finite-difference scheme for the time derivatives and the symmetric Suzuki-Trotter decomposition of the time-evolution operators. Assuming a k -local Hamiltonian \hat{H} composed of $O(N)$ terms, the number of the gates required for approximating $\hat{\mathcal{H}}^n$ in the quantum power method is $O(nkN)$, where N is the system size (i.e., the number of qubits). We numerically demonstrate that the quantum power method can control the systematic errors, due to the finite-difference scheme for the time derivatives and the Suzuki-Trotter decomposition of the time-evolution operators, in approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ with n as large as 100 for N up to 24. We apply the quantum power method to generate a Krylov subspace and perform, using noiseless numerical simulations, the multireference Krylov-subspace diagonalization for a one-dimensional spin-1/2 Heisenberg model with various reference states including those obtained by the VQE scheme. We find that the estimated ground-state energy as well as the ground-state fidelity are significantly

improved with increasing the power n , thus providing a way to systematically improve the VQE scheme. We also apply the Krylov-subspace diagonalization combined with the quantum power method to a Fermi-Hubbard model to demonstrate that the quantum power method remains effective even when the Hamiltonian in the qubit representation is not local. Furthermore, we briefly outline other applications of the quantum power method.

The rest of the paper is organized as follows. In Sec. II, we first summarize the main ideas and formulas of the quantum power method and list in a table major symbols used throughout the paper. In Sec. III, we provide the derivations of the main formulas and the technical details of the quantum power method, which includes the description of the central-finite-difference scheme for the time derivatives, basic properties of the approximated Hamiltonian power, and the Suzuki-Trotter decomposition of the time-evolution operators. In Sec. IV, we review the Krylov-subspace-diagonalization scheme for an application of the quantum power method. In Sec. V, we numerically demonstrate the quantum power method by considering the spin-1/2 Heisenberg model on a one-dimensional periodic chain as an example. After defining the spin-1/2 Heisenberg model, we first numerically show that the systematic errors in the quantum power method are well controlled to be essentially exact. We then present numerical results of the Krylov-subspace diagonalization combined with the quantum power method. The paper is summarized with discussions in Sec. VI. The Krylov-subspace-diagonalization scheme based on the quantum power method is compared with other algorithms reported recently and the distinctions of our method are highlighted in Appendix A. The Krylov-subspace diagonalization combined with the quantum power method is also numerically demonstrated for the Fermi-Hubbard model in Appendix B. Explicit forms of the higher-order symmetric Suzuki-Trotter decompositions generalized for multipartitioned Hamiltonians and their error analysis are provided in Appendix C. An alternative formalism of approximating the Hamiltonian power is discussed in Appendix D. For other applications of the quantum power method, some properties of the moments and cumulants are discussed in the context of the quantum power method, and the connected-moment expansion (CMX) for the imaginary-time evolution is demonstrated by numerical simulations in Appendix E. The Lanczos method with an emphasis on its connection to the moments is also described in Appendix F. Throughout the paper, we set the reduced Planck constant $\hbar = 1$.

II. MAIN FORMULAS

Here, we summarize the main ideas and formulas of the quantum power method. Figure 1 illustrates an overview of the formalism for the quantum power method based

on the higher-order derivative of the time-evolution operator $\hat{U}(t)$, which is decomposed approximately using the symmetric Suzuki-Trotter decomposition. Table I lists major symbols used in this paper. The derivations of the main formulas and the technical details are described in Sec. III.

A. Main ideas

Let $\hat{\mathcal{H}}$ be the time-independent Hamiltonian. Then, the time-evolution operator $\hat{U}(t)$ is given by

$$\hat{U}(t) = e^{-i\hat{\mathcal{H}}t} = \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} \hat{\mathcal{H}}^n, \quad (1)$$

where time t is real. The quantum power method proposed here is based simply by noticing that the n th power of the

Hamiltonian, $\hat{\mathcal{H}}^n$, is given by the n th derivative of the time-evolution operator at $t = 0$, i.e.,

$$\hat{\mathcal{H}}^n = i^n \left. \frac{d^n \hat{U}(t)}{dt^n} \right|_{t=0}. \quad (2)$$

The main purpose of this paper is to formulate, on the basis of the relation in Eq. (2), a quantum-classical hybrid scheme, which scales polynomially in both the number N of qubits and the power n for evaluating approximately the Hamiltonian power $\hat{\mathcal{H}}^n$ in a controlled manner. To this end, below we show that the Hamiltonian power $\hat{\mathcal{H}}^n$ can be approximated by a linear combination of unitaries under a controlled accuracy, if the central finite-difference scheme for the time derivative in Eq. (2) and the symmetric Suzuki-Trotter decomposition of the time-evolution operator $\hat{U}(t)$ are employed (see Fig. 1).

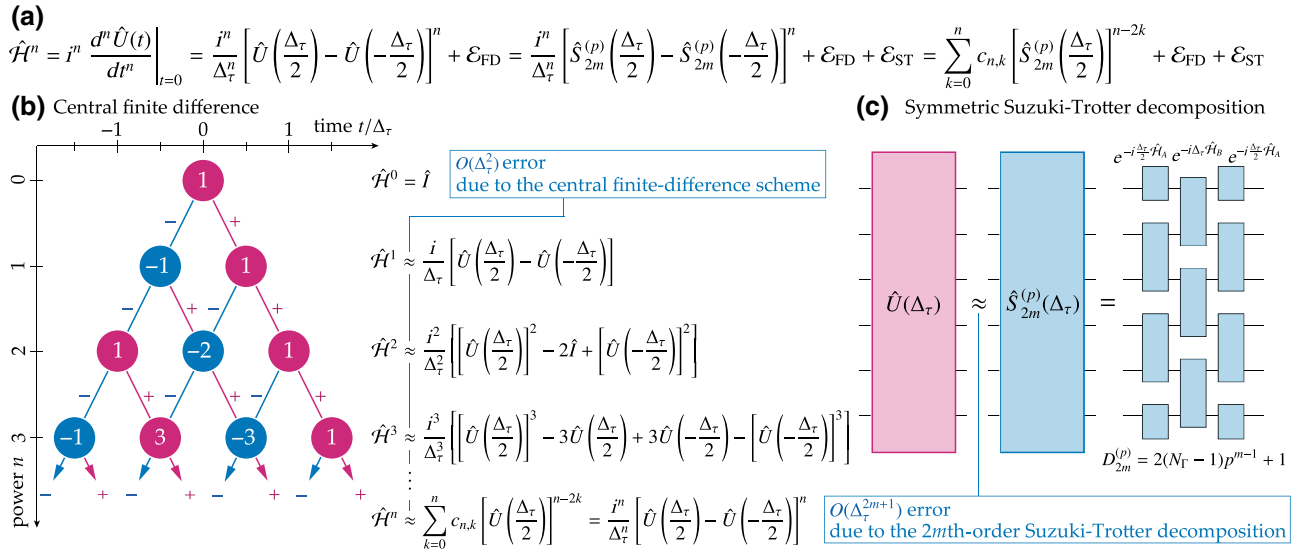


FIG. 1. Overview of the quantum power method proposed here. (a) The Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated as a linear combination of the time-evolution operators $[\hat{U}(\Delta_\tau/2)]^{n-2k}$ for $k = 0, 1, \dots, n$, in which each $\hat{U}(\Delta_\tau/2)$ is further decomposed into $\hat{S}_{2m}^{(p)}(\Delta_\tau/2)$ using the symmetric Suzuki-Trotter decomposition. Here, Δ_τ is a small time interval, and thus real positive number. \mathcal{E}_{FD} and \mathcal{E}_{ST} denote systematic errors due to the finite-difference scheme for the time derivatives and the symmetric Suzuki-Trotter decomposition of the time-evolution operators, respectively. (b) An illustration of the central-finite-difference scheme for the n th-order derivative of the time-evolution operator $\hat{U}(t)$ at $t = 0$. Pascal's triangle with an alternating sign in time t and power n provides coefficients $c_{n,k}$ of a linear combination of the time-evolution operators that approximates the Hamiltonian power $\hat{\mathcal{H}}^n$. The systematic error due to the finite-difference scheme is $\mathcal{E}_{\text{FD}} \sim O(\Delta_\tau^2)$. (c) A quantum circuit for the $2m$ th-order symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ of the time-evolution operator $\hat{U}(\Delta_\tau) = e^{-i\hat{\mathcal{H}}\Delta_\tau}$ with the systematic error of $O(\Delta_\tau^{2m+1})$. The systematic error \mathcal{E}_{ST} due to the Suzuki-Trotter decomposition for approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ in (a) is $O(\Delta_\tau^{2m})$ because of the factor $1/\Delta_\tau^n$ in $c_{n,k}$. $D_{2m}^{(p)} (= 2(N_\Gamma - 1)p^{m-1} + 1)$ is the circuit depth of a single $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ for a Hamiltonian $\hat{\mathcal{H}}$ subdivided into N_Γ parts, and p is typically an $O(1)$ integer parameter for the symmetric Suzuki-Trotter decomposition, independent of the number N of qubits. The figure refers to $m = 1, p = 3$, and $N = 6$ for a 2-local Hamiltonian with $N_\Gamma = 2$. The r th-order Richardson extrapolation improves systematically the systematic errors as $\mathcal{E}_{\text{FD}} \sim O(\Delta_\tau^{2r+2})$ and $\mathcal{E}_{\text{ST}} \sim O(\Delta_\tau^{2m+2r})$ at the expense of increasing the number $(r+1)(n+1)$ of terms in the linear combination. This implies that the lowest-order symmetric Suzuki-Trotter decomposition with $m = 1$ is adequate to control these systematic errors consistently. The number of gates, each of which is indicated by a small blue rectangle in (c), required to approximately represent the Hamiltonian power \mathcal{H}^n scales as $O(nkN)$ for a k -local Hamiltonian with a prefactor $D_{2m}^{(p)}$.

TABLE I. Major symbols used in this paper.

Symbol	Type	Description	Defining equation(s)
N	integer	number of qubits	...
t	real number	time	Eq. (1)
Δ_τ	real number	time interval	Eqs. (3) and (20)
\hat{U}	unitary operator	exact time-evolution operator	Eq. (1)
$\hat{S}_{2m}^{(p)}$	unitary operator	$2m$ th-order symmetric Suzuki-Trotter decomposition of \hat{U}	Eqs. (7), (30), and (33)
p	integer	parameter in $\hat{S}_{2m}^{(p)}$ determining its accuracy for given $m, p \geq 3$ and odd	Eq. (33)
$\hat{\mathcal{H}}$	Hermitian operator	time-independent Hamiltonian	Eqs. (1), (65), (B1), and (B6)
$\hat{\mathcal{H}}^n$	Hermitian operator	exact Hamiltonian power	Eq. (2)
$\hat{\mathcal{H}}_{\text{ST}}^n$	Hermitian operator	approximated Hamiltonian power	Eqs. (3) and (5)
$\hat{\mathcal{H}}_{\text{ST}(r)}^n$	Hermitian operator	approximated Hamiltonian power with the r th-order Richardson extrapolation	Eq. (14)
$\{c_{n,k}\}_{k=0}^n$	complex number	coefficients appearing in $\hat{\mathcal{H}}_{\text{ST}}^n$	Eq. (6)
N_Γ	integer	number of noncommuting parts in $\hat{\mathcal{H}}$	Eq. (17)
$D_{2m}^{(p)}$	integer	number of noncommuting exponentials in $\hat{S}_{2m}^{(p)}$	Eq. (18)
$\{s_i\}_{i=1}^{D_{2m}^{(p)}}$	real number	coefficients appearing in exponents of $\hat{S}_{2m}^{(p)}$, available via program in Listing 1	Eq. (C1)
\mathcal{K}	vector subspace	(block) Krylov subspace	Eq. (39)
M_B	integer	block size (i.e., number of reference states) in block Krylov subspace \mathcal{K}	Eq. (39)
\mathbf{H}	Hermitian matrix	matrix representation of $\hat{\mathcal{H}}$ in (block) Krylov subspace \mathcal{K}	Eq. (44)
\mathbf{S}	Hermitian matrix	overlap matrix in (block) Krylov subspace \mathcal{K}	Eq. (45)
$ \Psi_0\rangle$	quantum state	exact ground state of $\hat{\mathcal{H}}$ with the exact ground-state energy E_0	Eq. (41)
$ \Psi_{\text{KS}}\rangle$	quantum state	approximate ground state of $\hat{\mathcal{H}}$ spanned in \mathcal{K} with the corresponding energy E_{KS}	Eq. (42)
d	real number	operator distance, $0 \leq d \leq 1$	Eq. (75)
J	real number	exchange interaction in the Heisenberg model and transfer integral in the Hubbard model	Eq. (65) and Eq. (B1)
U_H	real number	on-site interaction in the Hubbard model	Eq. (B1)

B. Quantum power method

1. Main formulas

By applying the central finite-difference scheme with a small time interval Δ_τ for the time derivative in Eq. (2) and the symmetric Suzuki-Trotter decomposition of the time-evolution operator, the Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) + O(\Delta_\tau^2) + O(\Delta_\tau^{2m}), \quad (3)$$

where

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) = \sum_{k=0}^n c_{n,k} \left[\hat{S}_{2m}^{(p)} \left(\frac{\Delta_\tau}{2} \right) \right]^{n-2k} \quad (4)$$

$$= \frac{i^n}{\Delta_\tau^n} \left[\hat{S}_{2m}^{(p)} \left(\frac{\Delta_\tau}{2} \right) - \hat{S}_{2m}^{(p)} \left(-\frac{\Delta_\tau}{2} \right) \right]^n \quad (5)$$

with

$$c_{n,k} = \frac{i^n}{\Delta_\tau^n} (-1)^k \binom{n}{k} \quad (6)$$

being coefficients for the central finite-difference scheme [see Fig. 1(b) for a graphical derivation of $c_{n,k}$]. Note that the coefficient $c_{n,k}$ depends on Δ_τ , and if its dependence is denoted explicitly as $c_{n,k}(\Delta_\tau)$, the coefficient satisfies that $c_{n,k}(\Delta_\tau) = (-1)^n c_{n,k}(-\Delta_\tau) = c_{n,k}^*(-\Delta_\tau)$.

In Eq. (3), $O(\Delta_\tau^2)$ represents the systematic error \mathcal{E}_{FD} due to the finite-difference scheme for the time derivatives, and $O(\Delta_\tau^{2m})$ denotes the systematic error \mathcal{E}_{ST} due to the Suzuki-Trotter decomposition of the time-evolution operators. $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ is the $2m$ th-order symmetric Suzuki-Trotter decomposition of $\hat{U}(\Delta_\tau)$, given in Eq. (33), and satisfies that

$$\hat{U}(\Delta_\tau) = \hat{S}_{2m}^{(p)}(\Delta_\tau) + O(\Delta_\tau^{2m+1}) \quad (7)$$

and

$$\left[\hat{S}_{2m}^{(p)}(\Delta_\tau)\right]^\dagger = \left[\hat{S}_{2m}^{(p)}(\Delta_\tau)\right]^{-1} = \hat{S}_{2m}^{(p)}(-\Delta_\tau). \quad (8)$$

The superscript p in $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ is an odd-integer parameter with $p \geq 3$ that determines prefactors of the residual terms in $O(\Delta_\tau^{2m+1})$ but does not change the order of the approximation in Δ_τ (for numerical demonstrations, see Appendix C). The order $O(\Delta_\tau^{2m})$ of the Suzuki-Trotter error \mathcal{E}_{ST} in Eq. (3) is decreased by one from the naively expected order $O(\Delta_\tau^{2m+1})$ because of the factor $1/\Delta_\tau^n$ in $c_{n,k}$. $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ is the central quantity in the quantum power method that approximates the Hamiltonian power $\hat{\mathcal{H}}^n$.

Three remarks are in order. First, Eq. (3) already reveals a remarkable advantage in the quantum power method: in order to control the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} with the same order of accuracy, it is enough to adopt the lowest-order Suzuki-Trotter decomposition with $m = 1$, independently of the power n . Second, Eq. (4) indicates that the Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated by a linear combination of the $n + 1$ Suzuki-Trotter-decomposed time-evolution operators. Third, Eq. (5) implies that $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ satisfies the law of exponents

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}_{\text{ST}}^1(\Delta_\tau)\right]^n. \quad (9)$$

Moreover, $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ is Hermitian and an even function of Δ_τ , i.e.,

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)\right]^\dagger = \hat{\mathcal{H}}_{\text{ST}}^n(-\Delta_\tau), \quad (10)$$

indicating that the systematic error in odd powers of Δ_τ is absent in Eq. (3).

2. Richardson extrapolation

The systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} in Eq. (3) can be controlled by varying the time interval Δ_τ . However, it is often practically useful to reduce the systematic errors by not taking too small Δ_τ in the algorithmic level. A better error estimate can be achieved by systematically eliminating lower-order errors in Eq. (3) with the Richardson extrapolation.

In the Richardson extrapolation, $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ and $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau/h)$ with some real h (such that $0 < h \neq 1$) are used to eliminate the leading terms of the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} simultaneously in Eq. (3) as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}_{\text{ST}(1)}^n(\Delta_\tau) + O(\Delta_\tau^4) + O(\Delta_\tau^{2m+2}), \quad (11)$$

where

$$\hat{\mathcal{H}}_{\text{ST}(1)}^n(\Delta_\tau) = \frac{h^2 \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau/h) - \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)}{h^2 - 1} \quad (12)$$

is the first-order Richardson extrapolation of $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$. Since $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ is an even function of Δ_τ , $\hat{\mathcal{H}}_{\text{ST}(1)}^n(\Delta_\tau)$ is also an even function of Δ_τ and thus the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} in odd powers of Δ_τ are absent in Eq. (11).

We can use the Richardson extrapolation recursively to further eliminate the leading terms of the systematic errors in Eq. (11). Namely, the r th-order Richardson extrapolation $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ of the approximated Hamiltonian power can be obtained recursively as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r}), \quad (13)$$

where

$$\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) = \frac{h^{2r} \hat{\mathcal{H}}_{\text{ST}(r-1)}^n(\Delta_\tau/h) - \hat{\mathcal{H}}_{\text{ST}(r-1)}^n(\Delta_\tau)}{h^{2r} - 1} \quad (14)$$

with $\hat{\mathcal{H}}_{\text{ST}(0)}^n(\Delta_\tau) \equiv \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ and therefore the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} are reduced to $O(\Delta_\tau^{2+2r})$ and $O(\Delta_\tau^{2m+2r})$, respectively, after the r th-order Richardson extrapolation. One can easily show that

$$\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)\right]^\dagger = \hat{\mathcal{H}}_{\text{ST}(r)}^n(-\Delta_\tau) \quad (15)$$

because $\hat{\mathcal{H}}_{\text{ST}(0)}^n(\Delta_\tau)$ is Hermitian and is an even function of Δ_τ , and therefore the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} in odd powers of Δ_τ are absent in Eq. (13). However, $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ is no longer the n th power of $\hat{\mathcal{H}}_{\text{ST}(r)}^1(\Delta_\tau)$, i.e., $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) \neq \left[\hat{\mathcal{H}}_{\text{ST}(r)}^1(\Delta_\tau)\right]^n$, when $r \geq 1$, but obviously $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}_{\text{ST}(r)}^1(\Delta_\tau)\right]^n + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r})$. In our numerical simulations, we set $h = 2$ when the Richardson extrapolation is used.

Since $\hat{\mathcal{H}}_{\text{ST}(0)}^n(\Delta_\tau)$ is a linear combination of $n + 1$ unitaries $\left\{[\hat{S}_{2m}^{(p)}(\Delta_\tau/2)]^{n-2k}\right\}_{k=0}^n$, $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ is a linear combination of $(r + 1)(n + 1)$ unitaries $\left\{\left\{[\hat{S}_{2m}^{(p)}(\Delta_\tau/2h^l)]^{n-2k}\right\}_{k=0}^n\right\}_{l=0}^r$. Equation (13) hence reveals another significant feature of the quantum power method that the lowest-order symmetric Suzuki-Trotter decomposition with $m = 1$ suffices to systematically and consistently eliminate the lower-order systematic errors in \mathcal{E}_{FD} and \mathcal{E}_{ST} with only a polynomial increase of computational complexity. In Sec. V B, we show by numerical simulations that these systematic errors in the approximated Hamiltonian power are well controlled with the time interval Δ_τ for the power n as large as 100.

For the application purpose of the quantum power method, it is important that the symmetry of the Hamiltonian $\hat{\mathcal{H}}$ is still respected in the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$. This is indeed the case in the quantum

power method because

$$\left[\hat{\mathcal{H}}, \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) \right] = O(\Delta_\tau^{2m+2r}). \quad (16)$$

Therefore, the symmetry of the Hamiltonian $\hat{\mathcal{H}}$ is preserved in the quantum power method within the systematic error \mathcal{E}_{ST} due to the Suzuki-Trotter decomposition that can be well controlled. Notice that there is no contribution from the systematic error \mathcal{E}_{FD} due to the finite-difference scheme for the time derivatives in the right-hand side of Eq. (16) because $\left[\hat{\mathcal{H}}, \hat{U}(\Delta_\tau) \right] = 0$.

3. Gate count

In the quantum power method, the Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated with $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$, which is a linear combination of $(r+1)(n+1)$ unitaries $\left\{ \left[\hat{S}_{2m}^{(p)}(\Delta_\tau/2h^l) \right]^{n-2k} \right\}_{k=0}^r$, i.e., Suzuki-Trotter-decomposed time-evolution operators, and each unitary is treated separately. Therefore, the gate count is determined by the number of gates required for describing $[\hat{S}_{2m}^{(p)}(\pm\Delta_\tau/2h^l)]^n$ in a quantum circuit because the number of gates required scales linearly with the power of $\hat{S}_{2m}^{(p)}(\Delta_\tau/2h^l)$ and is independent of the argument.

The number of the noncommuting exponentials in $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ corresponds to the circuit depth of a quantum circuit for a single time-evolution operator $\hat{U}(\Delta_\tau)$ approximated by $\hat{S}_{2m}^{(p)}(\Delta_\tau)$, and thus it gives a prefactor for the gate count. Let us assume that the Hamiltonian $\hat{\mathcal{H}}$ can be divided into N_Γ parts as

$$\hat{\mathcal{H}} = \underbrace{\hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B + \hat{\mathcal{H}}_C + \cdots + \hat{\mathcal{H}}_Z}_{N_\Gamma \text{ terms}}, \quad (17)$$

where generally $[\hat{\mathcal{H}}_\Gamma, \hat{\mathcal{H}}_{\Gamma'}] \neq 0$ if $\Gamma \neq \Gamma'$ but terms within each $\hat{\mathcal{H}}_\Gamma$ commute to each other (here, $\Gamma, \Gamma' = A, B, \dots, Z$). As derived in Sec. III B 3, the number $D_{2m}^{(p)}$ of noncommuting exponentials in $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ is given by

$$D_{2m}^{(p)} = 2(N_\Gamma - 1)p^{m-1} + 1. \quad (18)$$

As illustrated in Fig. 1(c), the simplest case with $m = 1$, $p = 3$, and $N_\Gamma = 2$, for which $D_2^{(p)} = 3$, a quantum circuit for a single $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ has the circuit depth $D_{2m}^{(p)}$, and thus the circuit depth required for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ is at most $O(n)$ with a prefactor $D_{2m}^{(p)}$. It should be noted here that, as far as the quantum power method is concerned, the lowest-order symmetric Suzuki-Trotter decomposition (i.e., $m = 1$) is sufficient.

Let N be the number of qubits. Assuming that a Hamiltonian $\hat{\mathcal{H}}$ is k local and consists of $O(N)$ terms, each of

which is a Pauli string of length at most k , the number of gates required for $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ is $O(kN)$ [61] with a prefactor $D_{2m}^{(p)}$. Therefore, the number of gates required for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ is $O(nkN)$ with a prefactor $D_{2m}^{(p)} = 2N_\Gamma - 1 \sim O(1)$, where $O(1)$ implies that the quantity is independent of n and N . For example, for the spin-1/2 Heisenberg model considered in Sec. V, the locality of the Hamiltonian is independent of the system size, i.e., $k \sim O(1)$, and hence the gate count for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ scales as $O(nN)$. On the other hand, when a fermionic Hamiltonian is considered, the locality of the Hamiltonian may depend on the system size N due to a fermion-to-qubit mapping such as the Jordan-Wigner transformation [62] or the Bravyi-Kitaev transformation [63] (see also Refs. [64–66]). The Jordan-Wigner transformation represents a fermionic operator with an $O(N)$ number of Pauli operators, i.e., $k \sim O(N)$, and hence the gate count for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ scales asymptotically as $O(nN^2)$. The Bravyi-Kitaev transformation represents a fermionic operator with an $O(\log N)$ number of Pauli operators, i.e., $k \sim O(\log N)$, and hence the gate count for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ scales asymptotically as $O(nN \log N)$.

As described above, the r th-order Richardson extrapolation does not alter the number of gates required, but the number of the Suzuki-Trotter-decomposed time-evolution operators in $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ increases as $(r+1)(n+1)$. Therefore, for example, to evaluate the expectation value of $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ with respect to a given state $|\psi\rangle$, the $(r+1)(n+1)$ number of state overlaps such as $\langle \psi | [\hat{S}_{2m}^{(p)}(\Delta_\tau/2h^l)]^{n-2k} | \psi \rangle$ have to be estimated. However, these quantities can be evaluated on quantum computers separately in parallel with respect to k and l .

4. Possible circuit structure for the linear combination of time-evolution operators

The form of the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ in Eq. (5) suggests a direct treatment of the linear combination of the Suzuki-Trotter-decomposed time-evolution operators with a single quantum circuit [67–69] that forms a simple recursive structure. Figure 2 shows such a circuit structure for probabilistically generating the state $\propto [\hat{S}_{2m}^{(p)}(\Delta_\tau/2) - \hat{S}_{2m}^{(p)}(-\Delta_\tau/2)]^n |\psi\rangle$, among 2^n superposed states, in the N register qubits along with n ancilla qubits. However, the probability for finding the desired state in the register qubits becomes exponentially small in general if n is large. Let us define $P_{b_1 b_2 \dots b_n}$ as the probability for finding a bit string $b_1 b_2 \dots b_n$ by measuring the n ancilla qubits ($b_k = 0$ or 1 for $1 \leq k \leq n$). Then the probability for finding the bit string $11 \dots 1$, which is relevant for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ [70], is given by

$$P_{11 \dots 1} = \frac{1}{4^n} (-1)^n \langle \psi | [\hat{S}_{2m}^{(p)}(\Delta_\tau/2) - \hat{S}_{2m}^{(p)}(-\Delta_\tau/2)]^{2n} | \psi \rangle. \quad (19)$$

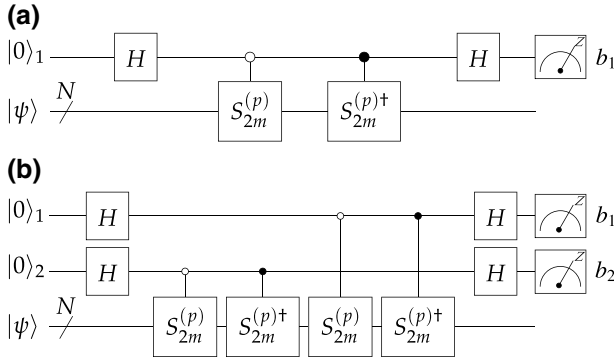


FIG. 2. Circuit with N register qubits and n ancilla qubits for probabilistically generating the state $\propto [\hat{S}_{2m}^{(p)}(\Delta_\tau/2) - \hat{S}_{2m}^{(p)}(-\Delta_\tau/2)]^n |\psi\rangle$ in the register qubits for (a) $n = 1$ and (b) $n = 2$. H , $S_{2m}^{(p)}$, and $S_{2m}^{(p)†}$ in the circuit denote the Hadamard gate, $\hat{S}_{2m}^{(p)}(\Delta_\tau/2)$, and $\hat{S}_{2m}^{(p)}(-\Delta_\tau/2)$, respectively. A controlled-unitary gate with a solid (open) circle indicates that the unitary gate is applied only if the control qubit is set to 1 (0). The probability $P_{b_1 b_2 \dots b_n}$ for finding the bit string $b_1 b_2 \dots b_n = 11 \dots 1$ in the ancilla qubits is given in Eq. (19).

If $|\psi\rangle$ were an eigenstate of $\hat{S}_{2m}^{(p)}(\Delta_\tau/2)$ with an eigenvalue $e^{i\lambda(\Delta_\tau)}$, it oscillates as $P_{11\dots 1} = [\sin \lambda(\Delta_\tau)]^{2n}$, but otherwise it is exponentially small. Therefore, as far as near-term applications with a limited number of gates are concerned, the linear combination of the Suzuki-Trotter-decomposed time-evolution operators is better treated with classical computers in the form of Eq. (4). However, we anticipate that, once a noiseless quantum computer is realized, the product form of Eq. (5) might have the advantage of robustness against loss of significance for small Δ_τ .

5. Summary of the proposed method

Figure 1 summarizes the quantum power method. In the quantum power method, the Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated to $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ represented as a linear combination of the $n + 1$ Suzuki-Trotter decomposed time-evolution operators $\{[\hat{S}_{2m}^{(p)}(\Delta_\tau/2)]^{n-2k}\}_{k=0}^n$. The systematic error \mathcal{E}_{FD} due to the finite-difference scheme for the time derivatives is $O(\Delta_\tau^2)$, and the systematic error \mathcal{E}_{ST} due to the Suzuki-Trotter decomposition of the time-evolution operators is $O(\Delta_\tau^{2m})$. These systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} can be both improved systematically with the r th-order Richardson extrapolation to $O(\Delta_\tau^{2+2r})$ and $O(\Delta_\tau^{2m+2r})$, respectively, by approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ with $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$, which is given as a linear combination of the $(r + 1)(n + 1)$ Suzuki-Trotter-decomposed time-evolution operators $\left\{ \left\{ [\hat{S}_{2m}^{(p)}(\Delta_\tau/2h^l)]^{n-2k} \right\}_{k=0}^n \right\}_{l=0}^r$. While the linear combination of the Suzuki-Trotter-decomposed time-evolution operators is treated classically,

each Suzuki-Trotter-decomposed time-evolution operator $[\hat{S}_{2m}^{(p)}(\Delta_\tau/2h^l)]^{n-2k}$ is evaluated on quantum computers.

C. Comparison with direct evaluation and classical computation

The direct evaluation of $\langle \psi | \hat{\mathcal{H}}^n | \psi \rangle$ requires the expectation values of $O(\min(N^n, 4^N))$ operators, possibly containing long strings of Pauli operators, provided that the Hamiltonian $\hat{\mathcal{H}}$ consists of $O(N)$ terms. Although the depth of the circuits for these terms is $O(1)$, the $O(\min(N^n, 4^N))$ measurements make the direct evaluation of $\langle \psi | \hat{\mathcal{H}}^n | \psi \rangle$ unfeasible as soon as the power n and the number N of qubits are large.

In classical computation, the computational complexity scales as $O(nN_D)$ for the evaluation of $\hat{\mathcal{H}}^n |\psi\rangle$, when the Hamiltonian $\hat{\mathcal{H}}$ is local and thus the Hamiltonian matrix is sparse. Here, N_D is the dimension of the Hilbert space, e.g., $N_D = 2^N$ for the spin-1/2 Heisenberg model. This implies that the computational complexity of the classical computation scales exponentially in N .

In the quantum power method proposed here, the gate count for approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ scales as $O(nkN)$ for a k -local Hamiltonian composed of $O(N)$ terms. In addition, the number of state overlaps required to evaluate is $(r + 1)(n + 1)$, which is polynomial in n and independent of N . Therefore, although it is approximate, the quantum power method is a potentially promising application for near-term quantum devices and would have a quantum advantage over the classical counterpart of the power method.

III. DERIVATIONS OF THE MAIN FORMULAS

Here, we provide the derivations of the main formulas in Sec. II and describe technical details of the quantum power method.

A. Hamiltonian power as a linear combination of unitary time-evolution operators

As shown in Eq. (2), the Hamiltonian power $\hat{\mathcal{H}}^n$ is given by the n th derivative of the time-evolution operator $\hat{U}(t)$ at $t = 0$. Here we show that, using the central finite-difference scheme for the time derivatives, the Hamiltonian power can be approximated by a linear combination of the time-evolution operators.

By introducing a small time interval Δ_τ , we replace the time derivative in Eq. (2) with the central finite difference as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}^n(\Delta_\tau) + O(\Delta_\tau^2), \quad (20)$$

where

$$\hat{\mathcal{H}}^n(\Delta_\tau) = \sum_{k=0}^n c_{n,k} \hat{U}\left(\left(\frac{n}{2} - k\right) \Delta_\tau\right) \quad (21)$$

and $c_{n,k}$ is given in Eq. (6). The systematic error $O(\Delta_\tau^2)$ in Eq. (20) is due to the finite differentiation and this is the same systematic error \mathcal{E}_{FD} in Eq. (3). Equations (20) and (21) thus indicate that the n th power of the Hamiltonian, $\hat{\mathcal{H}}^n$, can be approximated with a controlled accuracy as a linear combination of the time-evolution operators evaluated at $n + 1$ different time variables.

