Quantum power method by a superposition of time-evolved states

Kazuhiro Seki\textsuperscript{1} and Seiji Yunoki\textsuperscript{1,2,3}

\textsuperscript{1}Computational Quantum Matter Research Team, RIKEN Center for Emergent Matter Science (CEMS), Saitama 351-0198, Japan
\textsuperscript{2}Computational Materials Science Research Team, RIKEN Center for Computational Science (R-CPS), Hyogo 650-0047, Japan
\textsuperscript{3}Computational Condensed Matter Physics Laboratory, RIKEN Cluster for Pioneering Research (CPR), Saitama 351-0198, Japan

(Dated: August 11, 2020)

We propose a quantum-classical hybrid algorithm of the power method, here dubbed quantum power method, to evaluate $\hat{H}^n|\psi\rangle$ with quantum computers, where $n$ is a nonnegative integer, $\hat{H}$ is a time-independent Hamiltonian of interest, and $|\psi\rangle$ is a quantum state. The quantum power method is formulated on the fact that the Hamiltonian power $\hat{H}^n$ can generally be approximated by a linear combination of time-evolution operators $U(t) = e^{-i\hat{H}t}$ at $n + 1$ different time ($t$) variables. The formalism is based on a time-discretized form of the higher-order derivative $\hat{H}^n = i^n d^n/dt^n_{\lambda=0}$ incorporated with the central-finite-difference scheme and the symmetric Suzuki-Trotter decomposition, by which the approximated Hamiltonian power retains its Hermiticity under a controlled accuracy. In the quantum power method, the Suzuki-Trotter decomposition is employed to evaluate each time-evolution operator $U(t)$ separately at different time $t$ close to zero on quantum computers, while a linear combination of these time-evolution operators is treated on classical computers. The number of gates required for approximating $\hat{H}^n$ is $O(Nn)$, where $N$ is the number of qubits assuming a local Hamiltonian $\hat{H}$, and the systematic errors due to the finite-difference scheme for the time derivatives and the Suzuki-Trotter decomposition of the time-evolution operators can be improved systematically by the Richardson extrapolation with a polynomial increase of quantum resources. We numerically demonstrate that the quantum power method can control the systematic errors in approximating the Hamiltonian power $\hat{H}^n$ to be essentially exact for $n$ as large as 100. For an application of the quantum power method, we combine this method with a multireference Krylov-subspace diagonalization scheme and show, by noiseless numerical simulations for a spin-1/2 Heisenberg model with the Hamiltonian power $\hat{H}^n$ up to $n = 11$ (but not limited), that the estimated ground-state energy and the ground-state fidelity over a variational-quantum-eigensolver scheme is systematically improved with increasing the power $n$. Furthermore, we outline other applications of the quantum power method, including several moment based methods, and numerically demonstrate the connected-cluster expansion for the imaginary-time evolution to compare the results of the multireference Krylov-subspace diagonalization. We also show that, with the quantum power method, the first and the second moments, $\langle \phi|\hat{H}|\psi\rangle$ and $\langle \phi|\hat{H}^2|\psi\rangle$, respectively, can be evaluated with the same amount of quantum resource. Although we focus on the Hamiltonian power, the quantum power method can generally be applied to any sparse Hermitian operator $A$ to evaluate $A^n$.

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 [1–3]. With classical computers, a repeated multiplication of a Hamiltonian $\hat{H}$ 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 [4] including the Lanczos method [5–10], and polynomial-expansion methods [11]. Such methods allow for calculating not only ground states but also dynamics [12–16] 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 [17]. 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 $\hat{H}^n$.

Recently, simulating quantum many-body systems with quantum computers [18–21] attracts great interest due to experimental realizations of and advances on quantum de-
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 $\dot{\mathcal{H}}^n = i^n dt \dot{U}(t)/dt|_{t=0}$ of the time-evolution operator $\dot{U}(t) = e^{-i\mathcal{H}t}$, by which the Hamiltonian power $\mathcal{H}^n$ is represented as a linear combination of $\dot{U}(t)$ 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 that the Hamiltonian $\mathcal{H}$ is local, the number of the gates required for approximating $\mathcal{H}^n$ in the quantum power method is $O(Nn)$ 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 $\mathcal{H}^n$ 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 briefly outline other applications of the quantum power method.