From the unitarity of the time-evolution operator and its accordance with the time-reversed evolution,

$$\left[\hat{U}(t)\right]^\dagger = \left[\hat{U}(t)\right]^{-1} = \hat{U}(-t), \quad (22)$$

it follows that the approximated Hamiltonian power $\hat{\mathcal{H}}^n(\Delta_\tau)$ is Hermitian and an even function of Δ_τ i.e.,

$$\hat{\mathcal{H}}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}^n(\Delta_\tau)\right]^\dagger = \hat{\mathcal{H}}^n(-\Delta_\tau). \quad (23)$$

In the last equality, we use that $c_{n,k}$ in Eq. (6) is an even (odd) function of Δ_τ when n is even (odd). Since $\hat{\mathcal{H}}^n(\Delta_\tau)$ is an even function of Δ_τ , the systematic error \mathcal{E}_{FD} in odd powers of Δ_τ is absent in Eq. (20). Moreover, with the multiplication law of the time-evolution operator $\hat{U}(t)\hat{U}(t') = \hat{U}(t+t')$, Eq. (21) can be written as

$$\begin{aligned} \hat{\mathcal{H}}^n(\Delta_\tau) &= \sum_{k=0}^n c_{n,k} \left[\hat{U}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-2k} \\ &= \sum_{k=0}^n c_{n,k} \left[\hat{U}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-k} \left[\hat{U}\left(-\frac{\Delta_\tau}{2}\right)\right]^k \\ &= \frac{i^n}{\Delta_\tau^n} \left[\hat{U}\left(\frac{\Delta_\tau}{2}\right) - \hat{U}\left(-\frac{\Delta_\tau}{2}\right)\right]^n. \end{aligned} \quad (24)$$

The last line in Eq. (24) indicates that the approximated Hamiltonian power $\hat{\mathcal{H}}^n(\Delta_\tau)$ satisfies a law of exponents

$$\hat{\mathcal{H}}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}^1(\Delta_\tau)\right]^n. \quad (25)$$

Namely, $\hat{\mathcal{H}}^n(\Delta_\tau)$ is exactly the n th power of $\hat{\mathcal{H}}^{n=1}(\Delta_\tau)$ for $n \geq 0$. In fact, Eq. (24) can be understood simply as

$$\hat{\mathcal{H}}^n = \left[i \frac{d\hat{U}(t)}{dt} \Big|_{t=0} \right]^n = \left[\hat{\mathcal{H}}^1(\Delta_\tau) \right]^n + O(\Delta_\tau^2). \quad (26)$$

The systematic error in Eq. (20) can be systematically improved by the Richardson extrapolation. Following the same procedure described in Sec. II B 2, the leading order of the systematic error can be eliminated recursively by the r th-order Richardson extrapolation as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}_{(r)}^n(\Delta_\tau) + O(\Delta_\tau^{2+2r}), \quad (27)$$

where

$$\hat{\mathcal{H}}_{(r)}^n(\Delta_\tau) = \frac{h^{2r} \hat{\mathcal{H}}_{(r-1)}^n(\Delta_\tau/h) - \hat{\mathcal{H}}_{(r-1)}^n(\Delta_\tau)}{h^{2r} - 1}, \quad (28)$$

with $\hat{\mathcal{H}}_{(0)}^n(\Delta_\tau) \equiv \hat{\mathcal{H}}^n(\Delta_\tau)$. Because $\hat{\mathcal{H}}^n(\Delta_\tau)$ satisfies Eq. (23), one can readily show that $\hat{\mathcal{H}}_{(r)}^n(\Delta_\tau)$ is also Hermitian and is an even function of Δ_τ , and therefore the systematic error \mathcal{E}_{FD} in odd powers of Δ_τ is absent in Eq. (27). Since $\hat{\mathcal{H}}_{(0)}^n(\Delta_\tau)$ is a linear combination of the time-evolution operators at $n + 1$ different times, $\hat{\mathcal{H}}_{(r)}^n(\Delta_\tau)$ is a linear combination of the time-evolution operators at $(r + 1)(n + 1)$ different times. Note also that $\hat{\mathcal{H}}_{(r)}^n(\Delta_\tau) \neq \left[\hat{\mathcal{H}}_{(r)}^1(\Delta_\tau)\right]^n$ for $r \geq 1$, but obviously $\hat{\mathcal{H}}_{(r)}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}_{(r)}^1(\Delta_\tau)\right]^n + O(\Delta_\tau^{2+2r})$.

There are three additional remarks regarding the properties of the approximated Hamiltonian power $\hat{\mathcal{H}}^n(\Delta_\tau)$. First, if a forward or backward, instead of the central, finite-difference scheme is employed in Eq. (21), the Hermiticity and the even dependence on Δ_τ of $\hat{\mathcal{H}}^n(\Delta_\tau)$ in Eq. (23) are both violated. Therefore, the central finite-difference scheme is a crucial choice. Second, when the time-evolution operator $\hat{U}(\Delta_\tau)$ is approximated by a Suzuki-Trotter decomposition, the corresponding Suzuki-Trotter error \mathcal{E}_{ST} appears in Eqs. (21) and (24). Since the implementation of a higher-order Suzuki-Trotter decomposition on quantum computers requires many layers of gates, it is essential to control \mathcal{E}_{ST} with a lower-order Suzuki-Trotter decomposition. Third, if a symmetric Suzuki-Trotter decomposition, which retains the equivalence between the inverse of the time evolution and the time-reversed evolution [the right-most equality in Eq. (22)], is employed to decompose the time-evolution operators in Eqs. (21) and (24), the resulting $\hat{\mathcal{H}}^n(\Delta_\tau)$ still satisfies the Hermiticity and the even dependence on Δ_τ , as given in Eq. (10). Therefore, it is important to adopt a symmetric Suzuki-Trotter decomposition (see Sec. III B 2 for details).

B. Suzuki-Trotter decomposition

The formalism described above in Sec. III A is based on the exact time-evolution operator $\hat{U}(t)$ in Eq. (1). However, on quantum computers, the time-evolution operator with its exponent composed of the sum of noncommuting operators usually has to be represented as a product

of time-evolution operators with each exponent composed of the sum of commuting operators. For this purpose, the Suzuki-Trotter decomposition is employed to approximately decompose the time-evolution operator.

In this regard, we should emphasize that one of the crucial steps for the successful quantum power method is to determine properly in which stage the time-evolution operators in $\hat{\mathcal{H}}^n(\Delta_\tau)$ should be approximated by the Suzuki-Trotter decomposition, either in Eq. (21) or in Eq. (24). Although Eqs. (21) and (24) are exactly the same if the exact time-evolution operators are used, they are no longer the same in general once the time-evolution operators are approximated. Therefore, there are at least two routes to formulate the quantum power method. As we discuss in details, these two approaches give us two different algorithms that scale differently in the power n . It turns out that when the power n is larger than 4, the algorithm formulated on the basis of Eq. (24) with the lowest-order symmetric Suzuki-Trotter decomposition is preferable, otherwise the formalism based on Eq. (21) with the higher-order symmetric Suzuki-Trotter decompositions is favored in terms of the gate counts.

To understand the difference of these two approaches, in this section, we briefly summarize a systematic construction of the higher-order symmetric Suzuki-Trotter decompositions [71–73] for the quantum power method.

1. Recursive construction of higher-order Suzuki-Trotter decompositions

We now describe a systematic construction of the symmetric Suzuki-Trotter decompositions. Let us define $x = -i\Delta_\tau$ to simplify the notation. The second-order symmetric decomposition $\hat{S}_2(\Delta_\tau)$ of the time-evolution operator $\hat{U}(\Delta_\tau)$ for the Hamiltonian $\hat{\mathcal{H}}$ of the form in Eq. (17) is given by

$$\hat{U}(\Delta_\tau) = e^{x\hat{\mathcal{H}}} = \hat{S}_2(\Delta_\tau) + O(\Delta_\tau^3), \quad (29)$$

where

$$\hat{S}_2(\Delta_\tau) = \underbrace{e^{\frac{x}{2}\hat{\mathcal{H}}_A} e^{\frac{x}{2}\hat{\mathcal{H}}_B} e^{\frac{x}{2}\hat{\mathcal{H}}_C} \dots}_{N_\Gamma-1 \text{ exponentials}} \underbrace{e^{x\hat{\mathcal{H}}_Z} \dots}_{N_\Gamma-1 \text{ exponentials}} \underbrace{e^{\frac{x}{2}\hat{\mathcal{H}}_C} e^{\frac{x}{2}\hat{\mathcal{H}}_B} e^{\frac{x}{2}\hat{\mathcal{H}}_A}}_{2N_\Gamma-1 \text{ exponentials}}. \quad (30)$$

Equation (30) can be derived by using the well-known decomposition $e^{x(\hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B)} = e^{(x/2)\hat{\mathcal{H}}_A} e^{x\hat{\mathcal{H}}_B} e^{(x/2)\hat{\mathcal{H}}_A} + O(\Delta_\tau^3)$ repeatedly, e.g., $e^{x(\hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B + \hat{\mathcal{H}}_C)} = e^{(x/2)\hat{\mathcal{H}}_A} e^{x(\hat{\mathcal{H}}_C + \hat{\mathcal{H}}_B)} e^{(x/2)\hat{\mathcal{H}}_A} + O(\Delta_\tau^3) = e^{(x/2)\hat{\mathcal{H}}_A} e^{(x/2)\hat{\mathcal{H}}_B} e^{x\hat{\mathcal{H}}_C} e^{(x/2)\hat{\mathcal{H}}_B} e^{(x/2)\hat{\mathcal{H}}_A} + O(\Delta_\tau^3)$. The subscript “2” implies that $\hat{S}_2(\Delta_\tau)$ correctly represents $\hat{U}(\Delta_\tau) = e^{x\hat{\mathcal{H}}}$ to $O(\Delta_\tau^2)$. It is readily found that $\hat{S}_2(\Delta_\tau)$

satisfies

$$[\hat{S}_2(\Delta)]^\dagger = \hat{S}_2(-\Delta_\tau) \quad (31)$$

and

$$\hat{S}_2(\Delta_\tau)\hat{S}_2(-\Delta_\tau) = \hat{S}_2(-\Delta_\tau)\hat{S}_2(\Delta_\tau) = \hat{I}, \quad (32)$$

where \hat{I} is the identity operator, and therefore $\hat{S}_2(\Delta_\tau)$ is unitary.

It is noteworthy that if we write $\hat{S}_2(\Delta_\tau)$ in the form $\hat{S}_2(\Delta_\tau) = \exp[x\hat{\mathcal{H}} + x^2\hat{R}_2 + x^3\hat{R}_3 + \dots]$, then the residual terms \hat{R}_k with k even are zero [72]. This can be confirmed as follows. Equation (32) indicates that $\hat{S}_2(\Delta_\tau)$ commutes with $\hat{S}_2(-\Delta_\tau) = [\hat{S}_2(\Delta_\tau)]^{-1}$, implying that $\hat{I} = \hat{S}_2(\Delta_\tau)\hat{S}_2(-\Delta_\tau) = \exp[2(x^2\hat{R}_2 + x^4\hat{R}_4 + x^6\hat{R}_6 + \dots)]$ for arbitrary $x (= -i\Delta_\tau)$. We thus obtain that $\hat{R}_2 = \hat{R}_4 = \hat{R}_6 = \dots = 0$. This property holds for the higher-order symmetric decompositions described below, as they satisfy the relation corresponding to Eq. (32) by construction [72].

Starting with $\hat{S}_2^{(p)}(\Delta_\tau) \equiv \hat{S}_2(\Delta_\tau)$, the higher-order decomposition $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ for $m \geq 2$ that satisfies Eq. (7) can be constructed recursively as

$$\begin{aligned} \hat{S}_{2m}^{(p)}(\Delta_\tau) &= [\hat{S}_{2m-2}^{(p)}(k_m^{(p)}\Delta_\tau)]^{(p-1)/2} \\ &\times \hat{S}_{2m-2}^{(p)}(\tilde{k}_m^{(p)}\Delta_\tau) [\hat{S}_{2m-2}^{(p)}(k_m^{(p)}\Delta_\tau)]^{(p-1)/2}, \end{aligned} \quad (33)$$

where $\tilde{k}_m^{(p)} = 1 - (p-1)k_m^{(p)}$, $k_m^{(p)} = [(p-1) - (p-1)^{1/(2m-1)}]^{-1}$, and p is an odd integer with $p \geq 3$ [74]. The superscript “ (p) ” implies that $\hat{S}_{2m}^{(p)}$ consists of a product of p $\hat{S}_{2m-2}^{(p)}$ ’s. The parameter $k_m^{(p)}$ is determined so as to eliminate the residual term $x^{2m-1}\hat{R}_{2m-1}$ in $\ln \hat{S}_{2m}^{(p)}(\Delta_\tau)$ and thus

$$\hat{S}_{2m}^{(p)}(\Delta_\tau) = \exp[x\hat{\mathcal{H}} + x^{2m+1}\hat{R}_{2m+1} + \dots]. \quad (34)$$

Namely, $k_m^{(p)}$ is the solution of $(p-1)[k_m^{(p)}]^{2m-1} + [\tilde{k}_m^{(p)}]^{2m-1} = 0$ under the condition $(p-1)k_m^{(p)} + \tilde{k}_m^{(p)} = 1$. It is obvious that $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ satisfies

$$[\hat{S}_{2m}^{(p)}(\Delta_\tau)]^\dagger = \hat{S}_{2m}^{(p)}(-\Delta_\tau). \quad (35)$$

Since $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ also satisfies

$$\hat{S}_{2m}^{(p)}(\Delta_\tau)\hat{S}_{2m}^{(p)}(-\Delta_\tau) = \hat{S}_{2m}^{(p)}(-\Delta_\tau)\hat{S}_{2m}^{(p)}(\Delta_\tau) = \hat{I}, \quad (36)$$

the residual terms of even power such as $x^{2m}\hat{R}_{2m}$ are absent in the exponent of $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ in Eq. (34), shown by the same

argument for $m = 1$. Some of the higher-order symmetric Suzuki-Trotter decompositions are explicitly provided in Appendix C 1. As shown in Appendix C 2, the parameter p affects the accuracy of the decomposition for a given m .

2. Unitarity and time-reversed evolution of $\hat{S}_{2m}^{(p)}(\Delta_\tau)$

As implied in Eq. (36), $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ retains not only the unitarity but also the equivalence between the inverse and the time-reversed evolution, as given in Eq. (8). Therefore, the Hermiticity and the even dependence on Δ_τ of $\hat{\mathcal{H}}^n(\Delta_\tau)$ in Eq. (23) are both retained even when the exact time-evolution operators in Eqs. (21) and (24) are approximated by simply replacing them with $\hat{S}_{2m}^{(p)}$'s. Indeed, the main formula of the quantum power method in Eq. (3) is obtained by replacing $\hat{U}(\Delta_\tau/2)$ with $\hat{S}_{2m}^{(p)}(\Delta_\tau/2)$ in Eq. (24) and the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ satisfies Eq. (10). The same relations are also satisfied for $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ after the r th-order Richardson extrapolation, as given in Eq. (15).

In contrast to the symmetric Suzuki-Trotter decomposition, an asymmetric Suzuki-Trotter decomposition $\hat{F}(\Delta_\tau)$, such as $\hat{F}(\Delta_\tau) = e^{x\hat{\mathcal{H}}_A} e^{x\hat{\mathcal{H}}_B} \dots e^{x\hat{\mathcal{H}}_Z}$, results in

$$\left[\hat{F}(\Delta_\tau)\right]^\dagger = \left[\hat{F}(\Delta_\tau)\right]^{-1} \neq \hat{F}(-\Delta_\tau). \quad (37)$$

Thus, $\hat{F}(\Delta_\tau)$ retains the unitarity but the inverse is no longer equivalent to the time-reversed evolution. In this case, either the Hermiticity or the even dependence on Δ_τ of $\hat{\mathcal{H}}^n(\Delta_\tau)$ in Eq. (23) is violated if the exact time-evolution operators in Eqs. (21) and (24) are approximated by \hat{F} 's. For example, if we consider an operator $\hat{H}_H(\Delta_\tau) = i[\hat{F}(\Delta_\tau) - \hat{F}^\dagger(\Delta_\tau)]/\Delta_\tau$ to approximate $i[\hat{U}(\Delta_\tau) - \hat{U}(-\Delta_\tau)]/\Delta_\tau$, it satisfies the Hermiticity but is no longer an even function of Δ_τ as $\hat{H}_H(\Delta_\tau) = [\hat{H}_H(\Delta_\tau)]^\dagger \neq \hat{H}_H(-\Delta_\tau)$. On the other hand, an operator $\hat{H}_E(\Delta_\tau) = i[\hat{F}(\Delta_\tau) - \hat{F}(-\Delta_\tau)]/\Delta_\tau$ is an even function of Δ_τ but no longer satisfies the Hermiticity as $\hat{H}_E(\Delta_\tau) = \hat{H}_E(-\Delta_\tau) \neq [\hat{H}_E(\Delta_\tau)]^\dagger$. Therefore, the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ is essential for the resulting Suzuki-Trotter approximated $\hat{\mathcal{H}}^n(\Delta_\tau)$ to retain both the Hermiticity and the even dependence on Δ_τ . Note that asymmetric Suzuki-Trotter decompositions and their connection to symmetric ones have been studied in Ref. [75].

3. Circuit depth for a single time-evolution operator approximated by the Suzuki-Trotter decomposition

We now consider the circuit depth $D_{2m}^{(p)}$ required for a single time-evolution operator $\hat{U}(\Delta_\tau)$ approximated by the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}(\Delta_\tau)$, as in

Eq. (7) [also see Fig. 1(c)]. We define $D_{2m}^{(p)}$ as the number of noncommuting exponentials appearing in $\hat{S}_{2m}^{(p)}(\Delta_\tau)$. The depth of $\hat{S}_2(\Delta_\tau)$ is thus given by $D_2^{(p)} = 2N_\Gamma - 1$, as explicitly shown in Eq. (30). Since $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ consists of a product of p $\hat{S}_{2m-2}^{(p)}$'s, the depth of $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ without contracting commuting exponentials is $pD_{2m-2}^{(p)}$. However, since $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ involves $p - 1$ products of two consecutive $\hat{S}_{2m-2}^{(p)}$'s, between which two commuting exponentials reside, $p - 1$ exponentials can be contracted. We thus obtain that $D_{2m}^{(p)} = pD_{2m-2}^{(p)} - (p - 1)$ or equivalently $D_{2m}^{(p)} - 1 = p[D_{2m-2}^{(p)} - 1]$. By using this relation recursively, we can find that

$$\begin{aligned} D_{2m}^{(p)} - 1 &= p \left[D_{2m-2}^{(p)} - 1 \right] \\ &= p^2 \left[D_{2m-4}^{(p)} - 1 \right] \\ &= \dots \\ &= p^{m-1} \left[D_2^{(p)} - 1 \right]. \end{aligned} \quad (38)$$

Substituting $D_2^{(p)} = 2N_\Gamma - 1$ in Eq. (38) yields Eq. (18). Recalling that p is a typically $O(1)$ integer parameter, the depth increases exponentially with m but is independent of the number N of qubits. Therefore, the lower-order Suzuki-Trotter decomposition is highly desirable to shallow the depth of a quantum circuit.

4. Two routes for quantum power method

While the time-evolution operators satisfy the multiplication law $\hat{U}(\Delta_\tau)\hat{U}(\Delta'_\tau) = \hat{U}(\Delta_\tau + \Delta'_\tau)$, this is no longer correct when the time-evolution operators are approximated by the Suzuki-Trotter decomposition, i.e., $\hat{S}_{2m}^{(p)}(\Delta_\tau)\hat{S}_{2m}^{(p)}(\Delta'_\tau) \neq \hat{S}_{2m}^{(p)}(\Delta_\tau + \Delta'_\tau)$. Therefore, it is crucial to carefully consider when the time-evolution operators in the approximated Hamiltonian power $\hat{\mathcal{H}}^n(\Delta_\tau)$ should be replaced with the symmetric Suzuki-Trotter decomposition, either in Eq. (21) or in Eq. (24). This implies that there exist two different routes to formulate the quantum power method. Indeed, these two approaches provide two different algorithms of the quantum power method that differ in the scaling of complexity but control the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} with essentially the same accuracy. The quantum power method formulated in Sec. II B 1 is based on Eq. (24) that scales much better when the power n is large. In Appendix D, an alternative algorithm is formulated on the basis of Eq. (21), which is favored when the power n is small (e.g., $n \leq 4$ when $p = 3$).

IV. KRYLOV-SUBSPACE DIAGONALIZATION

As an application of the quantum power method, here we consider the Krylov-subspace diagonalization. We first

define a block Krylov subspace and review the subspace-diagonalization scheme [76]. We then describe how the quantum power method is combined with the Krylov-subspace diagonalization. Other applications of the quantum power method are outlined in Appendices E and F.

A. Block Krylov subspace

The block Krylov subspace of the Hamiltonian $\hat{\mathcal{H}}$ with reference states $\{|q_k\rangle\}_{k=1}^{M_B}$ is given as

$$\mathcal{K}_n(\hat{\mathcal{H}}, \{|q_k\rangle\}_{k=1}^{M_B}) = \text{span}\left(|q_1\rangle, \dots, |q_{M_B}\rangle, \hat{\mathcal{H}}|q_1\rangle, \dots, \hat{\mathcal{H}}|q_{M_B}\rangle, \dots, \hat{\mathcal{H}}^{n-1}|q_1\rangle, \dots, \hat{\mathcal{H}}^{n-1}|q_{M_B}\rangle\right), \quad (39)$$

where we call $M_B \geq 1$ the block size. We should note that the reference states $\{|q_k\rangle\}_{k=1}^{M_B}$ do not have to be orthogonal to each other but they are linearly independent. If $M_B = 1$, $\mathcal{K}_n(\hat{\mathcal{H}}, \{|q_k\rangle\}_{k=1}^{M_B})$ reduces to the conventional Krylov subspace. By defining

$$|u_i\rangle = \hat{\mathcal{H}}^{l-1}|q_k\rangle, \quad (40)$$

with $i = k + (l-1)M_B$ and $l = 1, 2, \dots, n$, the block Krylov subspace can be written simply as $\mathcal{K}_n(\hat{\mathcal{H}}, \{|q_k\rangle\}_{k=1}^{M_B}) = \text{span}\left(\{|u_i\rangle\}_{i=1}^{nM_B}\right)$.

B. Rayleigh-Ritz technique

Suppose that the ground state $|\Psi_0\rangle$ of the Hamiltonian $\hat{\mathcal{H}}$, satisfying

$$\hat{\mathcal{H}}|\Psi_0\rangle = E_0|\Psi_0\rangle \quad (41)$$

with E_0 being the ground-state energy, should be approximated with the (nonorthonormal) basis states $\{|u_i\rangle\}_{i=1}^{nM_B}$ in $\mathcal{K}_n(\hat{\mathcal{H}}, \{|q_k\rangle\}_{k=1}^{M_B})$ as

$$|\Psi_0\rangle \approx |\Psi_{\text{KS}}\rangle \equiv \sum_{i=1}^{nM_B} v_i |u_i\rangle, \quad (42)$$

where $\{v_i\}_{i=1}^{nM_B}$ are the expansion coefficients to be determined.

The expansion coefficients $\{v_i\}_{i=1}^{nM_B}$ can be determined by minimizing the energy expectation value $\langle \Psi_{\text{KS}} | \hat{\mathcal{H}} | \Psi_{\text{KS}} \rangle$ under the constraint $\langle \Psi_{\text{KS}} | \Psi_{\text{KS}} \rangle = 1$. To this end, let us

define the following function:

$$\begin{aligned} \mathcal{F}(\mathbf{v}, \mathbf{v}^*) &= \langle \Psi_{\text{KS}} | \hat{\mathcal{H}} | \Psi_{\text{KS}} \rangle - \epsilon (\langle \Psi_{\text{KS}} | \Psi_{\text{KS}} \rangle - 1) \\ &= \mathbf{v}^\dagger \mathbf{H} \mathbf{v} - \epsilon (\mathbf{v}^\dagger \mathbf{S} \mathbf{v} - 1) \\ &= \sum_{ij} v_i^* (H_{ij} - \epsilon S_{ij}) v_j + \epsilon, \end{aligned} \quad (43)$$

where ϵ is a Lagrange multiplier, $[\mathbf{v}]_i = v_i$,

$$[\mathbf{H}]_{ij} = H_{ij} = \langle u_i | \hat{\mathcal{H}} | u_j \rangle \quad (44)$$

is the subspace Hamiltonian matrix, and

$$[\mathbf{S}]_{ij} = S_{ij} = \langle u_i | u_j \rangle \quad (45)$$

is the subspace overlap matrix. Then, the condition $\partial \mathcal{F} / \partial v_i^* = 0$ for $1 \leq i \leq nM_B$ yields a generalized eigenvalue problem

$$\mathbf{H} \mathbf{v} = \epsilon \mathbf{S} \mathbf{v}. \quad (46)$$

Since both \mathbf{H} and \mathbf{S} are Hermitian, the condition $\partial \mathcal{F} / \partial v_i = 0$ for $1 \leq i \leq nM_B$ yields the same equation. The lowest eigenvalue ϵ and the corresponding eigenvector \mathbf{v} in Eq. (46) provide an approximation to the ground-state energy E_0 and the expansion coefficients $\{v_i\}_{i=1}^{nM_B}$ in Eq. (42), respectively. Note that when $M_B = 1$, the matrices \mathbf{H} and \mathbf{S} correspond to the Hankel matrices \mathcal{M}_{n-1} and \mathcal{L}_{n-1} , respectively, defined in Eqs. (F6) and (F5).

Since \mathbf{S} is a Hermitian matrix, it can be diagonalized by a unitary matrix \mathbf{V} as

$$\mathbf{V}^\dagger \mathbf{S} \mathbf{V} = \mathbf{s}, \quad (47)$$

where \mathbf{s} is the diagonal matrix that contains the eigenvalues of \mathbf{S} . Note that $\mathbf{s} > 0$ because \mathbf{S} is a Gram matrix and hence is positive definite. By using a matrix

$$\mathbf{W} = \mathbf{V} \mathbf{s}^{-1/2}, \quad (48)$$

Eq. (46) can be transformed to a standard Hermitian eigenvalue problem of the form

$$\mathbf{T} \mathbf{q} = \epsilon \mathbf{q}, \quad (49)$$

where

$$\mathbf{T} \equiv \mathbf{W}^\dagger \mathbf{H} \mathbf{W} \quad (50)$$

and $\mathbf{q} = \mathbf{W}^{-1} \mathbf{v}$. Thus, by solving the eigenvalue problem of Eq. (49), one can obtain ϵ and $\mathbf{v} = \mathbf{W} \mathbf{q}$. The eigenvector \mathbf{v} with the lowest eigenvalue ϵ provides the coefficients in the approximate ground state $|\Psi_{\text{KS}}\rangle$ [see Eq. (42)] with its

energy E_{KS} of the Hamiltonian $\hat{\mathcal{H}}$ in the Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}, \{|q_k\rangle\}_{k=1}^{M_B})$.

We note that if we use the Cholesky decomposition $\mathbf{S} = \mathbf{R}^\dagger \mathbf{R}$ with \mathbf{R} being an upper-triangular matrix, instead of the eigen decomposition in Eq. (47), \mathbf{T} reduces to the tridiagonal matrix in the Lanczos method when $M_B = 1$ [76].

C. Quantum-classical-hybrid Krylov-subspace method

Considering the Rayleigh-Ritz technique in a quantum-classical-hybrid computation, it is suited for quantum hardware to evaluate the matrix elements of \mathbf{H} in Eq. (44) and \mathbf{S} in Eq. (45), because the states $\{|u_i\rangle\}_{i=1}^{nM_B}$ are defined on the Hilbert space of $N_D = 2^N$ dimensions, for example, for the spin-1/2 Heisenberg model. On the other hand, the eigenvalue problem in the nM_B -dimensional block Krylov subspace given in Eq. (46) or Eq. (49) can be solved on classical computers, assuming that the Krylov subspace approximates reasonably well the eigenspace of the ground state with relatively small n and M_B , despite that the dimension N_D of the full Hilbert space could be much larger than nM_B . This feature is shared with other quantum-classical-hybrid subspace-diagonalization schemes reported previously [55–60].

We can now approximate the Hamiltonian power $\hat{\mathcal{H}}^{l-1}$ appearing in the Krylov-subspace basis $|u_i\rangle$ given in Eq. (40) as

$$|u_i\rangle = |\tilde{u}_i\rangle + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r}), \quad (51)$$

where

$$|\tilde{u}_i\rangle = \hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta_\tau)|q_k\rangle \quad (52)$$

with $i = k + (l-1)M_B$ for $1 \leq k \leq M_B$ and $1 \leq l \leq n$. Note that the systematic errors in Eq. (51) are absent when $l=1$. As described in Sec. II B, to approximate the Hamiltonian power $\hat{\mathcal{H}}^{l-1}$ by $\hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta_\tau)$ as in Eq. (52), the Suzuki-Trotter-decomposed time-evolution operators $\hat{S}_{2m}^{(p)}(\pm\Delta_\tau/2)$ have to be applied at most $l-1$ times to a state $|q_k\rangle$. This implies that the circuit depth required for constructing the block Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|q_k\rangle\}_{k=1}^{M_B})$ is at most $O(n)$ with a prefactor of $D_{2m}^{(p)}$. The circuit depth does not depend on the order r of the Richardson extrapolation.

With the basis states defined in Eq. (51), the subspace Hamiltonian matrix and the overlap matrix are approximated, respectively, as

$$H_{ij} = \tilde{H}_{ij} + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r}) \quad (53)$$

and

$$S_{ij} = \tilde{S}_{ij} + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r}), \quad (54)$$

where

$$\begin{aligned} [\tilde{\mathbf{H}}]_{ij} &= \tilde{H}_{ij} = \langle \tilde{u}_i | \hat{\mathcal{H}} | \tilde{u}_j \rangle \\ &= \langle q_k | \hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta_\tau) \hat{\mathcal{H}} \hat{\mathcal{H}}_{\text{ST}(r)}^{l'-1}(\Delta_\tau) | q_{k'} \rangle \end{aligned} \quad (55)$$

and

$$[\tilde{\mathbf{S}}]_{ij} = \tilde{S}_{ij} = \langle \tilde{u}_i | \tilde{u}_j \rangle = \langle q_k | \hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta_\tau) \hat{\mathcal{H}}_{\text{ST}(r)}^{l'-1}(\Delta_\tau) | q_{k'} \rangle, \quad (56)$$

with $i = k + (l-1)M_B$ and $j = k' + (l'-1)M_B$ for $1 \leq k, k' \leq M_B$ and $1 \leq l, l' \leq n$ in the block Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|q_k\rangle\}_{k=1}^{M_B})$. Here, the Hermiticity of the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta_\tau)$ in Eq. (15) is used. Note also that $\hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta_\tau) \hat{\mathcal{H}}_{\text{ST}(r)}^{l'-1}(\Delta_\tau) \neq \hat{\mathcal{H}}_{\text{ST}(r)}^{l+l'-2}(\Delta_\tau)$ for $r \geq 1$, but this equation is satisfied when $r = 0$.

More specifically, \tilde{H}_{ij} and \tilde{S}_{ij} in terms of $\hat{S}_{2m}^{(p)}(\pm\Delta_\tau/2)$ without the Richardson extrapolation are given, respectively, as

$$\begin{aligned} \tilde{H}_{ij} &= \sum_{v=0}^{l-1} \sum_{v'=0}^{l'-1} c_{l-1,v}^* c_{l'-1,v'} \langle q_k | \left[\hat{S}_{2m}^{(p)}\left(-\frac{\Delta_\tau}{2}\right) \right]^{l-1-2v} \\ &\quad \times \hat{\mathcal{H}} \left[\hat{S}_{2m}^{(p)}\left(\frac{\Delta_\tau}{2}\right) \right]^{l'-1-2v'} | q_{k'} \rangle \end{aligned} \quad (57)$$

and

$$\begin{aligned} \tilde{S}_{ij} &= \sum_{v=0}^{l-1} \sum_{v'=0}^{l'-1} c_{l-1,v}^* c_{l'-1,v'} \langle q_k | \left[\hat{S}_{2m}^{(p)}\left(-\frac{\Delta_\tau}{2}\right) \right]^{l-1-2v} \\ &\quad \times \left[\hat{S}_{2m}^{(p)}\left(\frac{\Delta_\tau}{2}\right) \right]^{l'-1-2v'} | q_{k'} \rangle. \end{aligned} \quad (58)$$

Note that the Suzuki-Trotter-decomposed time-evolution operators $[\hat{S}_{2m}^{(p)}(-\Delta_\tau/2)]^{l-1-2v}$ and $[\hat{S}_{2m}^{(p)}(\Delta_\tau/2)]^{l'-1-2v'}$ in Eq. (58) can be combined exactly as in the form shown in Eq. (64) when the Richardson extrapolation is not used. However, here we deliberately do not combine these two terms because it is helpful when the extension for the r th-order Richardson extrapolation is considered. Assuming that $\hat{\mathcal{H}}$ consists of $O(N)$ local terms, the number of state overlaps required for constructing all matrix elements of $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$ is $O(n^2 M_B^2 N)$ and $O(n^2 M_B^2)$, respectively. If the r th-order Richardson extrapolation is employed, the number of state overlaps to be evaluated is increased by a factor of $(r+1)^2$. The state overlaps in Eqs. (57) and (58) can be evaluated with an Hadamard-test-like circuit, for example [50,77–79].

However, for the purpose of solving the generalized eigenvalue problem in Eq. (46) or the corresponding standard eigenvalue problem in Eq. (49), one could evaluate

the matrix elements in Eqs. (44) and (45) more directly as

$$H_{ij} = \tilde{H}'_{ij} + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r}) \quad (59)$$

and

$$S_{ij} = \tilde{S}'_{ij} + O(\Delta_\tau^{2+2r}) + O(\Delta_\tau^{2m+2r}), \quad (60)$$

where

$$[\tilde{\mathbf{H}}']_{ij} = \tilde{H}'_{ij} = \langle q_k | \hat{\mathcal{H}}_{\text{ST}(r)}^{l+l'-1}(\Delta_\tau) | q_{k'} \rangle \quad (61)$$

and

$$[\tilde{\mathbf{S}}']_{ij} = \tilde{S}'_{ij} = \langle q_k | \hat{\mathcal{H}}_{\text{ST}(r)}^{l+l'-2}(\Delta_\tau) | q_{k'} \rangle, \quad (62)$$

with $i = k + (l-1)M_B$ and $j = k' + (l'-1)M_B$ for $1 \leq k, k' \leq M_B$ and $1 \leq l, l' \leq n$ in the block Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|q_k\rangle\}_{k=1}^{M_B})$. To be more specific, the matrix elements of $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$ for $r=0$, i.e., without the Richardson extrapolation, are given as

$$\tilde{H}'_{ij} = \sum_{v=0}^{l+l'-1} c_{l+l'-1,v} \langle q_k | \left[\hat{S}_{2m}^{(p)} \left(\frac{\Delta_\tau}{2} \right) \right]^{l+l'-1-2v} | q_{k'} \rangle \quad (63)$$

and

$$\tilde{S}'_{ij} = \sum_{v=0}^{l+l'-2} c_{l+l'-2,v} \langle q_k | \left[\hat{S}_{2m}^{(p)} \left(\frac{\Delta_\tau}{2} \right) \right]^{l+l'-2-2v} | q_{k'} \rangle. \quad (64)$$

The number of state overlaps required for constructing all matrix elements of both $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$ is thus $O(nM_B^2)$. If the r th-order Richardson extrapolation is employed, the number of state overlaps to be evaluated is increased by a factor of $(r+1)$.

Therefore, the approach based on Eqs. (61) and (62) is better than that based on Eqs. (55) and (56) in the sense that fewer state overlaps are required to approximately solve the Krylov-subspace diagonalization. However, although these two approaches are equivalent within the systematic errors, the approach based on Eqs. (61) and (62) loses the exact meaning of the variational principle for the ground state obtained by solving the (generalized) eigenvalue problem (also see Ref. [51]). This is because the approach based on Eqs. (55) and (56) respects the subspace structure, which is generated by the Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|q_k\rangle\}_{k=1}^{M_B})$, as opposed to the other approach, and thus the lowest eigenstate of the (generalized) eigenvalue problem with the matrix elements in Eqs. (55) and (56) satisfies exactly the variational principle. In addition, we find that the approach based on Eqs. (55) and (56) is more stable for numerical simulations. Therefore, we adopt the approach based on Eqs. (55)

and (56) in our numerical simulations shown in this paper unless otherwise stated. In Appendix A, our method described here is compared with other algorithms for the Krylov-subspace diagonalization.