The rest of the paper is organized as follows. In Sec. II, we introduce the spin-1/2 Heisenberg model on a one-dimensional periodic chain. In Sec. III, we formulate the quantum power method by describing the central-finite-difference scheme for the time derivatives, basic properties of the approximated Hamiltonian power, the Richardson extrapolation to systematically eliminate the systematic errors, and the Suzuki-Trotter decomposition for 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 first numerically demonstrate that the systematic errors in the quantum power method are well controlled to be essentially exact and then show numerical results of the Krylov-subspace diagonalization combined with the quantum power method for the spin-1/2 Heisenberg model. The paper is summarized with discussions in Sec. VI. Explicit forms of the higher-order symmetric Suzuki-Trotter decompositions and their error analysis are provided in Appendix A. An alternative formalism of approximating the Hamiltonian power is discussed in Appendix B. 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-cluster expansion (CMX) for the imaginary-time evolution is demonstrated by numerical simulations in Appendix C. The Lanczos method with an emphasis on its connection to the moments is also described in Appendix D. Throughout the paper, we set $\hbar = 1$.

II. MODEL

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

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

where $J > 0$ is the antiferromagnetic 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. $\hat{P}_{ij}$ is the swap operator which acts on the $i$th and $j$th qubits as $\hat{P}_{ij} |a\rangle |b\rangle = |b\rangle |a\rangle$. In the Hamiltonian in Eq. (1), the constant (identity) term $\hat{I}_i \hat{I}_j$ is added to the conventional Heisenberg Hamiltonian and thus $\mathcal{H}$ is simply a sum of swap operators. Indeed, the second equality in Eq. (1) follows from the identity

$$\hat{I}_i \hat{I}_j + \hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j + \hat{Z}_i \hat{Z}_j = 2 \hat{P}_{ij}$$

for $i \neq j$.

We consider the Hamiltonian $\mathcal{H}$ 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

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

where $i + 1$ in the subscript should be read as 1 if $i = N$ due to the periodic-boundary conditions. For the later use in the Suzuki-Trotter decomposition of the time-evolution operator, we divide the Hamiltonian into two parts as

$$\mathcal{H} = \mathcal{H}_A + \mathcal{H}_B,$$

with

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

and

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

III. FORMALISM

In this section, we formulate the quantum power method. Figure 1 illustrates an overview of the formalism based on the higher-order derivative of the time-evolution operator, which is decomposed approximately using the symmetric Suzuki-Trotter decomposition.

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

The time-evolution operator $\hat{U}(t)$ of the time-independent Hamiltonian $\hat{H}$ at time $t$ is given by

$$\hat{U}(t) = e^{-i\hat{H}t} = \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} \hat{H}^n,$$  (7)

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

$$\hat{H}^n = i^n \left. \frac{d^n \hat{U}(t)}{dt^n} \right|_{t=0}.$$  (8)

By introducing a small time interval $\Delta_r$, we replace the time derivative in Eq. (8) with the central finite-difference as

$$\hat{H}^n = \hat{H}^n(\Delta_r) + O(\Delta_r^2),$$  (9)

where

$$\hat{H}^n(\Delta_r) = \sum_{k=0}^{n} c_{n,k} \hat{U}(\frac{n}{2} - k) \Delta_r$$  (10)

and

$$c_{n,k} = \frac{i^n}{\Delta_r^n} (-1)^k \binom{n}{k}.$$  (11)

A derivation of the coefficients $c_{n,k}$ is shown in Fig. 1(b). Equations (9) and (10) imply that the $n$th power of the Hamiltonian, $\hat{H}^n$, can be approximated with a controlled accuracy as a linear combination of the time-evolution operators at $n+1$ different time variables.

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

$$[\hat{U}(t)]^\dagger = [\hat{U}(t)]^{-1} = \hat{U}(-t),$$  (12)

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

$$\hat{H}^n(\Delta_r) = [\hat{H}^n(\Delta_r)]^\dagger = \hat{H}^n(-\Delta_r).$$  (13)

In the last equality, we have used that $c_{n,k}$ in Eq. (11) is an even (odd) function of $\Delta_r$ when $n$ is even (odd). Since $\hat{H}^n(\Delta_r)$ is an even function of $\Delta_r$, the systematic error $E_{\text{FD}}$ in odd powers of $\Delta_r$ is absent in Eq. (9). Moreover, with the multiplication law of the time-evolution operator $\hat{U}(t)\hat{U}(t') = \hat{U}(t + t')$, Eq. (10) can be written as

$$\hat{H}^n(\Delta_r) = \sum_{k=0}^{n} c_{n,k} \left( \hat{U}(\frac{\Delta_r}{2}) \right)^{n-k} \left( \hat{U}(\frac{-\Delta_r}{2}) \right)^k$$

$$= \frac{i^n}{\Delta_r^n} \left[ \hat{U}(\frac{\Delta_r}{2}) - \hat{U}(\frac{-\Delta_r}{2}) \right]^n.$$  (14)

The last line in Eq. (14) indicates that the approximated Hamiltonian power $\hat{H}^n(\Delta_r)$ satisfies a law of exponents

$$\hat{H}^n(\Delta_r) = [\hat{H}^1(\Delta_r)]^n.$$  (15)

Namely, $\hat{H}^n(\Delta_r)$ is exactly the $n$th power of $\hat{H}^1(\Delta_r)$ for $n \geq 0$. In fact, Eq. (14) can be understood simply as

$$\hat{H}^n = \left[ \frac{d\hat{U}(t)}{dt} \right]_{t=0}^n = [\hat{H}^1(\Delta_r)]^n + O(\Delta_r^2).$$  (16)

B. Richardson extrapolation

As we shall discuss in Sec. V A, the systematic error $E_{\text{FD}}$ due to the finite-difference scheme for the time derivatives in Eq. (9) can be controlled by varying the time discretization step $\Delta_r$. However, it is often practically useful to reduce the systematic error $E_{\text{FD}}$ by not taking too small $\Delta_r$ in the algorithmic level. The Richardson extrapolation can achieve a better error estimate by systematically eliminating lower-order errors in Eq. (9).

In the Richardson extrapolation, $\hat{H}^n(\Delta_r)$ and $\hat{H}^n(\Delta_r/h)$ with some real $h$ (such that $0 < h \neq 1$) are used to eliminate the $O(\Delta_r^2)$ error in Eq. (9) as

$$\hat{H}^n = \hat{H}^n_{(1)}(\Delta_r) + O(\Delta_r^2),$$  (17)

where

$$\hat{H}^n_{(1)}(\Delta_r) = \frac{h^2 \hat{H}^n(\Delta_r/h) - \hat{H}^n(\Delta_r)}{h^2 - 1}$$  (18)

is the first-order Richardson extrapolation of $\hat{H}^n(\Delta_r)$. Since $\hat{H}^n(\Delta_r)$ is an even function of $\Delta_r$, $\hat{H}^n_{(1)}(\Delta_r)$ is also an even function of $\Delta_r$ and thus the systematic error $E_{\text{FD}}$ in odd powers of $\Delta_r$ is absent in Eq. (17).

We can use the Richardson extrapolation recursively to further eliminate the $O(\Delta_r^2)$ error in Eq. (17). Namely, the $r$th order Richardson extrapolation can be obtained recurrently as

$$\hat{H}^n = \hat{H}^n_{(r)}(\Delta_r) + O(\Delta_r^{2r})$$  (19)
\[ \hat{H}^n = \frac{i^n}{\Delta t^n} \left( \hat{U}(t) - \hat{U}(0) \right) \]

The formalism so far is based on the exact time-evolution operator, i.e., \( \hat{U}(t) \). For 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 \( c_{\Delta t} \) of a linear combination of the time-evolution operators that approximates the Hamiltonian power \( \hat{H}^n \). The systematic error due to the finite-difference scheme is \( \mathcal{E}_{FD} \). In our numerical simulations, it is found that the lowest-order symmetric Suzuki-Trotter decomposition with \( m = 1 \) is adequate to control these systematic errors consistently. The number of gates, indicated by small blue rectangles in (c), required to approximately represent the Hamiltonian power \( \hat{H}^n \) scales as \( O(Nn) \).

\[
\hat{H}_{(r)}(\Delta t) = \frac{1}{\Delta t} \left( \hat{U}(\Delta t/2) - \hat{U}(-\Delta t/2) \right)
\]

with \( \hat{H}_{(r)}(\Delta t) = \hat{H}^n(\Delta t) \), and therefore the systematic error \( \mathcal{E}_{FD} \) is reduced to \( O(\Delta t^{2r+2}) \) after the \( r \)th order Richardson extrapolation.