V. NUMERICAL DEMONSTRATION

In this section, we demonstrate the quantum power method by numerically simulating the spin-1/2 Heisenberg model. We first define the Hamiltonian of the Heisenberg model, and then show how the quantum power method can control the systematic errors in approximating the Hamiltonian power $\hat{\mathcal{H}}^n$. Next, as an application of the quantum power method, we show the numerical results of the multireference Krylov-subspace diagonalization combined with the quantum power method for the Heisenberg model. The numerical results of the multireference Krylov-subspace diagonalization combined with the quantum power method for the Fermi-Hubbard model, which involves more technical details, are also provided in Appendix B.

A. Heisenberg model

The spin-1/2 Heisenberg model is described by the following Hamiltonian:

$$\hat{\mathcal{H}} = \frac{J}{4} \sum_{\langle i,j \rangle} (\hat{I}_i \hat{I}_j + \hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j + \hat{Z}_i \hat{Z}_j) = \frac{J}{2} \sum_{\langle i,j \rangle} \hat{\mathcal{P}}_{ij}, \quad (65)$$

where $J > 0$ is the antiferromagnetic (AFM) exchange interaction, $\langle i,j \rangle$ runs over all nearest-neighbor pairs of qubits i and j connected with the exchange interaction J , and $\{\hat{X}_i, \hat{Y}_i, \hat{Z}_i\}$ and \hat{I}_i are the Pauli operators and the identity operator acting on the i th qubit. $\hat{\mathcal{P}}_{ij}$ is the SWAP operator, which acts on the i th and j th qubits as $\hat{\mathcal{P}}_{ij} |a\rangle_i |b\rangle_j = |b\rangle_i |a\rangle_j$. In the Hamiltonian in Eq. (65), the constant (identity) term $\hat{I}_i \hat{I}_j$ is added to the conventional Heisenberg Hamiltonian and thus $\hat{\mathcal{H}}$ is simply a sum of SWAP operators. Indeed, the second equality in Eq. (65) follows from the identity

$$\hat{I}_i \hat{I}_j + \hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j + \hat{Z}_i \hat{Z}_j = 2\hat{\mathcal{P}}_{ij} \quad (66)$$

for $i \neq j$.

We consider the Hamiltonian $\hat{\mathcal{H}}$ on a one-dimensional periodic chain with N sites (i.e., N qubits), and assume that N is even. Then, the Hamiltonian can be written as

$$\hat{\mathcal{H}} = \frac{J}{2} \sum_{i=1}^N \hat{\mathcal{P}}_{i,i+1}, \quad (67)$$

where $i+1$ in the subscript should be read as 1 if $i=N$ due to the periodic boundary conditions. For the use in the

Suzuki-Trotter decomposition of the time-evolution operator, we divide the Hamiltonian into two parts (i.e., $N_\Gamma = 2$) as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B, \quad (68)$$

with

$$\hat{\mathcal{H}}_A = \frac{J}{2} \sum_{i=1}^{N/2} \hat{\mathcal{P}}_{2i,2i+1} \quad (69)$$

and

$$\hat{\mathcal{H}}_B = \frac{J}{2} \sum_{i=1}^{N/2} \hat{\mathcal{P}}_{2i-1,2i}. \quad (70)$$

Notice that $[\hat{\mathcal{P}}_{2i,2i+1}, \hat{\mathcal{P}}_{2j,2j+1}] = [\hat{\mathcal{P}}_{2i-1,2i}, \hat{\mathcal{P}}_{2j-1,2j}] = 0$, where $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the commutator of two operators \hat{A} and \hat{B} .

For the one-dimensional spin-1/2 Heisenberg model $\hat{\mathcal{H}} = \hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B$ given in Eqs. (65) and (68), the time-evolution operator $\hat{U}(t) = e^{-i\hat{\mathcal{H}}t}$ is constituted by two elementary time-evolution operators associated with $\hat{\mathcal{H}}_A$ and $\hat{\mathcal{H}}_B$. Let us first introduce the exponential-SWAP (e -SWAP) gate \hat{U}_{ij} [80–84]

$$\hat{U}_{ij}(\theta) = \exp(-i\theta\hat{\mathcal{P}}_{ij}/2), \quad (71)$$

where θ is a real-valued parameter. The e -SWAP gate, which is equivalent to the SWAP^α gate up to a two-qubit global phase factor [43,85–87], is depicted schematically in Fig. 1(c) as a blue rectangular extended over two qubits. The gate corresponding to Eq. (71) can be implemented with three CNOT gates and few single-qubit rotations [88–90]. The time-evolution operators of $\hat{\mathcal{H}}_A$ and $\hat{\mathcal{H}}_B$ are given, respectively, by

$$\exp(-i\hat{\mathcal{H}}_A t) = \prod_{i=1}^{N/2} \hat{U}_{2i,2i+1}(tJ) \quad (72)$$

and

$$\exp(-i\hat{\mathcal{H}}_B t) = \prod_{i=1}^{N/2} \hat{U}_{2i-1,2i}(tJ). \quad (73)$$

Since $[\hat{U}_{2i,2i+1}, \hat{U}_{2j,2j+1}] = 0$ and $[\hat{U}_{2i-1,2i}, \hat{U}_{2j-1,2j}] = 0$ for $i \neq j$, the order of the product is not relevant in Eqs. (72) and (73). Figure 1(c) illustrates a typical circuit structure that approximates the time-evolution operator $\hat{U}(\Delta_\tau)$, consisting of a product of $\exp(-i\hat{\mathcal{H}}_A \Delta_\tau s_i)$'s and $\exp(-i\hat{\mathcal{H}}_B \Delta_\tau s_i)$'s with real parameters $\{s_i\}$. As in

Eq. (30) for the general case, the lowest-order symmetric Suzuki-Trotter decomposition of $\hat{U}(\Delta_\tau)$ for the bipartitioned Hamiltonian $\hat{\mathcal{H}} = \hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B$ is given by

$$\hat{S}_2(\Delta_\tau) = e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_A} e^{-i\Delta_\tau\hat{\mathcal{H}}_B} e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_A}. \quad (74)$$

B. Degree of approximation

We first examine quantitatively how the Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated by $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$. For this purpose, we define a distance $d(\hat{A}, \hat{B})$ between operators \hat{A} and \hat{B} as

$$d(\hat{A}, \hat{B}) = \sqrt{1 - \frac{|\langle \hat{A}, \hat{B} \rangle_F|}{\|\hat{A}\|_F \|\hat{B}\|_F}}, \quad (75)$$

where $\langle \hat{A}, \hat{B} \rangle_F$ denotes the Frobenius inner product between \hat{A} and \hat{B} defined by

$$\langle \hat{A}, \hat{B} \rangle_F = \text{Tr}[\hat{A}^\dagger \hat{B}] \quad (76)$$

and $\|\hat{A}\|_F$ denotes the Frobenius norm of \hat{A} , i.e.,

$$\|\hat{A}\|_F = \sqrt{\text{Tr}[\hat{A}^\dagger \hat{A}]}. \quad (77)$$

Note that $\langle \hat{A}, \hat{A} \rangle_F = \|\hat{A}\|_F^2$, $0 \leq |\langle \hat{A}, \hat{B} \rangle_F| \leq \|\hat{A}\|_F \|\hat{B}\|_F$, $0 \leq d(\hat{A}, \hat{B}) \leq 1$, $d(\hat{A}, \hat{B}) = d(a\hat{A}, b\hat{B})$ with a and b being nonzero complex numbers, and $d(\hat{A}, \hat{B}) = 0$ if and only if $\hat{A} = \hat{B}$. We compute the distance $d(\hat{A}, \hat{B})$ for $\hat{A} = \hat{\mathcal{H}}^n$ and $\hat{B} = \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ given in Eq. (4) for $r = 0$ and Eq. (14) for $r \geq 1$. The Hamiltonian $\hat{\mathcal{H}}$ is for the spin-1/2 Heisenberg model on an N -qubit ring given in Eq. (65).

Evaluating the distance is costly as it demands matrix-matrix multiplications or diagonalizations. To avoid such costly operations, we employ a stochastic evaluation of the trace as [91–95]

$$\text{Tr}[\hat{\mathcal{X}}] = \lim_{R \rightarrow \infty} \frac{1}{R} \sum_{\zeta=1}^R \langle \phi_\zeta | \hat{\mathcal{X}} | \phi_\zeta \rangle, \quad (78)$$

where $\hat{\mathcal{X}} \in \{\hat{A}^\dagger \hat{A}, \hat{B}^\dagger \hat{B}, \hat{A}^\dagger \hat{B}\}$ and

$$|\phi_\zeta\rangle = \sum_x e^{i\phi_\zeta(x)} |x\rangle \quad (79)$$

is a random-phase state with $\{|x\rangle\}$ being a complete orthonormal basis set such that $\langle x|x'\rangle = \delta_{xx'}$ and $\phi_\zeta(x)$ being a random variable drawn uniformly from $[0, 2\pi)$.

Note that $\langle \phi_\zeta | \phi_\zeta \rangle = 2^N$, i.e., the dimension N_D of the Hilbert space. We choose $\{|x\rangle\}$ as the orthonormal basis set that diagonalizes the local Pauli Z operators. The stochastic evaluation of the trace in Eq. (78) requires only sparse matrix-vector multiplications and a single inner-product calculation for each set ζ of random numbers $\phi_\zeta(x)$, if $\hat{\mathcal{X}}$ is represented as a product of sparse matrices, which is indeed the case here. Instead of taking the limit $R \rightarrow \infty$, we fix $R = 16$ for $N \geq 12$ and $R = 256$ for $N = 10$ and estimate error bars. Since $\langle \phi_\zeta | \hat{A}^\dagger \hat{A} | \phi_\zeta \rangle$, $\langle \phi_\zeta | \hat{B}^\dagger \hat{B} | \phi_\zeta \rangle$, and $\langle \phi_\zeta | \hat{A}^\dagger \hat{B} | \phi_\zeta \rangle$ for $\hat{A} = \hat{\mathcal{H}}^n$ and $\hat{B} = \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ are highly correlated to each other, error bars of $d(\hat{A}, \hat{B})$ must be estimated using the corresponding 3×3 covariance matrix.

Figure 3 shows the distance as a function of Δ_τ for $n = 1, 2$, and 3 with $N = 10, 12, 14, 16, 18, 20, 22$, and 24 using the symmetric Suzuki-Trotter decompositions \hat{S}_2 and $\hat{S}_4^{(3)}$. Figures 3(a)–3(c) show the results without the Richardson extrapolation ($r = 0$). Since the leading systematic error in $\hat{\mathcal{H}}_{\text{ST}(r=0)}^n(\Delta_\tau)$ is $O(\Delta_\tau^2)$, the distance scales almost linearly

in Δ_τ^2 for each N . The distance simply increases with increasing N and n . Figures 3(d)–3(f) show the results with the first-order Richardson extrapolation ($r = 1$). For each n , the distance with the Richardson extrapolation is an order of magnitude smaller than that without the Richardson extrapolation. The leading systematic error in $\hat{\mathcal{H}}_{\text{ST}(r=1)}^n(\Delta_\tau)$ is $O(\Delta_\tau^4)$, and the distance indeed scales almost linearly in Δ_τ^4 . As expected from Eq. (13), essentially no difference can be found between the results with \hat{S}_2 and $\hat{S}_4^{(3)}$, indicated, respectively, by empty and filled symbols in Fig. 3. These results clearly demonstrate that the systematic errors in approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ are well controlled.

Figure 4(a) shows the n dependence of the distance for $N = 24$ with various values of Δ_τ calculated using the lowest-order symmetric Suzuki-Trotter decomposition \hat{S}_2 . The distance first increases with n and tends to saturate at $n \sim 100$. It is remarkable to find in Fig. 4(b) that, even with the large power exponents as large as $n = 100$, the linear

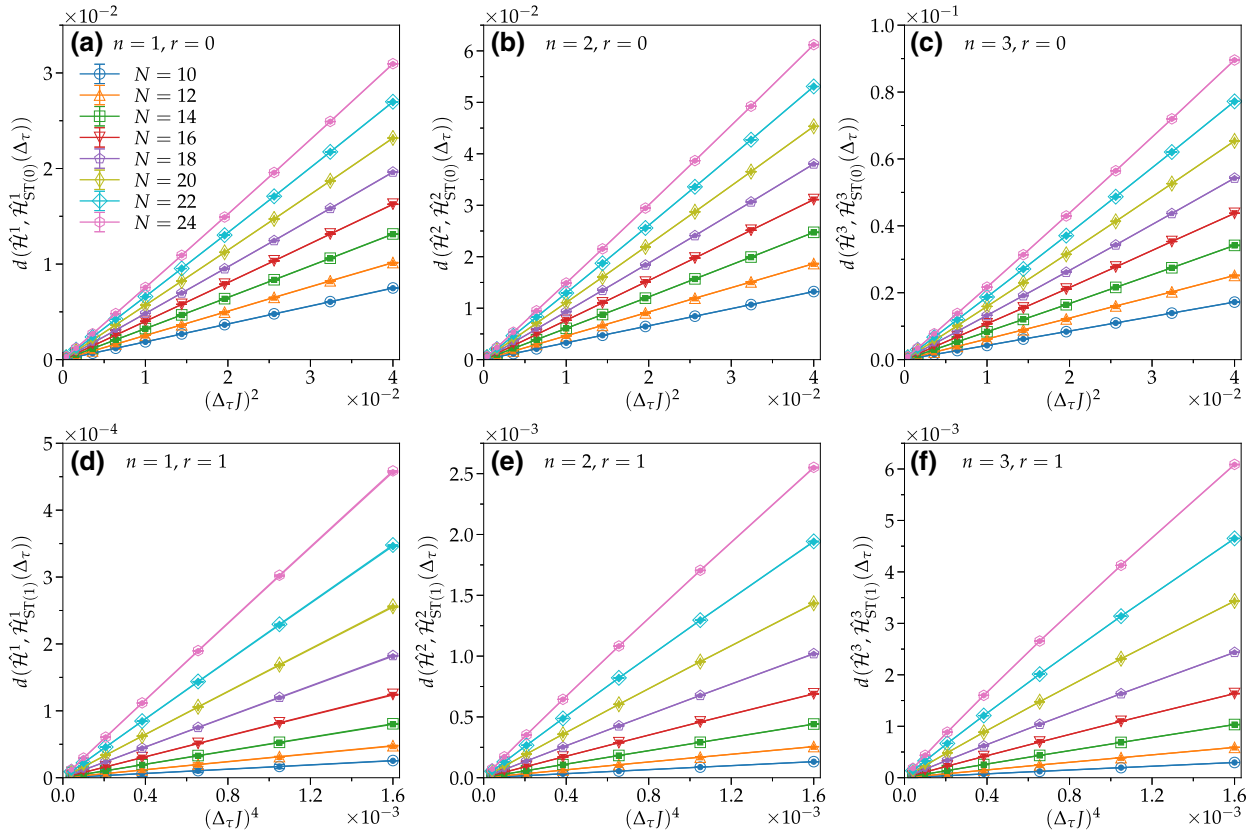


FIG. 3. (a)–(c) Distance $d(\hat{\mathcal{H}}^n, \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau))$ between the exact Hamiltonian power $\hat{\mathcal{H}}^n$ and the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ given in Eq. (4) without the Richardson extrapolation ($r = 0$) as a function of Δ_τ^2 for (a) $n = 1$, (b) $n = 2$, and (c) $n = 3$. (d)–(f) Same as (a)–(c) but with the first-order Richardson extrapolation ($r = 1$) given in Eq. (14) as a function of Δ_τ^4 . The Hamiltonian $\hat{\mathcal{H}}$ is for the spin-1/2 Heisenberg model on an N -qubit ring given in Eq. (65). The symmetric Suzuki-Trotter decompositions \hat{S}_2 (empty symbols) and $\hat{S}_4^{(3)}$ (filled small symbols) are used. However, these results are on top of each other, as is expected. The error bar indicates the standard error of the mean. The solid lines are a guide for the eye. Note that each panel employs a different scale in the y axis.

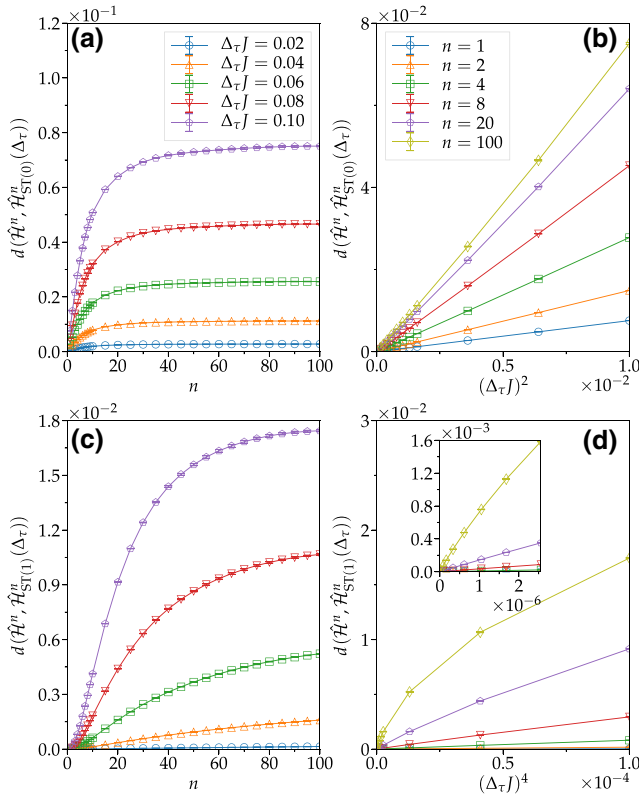


FIG. 4. (a) Distance $d(\hat{\mathcal{H}}^n, \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau))$ between the exact Hamiltonian power $\hat{\mathcal{H}}^n$ and the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ given in Eq. (4) without the Richardson extrapolation ($r = 0$) as a function of n for different values of Δ_τ . (b) Same as (a) but as a function of Δ_τ^2 for different values of the power n . (c) Same as (a) but with the first-order Richardson extrapolation ($r = 1$) given in Eq. (14). (d) Same as (c) but as a function of Δ_τ^4 for different values of the power n . The inset in (d) shows the enlarged plot for $\Delta_\tau J \leq 0.04$. The Hamiltonian $\hat{\mathcal{H}}$ is for the spin-1/2 Heisenberg model on an $N = 24$ qubit ring given in Eq. (65). The lowest-order symmetric Suzuki-Trotter decomposition \hat{S}_2 is used. The error bar indicates the standard error of the mean. The solid lines are a guide for the eye. Note that each panel employs a different scale in the y axis.

dependence of the distance on Δ_τ^2 remains in a wide range of Δ_τ ($\Delta_\tau J \lesssim 0.1$) and the distance is smoothly extrapolated to zero in the limit of $\Delta_\tau \rightarrow 0$, clearly demonstrating the controlled accuracy of the quantum power method. Figures 4(c) and 4(d) show the same results but obtained by using the first-order Richardson extrapolation ($r = 1$), for which the systematic errors in approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ are expected to be $O(\Delta_\tau^4)$. Indeed, our numerical simulations find the linear dependence of the distance on Δ_τ^4 for at least $\Delta_\tau J \lesssim 0.05$ when $n = 100$ [see the inset in Fig. 4(d)]. Notice also that the distance itself becomes smaller by a factor of approximately 5 even for

large n when the first-order Richardson extrapolation is employed.

C. Krylov-subspace diagonalization

We now perform numerical simulations of the Krylov-subspace diagonalization combined with the quantum power method to calculate the ground-state energy and fidelity of the spin-1/2 Heisenberg model described by the Hamiltonian $\hat{\mathcal{H}}$ in Eq. (65) on a periodic chain of $N = 16$ sites (i.e., qubits).

Considering the Krylov-subspace diagonalization as an application of the quantum power method on near-term quantum computers, it is crucial to reduce the circuit depth. As discussed in Secs. II B and IV C, the depth of the circuit required for constructing the block Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|q_k\rangle\}_{k=1}^{M_B})$ scales as $O(n)$ with a prefactor $D_{2m}^{(p)}$. Since m and p in the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}$ can be set to the minimum values $m = 1$ and $p = 3$, at least for the system sizes examined in the previous section including $N = 16$, the primary objective here is to reduce the power n . For this purpose, we first describe the selection of the reference states, aiming that the block Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|q_k\rangle\}_{k=1}^{M_B})$ spanned by these reference states can approximate reasonably well the target subspace, which in the present case is the eigenspace of the ground state of $\hat{\mathcal{H}}$. Then, we show by numerical simulations how the selection of the reference states affects the convergence to the ground state with n .

1. Selection of reference states

Equation (42) suggests that the ground state $|\Psi_0\rangle$ can be well approximated if the reference states $\{|q_k\rangle\}_{k=1}^{M_B}$ are chosen so that these states have substantial overlap with the exact ground state. Therefore, as the reference states, we introduce the following product states for the subspace diagonalization:

$$|q_1\rangle = |\Phi_A\rangle = \otimes_{i=1}^{N/2} |s_{2i,2i+1}\rangle, \quad (80)$$

$$|q_2\rangle = |\Phi_B\rangle = \otimes_{i=1}^{N/2} |s_{2i-1,2i}\rangle, \quad (81)$$

$$|q_3\rangle = |X_{\text{AFM}1}\rangle = \otimes_{i=1}^{N/2} |+\rangle_{2i-1} |-\rangle_{2i}, \quad (82)$$

$$|q_4\rangle = |X_{\text{AFM}2}\rangle = \otimes_{i=1}^{N/2} |+\rangle_{2i} |-\rangle_{2i+1}, \quad (83)$$

$$|q_5\rangle = |Y_{\text{AFM}1}\rangle = \otimes_{i=1}^{N/2} |R\rangle_{2i-1} |L\rangle_{2i}, \quad (84)$$

$$|q_6\rangle = |Y_{\text{AFM}2}\rangle = \otimes_{i=1}^{N/2} |R\rangle_{2i} |L\rangle_{2i+1}, \quad (85)$$

$$|q_7\rangle = |Z_{\text{AFM}1}\rangle = \otimes_{i=1}^{N/2} |0\rangle_{2i-1} |1\rangle_{2i}, \quad (86)$$

$$|q_8\rangle = |Z_{\text{AFM}2}\rangle = \otimes_{i=1}^{N/2} |0\rangle_{2i} |1\rangle_{2i+1}. \quad (87)$$

Here $|s_{ij}\rangle = (1/\sqrt{2})(|0\rangle_i |1\rangle_j - |1\rangle_i |0\rangle_j)$ is the spin-singlet state and is an eigenstate of the SWAP operator \hat{P}_{ij}

with eigenvalue -1 . It is also known as one of the Bell states. $|+\rangle_i = (1/\sqrt{2})(|0\rangle_i + |1\rangle_i)$ and $|-\rangle_i = (1/\sqrt{2})(|0\rangle_i - |1\rangle_i)$ are the eigenstates of \hat{X}_i with eigenvalues ± 1 , $|R\rangle_i = (1/\sqrt{2})(|0\rangle_i + i|1\rangle_i)$ and $|L\rangle_i = (1/\sqrt{2})(|0\rangle_i - i|1\rangle_i)$ are the eigenstates of \hat{Y}_i with eigenvalues ± 1 , and $|0\rangle_i$ and $|1\rangle_i$ are the eigenstates of \hat{Z}_i with eigenvalues ± 1 . $|\Phi_A\rangle$ and $|\Phi_B\rangle$ are the ground states of $\hat{\mathcal{H}}_A$ and $\hat{\mathcal{H}}_B$, respectively, while others are the AFM Néel states that are the ground states when a mean-field theory is applied to the Hamiltonian. These product states are expected to have a sizable overlap with the exact ground state (also see Fig. 6) and, moreover, they are easy to be prepared from $|0\rangle^{\otimes N}$ with appropriate combinations of Pauli, Hadamard, phase, and CNOT gates.

Another relevant candidate might be a variational state that has a substantial overlap with the ground state. We thus introduce

$$|q_9\rangle = |\Psi_{\text{VQE}}\rangle \quad (88)$$

as another reference state, where $|\Psi_{\text{VQE}}\rangle$ is an approximate ground state prepared with a VQE scheme. Specifically, we choose $|\Psi_{\text{VQE}}\rangle$ as a resonating-valence-bond-type wave function without the symmetry projection operator, containing 64 optimized variational parameters for $N = 16$ that do not reflect the spatial symmetry of the Hamiltonian, as reported in Ref. [96]. While the exact ground-state energy is $E_0/NJ = -0.196393522$, our variational state $|\Psi_{\text{VQE}}\rangle$ has the variational energy $\langle \Psi_{\text{VQE}} | \hat{\mathcal{H}} | \Psi_{\text{VQE}} \rangle / NJ = -0.1885$ (also see Fig. 5) and the ground-state fidelity $|\langle \Psi_0 | \Psi_{\text{VQE}} \rangle|^2 = 0.771$ (also see Fig. 6).

In our previous study [96], we have shown that restoration of the spatial symmetry that is broken by a circuit ansatz greatly improves the ground-state-energy estimation as well as the ground-state fidelity. Motivated by this finding, we introduce another set of the reference states $\{|\bar{q}_k\rangle\}_{k=1}^N$ with

$$|\bar{q}_k\rangle = \hat{\mathcal{T}}_{k-1} |\Psi_{\text{VQE}}\rangle, \quad (89)$$

where $\hat{\mathcal{T}}_k$ is a unitary operator representing the one-dimensional k -lattice-space translation with $\hat{\mathcal{T}}_0 = \hat{\mathcal{T}}$, and $|\Psi_{\text{VQE}}\rangle$ is the same state given in Eq. (88). With this set of the reference states, the translational symmetry that is broken in the apparent circuit structure of $|\Psi_{\text{VQE}}\rangle$ can be restored as a linear combination of the states in the block Krylov subspace, without applying a projection operator to $|\Psi_{\text{VQE}}\rangle$. For example, a simple sum of these N reference states $\{|\bar{q}_k\rangle\}_{k=1}^N$, i.e., $\sum_{k=1}^N |\bar{q}_k\rangle$, is translationally symmetric with momentum zero.

The reference states $|\Phi_A\rangle$, $|\Phi_B\rangle$, $|\Psi_{\text{VQE}}\rangle$, and $\{|\bar{q}_k\rangle\}_{k=1}^N$ introduced above are all spin-singlet states, i.e., the total spin and the Z component of the total spin being zero, while the X , Y , and Z components of the total spin are

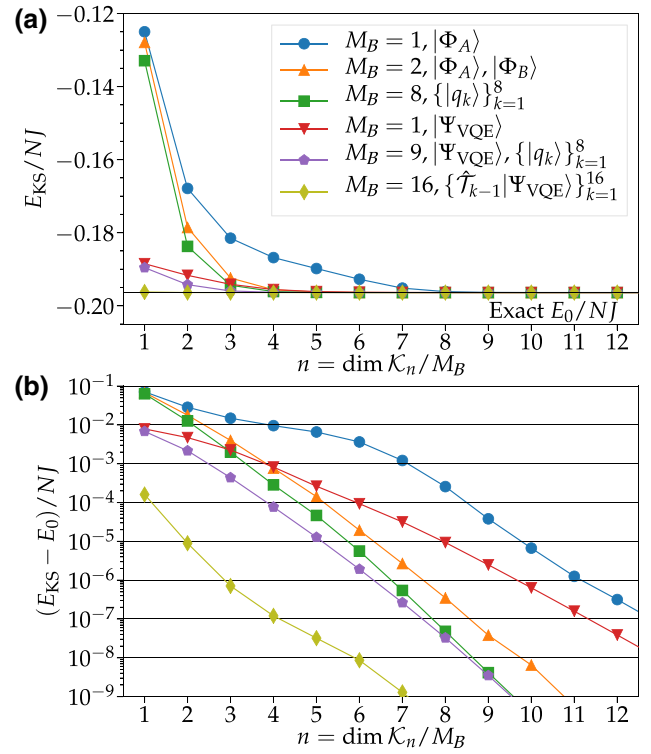


FIG. 5. (a) Ground-state energy E_{KS} for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring as a function of the dimension of the Krylov subspace \mathcal{K}_n per block size M_B , $n = \dim \mathcal{K}_n/M_B$, with various sets of the reference states. The horizontal line indicates the exact ground-state energy E_0 . The results are obtained with $\Delta_\tau J = 0.05$, $r = 1$, $m = 1$, and $p = 3$. (b) Same as (a) but a semilog plot of the energy difference $E_{\text{KS}} - E_0$ as a function of n .

zero for the reference states $|X_{\text{AFM } 1(2)}\rangle$, $|Y_{\text{AFM } 1(2)}\rangle$, and $|Z_{\text{AFM } 1(2)}\rangle$, respectively. Because the Hamiltonian $\hat{\mathcal{H}}$ considered here is spin $SU(2)$ symmetric and the quantum power method preserves the Hamiltonian symmetry as shown in Eq. (16), the Krylov subspace generated from these reference states remains in the same symmetry sector of the Hilbert space as the reference states. We select these reference states because it is known that the ground state of the spin-1/2 Heisenberg model considered here is spin singlet [97,98].

2. Ground-state energy and fidelity

Figures 5 and 6 show the estimated ground-state energy E_{KS} and the ground-state fidelity $F = |\langle \Psi_0 | \Psi_{\text{KS}} \rangle|^2$, obtained by solving Eq. (49), as a function of $n = \dim \mathcal{K}_n/M_B$, i.e., the dimension of the Krylov subspace \mathcal{K}_n per block size M_B . Note that $\hat{\mathcal{H}}_{\text{ST}(r)}^{n-1}(\Delta_\tau)$ is the maximum approximated Hamiltonian power multiplied to the reference states when the Krylov subspace $\mathcal{K}_n \left(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta_\tau), \{|\bar{q}_k\rangle\}_{k=1}^{M_B} \right)$ is constructed in Eq. (39).

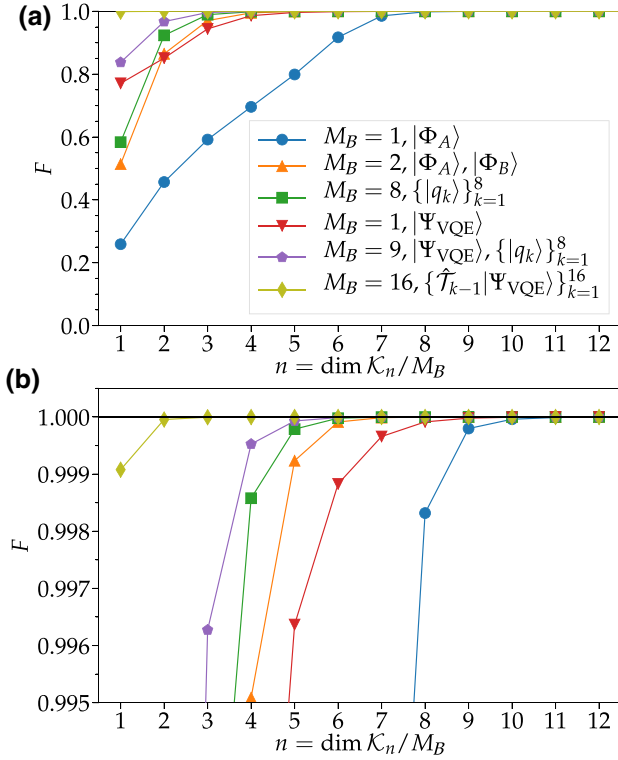


FIG. 6. (a) Same as Fig. 5(a) but for the ground-state fidelity $F = |\langle \Psi_0 | \Psi_{\text{KS}} \rangle|^2$. Here, $|\Psi_0\rangle$ is the exact ground state and $|\Psi_{\text{KS}}\rangle$ is the approximate ground state with the corresponding energy E_{KS} shown in Fig. 5. Both states are assumed to be normalized. (b) Enlarged plot of (a).

Here, the Krylov-subspace Hamiltonian matrix $[\tilde{\mathbf{H}}]_{ij}$ and the overlap matrix $[\tilde{\mathbf{S}}]_{ij}$ are computed as $\langle \tilde{u}_i | \hat{\mathcal{H}} | \tilde{u}_j \rangle$ and $\langle \tilde{u}_i | \tilde{u}_j \rangle$ in Eqs. (55) and (56), respectively. The first-order Richardson extrapolation ($r = 1$) and the lowest-order symmetric Suzuki-Trotter decomposition \hat{S}_2 are used for $\{\hat{\mathcal{H}}_{\text{ST}(r)}^l(\Delta\tau)\}_{l=1}^{n-1}$ with $\Delta\tau J = 0.05$, in which the systematic errors are practically negligible for our purpose, as discussed later in Fig. 7 (also see Figs. 3 and 4).

Let us first focus on the results for $n = 1$, where no Hamiltonian power is incorporated in the Krylov subspace. It is not surprising to find that the energy and the fidelity are substantially improved if the reference states include the VQE state $|\Psi_{\text{VQE}}\rangle$. The improvement is even more significant if we incorporate the spatially translated VQE states $\{|\tilde{q}_k\rangle\}_{k=1}^N$. Note that, if $M_B = 1$, the energies plotted at $n = 1$ are merely the expectation values of $\hat{\mathcal{H}}$ with respect to the corresponding reference state, e.g., $\langle \Phi_A | \hat{\mathcal{H}} | \Phi_A \rangle / NJ = -0.125$ and $\langle \Psi_{\text{VQE}} | \hat{\mathcal{H}} | \Psi_{\text{VQE}} \rangle / NJ = -0.1885$. The multireference scheme with $M_B > 1$ further decreases the energy and improves the fidelity without applying the Hamiltonian power to the reference states.

With increasing n , the energy decreases monotonically (Fig. 5) and the fidelity keeps increasing towards 1 (Fig. 6),

implying that the ground-state estimation can be improved systematically over a chosen set of reference states without any parameter optimization, at a cost of the linearly increasing circuit depth that scales at most as $2n - 1$. The nearly linear behavior of $E_{\text{KS}} - E_0$ in the semilog plot shown in Fig. 5(b) suggests the exponential convergence to the exact ground-state energy as a function of n , as in the Lanczos method [11]. Notice also that the energy as well as the fidelity for $M_B = 16$ is consistently better than those for $M_B \leq 9$ for every n . Moreover, the slope in the semilog plot of $E_{\text{KS}} - E_0$ and also the slope of the fidelity tend to be steeper for $M_B > 1$ than for $M_B = 1$ [see Figs. 5(b) and 6(a)], implying that the convergence towards the ground state is improved more efficiently in the multireference scheme with $M_B > 1$. Interestingly, even if $|\Psi_{\text{VQE}}\rangle$ is not included in a set of reference states, the multireference schemes with $M_B = 2$ and $M_B = 8$ surpass the scheme including only $|\Psi_{\text{VQE}}\rangle$ with $M_B = 1$ at $n = 5$ and 3, respectively, in terms of the ground-state energy E_{KS} . Therefore, the multireference scheme with $M_B > 1$ works effectively for reducing n and hence the number of gates in a circuit, even if simple product states with no variational parameters are chosen for the reference states. Table II summarizes the minimum dimension n of the Krylov subspace per block size and the corresponding circuit depth required for converging the ground-state energy E_{KS} with an accuracy $(E_{\text{KS}} - E_0)/NJ \leq 10^{-4}$ for $N = 16$. Note here that the commuting exponentials in $[\hat{S}_2(\pm\Delta\tau/2)]^{n-1}$ are contracted when the circuit depth is counted.