\( \hat{H}_{(r)}(\Delta t) \) is Hermitian because \( \hat{H}_{(r)}(\Delta t) \) is Hermitian. Since \( \hat{H}_{(r)}(\Delta t) \) is a linear combination of \( n+1 \) unitaries, \( \hat{H}_{(r)}(\Delta t) \) is a linear combination of \( (r+1)(n+1) \) unitaries. Note also that \( \hat{H}_{(r)}(\Delta t) \) is no longer the nth order power of \( \hat{H}_{(r)}(\Delta t) \), i.e., \( \hat{H}_{(r)}(\Delta t) \neq \left[ \hat{H}_{(r)}(\Delta t) \right]^n \), when \( r \geq 1 \), but obviously \( \hat{H}_{(r)}(\Delta t) = [\hat{H}_{(r)}(\Delta t)]^n + O(\Delta t^{2r+2}) \). In our numerical simulations, we choose \( h = 2 \) when the Richardson extrapolation is used.

Three additional remarks are in order regarding the properties of the approximated Hamiltonian power \( \hat{H}^n(\Delta t) \). First, if a forward or backward, instead of central, finite-difference scheme is employed in Eq. (10), the Hermiticity and the even dependence on \( \Delta t \) of \( \hat{H}^n(\Delta t) \) in Eq. (13) are both violated. Therefore, the central finite-difference scheme is a crucial choice. Second, when the time-evolution operator \( \hat{U}(\Delta t) \) is approximated by a Suzuki-Trotter decomposition, the corresponding Suzuki-Trotter error \( \mathcal{E}_{ST} \) appears in Eqs. (10) and (14). 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. (12)], is employed to decompose the time-evolution operators in Eqs. (10) and (14), the resulting \( \hat{H}^n(\Delta t) \) still satisfies the Hermiticity and the even dependence on \( \Delta t \). Therefore, it is important to adopt a symmetric Suzuki-Trotter decomposition (see Sec. III C 3 for details).

### C. Suzuki-Trotter decomposition

The formalism so far is based on the exact time-evolution operator \( \hat{U}(t) \) in Eq. (7). However, on quantum computers, the time-evolution operator with its exponent composed of...
the sum of non-commuting 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.

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 $\hat{H}(\Delta_r)$ should be approximated by the Suzuki-Trotter decomposition, either in Eq. (10) or in Eq. (14). Although Eqs. (10) and (14) 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 shall 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 four, the algorithm formulated on the basis of Eq. (14) with the lowest-order symmetric Suzuki-Trotter decomposition is favored in terms of the gate counts (see Appendix B).

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 [58–60] for the quantum power method.

1. Constituting elements and circuit structure

To be more concrete, we consider the case of the spin-1/2 Heisenberg model $\hat{H} = \hat{H}_A + \hat{H}_B$ given in Eqs. (1) and (4), where the time-evolution operator $\hat{U}(t) = e^{-i\hat{H}t}$ is constituted by two elementary time-evolution operators associated with $\hat{H}_A$ and $\hat{H}_B$. Let us first introduce the exponential-swap (e-swap) gate $\hat{U}_{i,j}$ [61–65]

$$
\hat{U}_{i,j}(\theta) = \exp(-i\theta \hat{P}_{i,j}/2),
$$

where $\theta$ is a real-valued parameter. The e-swap gate, which is equivalent to the swap$^\dagger$ gate up to a two-qubit global phase factor [41, 66–68], is depicted schematically in Fig. 1(c) as a blue rectangular extended over two qubits. The gate corresponding to Eq. (21) can be implemented with three CNOT gates and few single-qubit rotations [69–71]. The time-evolution operators of $\hat{H}_A$ and $\hat{H}_B$ are given respectively by

$$
\exp(-i\hat{H}_At) = \prod_{i=1}^{N/2} \hat{U}_{2i,2i+1}(tJ) \quad (22)
$$

and

$$
\exp(-i\hat{H}_Bt) = \prod_{i=1}^{N/2} \hat{U}_{2i-1,2i}(tJ). \quad (23)
$$

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. (22) and (23). As described in the following, Fig. 1(c) illustrates a typical circuit structure that approximates the time-evolution operator $\hat{U}(\Delta_r)$, consisting of a product of $\exp(-i\hat{H}_A\Delta_r)$’s and $\exp(-i\hat{H}_B\Delta_r)$’s with real parameters $\{s_i\}$.