3. $\Delta\tau$ dependence of the ground-state energy and extrapolation to $\Delta\tau \rightarrow 0$

Finally, we describe a strategy for performing the Krylov-subspace diagonalization combined with the

TABLE II. Minimum dimension n of the Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta\tau), \{|q_k\rangle\}_{k=1}^{M_B})$ per block size M_B necessary to converge the ground-state energy E_{KS} with an accuracy $(E_{\text{KS}} - E_0)/NJ \leq 10^{-4}$ for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring. The third column indicates the maximum circuit depth to generate the corresponding Krylov subspace basis. Note that eight additional layers are required to prepare the VQE state $|\Psi_{\text{VQE}}\rangle$ [96] and few additional gate operations are necessary to generate each state of $\{|q_k\rangle\}_{k=1}^8$, which are not counted in the maximum circuit depth listed in the third column.

Reference state(s)	$n = \dim \mathcal{K}_n / M_B$	Circuit depth
$M_B = 1, \Phi_A\rangle$	9	17
$M_B = 2, \Phi_A\rangle, \Phi_B\rangle$	6	11
$M_B = 8, \{ q_k\rangle\}_{k=1}^8$	5	9
$M_B = 1, \Psi_{\text{VQE}}\rangle$	7	13
$M_B = 9, \Psi_{\text{VQE}}\rangle, \{ q_k\rangle\}_{k=1}^8$	4	7
$M_B = 16, \{\hat{T}_{k-1} \Psi_{\text{VQE}}\rangle\}_{k=1}^{16}$	2	3

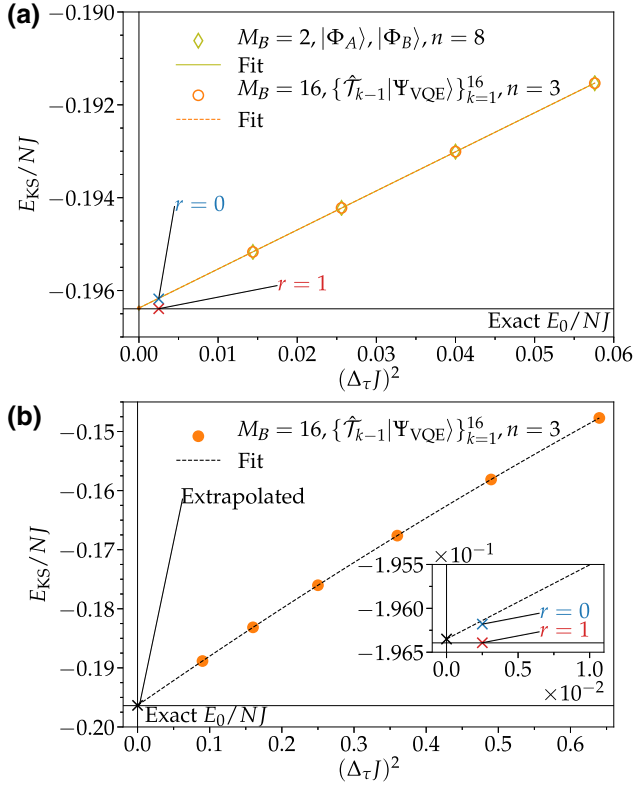


FIG. 7. (a) Ground-state energy E_{KS} evaluated for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring by setting the Krylov-subspace parameters $(M_B, n) = (2, 8)$ (diamonds) and $(16, 3)$ (circles) at $\Delta_\tau J = 0.12, 0.16, 0.2,$ and 0.24 without the Richardson extrapolation. The lines are fits to the two sets of data obtained separately by assuming the form $E_{KS}/NJ = a(\Delta_\tau J)^2 + b$ with a and b being fitting parameters. (b) Same as (a) but for $(M_B, n) = (16, 3)$ at $\Delta_\tau J = 0.3, 0.4, 0.5, 0.6, 0.7,$ and 0.8 without the Richardson extrapolation. The dashed line is a fit to the data obtained by assuming the form $E_{KS}/NJ = a(\Delta_\tau J)^4 + b(\Delta_\tau J)^2 + c$ with $a, b,$ and c being fitting parameters, and the extrapolated value at $\Delta_\tau = 0$ is indicated by black cross. The inset is an enlarged plot. For comparison, the results for $(M_B, n) = (16, 3)$ at $\Delta_\tau J = 0.05$ with $r = 0$ and 1 are also indicated by blue and red crosses, respectively, in (a) and the inset of (b). Note that the result with $r = 1$ corresponds to that shown in Fig. 5. The horizontal line indicates the exact ground-state energy E_0 .

quantum power method using a noisy near-term quantum computer. On a noisy quantum computer, a reasonably large Δ_τ should be used in order to evaluate, without being buried in the noise, the approximated Hamiltonian power formulated on the basis of the central finite differentiation. For instance, the expectation value of the difference $\hat{S}_2(\Delta_\tau/2) - \hat{S}_2(-\Delta_\tau/2)$ should be substantially more significant than the noise. However, as demonstrated above, the quantum power method can well control the systematic errors in approximating Hamiltonian power, and therefore one can accurately extrapolate the results to the limit of $\Delta_\tau \rightarrow 0$ even when a few results are available

for relatively large values of Δ_τ (also see, e.g., Refs. [99,100]).

Figure 7(a) shows the numerical results of the ground-state energy E_{KS} evaluated by setting the Krylov-subspace parameters $(M_B, n) = (2, 8)$ and $(16, 3)$ without the Richardson extrapolation at $\Delta_\tau J = 0.12, 0.16, 0.2,$ and 0.24 , which are larger than $\Delta_\tau J = 0.05$ used in Figs. 5 and 6. For the results shown in Fig. 7, we adopt the approach based on Eqs. (61) and (62) that requires a fewer number of state overlaps than that based on Eqs. (55) and (56). The lines are fits to the data assuming the form $E_{KS}/NJ = a(\Delta_\tau J)^2 + b$ with a and b being fitting parameters determined by the least-squares method. As shown in Fig. 7(a), the results are correctly extrapolated to the exact energy within 2–3 standard deviations of the fitting error (the error bars are not visible in the scale of the figure). It is more striking to find in Fig. 7(b) that a similar extrapolation, assuming the form $E_{KS}/NJ = a(\Delta_\tau J)^4 + b(\Delta_\tau J)^2 + c$ with $a, b,$ and c being fitting parameters, is also satisfactory even when much larger values of Δ_τ , as large as $\Delta_\tau J = 0.8$, are employed to evaluate the ground-state energy E_{KS} .

These results corroborate that the systematic errors due to a finite time interval Δ_τ are well controlled also for the quantities evaluated in the Krylov-subspace diagonalization, thus allowing us to extrapolate the results obtained for relatively large values of Δ_τ to the limit of $\Delta_\tau \rightarrow 0$. This could provide a promising error-mitigation strategy on a noisy near-term quantum computer. We should remark that, although the example shown here is the energy of the small system ($N = 16$), the Suzuki-Trotter error is known to be well controlled even in larger systems not only for the energy but also for other quantities in general (for example, see Ref. [101]).

VI. CONCLUSION AND DISCUSSION

We have proposed the quantum power method that approximates the Hamiltonian power $\hat{\mathcal{H}}^n$ with a linear combination of the time-evolution operators. The key ingredients of the quantum power method are the central-finite-difference scheme for the time derivatives and the symmetric Suzuki-Trotter decomposition of the time-evolution operators, both of which are essential to retain the Hermiticity and the even parity in Δ_τ of the approximated Hamiltonian power $\hat{\mathcal{H}}_{ST}^n(\Delta_\tau)$, i.e., $\hat{\mathcal{H}}_{ST}^n(\Delta_\tau) = [\hat{\mathcal{H}}_{ST}^n(\Delta_\tau)]^\dagger = \hat{\mathcal{H}}_{ST}^n(-\Delta_\tau)$. The systematic errors due to the finite differentiation and the Suzuki-Trotter decomposition are well controlled in the quantum power method as $\mathcal{E}_{FD} \sim O(\Delta_\tau^2)$ and $\mathcal{E}_{ST} \sim O(\Delta_\tau^{2m})$, respectively. The number of gates required for approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ is $O(nkN)$, where N is the number of qubits and a k -local Hamiltonian $\hat{\mathcal{H}}$ in the qubit representation composed of $O(N)$ terms is assumed, and thus it is at

most $O(nN^2)$ for a fermion Hamiltonian when the Jordan-Wigner transformation is used (see Appendix B). This should be contrasted to the classical power method, in which the computational complexity scales exponentially in N .

The r th-order Richardson extrapolation can be adopted to systematically improve the systematic errors as $\mathcal{E}_{\text{FD}} \sim O(\Delta\tau^{2+2r})$ and $\mathcal{E}_{\text{ST}} \sim O(\Delta\tau^{2m+2r})$ in the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)$, without increasing the number of gates required in each quantum circuit, although the number of terms in the linear combination, which can be treated classically, increases by the factor $r + 1$. Thus, both with and without the Richardson extrapolation, the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} can be consistently treated with the lowest-order Suzuki-Trotter decomposition with $m = 1$, independently of the power n , which reduces significantly the circuit depth. This is in sharp contrast to the algorithm that requires the higher-order Suzuki-Trotter decomposition with increasing the power n (see Appendix D). Therefore, the quantum power method proposed here is a potentially promising algorithm for near-term quantum devices.

By numerical simulations, we have tested the quantum power method and found that the Hamiltonian power $\hat{\mathcal{H}}^n$ for the spin-1/2 Heisenberg model can be well approximated by $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)$ with the controlled accuracy to be essentially exact for the power n up to 100 and N as large as 24 qubits, corresponding to the Hilbert space dimension $N_D = 2^N \approx 10^7$.

As an application of the quantum power method, we have demonstrated, with noiseless numerical simulations, the multireference Krylov-subspace diagonalization combined with the quantum power method for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring to evaluate the ground-state energy and the ground-state fidelity. Considering the Hamiltonian power $\hat{\mathcal{H}}^n$ up to $n = 11$, we have shown that the multireference Krylov-subspace-diagonalization scheme with the block size $M_B > 1$ greatly accelerate the convergence to the ground state, even when simple parameter-free product states are employed for the reference states. We have also found that the Krylov-subspace-diagonalization scheme with $M_B = 1$, corresponding to a quantum version of the standard Lanczos method [76], improves the ground-state energy of the VQE state $|\Psi_{\text{VQE}}\rangle$ almost exponentially with increasing n . Thus, the Krylov-subspace diagonalization combined with the quantum power method, which satisfies the variational principle by definition, can provide a systematic way to further improve a VQE state that has already a reasonable overlap with an exact ground state. This is a quantum analog to the Lanczos iteration scheme in the variational Monte Carlo method on classical computers [18], but here one can treat higher powers of the Hamiltonian on quantum computers.

We have also demonstrated by numerical simulations the multireference Krylov-subspace diagonalization combined with the quantum power method for the Fermi-Hubbard model on a 4×2 lattice. In this case, the Hamiltonian can be divided into $N_\Gamma = 4$ parts, instead of $N_\Gamma = 2$ parts for the case of the one-dimensional spin-1/2 Heisenberg model. Moreover, the Hamiltonian mapped from the fermion representation to the qubit representation on $N = 16$ qubits contains terms with long-range Pauli strings. Even in this case, considering the Hamiltonian power $\hat{\mathcal{H}}^n$ up to $n = 17$, we have found numerically that the ground-state energy converges almost exponentially with n (see Appendix B).

Although we have simulated only the ground-state energy, the expectation value of other observables that commute with the Hamiltonian $\hat{\mathcal{H}}$ can be evaluated similarly. When an observable $\hat{\mathcal{O}}$ does not commute with the Hamiltonian $\hat{\mathcal{H}}$, the expectation value with respect to the approximate ground state $|\Psi_0\rangle \approx |\tilde{\Psi}_{\text{KS}}\rangle \equiv \sum_{i=1}^{nM_B} v_i |\tilde{u}_i\rangle$, with the coefficients v_i already determined by solving Eq. (46) or Eq. (49) in the block Krylov subspace $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta\tau), \{|q_k\rangle\}_{k=1}^{M_B}) = \text{span}(\{|\tilde{u}_i\rangle\}_{i=1}^{nM_B})$, can also be evaluated as

$$\begin{aligned} \langle \Psi_0 | \hat{\mathcal{O}} | \Psi_0 \rangle &\approx \sum_{i=1}^{nM_B} \sum_{j=1}^{nM_B} v_i^* v_j \langle \tilde{u}_i | \hat{\mathcal{O}} | \tilde{u}_j \rangle \\ &= \sum_{i=1}^{nM_B} \sum_{j=1}^{nM_B} v_i^* v_j \sum_{\nu=0}^{l-1} \sum_{\nu'=0}^{l'-1} c_{l-1,\nu}^* c_{l'-1,\nu'} \\ &\quad \times \langle q_k | \left[\hat{\mathcal{S}}_{2m}^{(p)} \left(-\frac{\Delta\tau}{2} \right) \right]^{l-1-2\nu} \hat{\mathcal{O}} \\ &\quad \times \left[\hat{\mathcal{S}}_{2m}^{(p)} \left(\frac{\Delta\tau}{2} \right) \right]^{l'-1-2\nu'} | q_{k'} \rangle, \end{aligned} \quad (90)$$

where $|\tilde{u}_i\rangle = \hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta\tau)|q_k\rangle$, as given in Eq. (52), and the explicit form of $\hat{\mathcal{H}}_{\text{ST}(r)}^{l-1}(\Delta\tau)$ with $r = 0$ is used in the second line. Here, $i = k + (l-1)M_B$ and $j = k' + (l'-1)M_B$ for $1 \leq k, k' \leq M_B$ and $1 \leq l, l' \leq n$.

Our numerical simulations clearly demonstrate a promising potential that the quantum power method combined with the multireference Krylov-subspace diagonalization enables us to perform systematic and optimization-free calculations for quantum many-body systems, which is suitable for near-term quantum computers. Other applications of the quantum power method include various moment-based methods, which are briefly outlined in Appendices E and F. In these appendices, we show that the power method can evaluate $\langle \Psi | \hat{\mathcal{H}}^2 | \Psi \rangle$ with exactly the same amount of resource that is required for $\langle \Psi | \hat{\mathcal{H}} | \Psi \rangle$, and therefore, for example, the energy variance σ^2

$= \langle \Psi | \hat{\mathcal{H}}^2 | \Psi \rangle - \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle^2$ can be easily obtained. Here, $|\Psi\rangle$ is a given quantum state. Using numerical simulations, we also demonstrate the CMX for the imaginary-time evolution. This formalism can be easily extended to other methods, e.g., the high-temperature series expansion [102].

Finally, we remark that the quantum power method proposed here can generally be applied to any sparse Hermitian operator \hat{A} . In this case, the n th power of \hat{A} is given as

$$\hat{A}^n = i^n \left. \frac{d^n \hat{V}(t)}{dt^n} \right|_{t=0} \quad (91)$$

with the generating function $\hat{V}(t) = e^{-i\hat{A}t}$. We can use the central finite-difference scheme for the time derivatives to represent \hat{A}^n as a linear combination of unitary operator $\hat{V}(t)$ at different time variables. The symmetric Suzuki-Trotter decomposition is then used to decompose each unitary operator $\hat{V}(t)$.

ACKNOWLEDGMENTS

Parts of numerical simulations are done on the HOKU-SAI supercomputer at RIKEN (Project ID G20015). This work is supported by Grant-in-Aid for Research Activity start-up (Grant No. JP19K23433) and Grant-in-Aid for Scientific Research (B) (Grant No. JP18H01183) from MEXT, Japan.

Note added.—Recently, a related study by Bespalova and Kyriienko has been reported [103].

APPENDIX A: COMPARISON WITH OTHER ALGORITHMS FOR THE KRYLOV-SUBSPACE DIAGONALIZATION

In this appendix, we summarize the distinctions between the Krylov-subspace-diagonalization scheme described in Sec. IV and other algorithms reported recently, i.e., the QLanczos method [45], the MRSQK algorithm [50], and the QFD method [51]. However, before making a comparison, we emphasize that the quantum power method proposed here allows us to evaluate the Hamiltonian power $\hat{\mathcal{H}}^n$ directly by approximating it with a linear combination of the time-evolution operators. Consequently, the quantum power method finds many possible applications, some of which are described in this paper, and the Krylov-subspace diagonalization is one of the promising examples.

1. Brief review of the methods

The QLanczos method is based on the quantum imaginary-time evolution (QITE) for generating states that span a Krylov subspace. Namely, the Krylov subspace to

be approximated in the QLanczos method is given as

$$\begin{aligned} \mathcal{K}(e^{-2\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle) \\ = \text{span} \left(|\psi\rangle, e^{-2\Delta\tau\hat{\mathcal{H}}}|\psi\rangle, e^{-4\Delta\tau\hat{\mathcal{H}}}|\psi\rangle, \dots \right), \end{aligned} \quad (A1)$$

where $|\psi\rangle$ is a given reference state and only even powers of $e^{-\Delta\tau\hat{\mathcal{H}}}$ appear because it simplifies the evaluation of matrix elements of the Hamiltonian $\hat{\mathcal{H}}$ over the basis states in the Krylov subspace \mathcal{K} and the overlap matrix. However, this is irrelevant for the discussion given here. Let us assume that a Hamiltonian $\hat{\mathcal{H}}$ is given as $\hat{\mathcal{H}} = \sum_m \hat{h}[m]$, where $\hat{h}[m]$ represents the m th string of Pauli operators. With the first-order Suzuki-Trotter decomposition, the imaginary-time evolution operator for an imaginary time $l\Delta\tau$ can be written as $e^{-l\Delta\tau\hat{\mathcal{H}}} = \left(\prod_m e^{-\Delta\tau\hat{h}[m]} \right)^l + O(l\Delta\tau^2)$, where l is an integer, assuming that terms $\hat{h}[m]$ with different m are generally not commutable.

The QITE approximates, for each Trotter step, a normalized short imaginary-time evolved state $|\text{ITE}\rangle \equiv e^{-\Delta\tau\hat{h}[m]}|\psi\rangle/\sqrt{c}$ with $c = \langle \psi | e^{-2\Delta\tau\hat{h}[m]} | \psi \rangle \approx 1 - 2\Delta\tau \langle \psi | \hat{h}[m] | \psi \rangle$ by a unitary evolved state $|\text{ITE}'\rangle \equiv e^{-i\Delta\tau\hat{A}[m]}|\psi\rangle$, where $\hat{A}[m]$ is a Hermitian operator of the form

$$\hat{A}[m] = \sum_{i_{k_1}, i_{k_2}, \dots, i_{k_D}} a_{i_{k_1} i_{k_2} \dots i_{k_D}} [m] \hat{\sigma}_{i_{k_1}} \hat{\sigma}_{i_{k_2}} \dots \hat{\sigma}_{i_{k_D}} \quad (A2)$$

$$\equiv \sum_I a_I [m] \hat{\sigma}_I. \quad (A3)$$

Here, $\hat{\sigma}_{i_k} \in \{\hat{I}_k, \hat{X}_k, \hat{Y}_k, \hat{Z}_k\}$ in Eq. (A2), i.e., the identity and Pauli operators at the k th qubit, and $\hat{\sigma}_I$ in Eq. (A3) represents a Pauli string of length D (D is a parameter and is called domain size) with $I = \{i_{k_1}, i_{k_2}, \dots, i_{k_D}\}$ and $1 \leq k_1 < k_2 < \dots < k_D \leq N$. The sum over I runs at most up to 4^D (see Supplementary Information of Ref. [45] for a precise counting). The coefficients $a_I [m]$ should be determined by minimizing $\| |\text{ITE}\rangle - |\text{ITE}'\rangle \|$, which yields, to the first order of $\Delta\tau$, the linear system $\mathbf{C}\mathbf{a} = \mathbf{b}$ with $[\mathbf{C}]_{IJ} = \langle \psi | \hat{\sigma}_I^\dagger \hat{\sigma}_J | \psi \rangle$, $[\mathbf{a}]_I = a_I [m]$, and $[\mathbf{b}]_I = -(i/\sqrt{c}) \langle \psi | \hat{\sigma}_I^\dagger \hat{h}[m] | \psi \rangle$. Thus the solution vector \mathbf{a} of the linear system gives the coefficients $a_I [m]$. By repeating the above procedure for all the Trotter steps, a state that approximates $e^{-l\Delta\tau\hat{\mathcal{H}}}|\psi\rangle/\sqrt{\langle \psi | e^{-2l\Delta\tau\hat{\mathcal{H}}} | \psi \rangle}$ to the first order of $\Delta\tau$ can be constructed.

In the QLanczos method, one may wish to increase the time interval $\Delta\tau$ to avoid the linear dependency of the bases generated in the Krylov subspace \mathcal{K} . However, there is a trade-off that the increase of $\Delta\tau$ increases the systematic error in approximating the imaginary-time evolution operator with the Suzuki-Trotter decomposition and also requires to generally enlarge the qubit domain size D in

approximating the imaginary-time evolution with a unitary evolution. Instead, if one chooses a small value of $\Delta\tau$, these difficulties may not occur. However, the states generated are most likely linearly dependent on other states nearby in the imaginary time and thus many iterations of l may be required.

In the QFD method and the MRSQK algorithm (with a single reference state), the Krylov subspace to be approximated is given as

$$\mathcal{K}(e^{-i\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle) = \text{span}\left(|\psi\rangle, e^{-i\Delta\tau\hat{\mathcal{H}}}|\psi\rangle, e^{-i2\Delta\tau\hat{\mathcal{H}}}|\psi\rangle, \dots\right), \quad (\text{A4})$$

where the time-evolution operators $e^{-i\Delta\tau\hat{\mathcal{H}}}$ are approximated by a Suzuki-Trotter decomposition in practice. The MRSQK algorithm allows for the use of not only a single state but also multiple states as the reference states. The ground (or target) state is approximated as a linear combination of the time-evolved states $\{e^{-i\Delta\tau\hat{\mathcal{H}}}|\psi\rangle\}$ and the coefficients for the linear combination is determined by the Rayleigh-Ritz variational principle, i.e., the subspace diagonalization.

For example, the time-evolved state $e^{-i\Delta\tau\hat{\mathcal{H}}}|\psi\rangle$ can be expanded as

$$e^{-i\Delta\tau\hat{\mathcal{H}}}|\psi\rangle = |\psi\rangle - i\Delta\tau\hat{\mathcal{H}}|\psi\rangle + O(\Delta\tau^2). \quad (\text{A5})$$

Therefore, if $\Delta\tau$ is too small, a distance between the states $|\psi\rangle$ and $e^{-i\Delta\tau\hat{\mathcal{H}}}|\psi\rangle$ would be so small that a linear-dependency problem may occur in the subspace diagonalization. On the other hand, if large $\Delta\tau$ is chosen in order to reduce the linear-dependency problem, the corresponding Suzuki-Trotter error for approximating $e^{-i\Delta\tau\hat{\mathcal{H}}}$ becomes large. Hence, it is desirable to find an optimum value of $\Delta\tau$, which is however unknown *a priori*. This issue is similar to the case of the QLanczos method. The MRSQK algorithm can improve the linear-dependency problem by virtue of the multireference states as compared to the single reference state. However, in essence, the same issue may remain because the expansion process of the Krylov subspace is still based on a short real-time evolution.

In our method, the Krylov subspace is expanded by generating states $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)|\psi\rangle$ from a reference state $|\psi\rangle$. Namely, the Krylov subspace to be approximated is given as

$$\mathcal{K}(\hat{\mathcal{H}}, |\psi\rangle) = \text{span}\left(|\psi\rangle, \hat{\mathcal{H}}|\psi\rangle, \hat{\mathcal{H}}^2|\psi\rangle, \dots\right), \quad (\text{A6})$$

and the Hamiltonian power $\hat{\mathcal{H}}^n$ is approximated by a linear combination of the time-evolution operators, $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)$, with an $O(\Delta\tau^{2+2r})$ error. Therefore, in contrast to the MRSQK algorithm and the QFD method, the Krylov

basis $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)|\psi\rangle$ itself is a linear combination of time-evolved states. The coefficients $c_{n,k}$ for the linear combination in $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)|\psi\rangle$ are already determined from the central finite-difference formula.

More explicitly, the state $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)|\psi\rangle$ can be written as

$$\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)|\psi\rangle = \hat{\mathcal{H}}^n|\psi\rangle + \Delta\tau^{2+2r}\hat{r}_n|\psi\rangle + O(\Delta\tau^{4+2r}), \quad (\text{A7})$$

where \hat{r}_n is some Hermitian operator representing the leading error (residual) term in $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)$. Thus, in our method, the linear-dependency problem is expected to be less severe than in the other methods in the sense that the new basis in the Krylov subspace, e.g., $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta\tau)|\psi\rangle$, is in general linearly independent of $|\psi\rangle$, irrespectively of the value of $\Delta\tau$ (unless $|\psi\rangle$ is an eigenstate, e.g., the ground state, of $\hat{\mathcal{H}}$, which is the condition that indicates the convergence). Indeed, we have numerically found that the method is quite stable against the values of $\Delta\tau$ (see Fig. 7 and also the next section).

2. Condition number and residual ground-state energy

While making a fair comparison of different methods is not straightforward in any case, it would be instructive to examine by numerical simulations how the Krylov subspaces $\mathcal{K}_n(e^{-\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle)$, $\mathcal{K}_n(e^{-i\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle)$, and $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta\tau), |\psi\rangle)$ are different from each other, because it gains further insight into the different Krylov-subspace approaches described above. This is precisely the purpose of this section.

To this end, we employ the imaginary-time version of the second-order Suzuki-Trotter approximation (for the case of $N_\Gamma = 2$)

$$e^{-l\Delta\tau\hat{\mathcal{H}}} = \left(e^{-\frac{\Delta\tau}{2}\hat{\mathcal{H}}_A} e^{-\Delta\tau\hat{\mathcal{H}}_B} e^{-\frac{\Delta\tau}{2}\hat{\mathcal{H}}_A}\right)^l + O(l\Delta\tau^3) \quad (\text{A8})$$

to generate $\mathcal{K}_n(e^{-\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle)$, instead of the QITE reported originally in Ref. [45]. With this treatment, we can remove arbitrariness in the QITE such as the choice of $\hat{A}[m]$ operators and the domain-size parameter D . We also employ the second-order Suzuki-Trotter decomposition $\hat{S}_2(\Delta\tau)$ to generate $\mathcal{K}_n(e^{-i\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle)$ and $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta\tau), |\psi\rangle)$, i.e., with a similar decomposition for $e^{-i\Delta\tau\hat{\mathcal{H}}}$ as in Eq. (A8). Therefore, the Suzuki-Trotter error is at the same order for all three cases. Hereafter, we simply refer to $\mathcal{K}_n(e^{-\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle)$ as imaginary-time evolution (ITE), $\mathcal{K}_n(e^{-i\Delta\tau\hat{\mathcal{H}}}, |\psi\rangle)$ as real-time evolution (RTE), and $\mathcal{K}_n(\hat{\mathcal{H}}_{\text{ST}(r)}(\Delta\tau), |\psi\rangle)$ as the quantum power method (QPM). We use the spin-1/2 Heisenberg model on an $N = 16$ qubit ring as the Hamiltonian $\hat{\mathcal{H}}$, and adopt the

reference state $|\psi\rangle = |\Phi_A\rangle$ in Eq. (80) and $|\psi\rangle = |\Psi_{\text{VQE}}\rangle$ in Eq. (88). The overlap between these states and the exact ground state $|\Psi_0\rangle$ is $|\langle\Psi_0|\Phi_A\rangle|^2 = 0.259$ and $|\langle\Psi_0|\Psi_{\text{VQE}}\rangle|^2 = 0.771$ (see Fig. 6). In the numerical simulations, we vary $\Delta_\tau J = 0.01, 0.1, 0.3, 0.5,$ and 0.8 .

The linear dependency of the basis states in a Krylov subspace can be examined by calculating the (Euclidean norm) condition number of the overlap matrix \mathbf{S} ,

$$\text{cond}(\mathbf{S}) \equiv \frac{s_{\max}}{s_{\min}}, \quad (\text{A9})$$

where s_{\max} (s_{\min}) is the maximum (minimum) singular value of the overlap matrix \mathbf{S} defined in each Krylov subspace [for example, see Eq. (45)]. To focus on the linear-dependency issue, we assume that the basis states are normalized, as it is always possible by replacing $|u_i\rangle$ with $|u_i\rangle/\sqrt{\langle u_i|u_i\rangle}$, where $|u_i\rangle$ is an unnormalized basis state in ITE and QPM. The Krylov-subspace diagonalization in ITE and QPM for the normalized basis states can be formulated simply by replacing the matrix elements H_{ij} [for definition, see Eq. (44)] and S_{ij} with $H_{ij}/\sqrt{S_{ii}S_{jj}}$ and $S_{ij}/\sqrt{S_{ii}S_{jj}}$, respectively, or equivalently in the matrix form as

$$\mathbf{H} \mapsto \delta\mathbf{H}\delta \quad (\text{A10})$$

and

$$\mathbf{S} \mapsto \delta\mathbf{S}\delta, \quad (\text{A11})$$

where

$$\delta = \text{diag}\left(\frac{1}{\sqrt{S_{11}}}, \frac{1}{\sqrt{S_{22}}}, \dots\right). \quad (\text{A12})$$

The reduction scheme of the condition number of a matrix with a transformation by diagonal matrices as in Eqs. (A10) and (A11) is a widely used procedure known as equilibration of a matrix [76]. Note that although the normalization of the basis states alters the condition number of the overlap matrix \mathbf{S} in general, the resulting variational ground-state energy E_{KS} does not depend on whether or not the basis states in the Krylov subspace are normalized, as it can readily be confirmed from Eqs. (46), (A10), and (A11). In addition, the eigenvector \mathbf{v} in Eq. (46) is simply given by replacing \mathbf{v} with $\delta^{-1}\mathbf{v}$ when the equilibration of the matrices in Eqs. (A10) and (A11) is made. If the basis states are orthonormalized, the overlap matrix is an identity matrix, and hence the condition number is 1. On the other hand, if the basis states become linearly dependent, \mathbf{S} has zero singular value(s), and hence the condition number diverges. We show the numerical results for $\text{cond}(\mathbf{S}) \leq 10^{13}$, which is nearly the limit of double-precision floating-point number.

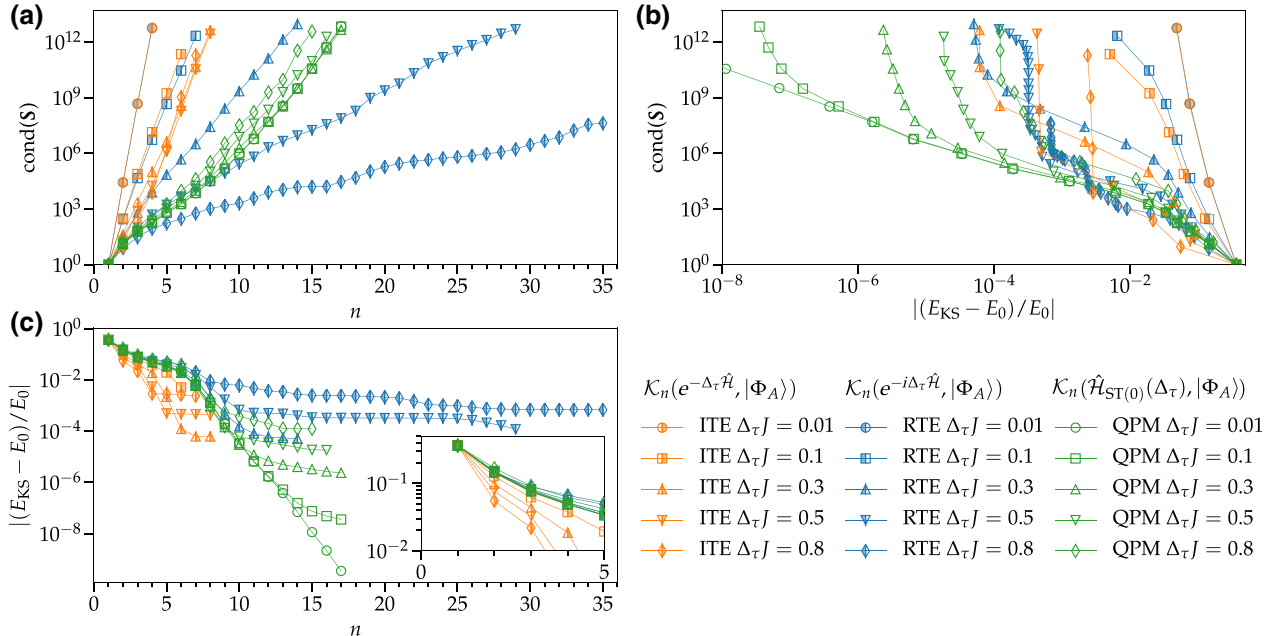


FIG. 8. (a) Condition number $\text{cond}(\mathbf{S})$ as a function of the subspace dimension n , (b) condition number $\text{cond}(\mathbf{S})$ as a function of the error in energy $|(E_{\text{KS}} - E_0)/E_0|$, and (c) error in energy $|(E_{\text{KS}} - E_0)/E_0|$ as a function of the subspace dimension n for three different Krylov subspaces (ITE, RTE, and QPM) with several values of Δ_τ . Here, the Hamiltonian $\hat{\mathcal{H}}$ is the spin-1/2 Heisenberg model on an $N = 16$ qubit ring with the exact ground-state energy E_0 and $|\psi\rangle = |\Phi_A\rangle$ is used as the reference state. For QPM, the Richardson extrapolation is not used ($r = 0$). Notice that the results for ITE and RTE with $\Delta_\tau J = 0.01$ are identical in this scale. The inset in (c) is the enlarged plot for small n .

Figure 8 shows the numerical results that summarize relations between the condition number $\text{cond}(\mathbf{S})$, the error in energy $|(E_{\text{KS}} - E_0)/E_0|$, and the dimension of the Krylov subspace n for several Δ_τ values with the reference state $|\psi\rangle = |\Phi_A\rangle$. As shown in Fig. 8(a), when Δ_τ is small, the condition number grows rapidly in n for ITE and RTE. With increasing Δ_τ , the condition number decreases significantly for RTE, while the decrease is less pronounced for ITE. On the other hand, the condition number for QPM is much less sensitive to Δ_τ than that for the other Krylov subspaces, even without the Richardson extrapolation ($r = 0$) [also see Fig. 10(a) for the results with the first-order Richardson extrapolation]. Remarkably, RTE shows the smaller condition number than QPM for $\Delta_\tau J = 0.5$ and 0.8 as a function of n . However, as we discuss below, the smaller condition number does not necessarily guarantee the better approximation to the ground state.

Figure 8(b) shows the condition number as a function of $|(E_{\text{KS}} - E_0)/E_0|$. It is found that, in all the Krylov subspaces, the condition number becomes larger as the error in energy becomes smaller. Such a behavior can be understood as follows. If a basis state $|u_i\rangle$ in the subspace is close to the exact ground state, $|u_i\rangle \sim |\Psi_0\rangle$, a good variational energy can be obtained, while the overlap matrix tends to be more ill conditioned because the generated states $|u_{i+1}\rangle$ can be almost parallel to $|u_i\rangle$. We should also note that the condition number as a function of $|(E_{\text{KS}} - E_0)/E_0|$ in Fig. 8(b) behaves rather similarly between ITE and RTE, especially when Δ_τ is small, and indeed it is almost identical when $\Delta_\tau J = 0.01$. As discussed above, this is expected

because these two Krylov subspaces should be equivalent in a region of small Δ_τ . Instead, the results for QPM are very different from those for ITE and RTE, even when Δ_τ is small. This is simply because the Krylov subspace generated in QPM is different from those in ITE and RTE, including in the limit of $\Delta_\tau \rightarrow 0$.

Figure 8(c) shows $|(E_{\text{KS}} - E_0)/E_0|$ as a function of n . While ITE gives the better energy than the others when n is small, it is difficult to reach larger n because of the large condition number. On the other hand, RTE can achieve the largest dimension of $n = 35$ (or more) when $\Delta_\tau J = 0.8$ owing to the smaller condition number. However, the significant improvement of $|(E_{\text{KS}} - E_0)/E_0|$ is not observed, despite the fact that the dimension of the subspace is substantially increased. We also find that both ITE and RTE give nearly the same error in energy at the maximum dimension n when the same Δ_τ value is used [also see Fig. 8(b)]. The results for QPM are located somewhere between those for ITE and RTE, and, depending on Δ_τ values, QPM can achieve the best accuracy in the ground-state energy among the three Krylov subspaces considered here. We should also note that the Δ_τ dependence of the results is somewhat scattered for ITE and RTE, but it is more systematic for QPM.

These features do not depend significantly on the choice of the reference state. Figure 9 summarizes the same results but for the reference state $|\psi\rangle = |\Psi_{\text{VQE}}\rangle$. Since the VQE state $|\Psi_{\text{VQE}}\rangle$ has a larger overlap with the exact ground state, a better convergence is generally expected in a Krylov-subspace diagonalization. Indeed, the error in

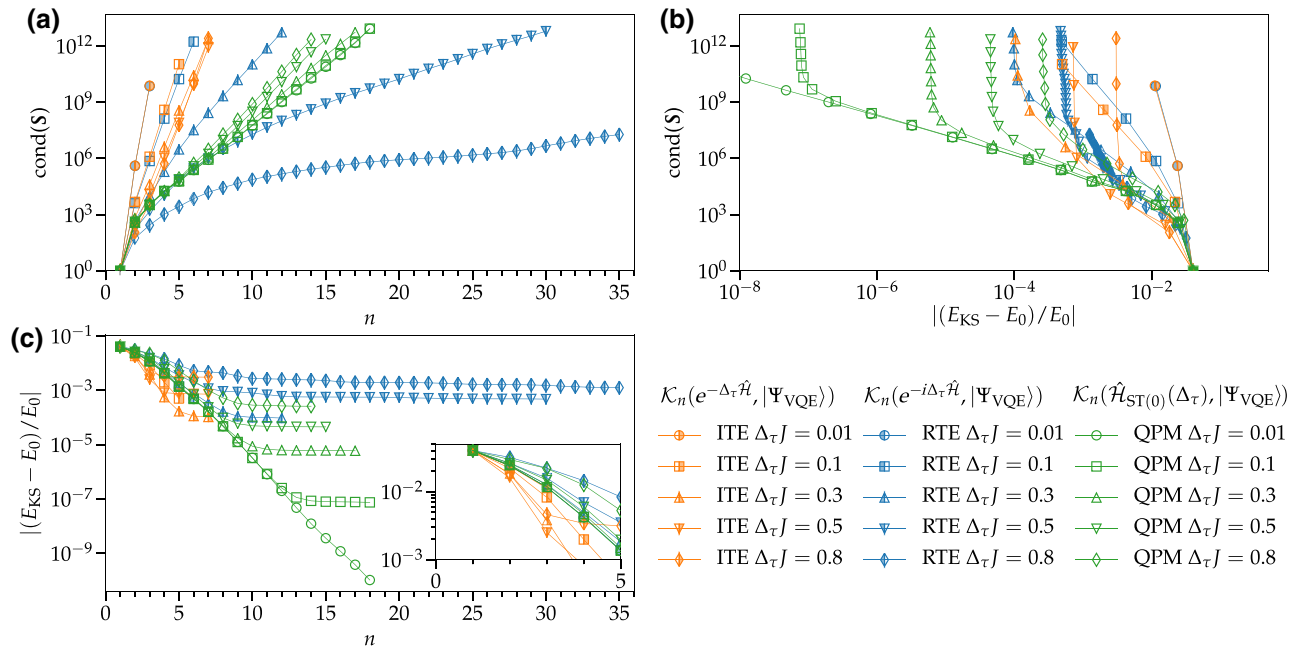


FIG. 9. Same as Fig. 8 but for the reference state $|\psi\rangle = |\Psi_{\text{VQE}}\rangle$. Notice that the results for ITE and RTE with $\Delta_\tau J = 0.01$ are identical in this scale.

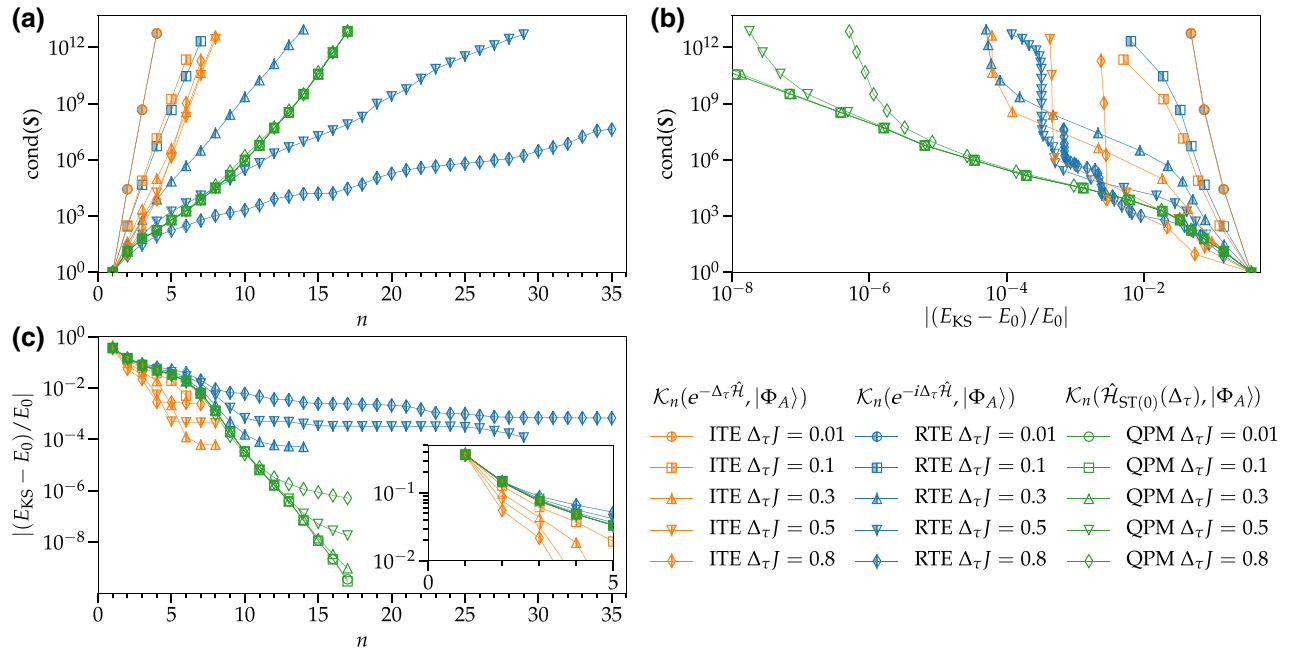


FIG. 10. Same as Fig. 8 but with the first-order Richardson extrapolation ($r = 1$) for QPM.

energy $|(E_{\text{KS}} - E_0)/E_0|$ at $n = 1$ is already one order of magnitude smaller than that for $|\psi\rangle = |\Phi_A\rangle$. Apart from this, the results are qualitatively the same as those in Fig. 8

As demonstrated in some details in Sec. V, the rather systematic dependence on Δ_τ for QPM allows us to perform the Richardson extrapolation with an increase of measurements, e.g., $(r + 1)^2 = 4$ times more for the first-order Richardson extrapolation. Figures 10 and 11 show

the same results as in Figs. 8 and 9 but with the first-order Richardson extrapolation for QPM. With the Richardson extrapolation, the QPM results almost collapse on a single curve (aside from the error in energy for large n and large Δ_τ), implying that the Δ_τ dependence becomes almost negligible. Although ITE and RTE may perform better than QPM in terms of the ground-state energy if an optimal Δ_τ value can be found, such an optimal value is not

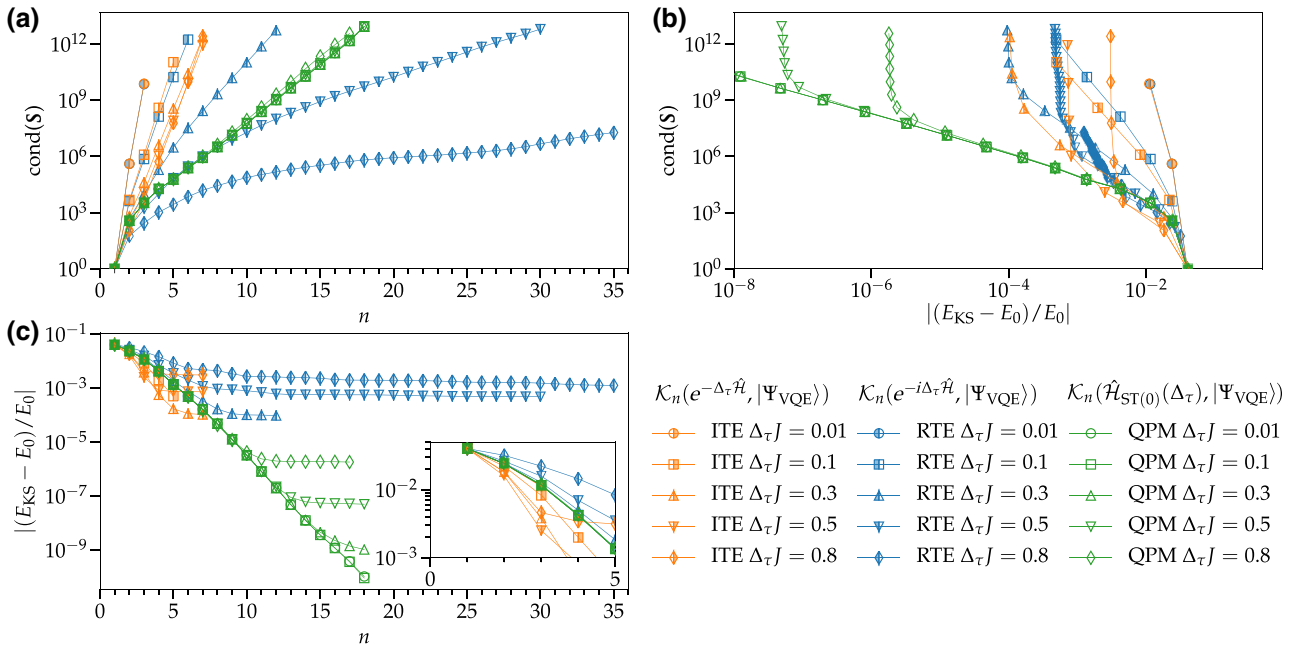


FIG. 11. Same as Fig. 9 but with the first-order Richardson extrapolation ($r = 1$) for QPM.

known *a priori*. In addition, considering that the exact solution is generally unknown, the systematic dependence on Δ_τ that guarantees the convergence to the exact is indispensable. We also note that performing an imaginary-time evolution on a quantum computer is rather involved when the QITE is used. Thus, the application of the quantum power method for the Krylov-subspace diagonalization would be a promising alternative for the Krylov-subspace approaches with quantum computers.

3. Summary

In short, as compared to the QLanczos method, our method is basically free from optimization of several parameters, such as the domain size D , the form of Hermitian operator $\hat{A}[m]$, and the coefficients $\mathbf{a}[m]$, which requires measurements of Pauli strings for \mathbf{C} and \mathbf{b} and solving the linear system $\mathbf{C}\mathbf{a} = \mathbf{b}$ of dimension approximately 4^D . The only but crucial parameter in our method is Δ_τ , which introduces the time-discretization error, but the systematic error is well controlled, as demonstrated numerically above and throughout this paper. In addition, our method can separate the choice of Δ_τ from the issue on the linear dependency of the basis states in the Krylov subspace, which is thus different from the other methods. In general, smaller Δ_τ approximates the ground state better. However, the well-controlled behavior of Δ_τ allows us to use large values of Δ_τ and then extrapolate the results to $\Delta_\tau \rightarrow 0$. This is also advantageous when a simulation on a noisy quantum computer is considered (see Sec. V C 3).

APPENDIX B: FERMION-HUBBARD MODEL

In this appendix, we demonstrate the Krylov-subspace diagonalization combined with the quantum power method for the spin-1/2 Fermi-Hubbard model on a square lattice with a ladderlike 4×2 cluster (i.e., $N = 16$ qubits) under open boundary conditions (Fig. 12). The Hamiltonian of the Fermi-Hubbard model is given by

$$\hat{\mathcal{H}} = -J \sum_{\sigma} \sum_{\langle i, j \rangle} (\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \text{H.c.}) + U_H \sum_{i=1}^{N/2} \left(\hat{n}_{i\uparrow} - \frac{1}{2} \right) \left(\hat{n}_{i\downarrow} - \frac{1}{2} \right), \quad (\text{B1})$$

where $\hat{c}_{i\sigma}^\dagger$ ($\hat{c}_{i\sigma}$) is a creation (annihilation) operator of a fermion at site i with spin σ ($=\uparrow, \downarrow$) and satisfies the canonical anticommutation relations $\{\hat{c}_{i\sigma}, \hat{c}_{j\sigma'}\} = \{\hat{c}_{i\sigma}^\dagger, \hat{c}_{j\sigma'}^\dagger\} = 0$ and $\{\hat{c}_{i\sigma}, \hat{c}_{j\sigma'}^\dagger\} = \delta_{ij} \delta_{\sigma\sigma'}$. $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$ is the density operator and $\langle i, j \rangle$ runs over all pairs of nearest-neighbor sites i and j on a square lattice. We assume that $J > 0$, $U_H > 0$, and the fermion density is 1, i.e., at half filling.

Before applying a fermion-to-qubit mapping, we first subdivide the Hamiltonian $\hat{\mathcal{H}}$ into noncommuting parts.

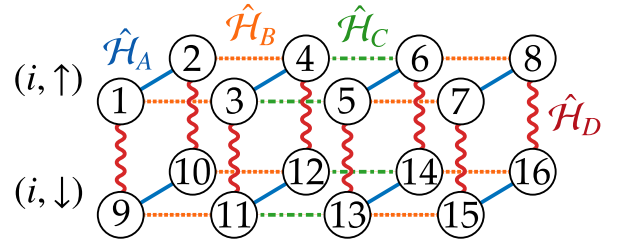


FIG. 12. The qubit indexing used for the Fermi-Hubbard model on a square lattice with a ladderlike 4×2 cluster under open boundary conditions. A circle with a number denotes a qubit and the lines between qubits indicate the terms of the subdivided Hamiltonians $\hat{\mathcal{H}}_A$, $\hat{\mathcal{H}}_B$, $\hat{\mathcal{H}}_C$, and $\hat{\mathcal{H}}_D$ in the fermion representation. Qubits 1–8 are assigned to fermions at site i ($1 \leq i \leq 8$) with spin \uparrow (upper layer), and qubits 9–16 are assigned to fermions at site i ($1 \leq i \leq 8$) with spin \downarrow (lower layer). The subdivided Hamiltonians $\hat{\mathcal{H}}_A$, $\hat{\mathcal{H}}_B$, and $\hat{\mathcal{H}}_C$ correspond to the hopping (J) terms, while $\hat{\mathcal{H}}_D$ corresponds to the interaction (U_H) terms, indicated, respectively, by the solid lines (blue), the dashed lines (orange), the dash-dotted lines (green), and the wavy lines (red).

For the 4×2 cluster, the Hamiltonian $\hat{\mathcal{H}}$ can be subdivided into $N_\Gamma = 4$ parts as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B + \hat{\mathcal{H}}_C + \hat{\mathcal{H}}_D, \quad (\text{B2})$$

where $\hat{\mathcal{H}}_A$ denotes the hopping terms along the rung direction, $\hat{\mathcal{H}}_B$ the hopping terms on the odd bonds along the leg direction, $\hat{\mathcal{H}}_C$ the hopping terms on the even bonds along the leg direction, and $\hat{\mathcal{H}}_D$ the on-site interaction terms, as schematically shown in Fig. 12. Note that all terms in each subdivided Hamiltonian $\hat{\mathcal{H}}_\Gamma$ commute with each other, although $[\hat{\mathcal{H}}_\Gamma, \hat{\mathcal{H}}_{\Gamma'}] \neq 0$ when $\Gamma \neq \Gamma'$. From Eq. (30), the lowest-order symmetric Suzuki-Trotter decomposition $\hat{S}_2(\Delta_\tau)$ of the time-evolution operator $\hat{U}(\Delta_\tau) = e^{-i\Delta_\tau \hat{\mathcal{H}}}$ can be given as

$$\hat{S}_2(\Delta_\tau) = e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_A} e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_B} e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_C} e^{-i\Delta_\tau \hat{\mathcal{H}}_D} \times e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_C} e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_B} e^{-i(\Delta_\tau/2)\hat{\mathcal{H}}_A}. \quad (\text{B3})$$

We now assign the fermion indexes to the qubit indexes as $(i, \uparrow) \mapsto i_\uparrow \equiv i$ and $(i, \downarrow) \mapsto i_\downarrow \equiv i + N/2$ for $1 \leq i \leq N/2$ [104]. The full indexing for the 4×2 cluster is given in Fig. 12. Let us now apply the Jordan-Wigner transformation to represent the fermion creation and annihilation operators by Pauli operators as [62,105]

$$\hat{c}_{i\sigma}^\dagger = \hat{\sigma}_{i\sigma}^- \prod_{k < i\sigma} \hat{Z}_k \quad (\text{B4})$$

and

$$\hat{c}_{i\sigma} = \prod_{k < i_\sigma} \hat{Z}_k \hat{\sigma}_{i_\sigma}^+ = \hat{\sigma}_{i_\sigma}^+ \prod_{k < i_\sigma} \hat{Z}_k, \quad (\text{B5})$$

where $\hat{\sigma}_i^\pm \equiv \frac{1}{2}(\hat{X}_i \pm i\hat{Y}_i)$. The fermion density operator is given by $\hat{n}_{i\sigma} = \hat{\sigma}_{i_\sigma}^- \hat{\sigma}_{i_\sigma}^+ = \frac{1}{2}(1 - \hat{Z}_{i_\sigma})$, implying that the single-particle state (i, σ) is occupied (unoccupied) if the i_σ th qubit state is $|1\rangle_{i_\sigma}$ ($|0\rangle_{i_\sigma}$). Now the Fermi-Hubbard Hamiltonian in the qubit representation reads

$$\begin{aligned} \hat{\mathcal{H}} = & -\frac{J}{2} \sum_{\sigma} \sum_{\langle i_\sigma, j_\sigma \rangle} (\hat{X}_{i_\sigma} \hat{X}_{j_\sigma} + \hat{Y}_{i_\sigma} \hat{Y}_{j_\sigma}) \hat{Z}_{\text{JW}, i_\sigma j_\sigma} \\ & + \frac{U_H}{4} \sum_{i=1}^{N/2} \hat{Z}_i \hat{Z}_{i+N/2}. \end{aligned} \quad (\text{B6})$$

Here, $\hat{Z}_{\text{JW}, ij} = \prod_{i \leq k \leq j} \hat{Z}_k$ is the Jordan-Wigner string for $i \leq j$, which supplies the fermion sign ± 1 , depending on

the population parity of the fermion occupation between the i th and j th qubits. Note that the subdivided Hamiltonians $\hat{\mathcal{H}}_\Gamma$ with $\Gamma = A, B, C, D$ are similarly transformed in the qubit representation (see Fig. 12) and apparently all terms in each $\hat{\mathcal{H}}_\Gamma$ in the qubit representation still commute with each other.

Figure 13 illustrates a circuit structure of the lowest-order Suzuki-Trotter-decomposed time-evolution operator $\hat{S}_2(\Delta_\tau)$ in Eq. (B3) for the Fermi-Hubbard model given in Eq. (B6) obtained by the Jordan-Wigner transformation. To implement $\hat{S}_2(\Delta_\tau)$ in a circuit, we first exponentiate the hopping term, i.e., $\exp[-i\theta(\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j) \hat{Z}_{\text{JW}, ij} / 2]$, by using a circuit representation of $\exp[-i\theta(\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j) / 2]$ (denoted as the XY gate in Fig. 13), which is composed of two CZ gates and six single-qubit rotations [44], sandwiched by $2(|i-j|-1)$ CZ gates [106] (also see Refs. [21, 88, 107–110] for other possible circuit realizations of the same operator). Here, the rotation angle is given by $\theta = -\Delta_\tau J / 2$. This gate is denoted as the K gate in Fig. 13. The interaction term is transformed into the Ising

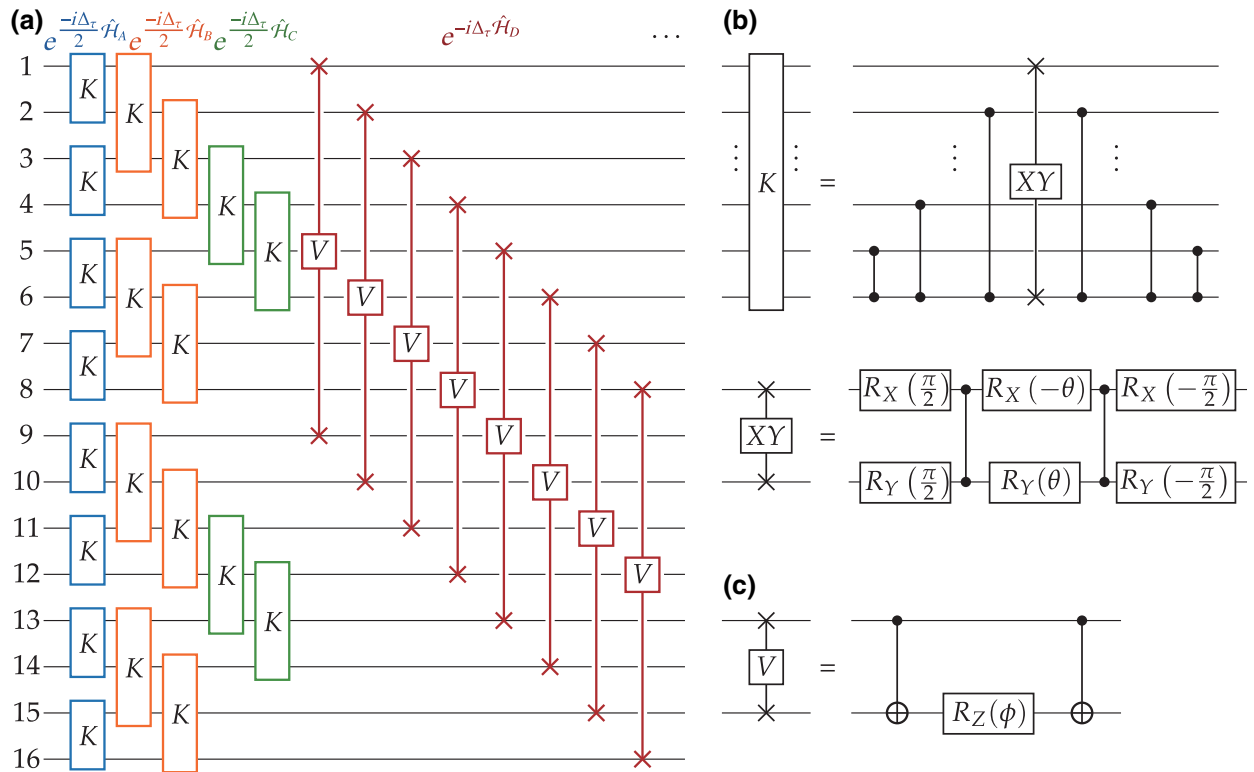


FIG. 13. (a) Circuit structure of the lowest-order Suzuki-Trotter-decomposed time-evolution operator $\hat{S}_2(\Delta_\tau)$ in Eq. (B3) for the Fermi-Hubbard model considered here in Eq. (B6). The gate K denotes the exponentiated hopping term defined in (b), while the gate V denotes the exponentiated interaction term defined in (c). Here, only the part corresponding to the first four exponentials is shown. (b) Decomposition of the K gate that represents $\exp[-i\theta(\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j) \hat{Z}_{\text{JW}, ij} / 2]$ operating at qubits i and j as well as all qubits between these two qubits. The gate XY represents $\exp[-i\theta(\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j) / 2]$ operating at qubits i and j . The gate $R_{X(Y)}(\theta)$ is given by $R_{X(Y)}(\theta) = \exp[-i\theta \hat{X}_i(\hat{Y}_i) / 2]$ operating at qubit i . (c) Decomposition of the V gate that represents $\exp[-i\phi \hat{Z}_i \hat{Z}_j / 2]$ operating at qubits i and j . The gate $R_Z(\phi)$ is given by $R_Z(\phi) = \exp[-i\phi \hat{Z}_i / 2]$ operating at qubit i . Here, $\theta = -\Delta_\tau J / 2$ and $\phi = \Delta_\tau U_H / 2$ for the Fermi-Hubbard model given in Eq. (B6).

interaction in Eq. (B6) and hence it can be exponentiated with a single-qubit rotation $\exp[-i\phi\hat{Z}_{i+N/2}/2]$ sandwiched by two $C_i\text{NOT}_{i+N/2}$ gates [61] (denoted as the V gate in Fig. 13). Here, the rotation angle is given by $\phi = \Delta_\tau U_H/2$.

As the reference states for the Krylov-subspace diagonalization, we consider the following product states:

$$|\Phi_A\rangle = \otimes_{i=1}^{N/2} |t_{2i-1}, 2i\rangle, \quad (\text{B7})$$

$$|Z_{\text{AFM}1}\rangle = \otimes_{i=1}^{N/4} |0\rangle_{2i-1} |1\rangle_{2i} \otimes_{i=1}^{N/4} |1\rangle_{2i-1+N/2} |0\rangle_{2i+N/2}, \quad (\text{B8})$$

$$|Z_{\text{AFM}2}\rangle = \otimes_{i=1}^{N/4} |1\rangle_{2i-1} |0\rangle_{2i} \otimes_{i=1}^{N/4} |0\rangle_{2i-1+N/2} |1\rangle_{2i+N/2}, \quad (\text{B9})$$

where $|t_{i,j}\rangle = (1/\sqrt{2})(|0\rangle_i |1\rangle_j + |1\rangle_i |0\rangle_j)$ is one of the Bell states and can be interpreted as a ‘‘spin-triplet state’’ in the qubit representation or as a ‘‘bonding state’’ in the fermion representation. $|\Phi_A\rangle$ is the ground state of the subdivided Hamiltonian $\hat{\mathcal{H}}_A$, because $\hat{\mathcal{H}}_A$ after the Jordan-Wigner transformation is merely a direct sum of two-site XY models with the ferromagnetic exchange interaction $-J < 0$, and hence the ground state is given by the direct product of $|t_{i,j}\rangle$. $|Z_{\text{AFM}1}\rangle$ and $|Z_{\text{AFM}2}\rangle$ are the Néel states (both in the qubit and fermion representations) with the staggered moments pointing alternatively along the spin- Z axis.

All these three states are within the subspace of the half filling and zero magnetization because they have $N/4 |0\rangle_{i_\sigma}$'s for $1 \leq i_\sigma \leq N/2$ and $N/4 |0\rangle_{i_\sigma}$'s for $N/2 + 1 \leq i_\sigma \leq N$. Moreover, these states can be easily generated from $|0\rangle^{\otimes N}$ with appropriate combinations of Pauli X , Hadamard, and CNOT gates. The particle number and magnetization are conserved even after applying the Hamiltonian power to these states. In addition to these three states, we adopt the ground state of $\hat{\mathcal{H}}$ at $U_H = 0$, $|\Psi_{U_H=0}\rangle$, as a reference state. Since $|\Psi_{U_H=0}\rangle$ is a Slater determinant, i.e., a particular case of fermionic Gaussian states, it can in principle be prepared on a quantum circuit with at most $O(N^2)$ gates [21,111–113]. $|\Psi_{U_H=0}\rangle$ is also within the subspace of the half filling and zero magnetization.

Figures 14 and 15 show the numerical results of the estimated ground-state energy E_{KS} and the ground-state fidelity $F = |\langle \Psi_0 | \Psi_{\text{KS}} \rangle|^2$, respectively, for the Fermi-Hubbard model with $U_H/J = 4$, obtained by the same procedures as in the case of the spin-1/2 Heisenberg model discussed in Sec. V C. Here, $|\Psi_0\rangle$ is the exact ground state and we set the time interval $\Delta_\tau J = 0.05$ with $r = 1$, $m = 1$, $p = 3$, and $N_\Gamma = 4$ for approximating the Hamiltonian power in the numerical simulations. The exact ground-state energy per site is $E_0/(NJ/2) = -1.626\,562\,894$. Note that N is the number of qubits and

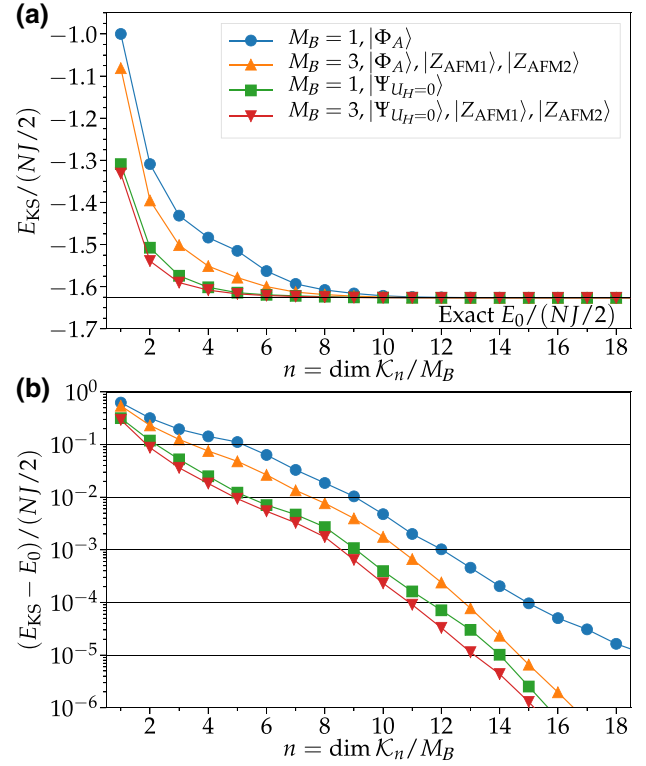


FIG. 14. Same as Fig. 5 but for the Fermi-Hubbard model with $U_H/J = 4$ at half filling.

the number of lattice sites of the Fermi-Hubbard model is given by $N/2$.

As shown in Figs. 14 and 15, the convergence to the ground state is improved with increasing the block size M_B , which is similar to the case of the spin-1/2 Heisenberg model found in Figs. 5 and 6. We also observe in Fig. 14(b) the exponential convergence of the energy with respect to n for any set of reference states. We can also notice in Fig. 15(a) that the noninteracting ground state $|\Psi_{U_H=0}\rangle$ has a significantly larger overlap with $|\Psi_0\rangle$ than $|\Phi_A\rangle$, and indeed the results obtained with $|\Psi_{U_H=0}\rangle$ shows the faster convergence than those with $|\Phi_A\rangle$.

The relatively slower convergence found here for the Fermi-Hubbard model as compared to the case of the spin-1/2 Heisenberg model, when only the simple product states other than $|\Psi_{U_H=0}\rangle$ are used, might be due to the smaller state overlaps between the reference states and the exact ground state. These results clearly demonstrate that the quantum power method can also be effective for fermion systems where a transformed Hamiltonian in the qubit representation involves more complex terms.

Finally, it should be noted that although the Bravyi-Kitaev transformation is known to scale asymptotically better than the Jordan-Wigner transformation in gate count,

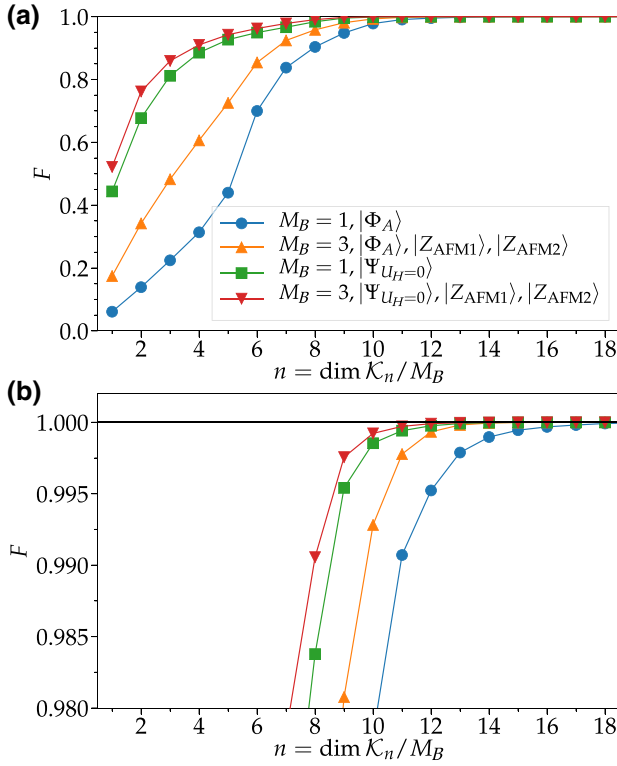


FIG. 15. Same as Fig. 6 but for the Fermi-Hubbard model with $U_H/J = 4$ at half filling.

the difference in efficiency between the two transformations is not obvious for small systems such as that considered here with $N = 16$ (see, for example, Refs. [114,115]). In addition, since the Jordan-Wigner transformation is based on the occupation basis of fermions, density-density interactions such as the Hubbard interaction can be written simply as Ising interactions [see Eq. (B6)], while Pauli strings might be involved in the interaction term when the Bravyi-Kitaev transformation is employed (see, for example, Refs. [64,66]). A comparative study of fermion-to-qubit mappings for the Hubbard model is, however, beyond the scope of this paper, and we have adopted the Jordan-Wigner transformation in favor of the simple form of the interaction term.