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

We now describe a systematic construction of the symmetric Suzuki-Trotter decompositions. Let us define $x = -i\Delta_r$ to simplify the notation. The second-order symmetric decomposition $\hat{S}_2(\Delta_r)$ of the time-evolution operator $\hat{U}(\Delta_r)$ is given by

$$
\hat{U}(\Delta_r) = e^{x\hat{H}_A+\hat{H}_B} = \hat{S}_2(\Delta_r) + O(\Delta_r^3), \quad (24)
$$

where

$$
\hat{S}_2(\Delta_r) = e^{ix\hat{H}_A}e^{ix\hat{H}_B}e^{ix\hat{H}_A} \quad (25)
$$

and $[\hat{H}_A, \hat{H}_B] \neq 0$. The subscript “2” implies that $\hat{S}_2(\Delta_r)$ correctly represent $\hat{U}(\Delta_r)$ to $O(\Delta_r^2)$. It is readily found that $\hat{S}_2(\Delta_r)$ satisfies

$$
\hat{S}_2(\Delta_r)\hat{S}_2(-\Delta_r) = \hat{S}_2(-\Delta_r)\hat{S}_2(\Delta_r) = \hat{I},
$$

where $\hat{I}$ is the identity operator. It is noteworthy that if we write $\hat{S}_2(\Delta_r)$ in the form

$$
\hat{S}_2(\Delta_r) = \exp\left[x(\hat{H}_A + \hat{H}_B) + x^2\hat{R}_2 + x^3\hat{R}_3 + \cdots \right],
$$

then the residual terms $\hat{R}_k$ with $k$ even are zero [59]. This can be confirmed as follows. Equation (26) indicates that $\hat{S}_2(\Delta_r)$ commutes with $\hat{S}_2(-\Delta_r) = [\hat{S}_2(\Delta_r)]^{-1}$, implying that $\hat{I} = \hat{S}_2(\Delta_r)\hat{S}_2(-\Delta_r) = \exp\left[2(x^2\hat{R}_2 + x^4\hat{R}_4 + x^6\hat{R}_6 + \cdots)\right]$ for arbitrary $x(-i\Delta_r)$. We thus obtain that $\hat{R}_2 = \hat{R}_4 = \hat{R}_6 = \cdots = 0$. This property holds for the higher-order symmetric decompositions described below, as they satisfy the relation corresponding to Eq. (26) by construction [59].

Starting with $\hat{S}_2(p)(\Delta_r) \equiv \hat{S}_2(\Delta_r)$, the higher-order decomposition $\hat{S}_{2m}(\Delta_r)$ for $m \geq 2$, satisfying

$$
\hat{U}(\Delta_r) = e^{x\hat{H}_A+\hat{H}_B} = \hat{S}_{2m}(\Delta_r) + O(\Delta_r^{2m+1}), \quad (27)
$$

can be constructed recursively as

$$
\hat{S}_{2m}(\Delta_r) = \left[\hat{S}_{2m-2}(k_m^{(p)}(\Delta_r))\right]^{(p-1)/2} \times \hat{S}_{2m-2}(k_m^{(p)}(\Delta_r))^\dagger \left[\hat{S}_{2m-2}(k_m^{(p)}(\Delta_r))\right]^{(p-1)/2}, \quad (28)
$$

where $k_m^{(p)} = 1 - (p-1)k_m^{(p)}$, $k_m^{(p)} = [(p-1)-(p-1)^{1/(2m-1)}]^{-1}$, and $p$ is an odd integer with $p \geq 3$ [72]. The superscript "(")" implies that $\hat{S}_{2m}$ consists of a product of $p \hat{S}_{2m-2}$’s. The parameter $k_m^{(p)}$ is determined so as to eliminate the residual term $x^{2m-1}\hat{R}_{2m-1}$ in $\ln \hat{S}_{2m}(\Delta_r)$ and thus