APPENDIX C: HIGHER-ORDER SYMMETRIC SUZUKI-TROTTER DECOMPOSITIONS $\hat{S}_{2m}^{(p)}(\Delta\tau)$

In this appendix, we provide a PYTHON program that generates coefficients required for the higher-order symmetric Suzuki-Trotter decompositions $\hat{S}_{2m}^{(p)}(\Delta\tau)$ introduced in Sec. III B 1, and examine numerically the systematic errors due to the Suzuki-Trotter decompositions when the time-evolution operator $\hat{U}(t)$ is approximated by a product of $\hat{S}_{2m}^{(p)}(\Delta\tau)$ with different parameters m and p . Note that m is an integer with $m \geq 1$ and p is an odd integer with $p \geq 3$.

1. Coefficients for higher-order Suzuki-Trotter decompositions

Listing 1 shows a PYTHON program that generates the coefficients $\{s_i\}_{i=1}^{D_{2m}^{(p)}}$ for a given set of parameters m and p in the symmetric Suzuki-Trotter decompositions $\hat{S}_{2m}^{(p)}(\Delta\tau)$:

$$\hat{S}_{2m}^{(p)}(\Delta\tau) = e^{xs_1\hat{\mathcal{H}}_A} e^{xs_2\hat{\mathcal{H}}_B} e^{xs_3\hat{\mathcal{H}}_C} \times \dots \times e^{xs_{D_{2m}^{(p)}-2}\hat{\mathcal{H}}_C} e^{xs_{D_{2m}^{(p)}-1}\hat{\mathcal{H}}_B} e^{xs_{D_{2m}^{(p)}}\hat{\mathcal{H}}_A}, \quad (C1)$$

where $x = -i\Delta\tau$ and $D_{2m}^{(p)} = 2(N_\Gamma - 1)p^{m-1} + 1$ as given in Eq. (18). The program includes an example for $m = 2$, $p = 5$, and $N_\Gamma = 2$. In this case, the symmetric Suzuki-Trotter decomposition has a form

$$\hat{S}_4^{(5)}(\Delta\tau) = e^{xs_1\hat{\mathcal{H}}_A} e^{xs_2\hat{\mathcal{H}}_B} e^{xs_3\hat{\mathcal{H}}_A} e^{xs_4\hat{\mathcal{H}}_B} \times e^{xs_5\hat{\mathcal{H}}_A} e^{xs_6\hat{\mathcal{H}}_B} e^{xs_7\hat{\mathcal{H}}_A} \times e^{xs_8\hat{\mathcal{H}}_B} e^{xs_9\hat{\mathcal{H}}_A} e^{xs_{10}\hat{\mathcal{H}}_B} e^{xs_{11}\hat{\mathcal{H}}_A} \quad (C2)$$

```

1 import numpy
2
3 def Suzuki_Trotter(m,p,NGamma):
4     s=numpy.array([0.5 for _ in range(2*NGamma-1)])
5     s[NGamma-1]=1.0
6     pl=int((p-1)/2)
7     for mm in range(m-1):
8         k0=1.0/(p-1-(p-1)**(1.0/(2*mm+3)))
9         k1=1.0-(p-1)*k0
10        sl=s[:-1]*k0
11        sl[0]=sl[0]*2.0
12        sl=numpy.concatenate([sl for _ in range(pl)
13        ])
14        sl[0]=sl[0]/2.0
15        s=s*k1
16        s[0]=s[0]+sl[0]
17        s[-1]=s[0]
18        s=numpy.concatenate([sl,s,sl[::-1]])
19    return s
20 # example
21 m=2
22 p=5
23 NGamma=2
24 s=Suzuki_Trotter(m,p,NGamma)
25 T=numpy.cumsum(s)
26 print('order =',2*m)
27 print('depth =',len(s))
28 print('sum(s) =',numpy.sum(s))
29 for i in range(len(s)):
30     print(i+1,s[i],T[i])
    
```

LISTING 1. A PYTHON program for generating the coefficients $\{s_i\}_{i=1}^{D_{2m}^{(p)}}$ in the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}$.

and the output of the program gives the 11 coefficients

$$\begin{aligned}
s_1 &= 0.207\,245\,385\,897\,187\,86, \\
s_2 &= 0.414\,490\,771\,794\,375\,7, \\
s_3 &= 0.414\,490\,771\,794\,375\,7, \\
s_4 &= 0.414\,490\,771\,794\,375\,7, \\
s_5 &= -0.121\,736\,157\,691\,563\,57, \\
s_6 &= -0.657\,963\,087\,177\,502\,8, \\
s_7 &= -0.121\,736\,157\,691\,563\,57, \\
s_8 &= 0.414\,490\,771\,794\,375\,7, \\
s_9 &= 0.414\,490\,771\,794\,375\,7, \\
s_{10} &= 0.414\,490\,771\,794\,375\,7, \\
s_{11} &= 0.207\,245\,385\,897\,187\,86.
\end{aligned}$$

By modifying lines 21–23 in the program, one can obtain $\{s_i\}_{i=1}^{D_{2m}^{(p)}}$ for other values of m , p , and N_Γ .

Notice that the coefficients $\{s_i\}_{i=1}^{D_{2m}^{(p)}}$ are symmetric, i.e.,

$$s_i = s_{D_{2m}^{(p)} - i + 1} \quad (\text{C3})$$

and satisfy the following sum rule:

$$\sum_{i=1}^{D_{2m}^{(p)}} s_i = N_\Gamma \quad (\text{C4})$$

for any m and p . Although it is sufficient to find the coefficients $\{s_i\}_{i=1}^{D_{2m}^{(p)}}$ for our purpose, the program can also output a cumulative sum T_i of the coefficient s_i defined as

$$T_i = \sum_{k=1}^i s_k. \quad (\text{C5})$$

By plotting T_i as a function of i [or $i/D_{2m}^{(p)}$] for several values of m with a fixed p , one can find a fractal feature appearing in the higher-order Suzuki-Trotter decompositions [71,74].

2. Numerical examination of a Suzuki-Trotter error

Here we numerically examine the systematic errors due to the Suzuki-Trotter decompositions when the time-evolution operator $\hat{U}(t)$ is approximated by a product of $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ with different parameters m and p . The Trotter formula [116–118] combined with $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ yields

$$\hat{U}(t) = \left[\hat{S}_{2m}^{(p)}(\Delta_\tau) \right]^M + O(t\Delta_\tau^{2m}), \quad (\text{C6})$$

where M is an integer such that $t = M\Delta_\tau$.

Figure 16 shows the real part of the difference between the exact propagator

$$K(t) = \langle \Psi_0 | \hat{U}(t) | \Psi_0 \rangle \quad (\text{C7})$$

and the approximated propagator

$$\tilde{K}(t) = \langle \Psi_0 | \left[\hat{S}_{2m}^{(p)}(\Delta_\tau) \right]^M | \Psi_0 \rangle, \quad (\text{C8})$$

i.e.,

$$\text{Re}\delta K(t) = \text{Re}\tilde{K}(t) - \text{Re}K(t), \quad (\text{C9})$$

with $\Delta_\tau J = 0.07$ and 0.1 for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring described by the Hamiltonian $\hat{\mathcal{H}}$ in Eq. (65), in which the Hamiltonian is subdivided into $N_\Gamma = 2$ parts. Here, $|\Psi_0\rangle$ is the exact ground state. The exact propagator is simply given by $K(t) = e^{-iE_0 t}$, where E_0 is the exact ground-state energy. As expected, when p is fixed, the error decreases by orders of magnitude with increasing m . It is also found that, when m is fixed, the error decreases by orders of magnitude with increasing p . Although we only show $\text{Re}\delta K(t)$, the imaginary part of the difference, $\text{Im}\delta K(t)$, behaves similarly.

We should emphasize here that while the deviation of the approximated propagator $\tilde{K}(t)$ from the exact one $K(t)$ becomes larger in the long time limit ($tJ \gg 1$), the quantum power method proposed here is formulated on the basis of the time-evolution operators $\hat{U}(t)$ at time t close to zero, for which the deviation is small. Therefore, this is another advantage of the quantum power method in controlling the Suzuki-Trotter error over other quantum algorithms that require the long-time dynamics approximately described by the Suzuki-Trotter-decomposed time-evolution operators.

Figure 17 shows $\text{Re}\delta K(t)$ divided by $(\Delta_\tau J)^{2m}$ for $\Delta_\tau J = 0.07, 0.11, 0.13,$ and 0.17 . As expected from Eq. (C6), the values of $\delta K(t)/(\Delta_\tau J)^{2m}$ for different Δ_τ are almost on the same curve. It is also found that the error decreases with increasing p for a fixed m , independently of Δ_τ . This suggests that the increase of p reduces the coefficient of the leading-error term by orders of magnitude. However, as shown in Eq. (18), p is the base of the exponential in $D_{2m}^{(p)}$, which determines the circuit depth. Thus, as far as noisy near-term quantum computers are concerned, $p = 3$ might be a more suitable value than $p \geq 5$.

We also examine the systematic errors due to the Suzuki-Trotter decompositions $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ in approximating the time-evolution operator for the Fermi-Hubbard model on a square lattice with a ladderlike 4×2 cluster under open boundary conditions, in which the Hamiltonian is subdivided into $N_\Gamma = 4$ parts, as described in Appendix B. Figure 18 shows the numerical results of $\text{Re}\delta K(t)$ divided by $(\Delta_\tau J)^{2m}$ for different values of Δ_τ for the Fermi-Hubbard model with $U_H/J = 4$ at half filling. As in the

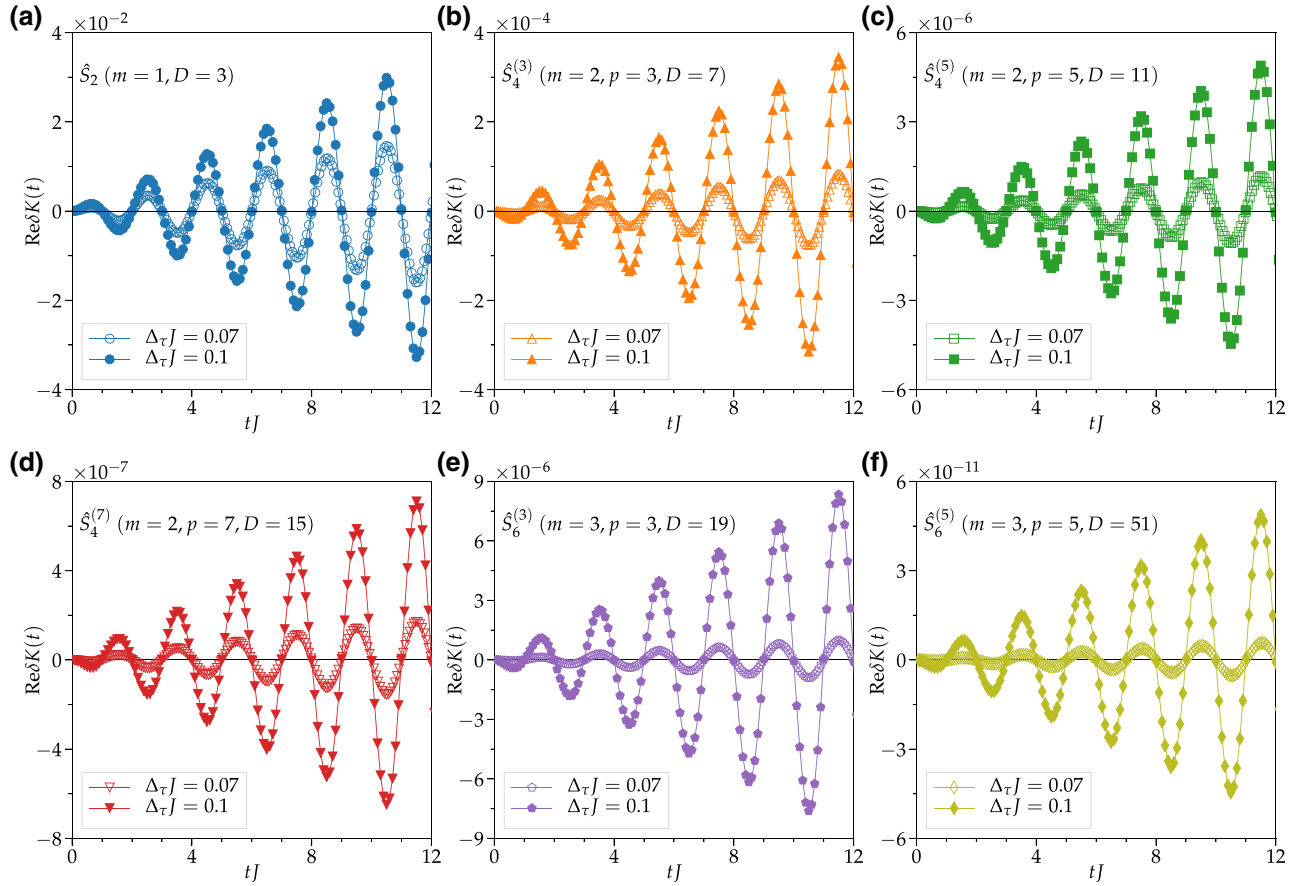


FIG. 16. Deviation of the real part of the propagator from the exact value, $\text{Re}\delta K(t)$, for $\Delta_\tau J = 0.07$ (empty symbols) and $\Delta_\tau J = 0.1$ (filled symbols) with different approximation schemes of the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ for the time-evolution operator. (a) $m = 1, D = 3$, (b) $m = 2, p = 3, D = 7$, (c) $m = 2, p = 5, D = 11$, (d) $m = 2, p = 7, D = 15$, (e) $m = 3, p = 3, D = 19$, and (f) $m = 3, p = 5, D = 51$, where $D = D_{2m}^{(p)}$ is the depth of a single $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ given in Eq. (18). Note that $\hat{S}_2(\Delta_\tau)$ corresponds to $\hat{S}_{2m}^{(p)}(\Delta_\tau)$ with $m = 1$ and $p = 3$. The results are for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring described by the Hamiltonian $\hat{\mathcal{H}}$ in Eq. (65) and $N_\Gamma = 2$. The solid lines are a guide for the eye.

case of the spin-1/2 Heisenberg model shown in Fig. 17, these values for different values of Δ_τ are almost on the same curve. In addition, the systematic errors decrease with increasing p for a fixed m , independently of Δ_τ , suggesting that the increase of p reduces the coefficient of the leading-error term by orders of magnitude also for $N_\Gamma = 4$.

Finally, we note that several exponential-product formulas, not limited to those found by Suzuki, up to the depth ≤ 11 with an error analysis can be found in Ref. [119]. Other error analysis of the Suzuki-Trotter decomposition devoted for quantum computing can be found in Refs. [120–122].

APPENDIX D: ANOTHER FORMALISM FOR APPROXIMATING THE HAMILTONIAN POWER

As discussed in Sec. III B, we can formulate at least two different algorithms for evaluating the Hamiltonian power $\hat{\mathcal{H}}^n$, depending on in which stage the time-evolution

operators in the approximated Hamiltonian power $\hat{\mathcal{H}}^n(\Delta_\tau)$ are replaced with the symmetric Suzuki-Trotter decomposition, either in Eq. (21) or in Eq. (24). In the quantum power method described in Sec. II B, the time-evolution operators in Eq. (24) are approximated by the symmetric Suzuki-Trotter decomposition. In this appendix, we describe the other formalism by approximating the time-evolution operators in Eq. (21) and show that the resulting algorithm scales differently from the one formulated in Sec. II B.

By incorporating the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}$ into the approximated Hamiltonian power $\hat{\mathcal{H}}^n(\Delta_\tau)$ in Eq. (21), the Hamiltonian power $\hat{\mathcal{H}}^n$ is now approximated as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) + O(\Delta_\tau^2) + \mathcal{E}_{\text{ST}}, \quad (\text{D1})$$

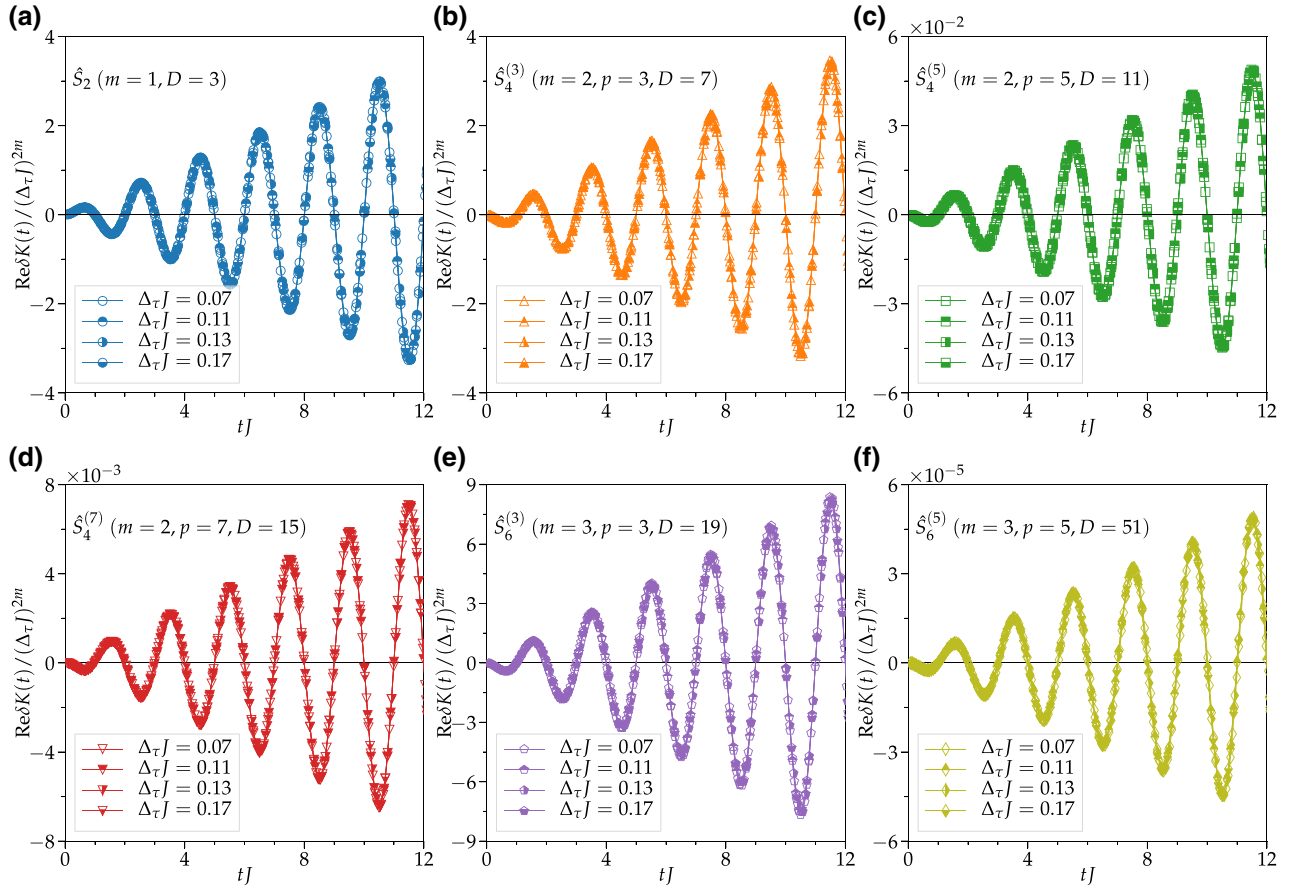


FIG. 17. Same as Fig. 16 but the deviation $\text{Re}\delta K(t)$ divided by $(\Delta_\tau J)^{2m}$ for several values of Δ_τ as indicated in the figures.

where

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) = \sum_{k=0}^n c_{n,k} \hat{S}_{2m}^{(p)}\left(\left(\frac{n}{2} - k\right) \Delta_\tau\right), \quad (\text{D2})$$

$O(\Delta_\tau^2)$ represents the systematic error \mathcal{E}_{FD} due to the finite-difference scheme for the time derivatives, and \mathcal{E}_{ST} denotes the systematic error due to the Suzuki-Trotter decomposition of the time-evolution operators, with the order of \mathcal{E}_{ST} being discussed below. We should emphasize here that $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) \neq \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ for $n \geq 2$, where $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ is defined in Eq. (4), because

$$\hat{S}_{2m}^{(p)}(\Delta_\tau) \hat{S}_{2m}^{(p)}(\Delta'_\tau) \neq \hat{S}_{2m}^{(p)}(\Delta_\tau + \Delta'_\tau) \quad (\text{D3})$$

for $\Delta_\tau \neq -\Delta'_\tau$, although the exact time-evolution operators satisfy the multiplication law $\hat{U}(\Delta_\tau)\hat{U}(\Delta'_\tau) = \hat{U}(\Delta_\tau + \Delta'_\tau)$. Note also that $\hat{\mathcal{H}}_{\text{ST}}^1(\Delta_\tau) = \hat{\mathcal{H}}_{\text{ST}}^1(\Delta_\tau)$.

We can readily confirm that $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ is Hermitian and an even function of Δ_τ , i.e.,

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) = \left[\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)\right]^\dagger = \hat{\mathcal{H}}_{\text{ST}}^n(-\Delta_\tau), \quad (\text{D4})$$

as in the case of $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ given in Eq. (10) and hence the systematic error \mathcal{E}_{ST} (as well as the systematic error \mathcal{E}_{FD} , see Sec. III A) in odd powers of Δ_τ is absent in Eq. (D1). We can also show that $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ does *not* satisfy the law of exponents, i.e.,

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) \neq \left[\hat{\mathcal{H}}_{\text{ST}}^1(\Delta_\tau)\right]^n \quad (\text{D5})$$

for $n \geq 2$, simply because of Eq. (D3), but only satisfies it approximately within the systematic errors. This is in sharp contrast to the case of $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$, which satisfies exactly the law of exponents in Eq. (9).

At first glance, one would tend to conclude that $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ in Eq. (D2) is more suitable to approximate the Hamiltonian power $\hat{\mathcal{H}}^n$ than $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ in Eq. (4), because each term in $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ contains a single $\hat{S}_{2m}^{(p)}$, not a product of multiple $\hat{S}_{2m}^{(p)}$'s as in $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$, thus expecting fewer gates in the circuit. However, the disadvantage of $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ in Eq. (D2) is that the higher-order Suzuki-Trotter decompositions are required for approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ with larger n .

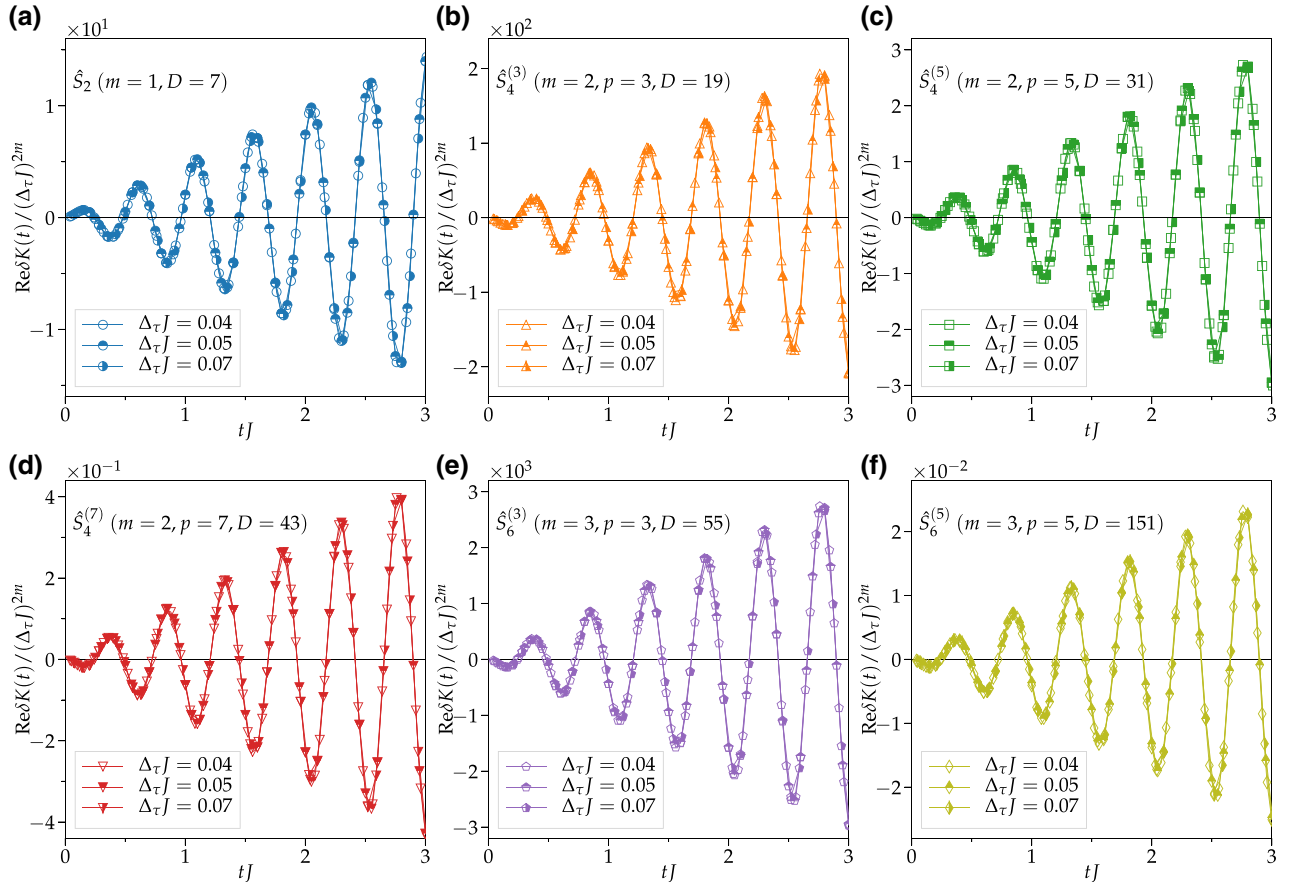


FIG. 18. Same as Fig. 17 but for the Fermi-Hubbard model on a square lattice with a 4×2 cluster under open boundary conditions described by the Hamiltonian $\hat{\mathcal{H}}$ in Eq. (B6). The other parameters of the model are $U_H/J = 4$ and $N = 16$ at half filling. $N_\Gamma = 4$ and $\Delta_\tau J = 0.04, 0.05$, and 0.07 are used.

This can be understood by recalling that $\hat{S}_{2m}^{(p)}(t)$ has a form of Eq. (34):

$$\hat{S}_{2m}^{(p)}(t) = \exp \left[-it\hat{\mathcal{H}} + (-it)^{2m+1}\hat{R}_{2m+1} + \dots \right]. \quad (\text{D6})$$

Accordingly, the higher-order derivative of $\hat{S}_{2m}^{(p)}(t)$ at $t = 0$ is given by

$$i^n \frac{d^n \hat{S}_{2m}^{(p)}(t)}{dt^n} \Big|_{t=0} = \hat{\mathcal{H}}^n \quad (\text{D7})$$

for $n \leq 2m$, but

$$i^n \frac{d^n \hat{S}_{2m}^{(p)}(t)}{dt^n} \Big|_{t=0} \neq \hat{\mathcal{H}}^n \quad (\text{D8})$$

for $n > 2m$. For example, if $n = 2m + 1$, the derivative reads

$$i^{2m+1} \frac{d^{2m+1} \hat{S}_{2m}^{(p)}(t)}{dt^{2m+1}} \Big|_{t=0} = \hat{\mathcal{H}}^{2m+1} + (2m+1)! \hat{R}_{2m+1}. \quad (\text{D9})$$

It is now important to notice that the right-hand side of Eq. (D2) corresponds to the central finite-difference approximation of $i^n [d^n \hat{S}_{2m}^{(p)}(t)/dt^n]_{t=0}$, i.e.,

$$i^n \frac{d^n \hat{S}_{2m}^{(p)}(t)}{dt^n} \Big|_{t=0} = \sum_{k=0}^n c_{n,k} \hat{\mathcal{S}}_{2m}^{(p)} \left(\left(\frac{n}{2} - k \right) \Delta_\tau \right) + O(\Delta_\tau^2). \quad (\text{D10})$$

In other words, the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ in Eq. (D2) is given by the higher-order derivative of $\hat{S}_{2m}^{(p)}(t)$ at $t = 0$ as

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau) = i^n \frac{d^n \hat{S}_{2m}^{(p)}(t)}{dt^n} \Big|_{t=0} + O(\Delta_\tau^2). \quad (\text{D11})$$

Now, it is obvious that the formalism in Eq. (D2) breaks down if $n > 2m$ because in this case, according to Eq. (D8), $\lim_{\Delta\tau \rightarrow 0} \hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) \neq \hat{\mathcal{H}}^n$, which contradicts Eq. (D1). Therefore,

$$2m \geq n \quad (\text{D12})$$

is required for approximating the Hamiltonian power $\hat{\mathcal{H}}^n$ by $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ under a controlled accuracy with the systematic error

$$\mathcal{E}_{\text{ST}} \sim O(\Delta\tau^2). \quad (\text{D13})$$

This is the most important difference from the algorithm described in Sec. II B, where the lowest-order Suzuki-Trotter decomposition with $m = 1$ is adequate for any power n .

There are two remarks in order. First, the approximated Hamiltonian power $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ in Eq. (5) can be considered as

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) = \left[i \frac{d\hat{S}_{2m}^{(p)}(t)}{dt} \Big|_{t=0} + O(\Delta\tau^2) \right]^n. \quad (\text{D14})$$

Therefore, the lowest-order symmetric Suzuki-Trotter decomposition with $m = 1$ is adequate to satisfy Eq. (D12) and indeed, as discussed in Sec. II B, it approximates the Hamiltonian power $\hat{\mathcal{H}}^n$ with the controlled accuracy. Second, although we have emphasized that the violation of the multiplication law $\hat{S}_{2m}^{(p)}(\Delta\tau/2)\hat{S}_{2m}^{(p)}(\Delta\tau/2) \neq \hat{S}_{2m}^{(p)}(\Delta\tau)$ is the essential point that distinguishes the two algorithms described here and in Sec. II B, this equation is satisfied within the systematic error, i.e.,

$$\hat{S}_{2m}^{(p)}\left(\frac{\Delta\tau}{2}\right)\hat{S}_{2m}^{(p)}\left(\frac{\Delta\tau}{2}\right) = \hat{S}_{2m}^{(p)}(\Delta\tau) + O(\Delta\tau^{2m+1}). \quad (\text{D15})$$

Accordingly, the two algorithms described here and in Sec. II B should be the same within the systematic error. In fact, the approximated Hamiltonian powers $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ and $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ in Eqs. (4) and (D2), respectively, are equivalent within the systematic error because

$$\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) = \sum_{k=0}^n c_{n,k} \left[\hat{S}_{2m}^{(p)}\left(\frac{\Delta\tau}{2}\right) \right]^{n-2k} \quad (\text{D16})$$

$$= \sum_{k=0}^n c_{n,k} \left[\hat{S}_{2m}^{(p)}\left(\left(\frac{n}{2} - k\right)\Delta\tau\right) + O(\Delta\tau^{2m+1}) \right] \quad (\text{D17})$$

$$= \hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) + O(\Delta\tau^{2m+1-n}), \quad (\text{D18})$$

provided that $2m + 1 > n$, which is consistent with Eq. (D12).

As an example, we show in Fig. 19 the expectation values $\langle \hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) \rangle$ and $\langle \hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) \rangle$ with respect to the quantum states $|\Phi_A\rangle$ and $|\Psi_{\text{VQE}}\rangle$ of the spin-1/2 Heisenberg model on an $N = 16$ qubit ring for the power $n = 3$. Here, a simplified notation of the expectation value

$$\langle \dots \rangle \equiv \langle \Psi | \dots | \Psi \rangle \quad (\text{D19})$$

is introduced with $|\Psi\rangle \in \{|\Phi_A\rangle, |\Psi_{\text{VQE}}\rangle\}$ given in Eqs. (80) and (88). According to Eqs. (D7)–(D9), $\langle \hat{\mathcal{H}}_{\text{ST}}^3(\Delta\tau) \rangle$ in the limit of $\Delta\tau \rightarrow 0$ should converge as

$$\lim_{\Delta\tau \rightarrow 0} \langle \hat{\mathcal{H}}_{\text{ST}}^3(\Delta\tau) \rangle = \langle \hat{\mathcal{H}}^3 \rangle \quad (\text{D20})$$

for $m \geq 2$, but

$$\lim_{\Delta\tau \rightarrow 0} \langle \hat{\mathcal{H}}_{\text{ST}}^3(\Delta\tau) \rangle = \langle \hat{\mathcal{H}}^3 + 3!\hat{R}_3 \rangle \quad (\text{D21})$$

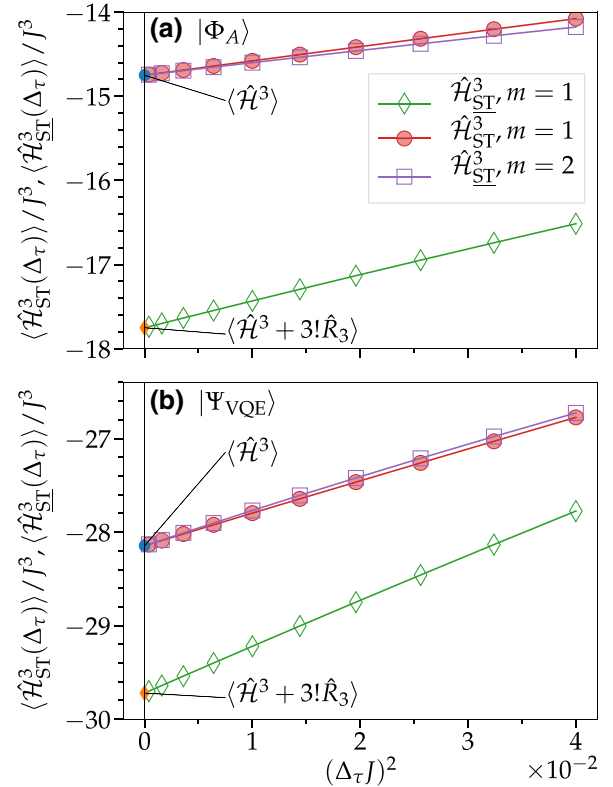


FIG. 19. $\langle \hat{\mathcal{H}}_{\text{ST}}^3(\Delta\tau) \rangle$ and $\langle \hat{\mathcal{H}}_{\text{ST}}^3(\Delta\tau) \rangle$ as a function of $\Delta\tau^2$ evaluated for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring, using different orders of the symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}$ with $m = 1, 2$ and $p = 3$. For a quantum state $|\Psi\rangle$, we choose (a) the singlet-pair product state $|\Phi_A\rangle$ in Eq. (80) and (b) the VQE state $|\Psi_{\text{VQE}}\rangle$ in Eq. (88). The exact values are indicated at $\Delta\tau = 0$ with the filled symbols (blue and orange).

for $m = 1$. Here, the explicit form of the residual term \hat{R}_3 in Eq. (D21) for $N_\Gamma = 2$ can be derived by using the Baker-Campbell-Hausdorff formula for \hat{S}_2 as [72,119,123]

$$\hat{R}_3 = -\frac{1}{24} \left[\hat{\mathcal{H}}_A, \left[\hat{\mathcal{H}}_A, \hat{\mathcal{H}}_B \right] \right] + \frac{1}{12} \left[\hat{\mathcal{H}}_B, \left[\hat{\mathcal{H}}_B, \hat{\mathcal{H}}_A \right] \right]. \quad (\text{D22})$$

The numerical results in Fig. 19 confirm Eqs. (D20) and (D21), as well as the expected behavior $\lim_{\Delta\tau \rightarrow 0} \langle \hat{\mathcal{H}}_{\text{ST}}^3(\Delta\tau) \rangle = \langle \hat{\mathcal{H}}^3 \rangle$ for $m = 1$. Note also that the linear convergence of these quantities to the exact values as a function of $\Delta\tau^2$ shown in Fig. 19 corroborates the systematic errors expected for $\langle \hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) \rangle$ in Eqs. (D1) and (D13) and for $\langle \hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau) \rangle$ in Eq. (3).