$$
\hat{S}_{2m}(\Delta_r) = \exp \left[ x(\hat{H}_A + \hat{H}_B) + x^{2m-1}\hat{R}_{2m-1} + \cdots \right]. \quad (29)
$$
Namely, \( k_{m}^{(p)} \) is the solution of \((p-1)k_{m}^{(p)} e^{-k_{m}^{(p)} \Delta} + k_{m}^{(p)} e^{k_{m}^{(p)} \Delta} = 0 \) under the condition \((p-1)\bar{k}_{m}^{(p)} + \bar{k}_{m}^{(p)} = 1 \). Since \( \hat{S}_{2m}^{(p)}(\Delta) \) satisfies
\[
\hat{S}_{2m}^{(p)}(\Delta)\hat{S}_{2m}^{(p)}(-\Delta) = \hat{S}_{2m}^{(p)}(-\Delta)\hat{S}_{2m}^{(p)}(\Delta) = \hat{I},
\]
the residual terms of even power such as \( x^{2m}\hat{R}_{2m} \) are absent in the exponent of \( \hat{S}_{2m}^{(p)}(\Delta) \) in Eq. (29), shown by the same argument for \( m = 1 \). Some of the higher-order symmetric Suzuki-Trotter decompositions are explicitly provided in Appendix A. As shown in Appendix A.2, the parameter \( p \) affects the accuracy of the decomposition for a given \( m \).

### 3. Unitarity and time-reversed evolution of \( \hat{S}_{2m}^{(p)}(\Delta) \)

As implied in Eq. (30), \( \hat{S}_{2m}^{(p)}(\Delta) \) retains not only the unitarity but also the equivalence between the inverse and time-reversed evolution, i.e.,

\[
\left[ \hat{S}_{2m}^{(p)}(\Delta) \right]^{-1} = \left[ \hat{S}_{2m}^{(p)}(-\Delta) \right]^{-1} = \hat{S}_{2m}^{(p)}(-\Delta).
\]

Therefore, the Hermiticity and the even dependence on \( \Delta \) of \( \hat{H}^{m}(\Delta) \) in Eq. (13) are both retained even when the exact time-evolution operators in Eqs. (10) and (14) are approximated by simply replacing them with \( \hat{S}_{2m}^{(p)} \)'s.

In contrast, an asymmetric Suzuki-Trotter decomposition \( \hat{F}(\Delta) \), such as \( \hat{F}(\Delta) = e^{i\hat{R}^{\prime}} e^{i\hat{R}} \), results in
\[
\left[ \hat{F}(\Delta) \right]^{-1} \neq \hat{F}(-\Delta).
\]

Thus, \( \hat{F}(\Delta) \) 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 \( \Delta \) of \( \hat{H}^{m}(\Delta) \) in Eq. (13) is violated if the exact time-evolution operators in Eqs. (10) and (14) are approximated by \( \hat{F}' \)'s. For example, if we consider an operator \( \hat{H}_{T}^{m}(\Delta) = \{i[\hat{F}(\Delta) - \hat{F}(-\Delta)]/\Delta \tau \) to approximate \( i[\hat{U}(\Delta) - \hat{U}(-\Delta)]/\Delta \tau \), it satisfies the Hermiticity but is no longer an even function of \( \Delta \); as \( \hat{H}_{T}^{m}(\Delta) = [\hat{H}_{T}^{m}(\Delta)]^{*} \neq \hat{H}_{T}^{m}(-\Delta) \). On the other hand, an operator \( \hat{H}_{E}^{m}(\Delta) = i[\hat{F}(\Delta) - \hat{F}(-\Delta)]/\Delta \) is an even function of \( \Delta \) but no longer satisfies the Hermiticity as \( \hat{H}_{E}^{m}(\Delta) = \hat{H}_{E}^{m}(-\Delta) \neq [\hat{H}_{E}^{m}(\Delta)]^{*} \). Therefore, the symmetric Suzuki-Trotter decomposition \( \hat{S}_{2m}^{(p)}(\Delta) \) is essential for the resulting Suzuki-Trotter approximation \( \hat{H}^{m}(\Delta) \) to retain both the Hermiticity and the even dependence on \( \Delta \). Note that asymmetric Suzuki-Trotter decompositions and their connection to symmetric ones have been studied in Ref. [73].

### 4. 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 \( \hat{U}(\Delta) \) approximated by the symmetric Suzuki-Trotter decomposition \( \hat{S}_{2m}^{(p)}(\Delta) \), as in Eq. (27) [also see Fig. 1(c)]. We define \( D_{2m}^{(p)} \) as the number of noncommuting exponentials appearing in \( \hat{S}_{2m}^{(p)}(\Delta) \). For example, the depth of \( \hat{S}_{2m}^{(p)}(\Delta) \) is \( D_{2m}^{(p)} = 3 \). Since \( \hat{S}_{2m}^{(p)}(\Delta) \) consists of a product of \( p \) \( \hat{S}_{2m}^{(p)} \)'s, the depth of \( \hat{S}_{2m}^{(p)}(\Delta) \) without contracting commuting exponentials is \( pD_{2m}^{(p)} \). However, since \( \hat{S}_{2m}^{(p)}(\Delta) \) involves \( p-1 \) products of two consecutive \( \hat{S}_{2m}^{(p)} \)'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
\[
D_{2m}^{(p)} - 1 = p^{[D_{2m-2}^{(p)} - 1]}
\]
\[
= \cdots
\]
\[
= p^{m-1}[D_{2}^{(p)} - 1].
\]

Substituting \( D_{2}^{(p)} = 3 \) in Eq. (33) yields that
\[
D_{2m}^{(p)} = 2p^{m-1} + 1.
\]

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.

### D. Quantum power method

While the time-evolution operators satisfy the multiplica-

tion law \( \hat{U}(\Delta_{2})\hat{U}(\Delta_{1}) = \hat{U}(\Delta_{2} + \Delta_{1}) \), this is no longer correct when the time-evolution operators are approximated by the Suzuki-Trotter decomposition, i.e., \( \hat{S}_{2m}^{(p)}(\Delta_{2})\hat{S}_{2m}^{(p)}(\Delta_{1}) \neq \hat{S}_{2m}^{(p)}(\Delta_{2} + \Delta_{1}) \). Therefore, it is crucial to carefully consider when the time-evolution operators in the approximated Hamiltonian power \( \hat{H}^{m}(\Delta) \) should be replaced with the symmetric Suzuki-Trotter decomposition, either in Eq. (10) or in Eq. (14), implying that there exist two different routes to formulate the quantum power method. As we shall show here and in Appendix B, these two approaches provide two different algorithms of the quantum power method that differ in the scaling of complexity but control the systematic errors \( \epsilon_{\text{SYD}} \) and \( \epsilon_{\text{ST}} \) with essentially the same accuracy. In this section, we formulate the quantum power method based on Eq. (14) that scales much better when the power \( n \) is large. In Appendix B, we describe an alternative algorithm formulated on the basis of Eq. (10), which is favored when the power \( n \) is small (\( n \ll 4 \)).

By incorporating the symmetric Suzuki-Trotter decomposition into the approximated Hamiltonian power \( \hat{H}^{m}(\Delta) \) in Eq. (14), the Hamiltonian power \( \hat{H}^{n} \) is finally approximated as
\[
\hat{H}^{n} = \hat{H}_{\text{ST}}^{n}(\Delta) + O(\Delta_{2}^{2}) + O(\Delta_{3}^{2}m),
\]
Here, we have used the fact that as in Eq. (31), but note that with combination of \( \hat{r} \) is the central quantity in the quantum power method that approximates the Hamiltonian power \( \hat{H}^n \) as

\[
\hat{H}^n_{ST}(\Delta_r) = \sum_{k=0}^{n} c_{n,k} \left[ \hat{S}^{(p)}_{2m}(\Delta_r/2) \right]^{n-2k} \tag{36}
\]

\[
= \frac{1}{\Delta_r^n} \left[ \hat{S}^{(p)}_{2m}(\Delta_r/2) - \hat{S}^{(p)}_{2m}(\Delta_r/2)^n \right]. \tag{37}
\]

Here, we have used the fact that \( [ \hat{S}^{(p)}_{2m}(\Delta_r/2) ]^{-1} = \hat{S}^{(p)}_{2m}(\Delta_r/2) \), as in Eq. (31), but note that \( \hat{S}^{(p)}_{2m}(\Delta_r/2) \neq \hat{S}^{(p)}_{2m}(\Delta_r) \). It should also be noticed that the order \( O(\Delta_r^{2m+1}) \) of the Suzuki-Trotter error \( E_{ST} \) in Eq. (35) is decreased by one from the naively expected order \( O(\Delta_r^{2m+1}) \) as in Eq. (27), because of the factor \( 1/\Delta_r^n \) in \( c_{n,k} \).

Equation (35) already reveals a remarkable advantage in the quantum power method formulated on the basis of Eq. (14): in order to control the systematic errors \( E_{FD} \) and \( E_{ST} \) with the same order of accuracy, it is enough to adopt the lowest-order Suzuki-Trotter decomposition with \( m = 1 \), independently of the power \( n \). This is in sharp contrast to the other formalism based on Eq. (10), in which the order \( m \) of the Suzuki-Trotter decomposition has to be increased with the power \( n \) and, as described in more details in Appendix B, essentially the complexity increases exponentially with the power \( n \).