Let us now discuss the gate count for approximating $\hat{\mathcal{H}}^n$ with $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$. As described above, Eq. (D12) sets the order of the Suzuki-Trotter decomposition such that $2m \geq n$, i.e., the smallest order m of the Suzuki-Trotter decomposition to evaluate $\hat{\mathcal{H}}^n$ being $m = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ is the ceiling function that returns the minimum integer larger than or equal to the argument. Therefore, assuming that a k -local Hamiltonian $\hat{\mathcal{H}}$ composed of $O(N)$ terms, the number of gates required for approximating $\hat{\mathcal{H}}^n$ with $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta)$ is $O(p^{n/2} kN)$ because the circuit depth $D_{2m}^{(p)}$ for the single Suzuki-Trotter-decomposed time-evolution operator $\hat{S}_{2m}^{(p)}$ is given by Eq. (18), and thus it increases exponentially in the power n . In contrast, as described in Sec. III B 3, the number of gates required for approximating $\hat{\mathcal{H}}^n$ with $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ is $O(nkN)$ with a prefactor $D_2^{(p)} = 2N_\Gamma - 1 \sim O(1)$, i.e., increasing polynomially in N and n .

This indicates that the algorithm based on $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ suffers from the exponential increase of the number of gates for large n . However, the algorithm based on $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ can be more favorable than that based on $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ when the power n is small. To be more specific, let us consider the case of $p = 3$ and $N_\Gamma = 2$. Then, the circuit depth for $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ is given by $D_{2\lceil n/2 \rceil}^{(3)} = 3, 3, 7, 7, 19, 19, 55, 55, 163, \dots$, while the largest circuit depth for $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ involving $[\hat{S}_2^{(p)}(\pm\Delta/2)]^n$ is $n(D_2^{(3)} - 1) + 1 = 2n + 1 = 3, 5, 7, 9, 11, 13, 15, 17, 19, \dots$, for the Hamiltonian power $n = 1, 2, 3, 4, 5, 6, 7, 8, 9, \dots$. Here, for the latter, the depth is counted by assuming that the commuting exponentials in $[\hat{S}_2^{(p)}(\pm\Delta/2)]^n$ are contracted (also see Table II). Therefore, in this case with $p = 3$ and $N_\Gamma = 2$, the algorithm based on $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ is more preferable than that based on $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ as long as the power $n \leq 4$. In fact, one can readily show that this is generally the case, irrespectively of the value of N_Γ , when $p = 3$ [124]. As shown in Appendix E 4, the algorithm based on $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ is

indeed particularly useful when the lowest-order moments are evaluated.

To apply the quantum power method formulated in this appendix to the Krylov-subspace-diagonalization scheme, it is crucial to reduce the maximum power n appearing in the formalism. Defining

$$|\tilde{u}_i\rangle = \hat{\mathcal{H}}_{\text{ST}}^{l-1}(\Delta\tau)|q_k\rangle \quad (\text{D23})$$

for the basis set generated in the block Krylov subspace $\mathcal{K}_n \left(\hat{\mathcal{H}}_{\text{ST}}(\Delta\tau), \{|q_k\rangle\}_{k=1}^{M_B} \right)$, the matrix elements \mathbf{H} and \mathbf{S} in Eqs. (44) and (45) are now approximated by replacing $|u_i\rangle$ with $|\tilde{u}_i\rangle$ as

$$\tilde{H}_{ij} = \langle \tilde{u}_i | \hat{\mathcal{H}} | \tilde{u}_j \rangle = \langle q_k | \hat{\mathcal{H}}_{\text{ST}}^{l-1}(\Delta\tau) \hat{\mathcal{H}} \hat{\mathcal{H}}_{\text{ST}}^{l-1}(\Delta\tau) | q_{k'} \rangle \quad (\text{D24})$$

and

$$\tilde{S}_{ij} = \langle \tilde{u}_i | \tilde{u}_j \rangle = \langle q_k | \hat{\mathcal{H}}_{\text{ST}}^{l-1}(\Delta\tau) \hat{\mathcal{H}}_{\text{ST}}^{l'-1}(\Delta\tau) | q_{k'} \rangle, \quad (\text{D25})$$

where $i = k + (l-1)M_B$ and $j = k' + (l'-1)M_B$ for $1 \leq k, k' \leq M_B$ and $1 \leq l, l' \leq n$. Similarly to Eqs. (55) and (56), the power exponents are distributed to the left and the right basis states.

To be more specific, \tilde{H}_{ij} and \tilde{S}_{ij} in terms of $\hat{S}_{2m}^{(p)}$ are given as

$$\begin{aligned} \tilde{H}_{ij} &= \sum_{v=0}^{l-1} \sum_{v'=0}^{l'-1} c_{l-1,v}^* c_{l'-1,v'} \\ &\times \langle q_k | \hat{S}_{2m}^{(p)}(-t_v^{(l-1)}) \hat{\mathcal{H}} \hat{S}_{2m}^{(p)}(t_{v'}^{(l'-1)}) | q_{k'} \rangle \end{aligned} \quad (\text{D26})$$

and

$$\begin{aligned} \tilde{S}_{ij} &= \sum_{v=0}^{l-1} \sum_{v'=0}^{l'-1} c_{l-1,v}^* c_{l'-1,v'} \\ &\times \langle q_k | \hat{S}_{2m}^{(p)}(-t_v^{(l-1)}) \hat{S}_{2m}^{(p)}(t_{v'}^{(l'-1)}) | q_{k'} \rangle, \end{aligned} \quad (\text{D27})$$

where $t_v^{(l-1)} = \left(\frac{l-1}{2} - v\right) \Delta\tau$ and Eq. (8) is used. The number of terms in Eqs. (D26) and (D27) is $O(Nl')$ and $O(l'l')$, respectively. Here, we assume that $\hat{\mathcal{H}}$ consists of $O(N)$ local terms. In total, $O(n^2 M_B^2 N)$ and $O(n^2 M_B^2)$ state overlaps are required to be evaluated for constructing all matrix elements of the $nM_B \times nM_B$ matrices $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$, respectively.

Finally, we note that the systematic errors \mathcal{E}_{FD} and \mathcal{E}_{ST} in Eqs. (D1) and (D13) can be improved systematically, without increasing the gate count of each circuit, by adopting

the Richardson extrapolation as

$$\hat{\mathcal{H}}^n = \hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) + O(\Delta_\tau^{2+2r}), \quad (\text{D28})$$

where $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ is the r th-order Richardson extrapolation of the approximated Hamiltonian power, i.e.,

$$\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau) = \frac{h^{2r} \hat{\mathcal{H}}_{\text{ST}(r-1)}^n(\Delta_\tau/h) - \hat{\mathcal{H}}_{\text{ST}(r-1)}^n(\Delta_\tau)}{h^{2r} - 1}, \quad (\text{D29})$$

with $\hat{\mathcal{H}}_{\text{ST}(0)}^n(\Delta_\tau) \equiv \hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$. Since $\hat{\mathcal{H}}_{\text{ST}(0)}^n(\Delta_\tau)$ is a linear combination of $n+1$ unitaries $\left\{ \hat{S}_{2m}^{(p)}[(n-2k)\Delta_\tau/2] \right\}_{k=0}^n$, $\hat{\mathcal{H}}_{\text{ST}(r)}^n(\Delta_\tau)$ is a linear combination of $(r+1)(n+1)$ unitaries $\left\{ \left\{ \hat{S}_{2m}^{(p)}[(n-2k)\Delta_\tau/2h^l] \right\}_{k=0}^n \right\}_{l=0}^r$.

APPENDIX E: MOMENT METHODS

In this appendix, we outline moment methods as other applications of the quantum power method to evaluate the moments and cumulants of the Hamiltonian. By using numerical simulations, we demonstrate the CMX for a short imaginary-time evolution and estimate the ground-state energy of the spin-1/2 Heisenberg model. These numerical results are compared with those obtained by the multireference Krylov-subspace diagonalization combined with the quantum power method discussed in Sec. V C. We also show that the quantum power method can particularly simply evaluate the lowest-order moments.

1. Moment and cumulant

The Feynman propagator with respect to a state $|\Psi\rangle$ can be written as

$$K(t) = \langle \hat{U}(t) \rangle = \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} \mu_n, \quad (\text{E1})$$

where $\hat{U}(t)$ is the time-evolution operator given in Eq. (1) and

$$\mu_n = \langle \hat{\mathcal{H}}^n \rangle \quad (\text{E2})$$

is the n th Hamiltonian moment. We also define the generating function $\Phi(t)$ of the cumulants $\{\kappa_n\}$ as

$$\Phi(t) \equiv \ln K(t) = \ln \langle e^{-i\hat{\mathcal{H}}t} \rangle \equiv \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} \kappa_n. \quad (\text{E3})$$

Thus, the n th moment μ_n and cumulant κ_n are given by the n th time derivative of generating functions $K(t)$ and $\Phi(t)$,

respectively, as

$$\mu_n = i^n \left. \frac{d^n K(t)}{dt^n} \right|_{t=0} \quad (\text{E4})$$

and

$$\kappa_n = i^n \left. \frac{d^n \Phi(t)}{dt^n} \right|_{t=0}. \quad (\text{E5})$$

We note that, recently, a method making use of the expectation value of the time-evolution operator has been proposed for evaluating eigenvalues of the Hamiltonian [125].

It should be noticed [126] that the n th moment μ_n can be expressed as

$$\mu_n = \kappa_n + \sum_{k=1}^{n-1} \binom{n-1}{k-1} \kappa_k \mu_{n-k} \quad (\text{E6})$$

and, equivalently, the n th cumulant κ_n can be expressed as

$$\kappa_n = \mu_n - \sum_{k=1}^{n-1} \binom{n-1}{k-1} \kappa_k \mu_{n-k}. \quad (\text{E7})$$

Therefore, from the moments $\{\mu_k\}_{k \leq n}$, one can obtain the cumulants $\{\kappa_k\}_{k \leq n}$ and vice versa. A remarkable difference between these two quantities is that the magnitude of the moment grows exponentially in n as $\mu_n \sim O(N^n)$, while the magnitude of the cumulant remains as $\kappa_n \sim O(N)$ [127].

2. Finite-difference approximation

By approximating the derivative in Eq. (E4) with the central-finite-difference method, we obtain that

$$\mu_n = \mu_n(\Delta_\tau) + O(\Delta_\tau^2), \quad (\text{E8})$$

where

$$\mu_n(\Delta_\tau) = \sum_{i=0}^n c_{n,i} K\left(\left(\frac{n}{2} - i\right) \Delta_\tau\right). \quad (\text{E9})$$

Using $K(-\Delta_\tau) = K(\Delta_\tau)^*$, $\mu_n(\Delta_\tau)$ for n odd and n even can be expressed, respectively, as

$$\mu_{2m+1}(\Delta_\tau) = 2i \sum_{i=0}^m c_{2m+1,i} \text{Im} K\left(\left(m + \frac{1}{2} - i\right) \Delta_\tau\right) \quad (\text{E10})$$

and

$$\mu_{2m}(\Delta_\tau) = c_{2m,m} + 2 \sum_{i=0}^{m-1} c_{2m,i} \text{Re} K\left((m-i) \Delta_\tau\right), \quad (\text{E11})$$

where $K(0) = 1$ and $c_{2m,m} = (1/\Delta_\tau^{2m}) \binom{2m}{m}$ are used in Eq. (E11). Thus, for obtaining $\mu_n(\Delta_\tau)$, it suffices to evaluate $K(t)$ at equally spaced $\lceil n/2 \rceil$ different points, where $\lceil \cdot \rceil$ denotes the ceiling function defined previously. Moreover, if $\{K(l\Delta_\tau/2)\}_{l=1}^{n-1}$ used for evaluating the moments $\{\mu_l(\Delta_\tau)\}_{l=1}^{n-1}$ are all stored, only $K(n\Delta_\tau/2)$ has to be evaluated for $\mu_n(\Delta_\tau)$. Therefore, for obtaining all n moments $\{\mu_l(\Delta_\tau)\}_{l=1}^n$, it is sufficient to evaluate the propagator at n different points, i.e., $\{K(l\Delta_\tau/2)\}_{l=1}^n$, only once.

We can apply the same argument for the cumulants. Using the central-finite-difference method, the n th cumulant is evaluated as

$$\kappa_n = \kappa_n(\Delta_\tau) + O(\Delta_\tau^2), \quad (\text{E12})$$

where

$$\kappa_n(\Delta_\tau) = \sum_{i=0}^n c_{n,i} \Phi\left(\left(\frac{n}{2} - i\right) \Delta_\tau\right). \quad (\text{E13})$$

Because $\Phi(-\Delta_\tau) = \Phi(\Delta_\tau)^*$, $\kappa_n(\Delta_\tau)$ for n odd and n even can be expressed, respectively, as

$$\kappa_{2m+1}(\Delta_\tau) = 2i \sum_{i=0}^m c_{2m+1,i} \text{Im}\Phi\left(\left(m + \frac{1}{2} - i\right) \Delta_\tau\right) \quad (\text{E14})$$

and

$$\kappa_{2m}(\Delta_\tau) = 2 \sum_{i=0}^{m-1} c_{2m,i} \text{Re}\Phi\left((m - i) \Delta_\tau\right), \quad (\text{E15})$$

where $\Phi(0) = \ln K(0) = 0$ is used in Eq. (E15). Note that, if we write the propagator as $K(t) = a(t)e^{i\varphi(t)}$ with $a(t)$ and $\varphi(t)$ real, then $\text{Re}\Phi(t) = \ln a(t)$ and $\text{Im}\Phi(t) = \varphi(t)$, implying that the cumulants with odd order are related to the phase of $K(t)$, while the cumulants with even order are related to the amplitude of $K(t)$. Recently, an efficient method for estimating the overlap amplitude of two pure states has been proposed [128]. Such a method might be utilized for evaluating the cumulants with even order.

3. Quantum power method for moment and cumulant

As shown explicitly in the previous section, the n th moment μ_n can be approximated as a linear combination of the Feynman propagator $K(t)$, i.e., the expectation value of the time-evolution operator $\hat{U}(t)$, evaluated at different time variables $t_i^{(n)} = (\frac{n}{2} - i) \Delta_\tau$ for $i = 0, 1, \dots, n$. Similarly, the n th cumulant κ_n can be approximated as a linear combination of $\Phi(t)$, i.e., logarithm of the Feynman propagator $K(t)$, evaluated at different time variables $t_i^{(n)}$. Therefore, an important quantity here is again the time-evolution operator $\hat{U}(t)$.

To implement on quantum computers, the time-evolution operator is further decomposed approximately by using the symmetric Suzuki-Trotter decomposition as in Eq. (7). However, at this point, it is crucially important to recall the argument given in Sec. III B 4 and Appendix D. Although the time-evolution operator $\hat{U}(t)$ evaluated at time $t_i^{(n)} = (\frac{n}{2} - i) \Delta_\tau$ satisfies that

$$\hat{U}\left(\left(\frac{n}{2} - i\right) \Delta_\tau\right) = \left[\hat{U}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-2i}, \quad (\text{E16})$$

and thus the approximated n th moment $\mu_n(\Delta_\tau)$ in Eq. (E9) is equivalent to

$$\mu_n(\Delta_\tau) = \sum_{i=0}^n c_{n,i} \left\langle \left[\hat{U}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-2i} \right\rangle, \quad (\text{E17})$$

these are no longer generally correct when the time-evolution operators are approximated by the Suzuki-Trotter decomposition, i.e.,

$$\hat{S}_{2m}^{(p)}\left(\left(\frac{n}{2} - i\right) \Delta_\tau\right) \neq \left[\hat{S}_{2m}^{(p)}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-2i}. \quad (\text{E18})$$

Therefore, the Feynman propagator $K(t_i^{(n)})$ in Eq. (E9) can be approximated either as

$$K(t_i^{(n)}) = \left\langle \hat{S}_{2m}^{(p)}\left(\left(\frac{n}{2} - i\right) \Delta_\tau\right) \right\rangle + O(\Delta_\tau^{2m+1}) \quad (\text{E19})$$

or

$$K(t_i^{(n)}) = \left\langle \left[\hat{S}_{2m}^{(p)}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-2i} \right\rangle + O(\Delta_\tau^{2m+1}). \quad (\text{E20})$$

If the Feynman propagator $K(t_i^{(n)})$ is approximated as in Eq. (E20), the n th moment μ_n is given by

$$\mu_n = \sum_{i=0}^n c_{n,i} \left\langle \left[\hat{S}_{2m}^{(p)}\left(\frac{\Delta_\tau}{2}\right)\right]^{n-2i} \right\rangle + O(\Delta_\tau^2) + O(\Delta_\tau^{2m}) \quad (\text{E21})$$

and thus the lowest-order symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}$ with $m = 1$ can be adopted (see Sec. II B). This approach is suitable for the calculations of higher-order moments and cumulants. On the other hand, if the Feynman propagator $K(t_i^{(n)})$ is approximated as in Eq. (E19), the higher-order symmetric Suzuki-Trotter decomposition $\hat{S}_{2m}^{(p)}$ is required. As discussed in Appendix D, in order to evaluate the n th moment μ_n with the controlled accuracy, the order of the symmetric Suzuki-Trotter

decomposition $\hat{S}_{2m}^{(p)}$ must be $2m \geq n$ [see Eq. (D12)]. In this case, the systematic error is $O(\Delta_\tau^2)$, i.e.,

$$\mu_n = \sum_{i=0}^n c_{n,i} \left\langle \hat{S}_{2m}^{(p)} \left(\left(\frac{n}{2} - i \right) \Delta_\tau \right) \right\rangle + O(\Delta_\tau^2). \quad (\text{E22})$$

Therefore, this approach is not suitable for large n but is more preferable than the other approach when $n \leq 4$, assuming $p = 3$ in the symmetric Suzuki-Trotter decomposition. Since the cumulant κ_n can be expressed in terms of the moments $\{\mu_k\}_{k \leq n}$ as in Eq. (E7), the same argument is applied for the cumulant.

4. First and second moments

The first and second moments are the most fundamental quantities for many practical purposes because $\mu_1 = \langle \hat{\mathcal{H}} \rangle$ is the average of the energy and $\mu_2 = \langle \hat{\mathcal{H}}^2 \rangle$ is related to the variance of the energy. The first moment $\langle \hat{\mathcal{H}} \rangle$ is directly evaluated by measuring each term of the Hamiltonian $\hat{\mathcal{H}}$ on quantum computers. Perhaps, $\langle \hat{\mathcal{H}}^2 \rangle$ could also be evaluated in the same way, although terms to be measured are increased by a factor of $O(N)$, assuming that a Hamiltonian $\hat{\mathcal{H}}$ is local. The quantum power method can provide an alternative approach to evaluate these quantities with the same amount of resource.

From Eqs. (E10) and (E11), we can approximate the first and second moments μ_1 and μ_2 as

$$\mu_1(\Delta_\tau) = -\frac{2}{\Delta_\tau} \text{Im} \left\langle \hat{U} \left(\frac{\Delta_\tau}{2} \right) \right\rangle \quad (\text{E23})$$

and

$$\mu_2(\Delta_\tau) = \frac{2}{\Delta_\tau^2} \left[1 - \text{Re} \left\langle \hat{U}(\Delta_\tau) \right\rangle \right], \quad (\text{E24})$$

respectively. This is already remarkable because the second moment μ_2 is also estimated simply by the expectation value of a single time-evolution operator. To evaluate these quantities on quantum computers, the time-evolution operator $\hat{U}(\Delta_\tau)$ is approximated by the lowest-order symmetric Suzuki-Trotter decomposition $\hat{S}_2(\Delta_\tau)$ (see Appendix E3). Therefore, in the quantum power method, the first and second moments μ_1 and μ_2 are estimated simply by evaluating $\text{Im} \left\langle \hat{S}_2 \left(\frac{\Delta_\tau}{2} \right) \right\rangle$ and $\text{Re} \left\langle \hat{S}_2(\Delta_\tau) \right\rangle$, i.e.,

$$\mu_1(\Delta_\tau) \approx -\frac{2}{\Delta_\tau} \text{Im} \left\langle \hat{S}_2 \left(\frac{\Delta_\tau}{2} \right) \right\rangle \quad (\text{E25})$$

and

$$\mu_2(\Delta_\tau) \approx \frac{2}{\Delta_\tau^2} \left[1 - \text{Re} \left\langle \hat{S}_2(\Delta_\tau) \right\rangle \right], \quad (\text{E26})$$

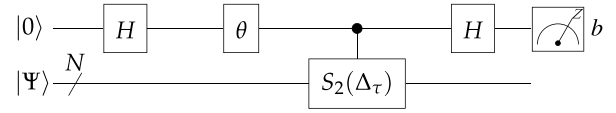


FIG. 20. Quantum circuit to evaluate $\text{Re}\langle \Psi | \hat{S}_2(\Delta_\tau) | \Psi \rangle$ or $\text{Im}\langle \Psi | \hat{S}_2(\Delta_\tau) | \Psi \rangle$. θ in the circuit denotes the phase gate such that $\hat{\theta}|0\rangle = |0\rangle$ and $\hat{\theta}|1\rangle = e^{i\theta}|1\rangle$. Since $P_0 - P_1 = \text{Re}[e^{i\theta} \langle \Psi | \hat{S}_2(\Delta_\tau) | \Psi \rangle]$, one can evaluate $\text{Re}\langle \Psi | \hat{S}_2(\Delta_\tau) | \Psi \rangle$ if $\theta = 0$ and $\text{Im}\langle \Psi | \hat{S}_2(\Delta_\tau) | \Psi \rangle$ if $\theta = -\pi/2$ from the difference of the probabilities P_0 and P_1 . Here, P_b is the probability for finding a bit b ($= 0, 1$) by measuring the ancilla qubit.

respectively. Although we have to introduce an ancilla qubit (see Fig. 20), $\mu_1 = \langle \hat{\mathcal{H}} \rangle$ and $\mu_2 = \langle \hat{\mathcal{H}}^2 \rangle$ can thus be estimated with exactly the same amount of resource. If noise in quantum devices is not destructively serious, this approach based on the quantum power method might be more suitable than the direct approach measuring all terms in $\hat{\mathcal{H}}$ and $\hat{\mathcal{H}}^2$.

Figure 21 shows the numerical results of μ_1 and μ_2 evaluated from Eqs. (E25) and (E26) for the spin-1/2 Heisenberg model defined in Eq. (65) with two different quantum states. We also show the results obtained by employing the first-order Richardson extrapolation, i.e.,

$$\mu_{n(1)}(\Delta_\tau) = \frac{h^2 \mu_n(\Delta_\tau/h) - \mu_n(\Delta_\tau)}{h^2 - 1} \quad (\text{E27})$$

for $n = 1$ and 2 , which expects that the systematic error scales as $O(\Delta_\tau^4)$, instead of $O(\Delta_\tau^2)$ without the Richardson extrapolation. Our numerical simulations clearly demonstrate that the systematic errors are well controlled and the results converge smoothly to the exact values in the limit of $\Delta_\tau \rightarrow 0$. The quantum power method for the first and second moments could be useful to, e.g., the energy variance minimization for optimizing a parametrized quantum circuit [129].

Here, we only consider the first and second moments, but the higher-order moments can be similarly evaluated. For example, the third and fourth moments are given as

$$\mu_3(\Delta_\tau) = \frac{2}{\Delta_\tau^3} \left[\text{Im} \left\langle \hat{U} \left(\frac{3\Delta_\tau}{2} \right) \right\rangle - 3 \text{Im} \left\langle \hat{U} \left(\frac{\Delta_\tau}{2} \right) \right\rangle \right] \quad (\text{E28})$$

and

$$\mu_4(\Delta_\tau) = \frac{2}{\Delta_\tau^4} \left[\text{Re} \left\langle \hat{U}(2\Delta_\tau) \right\rangle - 4 \text{Re} \left\langle \hat{U}(\Delta_\tau) \right\rangle + 3 \right], \quad (\text{E29})$$

respectively. To implement these on quantum computers, $\hat{U}(\Delta_\tau)$ is now approximated by using the higher-order Suzuki-Trotter decomposition $\hat{S}_4^{(p)}(\Delta_\tau)$ with $m = 2$ (also see Fig. 19), which is still affordable.

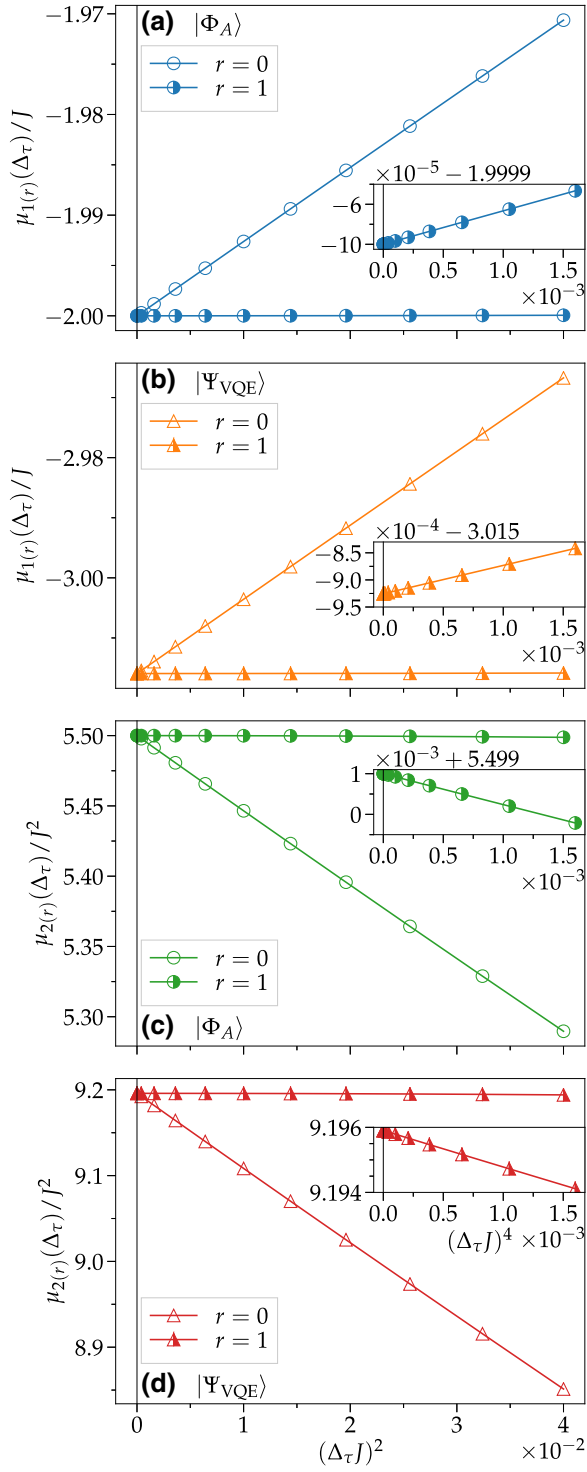


FIG. 21. (a),(b) The first moment μ_1 and (c),(d) the second moment μ_2 as a function of $\Delta\tau^2$ evaluated from Eqs. (E25) and (E26), respectively, by numerical simulations for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring. For a quantum state $|\Psi\rangle$, we choose (a),(c) the singlet-pair product state $|\Phi_A\rangle$ in Eq. (80) and (b),(d) the VQE state $|\Psi_{\text{VQE}}\rangle$ in Eq. (88). The results obtained by the first-order Richardson extrapolation ($r = 1$) are also plotted. The insets show the same results for the first-order Richardson extrapolation but plotted against $\Delta\tau^4$. The exact values are indicated at $\Delta\tau = 0$ with the filled symbols.

5. Imaginary-time evolution

For an application of the cumulants, we now consider the imaginary-time evolution of a quantum state $|\Psi\rangle$, i.e.,

$$|\Psi(\tau)\rangle = \frac{e^{-\tau\hat{\mathcal{H}}/2}|\Psi\rangle}{\sqrt{\langle\Psi|e^{-\tau\hat{\mathcal{H}}|\Psi\rangle}} \quad (\text{E30})$$

for τ real. We introduce a simplified notation for the imaginary-time-dependent expectation value as $\langle\cdots\rangle_\tau \equiv \langle\Psi(\tau)|\cdots|\Psi(\tau)\rangle$. Then, the energy expectation value with respect to $|\Psi(\tau)\rangle$ is given as

$$\begin{aligned} E(\tau) &= \langle\hat{\mathcal{H}}\rangle_\tau \\ &= \frac{\langle\Psi|e^{-\tau\hat{\mathcal{H}}/2}\hat{\mathcal{H}}e^{-\tau\hat{\mathcal{H}}/2}|\Psi\rangle}{\langle\Psi|e^{-\tau\hat{\mathcal{H}}|\Psi\rangle}} \\ &= \frac{\langle\hat{\mathcal{H}}e^{-\tau\hat{\mathcal{H}}}\rangle}{\langle e^{-\tau\hat{\mathcal{H}}}\rangle}. \end{aligned} \quad (\text{E31})$$

Observing that $E(\tau) = -(d/d\tau) \ln\langle e^{-\tau\hat{\mathcal{H}}}\rangle = -(d/d\tau) \sum_{n=0}^{\infty} [(-\tau)^n/n!]\kappa_n$, the CMX of the energy is given as [126]

$$E(\tau) = \sum_{n=0}^{\infty} \frac{(-\tau)^n}{n!} \kappa_{n+1}. \quad (\text{E32})$$

Figure 22 shows the exact $E(\tau)$ and the CMX of the energy truncated at the n_{max} th cumulant

$$E_{n_{\text{max}}}(\tau) = \sum_{n=0}^{n_{\text{max}}-1} \frac{(-\tau)^n}{n!} \kappa_{n+1} \quad (\text{E33})$$

for the spin-1/2 Heisenberg model defined in Eq. (65), where the VQE state $|\Psi_{\text{VQE}}\rangle$ in Eq. (88) is selected for the quantum state $|\Psi\rangle$ in Eq. (E30). The energy $E(\tau)$ with the exact imaginary-time evolution decreases monotonically in τ , because the first derivative of $E(\tau)$ is minus of the energy fluctuation [126]

$$\frac{dE(\tau)}{d\tau} = -\left(\langle\hat{\mathcal{H}}^2\rangle_\tau - \langle\hat{\mathcal{H}}\rangle_\tau^2\right) \leq 0, \quad (\text{E34})$$

where the equality satisfies if and only if $|\Psi_{\text{VQE}}(\tau)\rangle$ is an exact eigenstate (e.g., the ground state) of $\hat{\mathcal{H}}$. On the other hand, due to the truncation of the series at finite order, $E_{n_{\text{max}}}(\tau)$ at large τ diverges to $-\infty$ for even $n_{\text{max}} \geq 2$ or to $+\infty$ for odd $n_{\text{max}} \geq 3$. Note that $E_2(\tau) = \kappa_1 - \kappa_2\tau$ is the tangent line of $E(\tau)$ at $\tau = 0$. We also find that the convergence of $E_{n_{\text{max}}}(\tau)$ to the exact ground-state energy E_0 with respect to the power exponents n_{max} in the cumulants required is rather slower, as compared with the Krylov-subspace diagonalization with either $M_B = 1$ or $M_B = 9$

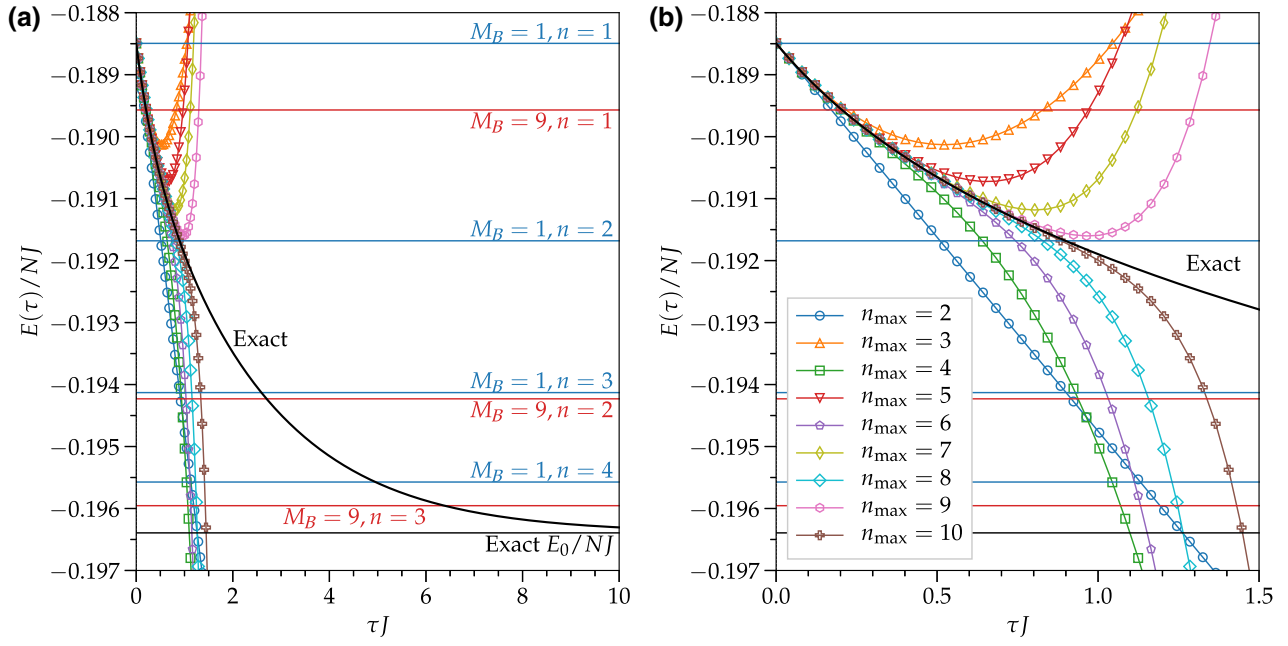


FIG. 22. (a) Energy expectation value $E(\tau)$ with respect to the exact imaginary-time evolution of $|\Psi_{\text{VQE}}\rangle$ (black solid line) and the CMX energy $E_{n_{\text{max}}}(\tau)$ with various truncation order n_{max} (symbols) as a function of the imaginary time τ for the spin-1/2 Heisenberg model on an $N = 16$ qubit ring given in Eq. (65). For comparison, the energies estimated with the Krylov-subspace diagonalization involving $|\Psi_{\text{VQE}}\rangle$ as a reference state with $M_B = 1$ for $1 \leq n \leq 4$ (blue horizontal lines) and with $M_B = 9$ for $1 \leq n \leq 3$ (red horizontal lines) are shown. These are the same results in Fig. 5. Here, $n = \dim \mathcal{K}_n / M_B$, i.e., the dimension of the Krylov subspace \mathcal{K}_n per block size M_B (for details, see Sec. V C). (b) Same as (a) but an enlarged plot for $0 \leq \tau J \leq 1.5$.

discussed in Sec. V C. This is not quite surprising because the form of $E(\tau)$ in Eq. (E33) is an expansion around $\tau = 0$, which is analogous to the high-temperature expansion.

Finally, it should be noted that recently several schemes different from the quantum power method to evaluate $\langle \hat{\mathcal{H}}^n \rangle$ are proposed and demonstrated with various CMX methods for quantum-chemistry Hamiltonians [130] and with the Lanczos method for Heisenberg Hamiltonians [131].

APPENDIX F: LANCZOS METHOD

In this appendix, we briefly outline the Lanczos method with an emphasis on its aspect as a moment method [132,133], i.e., a potential application of the quantum power method.

1. Lanczos tridiagonal matrix and Hamiltonian moment

The Lanczos method generates a sequence of orthonormalized states $\{|q_i\rangle\}$, satisfying $\langle q_i | q_j \rangle = \delta_{ij}$, from an initial (reference) state $|q_1\rangle = |\Psi\rangle$ recursively as

$$\hat{\mathcal{H}}|q_i\rangle = \beta_{i-1}|q_{i-1}\rangle + \alpha_i|q_i\rangle + \beta_i|q_{i+1}\rangle, \quad (\text{F1})$$

with $\alpha_i = \langle q_i | \hat{\mathcal{H}} | q_i \rangle$, $\beta_i = \langle q_i | \hat{\mathcal{H}} | q_{i+1} \rangle$, $\beta_0 \equiv 0$, and $|q_0\rangle \equiv 0$. After obtaining $\{|q_i\rangle\}_{i=1}^n$, the Hamiltonian $\hat{\mathcal{H}}$ can be

represented as a tridiagonal matrix $[\mathbf{T}_n]_{ij} = \langle q_i | \hat{\mathcal{H}} | q_j \rangle$ as

$$\mathbf{T}_n = \begin{bmatrix} \alpha_1 & \beta_1 & & & & & & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & & & & & & \\ & \beta_2 & \alpha_3 & \ddots & & & & & & & \\ & & \ddots & \ddots & & & & & & & \\ & & & & \ddots & \ddots & & & & & \\ & & & & & & \beta_{n-1} & & & & \\ & & & & & & \beta_{n-1} & \alpha_n & & & \end{bmatrix}. \quad (\text{F2})$$

The matrix elements $\{\alpha_i\}$ and $\{\beta_i\}$ can also be constructed recursively using the Hamiltonian moments [132, 133]. Following Ref. [132], $\{\alpha_i\}$ and $\{\beta_i\}$ are given in terms of $\{\mu_n\}$ as

$$\alpha_i = \left(\frac{\mathcal{L}_{i-1}}{\mathcal{L}_{i-2}} \right) \left(\frac{\mathcal{M}_{i-2}}{\mathcal{M}_{i-3}} \right)^{-1} + \left(\frac{\mathcal{M}_{i-1}}{\mathcal{M}_{i-2}} \right) \left(\frac{\mathcal{L}_{i-1}}{\mathcal{L}_{i-2}} \right)^{-1} \quad (\text{F3})$$

and

$$\beta_i^2 = \left(\frac{\mathcal{L}_i}{\mathcal{L}_{i-1}} \right) \left(\frac{\mathcal{L}_{i-1}}{\mathcal{L}_{i-2}} \right)^{-1}, \quad (\text{F4})$$

where $\mathcal{L}_n \equiv \det \mathcal{L}_n$ and $\mathcal{M}_n \equiv \det \mathcal{M}_n$ are determinants of $(n+1) \times (n+1)$ Hankel matrices defined, respectively, as $[\mathcal{L}_n]_{ij} = \mu_{i+j-2}$ and $[\mathcal{M}_n]_{ij} = \mu_{i+j-1}$, or more

explicitly

$$\mathcal{L}_n = \begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_{n-1} & \mu_n \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_n & \mu_{n+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{n+1} & \mu_{n+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{n-1} & \mu_n & \mu_{n+1} & \cdots & \mu_{2n-2} & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n-1} & \mu_{2n} \end{bmatrix} \quad (\text{F5})$$

for $n \geq 0$ and

$$\mathcal{M}_n = \begin{bmatrix} \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_n & \mu_{n+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{n+1} & \mu_{n+2} \\ \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{n+2} & \mu_{n+3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n-1} & \mu_{2n} \\ \mu_{n+1} & \mu_{n+2} & \mu_{n+3} & \cdots & \mu_{2n} & \mu_{2n+1} \end{bmatrix} \quad (\text{F6})$$

for $n \geq 0$. Equations (F3) and (F4) hold for $i \geq 1$ provided that \mathcal{M}_n and \mathcal{L}_n with negative indices are defined as $\mathcal{M}_{-1} = 1$, $\mathcal{M}_{-2} = 0$, and $\mathcal{L}_{-1} = 1$. The Hankel matrices \mathcal{L}_{n-1} and \mathcal{M}_{n-1} are identical, respectively, to \mathbf{S} in Eq. (45) and \mathbf{H} in Eq. (44) if $M_B = 1$. It is noticed in Eqs. (F3) and (F4) that the Lanczos matrix elements α_i and β_i are expressed in terms of the ratios of the Hankel determinants whose matrix dimensions differ only by 1. The particular structure of the Hankel matrices \mathcal{L}_n and \mathcal{M}_n allows us to evaluate the ratios of the determinants appearing in Eqs. (F3) and (F4) recursively, as described in Appendix F 2.

It is instructive to give the explicit forms of the first few matrix elements of \mathbf{T}_n . The first three matrix elements required for constructing the 2×2 matrix \mathbf{T}_2 are given by

$$\alpha_1 = \langle \hat{\mathcal{H}} \rangle, \quad (\text{F7})$$

$$\beta_1 = \sqrt{\langle \hat{\mathcal{H}}^2 \rangle - \langle \hat{\mathcal{H}} \rangle^2}, \quad (\text{F8})$$

$$\alpha_2 = \frac{\langle \hat{\mathcal{H}}^3 \rangle - 2\langle \hat{\mathcal{H}}^2 \rangle \langle \hat{\mathcal{H}} \rangle + \langle \hat{\mathcal{H}} \rangle^3}{\langle \hat{\mathcal{H}}^2 \rangle - \langle \hat{\mathcal{H}} \rangle^2}, \quad (\text{F9})$$

where $\langle \cdots \rangle = \langle q_1 | \cdots | q_1 \rangle$. Therefore, α_1 and β_1^2 are the energy expectation value and the energy variance with respect to the initial state $|q_1\rangle$, respectively.

2. Ratio of Hankel determinants

We now describe a way to calculate recursively the ratio of the determinants appearing in Eqs. (F3) and (F4). Let us first review the determinant and the matrix-inversion formulas for general matrices. Let \mathbf{A}_n be an $n \times n$ matrix, \mathbf{b} be an $n \times 1$ matrix, \mathbf{c} be an $n \times 1$ matrix, and d be

a 1×1 matrix (i.e., a scalar), and let us consider an $(n+1) \times (n+1)$ matrix \mathbf{A}_{n+1} of the form

$$\mathbf{A}_{n+1} = \begin{bmatrix} \mathbf{A}_n & \mathbf{b} \\ \mathbf{c}^T & d \end{bmatrix}. \quad (\text{F10})$$

If we define

$$r = d - \mathbf{c}^T \mathbf{A}_n^{-1} \mathbf{b}, \quad (\text{F11})$$

the determinant of \mathbf{A}_{n+1} is given by

$$\det \mathbf{A}_{n+1} = \det \begin{bmatrix} \mathbf{A}_n & \mathbf{b} \\ \mathbf{c}^T & d \end{bmatrix} = r \det \mathbf{A}_n, \quad (\text{F12})$$

and the inverse \mathbf{A}_{n+1}^{-1} is given by

$$\mathbf{A}_{n+1}^{-1} = \begin{bmatrix} \mathbf{A}_n & \mathbf{b} \\ \mathbf{c}^T & d \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_n^{-1} + (\mathbf{A}_n^{-1} \mathbf{b})(\mathbf{c}^T \mathbf{A}_n^{-1})/r & -\mathbf{A}_n^{-1} \mathbf{b}/r \\ -\mathbf{c}^T \mathbf{A}_n^{-1}/r & 1/r \end{bmatrix}. \quad (\text{F13})$$

Now we apply the above formulas to recursively evaluate the ratios of the determinants of \mathcal{L}_n and \mathcal{L}_{n-1} . Due to its particular structure, \mathcal{L}_n can be expressed in terms of \mathcal{L}_{n-1} as

$$\mathcal{L}_n = \begin{bmatrix} \mathcal{L}_{n-1} & \mathbf{m}_n \\ \mathbf{m}_n^T & \mu_{2n} \end{bmatrix} \quad (\text{F14})$$

with the following n -dimensional vector:

$$\mathbf{m}_n^T = (\mu_n, \mu_{n+1}, \dots, \mu_{2n-1}). \quad (\text{F15})$$

From the formula in Eq. (F12), the ratio of the determinants is given by

$$\frac{\mathcal{L}_n}{\mathcal{L}_{n-1}} = \frac{\det \mathcal{L}_n}{\det \mathcal{L}_{n-1}} = r_n \quad (\text{F16})$$

with

$$r_n = \mu_{2n} - \mathbf{m}_n^T \mathcal{L}_{n-1}^{-1} \mathbf{m}_n, \quad (\text{F17})$$

which involves the inverse \mathcal{L}_{n-1}^{-1} whose dimension is less than that of \mathcal{L}_n^{-1} by 1.

The inverse matrix \mathcal{L}_n^{-1} can be calculated using Eq. (F13). Starting with

$$\mathcal{L}_0^{-1} = \mu_0^{-1}, \quad (\text{F18})$$

\mathcal{L}_n^{-1} for $n \geq 1$ can be constructed from \mathcal{L}_{n-1}^{-1} and \mathbf{m}_n recursively as

$$\begin{aligned} \mathcal{L}_n^{-1} &= \begin{bmatrix} \mathcal{L}_{n-1}^{-1} & \mathbf{m}_n \\ \mathbf{m}_n^T & \mu_{2n} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathcal{L}_{n-1}^{-1} + (\mathcal{L}_{n-1}^{-1} \mathbf{m}_n)(\mathcal{L}_{n-1}^{-1} \mathbf{m}_n)^T / r_n & -\mathcal{L}_{n-1}^{-1} \mathbf{m}_n / r_n \\ -(\mathcal{L}_{n-1}^{-1} \mathbf{m}_n)^T / r_n & 1 / r_n \end{bmatrix}, \end{aligned} \quad (\text{F19})$$

where $(\mathcal{L}_n^{-1})^T = \mathcal{L}_n^{-1}$ is used. Thus, starting with the known \mathcal{L}_0^{-1} and using Eqs. (F17) and (F19), one can obtain $\{r_n\}$ recursively as $\mathcal{L}_0^{-1} \rightarrow r_1 \rightarrow \mathcal{L}_1^{-1} \rightarrow r_2 \rightarrow \mathcal{L}_2^{-1} \rightarrow r_3 \rightarrow \dots$.

It should be noted that Eq. (F17) involves a matrix-vector multiplication and, in addition, Eq. (F19) involves a rank-1 update. Therefore, the complexity for computing the ratio of determinants in Eq. (F16) is $O(n^2)$. This is more efficient when n is large because the direct calculation of a determinant from scratch, e.g., by using the LU decomposition, requires $O(n^3)$ operations. Noticing that $[\mathcal{L}_n]_{ij} = \mu_{i+j-2}$ while $[\mathcal{M}_n]_{ij} = \mu_{i+j-1}$, the similar recursive formula for \mathcal{M}_n can be readily derived simply by replacing the indexes for the moments in the above as $\{\mu_i\}_{i=0}^{2n} \rightarrow \{\mu_{i+1}\}_{i=0}^{2n}$.

-
- [1] J. P. F. LeBlanc *et al.* (Simons Collaboration on the Many-Electron Problem), Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms, *Phys. Rev. X* **5**, 041041 (2015).
- [2] M. Motta *et al.* (Simons Collaboration on the Many-Electron Problem), Towards the Solution of the Many-Electron Problem in Real Materials: Equation of State of the Hydrogen Chain with State-of-the-Art Many-Body Methods, *Phys. Rev. X* **7**, 031059 (2017).
- [3] M. Motta, C. Genovese, F. Ma, Z.-H. Cui, R. Sawaya, G. K.-L. Chan, N. Chepiga, P. Helms, C. Jiménez-Hoyos, A. J. Millis, U. Ray, E. Ronca, H. Shi, S. Sorella, E. M. Stoudenmire, S. R. White, and S. Zhang (Simons Collaboration on the Many-Electron Problem), Ground-State Properties of the Hydrogen Chain: Dimerization, Insulator-to-Metal Transition, and Magnetic Phases, *Phys. Rev. X* **10**, 031058 (2020).
- [4] J. J. Eriksen *et al.*, The ground state electronic energy of benzene, *J. Phys. Chem. Lett.* **11**, 8922 (2020).
- [5] J. Liesen and Z. Strakoš, *Krylov Subspace Methods* (Oxford University Press, Oxford, 2013), Chap. 3.
- [6] E. Dagotto, Correlated electrons in high-temperature superconductors, *Rev. Mod. Phys.* **66**, 763 (1994).
- [7] J. Jaklič and P. Prelošek, Lanczos method for the calculation of finite-temperature quantities in correlated systems, *Phys. Rev. B* **49**, 5065 (1994).
- [8] J. Jaklič and P. Prelošek, Finite-temperature properties of doped antiferromagnets, *Adv. Phys.* **49**, 1 (2000).
- [9] A. Weiße and H. Fehske, in *Computational Many-Particle Physics*, edited by H. Fehske, R. Schneider, and A. Weiße (Springer, Berlin, Heidelberg, 2008), Vol. 739, p. 529.
- [10] P. Prelošek and J. Bonča, in *Strongly Correlated Systems: Numerical Methods*, edited by A. Avella and F. Mancini (Springer, Berlin, Heidelberg, 2013), p. 1.
- [11] E. Koch, in *Many-Body Methods for Real Materials*, Schriften des Forschungszentrums Jülich Modeling and Simulation Vol. 9, edited by E. Pavarini and S. Zhang (Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag, Jülich, 2019), Chap. 7.
- [12] A. Weiße and H. Fehske, in *Computational Many-Particle Physics*, edited by H. Fehske, R. Schneider, and A. Weiße (Springer, Berlin, Heidelberg, 2008), p. 545.
- [13] H. Tal-Ezer and R. Kosloff, An accurate and efficient scheme for propagating the time dependent Schrödinger equation, *J. Chem. Phys.* **81**, 3967 (1984).
- [14] T. J. Park and J. C. Light, Unitary quantum time evolution by iterative Lanczos reduction, *J. Chem. Phys.* **85**, 5870 (1986).
- [15] A. Vijay and H. Metiu, A polynomial expansion of the quantum propagator, the Green's function, and the spectral density operator, *J. Chem. Phys.* **116**, 60 (2002).
- [16] T. Iitaka and T. Ebisuzaki, Algorithm for Linear Response Functions at Finite Temperatures: Application to ESR Spectrum of $s = \frac{1}{2}$ Antiferromagnet Cu Benzoate, *Phys. Rev. Lett.* **90**, 047203 (2003).
- [17] N. Mohankumar and S. M. Auerbach, On time-step bounds in unitary quantum evolution using the Lanczos method, *Comput. Phys. Commun.* **175**, 473 (2006).
- [18] S. Sorella, Generalized Lanczos algorithm for variational quantum Monte Carlo, *Phys. Rev. B* **64**, 024512 (2001).
- [19] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [20] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, *Science* **309**, 1704 (2005).
- [21] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, Solving strongly correlated electron models on a quantum computer, *Phys. Rev. A* **92**, 062318 (2015).
- [22] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [23] B. Bauer, S. Bravyi, M. Motta, and G. Kin-Lic Chan, Quantum algorithms for quantum chemistry and quantum materials science, *Chem. Rev.* **120**, 12685 (2020).
- [24] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, Coherent control of macroscopic quantum states in a single-Cooper-pair box, *Nature* **398**, 786 (1999).
- [25] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, Linear optical quantum computing with photonic qubits, *Rev. Mod. Phys.* **79**, 135 (2007).
- [26] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Quantum computers, *Nature* **464**, 45 (2010).

- [27] Z.-L. Xiang, S. Ashhab, J. Q. You, and F. Nori, Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems, *Rev. Mod. Phys.* **85**, 623 (2013).
- [28] R. Barends *et al.*, Superconducting quantum circuits at the surface code threshold for fault tolerance, *Nature* **508**, 500 (2014).
- [29] J. M. Chow, J. M. Gambetta, E. Magesan, D. W. Abraham, A. W. Cross, B. R. Johnson, N. A. Masluk, C. A. Ryan, J. A. Smolin, S. J. Srinivasan, and M. Steffen, Implementing a strand of a scalable fault-tolerant quantum computing fabric, *Nat. Commun.* **5**, 4015 (2014).
- [30] J. Kelly *et al.*, State preservation by repetitive error detection in a superconducting quantum circuit, *Nature* **519**, 66 (2015).
- [31] D. Ristè, S. Poletto, M.-Z. Huang, A. Bruno, V. Vesterinen, O.-P. Saira, and L. DiCarlo, Detecting bit-flip errors in a logical qubit using stabilizer measurements, *Nat. Commun.* **6**, 6983 (2015).
- [32] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
- [33] W. Asavanant, Y. Shiozawa, S. Yokoyama, B. Charoensombutamon, H. Emura, R. N. Alexander, S. Takeda, J.-i. Yoshikawa, N. C. Menicucci, H. Yonezawa, and A. Furusawa, Generation of time-domain-multiplexed two-dimensional cluster state, *Science* **366**, 373 (2019).
- [34] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [35] 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).
- [36] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [37] P. J. J. O'Malley *et al.*, Scalable Quantum Simulation of Molecular Energies, *Phys. Rev. X* **6**, 031007 (2016).
- [38] 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).
- [39] 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).
- [40] J. Li, X. Yang, X. Peng, and C.-P. Sun, Hybrid Quantum-Classical Approach to Quantum Optimal Control, *Phys. Rev. Lett.* **118**, 150503 (2017).
- [41] G. Mazzola, P. J. Ollitrault, P. K. Barkoutsos, and I. Tavernelli, Nonunitary Operations for Ground-State Calculations in Near-Term Quantum Computers, *Phys. Rev. Lett.* **123**, 130501 (2019).
- [42] F. Arute *et al.*, Hartree-Fock on a superconducting qubit quantum computer, *Science* **369**, 1084 (2020).
- [43] J.-G. Liu, Y.-H. Zhang, Y. Wan, and L. Wang, Variational quantum eigensolver with fewer qubits, *Phys. Rev. Res.* **1**, 023025 (2019).
- [44] M. Foss-Feig, D. Hayes, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, J. M. Pino, and A. C. Potter, Holographic quantum algorithms for simulating correlated spin systems, [arXiv:2005.03023](https://arxiv.org/abs/2005.03023) [quant-ph] (2020).
- [45] 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.* **16**, 205 (2019).
- [46] 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).
- [47] H. Nishi, T. Kosugi, and Y. Matsushita, Implementation of quantum imaginary-time evolution method on NISQ devices: Nonlocal approximation, [arXiv:2005.12715](https://arxiv.org/abs/2005.12715) [quant-ph] (2020).
- [48] N. Gomes, F. Zhang, N. F. Berthussen, C.-Z. Wang, K.-M. Ho, P. P. Orth, and Y. Yao, Efficient step-merged quantum imaginary time evolution algorithm for quantum chemistry, *J. Chem. Theory Comput.* **16**, 6256 (2020).
- [49] K. Yeter-Aydeniz, G. Siopsis, and R. C. Pooser, Scattering in the Ising Model Using Quantum Lanczos Algorithm, [arXiv:2008.08763](https://arxiv.org/abs/2008.08763) [quant-ph] (2020).
- [50] N. H. Stair, R. Huang, and F. A. Evangelista, A multireference quantum Krylov algorithm for strongly correlated electrons, *J. Chem. Theory Comput.* **16**, 2236 (2020).
- [51] R. M. Parrish and P. L. McMahon, Quantum Filter Diagonalization: Quantum Eigendecomposition without Full Quantum Phase Estimation, [arXiv:1909.08925](https://arxiv.org/abs/1909.08925) [quant-ph] (2019).
- [52] O. Kyriienko, Quantum inverse iteration algorithm for programmable quantum simulators, *npj Quantum Inf.* **6**, 7 (2020).
- [53] A. M. Childs, R. Kothari, and R. D. Somma, Quantum algorithm for systems of linear equations with exponentially improved dependence on precision, *SIAM J. Comput.* **46**, 1920 (2017).
- [54] T. Liu, J.-G. Liu, L. Wang, and H. Fan, Probabilistic Nonunitary Gate in Imaginary Time Evolution, [arXiv:2006.09726](https://arxiv.org/abs/2006.09726) [quant-ph] (2020).
- [55] 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).
- [56] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, Computation of Molecular Spectra on a Quantum Processor with an Error-Resilient Algorithm, *Phys. Rev. X* **8**, 011021 (2018).
- [57] R. M. Parrish, E. G. Hohenstein, P. L. McMahon, and T. J. Martínez, Quantum Computation of Electronic Transitions Using a Variational Quantum Eigensolver, *Phys. Rev. Lett.* **122**, 230401 (2019).
- [58] K. M. Nakanishi, K. Mitarai, and K. Fujii, Subspace-search variational quantum eigensolver for excited states, *Phys. Rev. Res.* **1**, 033062 (2019).
- [59] K. Heya, K. M. Nakanishi, K. Mitarai, and K. Fujii, Subspace Variational Quantum Simulator, [arXiv:1904.08566](https://arxiv.org/abs/1904.08566) [quant-ph] (2019).
- [60] W. J. Huggins, J. Lee, U. Baek, B. O'Gorman, and K. B. Whaley, A non-orthogonal variational quantum eigensolver, *New J. Phys.* **22**, 073009 (2020).

- [61] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000).
- [62] P. Jordan and E. Wigner, Über das paulische Äquivalenzverbot, *Z. Phys.* **47**, 631 (1928).
- [63] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, *Ann. Phys.* **298**, 210 (2002).
- [64] J. T. Seeley, M. J. Richard, and P. J. Love, The Bravyi-Kitaev transformation for quantum computation of electronic structure, *J. Chem. Phys.* **137**, 224109 (2012).
- [65] A. Tranter, S. Sofia, J. Seeley, M. Kaicher, J. McClean, R. Babbush, P. V. Coveney, F. Mintert, F. Wilhelm, and P. J. Love, The Bravyi-Kitaev transformation: Properties and applications, *Int. J. Quantum Chem.* **115**, 1431 (2015).
- [66] V. Havlíček, M. Troyer, and J. D. Whitfield, Operator locality in the quantum simulation of fermionic models, *Phys. Rev. A* **95**, 032332 (2017).
- [67] A. M. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, *Quantum Inf. Comput.* **12**, 901 (2012).
- [68] T. Kosugi and Y. Matsushita, Construction of green's functions on a quantum computer: Quasiparticle spectra of molecules, *Phys. Rev. A* **101**, 012330 (2020).
- [69] T. Kosugi and Y.-i. Matsushita, Linear-response functions of molecules on a quantum computer: Charge and spin responses and optical absorption, *Phys. Rev. Res.* **2**, 033043 (2020).
- [70] The expectation value of $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta\tau)$ with an odd power n might be evaluated with a Hadamard-test-like circuit by introducing an additional ancilla qubit.
- [71] M. Suzuki, Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations, *Phys. Lett. A* **146**, 319 (1990).
- [72] H. Yoshida, Construction of higher order symplectic integrators, *Phys. Lett. A* **150**, 262 (1990).
- [73] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, *J. Math. Phys.* **32**, 400 (1991).
- [74] N. Hatano and M. Suzuki, Finding exponential product formulas of higher orders, *Lect. Notes Phys.* **679**, 37 (2005).
- [75] M. Suzuki, General nonsymmetric higher-order decomposition of exponential operators and symplectic integrators, *J. Phys. Soc. Jpn.* **61**, 3015 (1992).
- [76] F. Chatelin, *Eigenvalues of Matrices* (SIAM, Philadelphia, 2012).
- [77] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, *Quantum Sci. Technol.* **4**, 014008 (2018).
- [78] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer, *Quantum Sci. Technol.* **4**, 045005 (2019).
- [79] 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).
- [80] D. Loss and D. P. DiVincenzo, Quantum computation with quantum dots, *Phys. Rev. A* **57**, 120 (1998).
- [81] D. P. DiVincenzo, D. Bacon, J. Kempe, G. Burkard, and K. B. Whaley, Universal quantum computation with the exchange interaction, *Nature* **408**, 339 (2000).
- [82] R. Brunner, Y.-S. Shin, T. Obata, M. Pioro-Ladrière, T. Kubo, K. Yoshida, T. Taniyama, Y. Tokura, and S. Tarucha, Two-Qubit Gate of Combined Single-Spin Rotation and Interdot Spin Exchange in a Double Quantum Dot, *Phys. Rev. Lett.* **107**, 146801 (2011).
- [83] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum principal component analysis, *Nat. Phys.* **10**, 631 (2014).
- [84] H.-K. Lau and M. B. Plenio, Universal Quantum Computing with Arbitrary Continuous-Variable Encoding, *Phys. Rev. Lett.* **117**, 100501 (2016).
- [85] H. Fan, V. Roychowdhury, and T. Szkopek, Optimal two-qubit quantum circuits using exchange interactions, *Phys. Rev. A* **72**, 052323 (2005).
- [86] S. Balakrishnan and R. Sankaranarayanan, Entangling characterization of SWAP^{1/m} and controlled unitary gates, *Phys. Rev. A* **78**, 052305 (2008).
- [87] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes, Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm, *npj Quantum Inf.* **6**, 10 (2020).
- [88] G. Vidal and C. M. Dawson, Universal quantum circuit for two-qubit transformations with three controlled-NOT gates, *Phys. Rev. A* **69**, 010301 (2004).
- [89] A. Chiesa, F. Tacchino, M. Grossi, P. Santini, I. Tavernelli, D. Gerace, and S. Carretta, Quantum hardware simulating four-dimensional inelastic neutron scattering, *Nat. Phys.* **15**, 455 (2019).
- [90] Our decomposition of the e -SWAP gate in Ref. [96] is not optimal in terms of the number of CNOT gates.
- [91] D. A. Drabold and O. F. Sankey, Maximum Entropy Approach for Linear Scaling in the Electronic Structure Problem, *Phys. Rev. Lett.* **70**, 3631 (1993).
- [92] A. Hams and H. De Raedt, Fast algorithm for finding the eigenvalue distribution of very large matrices, *Phys. Rev. E* **62**, 4365 (2000).
- [93] T. Iitaka and T. Ebisuzaki, Random phase vector for calculating the trace of a large matrix, *Phys. Rev. E* **69**, 057701 (2004).
- [94] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, The kernel polynomial method, *Rev. Mod. Phys.* **78**, 275 (2006).
- [95] K. Seki and S. Yunoki, Thermodynamic properties of an $S = \frac{1}{2}$ ring-exchange model on the triangular lattice, *Phys. Rev. B* **101**, 235115 (2020).
- [96] K. Seki, T. Shirakawa, and S. Yunoki, Symmetry-adapted variational quantum eigensolver, *Phys. Rev. A* **101**, 052340 (2020).
- [97] W. Marshall, Antiferromagnetism, *Proc. R. Soc. London, Ser. A. Math. Phys. Sci.* **232**, 48 (1955).
- [98] E. Lieb and D. Mattis, Ordering energy levels of interacting spin systems, *J. Math. Phys.* **3**, 749 (1962).
- [99] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, *Phys. Rev. X* **7**, 021050 (2017).
- [100] S. Endo, S. C. Benjamin, and Y. Li, Practical Quantum Error Mitigation for Near-Future Applications, *Phys. Rev. X* **8**, 031027 (2018).

- [101] K. Seki and S. Sorella, Benchmark study of an auxiliary-field quantum Monte Carlo technique for the Hubbard model with shifted-discrete Hubbard-Stratonovich transformations, *Phys. Rev. B* **99**, 144407 (2019).
- [102] J. Oitmaa, C. Hamer, and W. Zheng, *Series Expansion Methods for Strongly Interacting Lattice Models* (Cambridge, New York, 2006).
- [103] T. A. Bespalova and O. Kyriienko, Hamiltonian operator approximation for energy measurement and ground state preparation, [arXiv:2009.03351](https://arxiv.org/abs/2009.03351) [quant-ph] (2020).
- [104] J.-M. Reiner, M. Marthaler, J. Braumüller, M. Weides, and G. Schön, Emulating the one-dimensional Fermi-Hubbard model by a double chain of qubits, *Phys. Rev. A* **94**, 032338 (2016).
- [105] S. Rodriguez, Linear antiferromagnetic chain, *Phys. Rev.* **116**, 1474 (1959).
- [106] 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).
- [107] P.-L. Dallaire-Demers and F. K. Wilhelm, Quantum gates and architecture for the quantum simulation of the Fermi-Hubbard model, *Phys. Rev. A* **94**, 062304 (2016).
- [108] J.-M. Reiner, S. Zanker, I. Schwenk, J. Leppäkangas, F. Wilhelm-Mauch, G. Schön, and M. Marthaler, Effects of gate errors in digital quantum simulations of fermionic systems, *Quantum Sci. Technol.* **3**, 045008 (2018).
- [109] P.-L. Dallaire-Demers, M. Stechly, J. F. Gonthier, N. T. Bashige, J. Romero, and Y. Cao, An application benchmark for fermionic quantum simulations, [arXiv:2003.01862](https://arxiv.org/abs/2003.01862) [quant-ph] (2020).
- [110] F. Arute *et al.*, Observation of separated dynamics of charge and spin in the Fermi-Hubbard model, [arXiv:2010.07965](https://arxiv.org/abs/2010.07965) [quant-ph] (2020).
- [111] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, *Phys. Rev. Lett.* **120**, 110501 (2018).
- [112] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions, *Phys. Rev. Appl.* **9**, 044036 (2018).
- [113] T. Shirakawa, K. Seki, and S. Yunoki, Discretized quantum adiabatic process for free fermions and comparison with the imaginary-time evolution, *Phys. Rev. Res.* **3**, 013004 (2021).
- [114] A. Tranter, P. J. Love, F. Mintert, and P. V. Coveney, A comparison of the Bravyi-Kitaev and Jordan-Wigner transformations for the quantum simulation of quantum chemistry, *J. Chem. Theory Comput.* **14**, 5617 (2018).
- [115] 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).
- [116] H. F. Trotter, On the product of semi-groups of operators, *Proc. Am. Math. Soc.* **10**, 545 (1959).
- [117] M. Suzuki, Relationship between d -Dimensional Quantum Spin Systems and $(d+1)$ -Dimensional Ising Systems: Equivalence, Critical Exponents and Systematic Approximants of the Partition Function and Spin Correlations, *Prog. Theor. Phys.* **56**, 1454 (1976).
- [118] M. Suzuki, Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems, *Comm. Math. Phys.* **51**, 183 (1976).
- [119] I. Omelyan, I. Mryglod, and R. Folk, Symplectic analytically integrable decomposition algorithms: Classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations, *Comput. Phys. Commun.* **151**, 272 (2003).
- [120] A. Papageorgiou and C. Zhang, On the efficiency of quantum algorithms for Hamiltonian simulation, *Quantum Inf. Process.* **11**, 541 (2012).
- [121] M. Heyl, P. Hauke, and P. Zoller, Quantum localization bounds Trotter errors in digital quantum simulation, *Sci. Adv.* **5**, eaau8342 (2019).
- [122] 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).
- [123] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, 2017).
- [124] This is because the circuit depth for $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ is given by $D_{2^{\lfloor n/2 \rfloor}}^{(3)} = 2(N_\Gamma - 1)3^{\lfloor n/2 \rfloor - 1} + 1$, while the largest circuit depth for $\hat{\mathcal{H}}_{\text{ST}}^n(\Delta_\tau)$ involving $[\hat{S}_2^{(p)}(\pm\Delta_\tau/2)]^n$ is $n(D_2^{(3)} - 1) + 1 = 2(N_\Gamma - 1)n + 1$, when $p = 3$.
- [125] R. D. Somma, Quantum eigenvalue estimation via time series analysis, *New J. Phys.* **21**, 123025 (2019).
- [126] D. Horn and M. Weinstein, The t expansion: A nonperturbative analytic tool for Hamiltonian systems, *Phys. Rev. D* **30**, 1256 (1984).
- [127] R. Kubo, Generalized cumulant expansion method, *J. Phys. Soc. Jpn.* **17**, 1100 (1962).
- [128] M. Fanizza, M. Rosati, M. Skotiniotis, J. Calsamiglia, and V. Giovannetti, Beyond the Swap Test: Optimal Estimation of Quantum State Overlap, *Phys. Rev. Lett.* **124**, 060503 (2020).
- [129] D.-B. Zhang, Z.-H. Yuan, and T. Yin, Variational quantum eigensolvers by variance minimization, [arXiv:2006.15781](https://arxiv.org/abs/2006.15781) [quant-ph] (2020).
- [130] K. Kowalski and B. Peng, Quantum simulations employing connected moments expansions, *J. Chem. Phys.* **153**, 201102 (2020).
- [131] H. J. Vallury, M. A. Jones, C. D. Hill, and L. C. L. Hollenberg, Quantum computed moments correction to variational estimates, *Quantum* **4**, 373 (2020).
- [132] N. S. Witte and L. C. L. Hollenberg, Plaquette expansion proof and interpretation, *Z. Phys. B* **95**, 531 (1994).
- [133] W. C. Haxton, K. M. Nolle, and K. M. Zurek, Piecewise moments method: Generalized Lanczos technique for nuclear response surfaces, *Phys. Rev. C* **72**, 065501 (2005).