Equation (37) indicates that \( \hat{H}^n_{ST}(\Delta_r) \) satisfies the law of exponents

\[
\hat{H}^n_{ST}(\Delta_r) = \left[ \hat{H}^1_{ST}(\Delta_r) \right]^n. \tag{38}
\]

Moreover, \( \hat{H}^n_{ST}(\Delta_r) \) is Hermitian and an even function of \( \Delta_r \), i.e.,

\[
\hat{H}^n_{ST}(\Delta_r) = \left[ \hat{H}^n_{ST}(\Delta_r) \right]^\dagger = \hat{H}^n_{ST}(-\Delta_r), \tag{39}
\]

indicating that the systematic error \( E_{ST} \) in odd powers of \( \Delta_r \) is absent in Eq. (35). Therefore, recalling that the systematic error \( E_{FD} \) in odd powers of \( \Delta_r \) is also absent (see Sec. III B), the Richardson extrapolation can eliminate the finite-difference error \( E_{FD} \) and the Suzuki-Trotter error \( E_{ST} \) simultaneously as

\[
\hat{H}^n = \hat{H}^n_{ST(r)}(\Delta_r) + O(\Delta_r^{2+2r}) + O(\Delta_r^{2m+2r}), \tag{40}
\]

where \( \hat{H}^n_{ST(r)}(\Delta_r) \) is the \( r \)th order Richardson extrapolation of the approximated Hamiltonian power, i.e.,

\[
\hat{H}^n_{ST(r)}(\Delta_r) = \frac{\hat{H}^n_{ST(r-1)}(\Delta_r/h) - \hat{H}^n_{ST(r-1)}(\Delta_r)}{h^{2r} - 1}, \tag{41}
\]

with \( \hat{H}^{n}_{ST0}(\Delta_r) \equiv \hat{H}^n_{ST}(\Delta_r) \). Since \( \hat{H}^n_{ST0}(\Delta_r) \) is a linear combination of \( n + 1 \) unitaries \( \hat{S}^{(p)}_{2m} \), \( \hat{H}^n_{STr}(\Delta_r) \) is a linear combination of \( (r+1)(n+1) \) unitaries \( \hat{S}^{(p)}_{2m} \). Equation (40) reveals another significant feature of the quantum power method that only polynomial resources with the lowest-order symmetric Suzuki-Trotter decomposition suffice to systematically and consistently eliminate the lower order systematic errors in \( E_{FD} \) and \( E_{ST} \). In Sec. V A, we will show by numerical simulations that these systematic errors in the approximated Hamiltonian power are well controlled with the time-discretization step \( \Delta_r \) 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 \( \hat{H} \) is still respected in the approximated Hamiltonian power \( \hat{H}^n_{ST(r)}(\Delta_r) \). This is indeed the case in the quantum power method formulated here because

\[
[\hat{H}, \hat{H}^n_{ST(r)}(\Delta_r)] = O(\Delta_r^{2m+2r}). \tag{42}
\]

Therefore, the symmetry of the Hamiltonian \( \hat{H} \) is preserved in the quantum power method within the systematic error \( E_{ST} \) due to the Suzuki-Trotter decomposition that can be well controlled. Notice that there is no contribution from the systematic error \( E_{FD} \) due to the finite-difference scheme of the time derivatives in the right hand side of Eq. (42) because \( [\hat{H}, \hat{H}^n_{ST(r)}(\Delta_r)] = 0 \).

Figure 1 summarizes the quantum power method formulated here. In the quantum power method, the Hamiltonian power \( \hat{H}^n \) is approximated to \( \hat{H}^n_{ST}(\Delta_r) \) represented as a linear combination of the \( n+1 \) Suzuki-Trotter decomposed time-evolution operators \( [\hat{S}^{(p)}_{2m}(\Delta_r/2)]^{n-2k} \). The systematic error \( E_{FD} \) due to the finite-difference scheme for the time derivatives is \( O(\Delta_r^7) \), and the systematic error \( E_{ST} \) due to the Suzuki-Trotter decomposition of the time-evolution operators is \( O(\Delta_r^{2m}) \). These systematic errors \( E_{FD} \) and \( E_{ST} \) can be both improved systematically by the \( r \)-th order Richardson extrapolation to \( O(\Delta_r^{2+2r}) \) and \( O(\Delta_r^{2m+2r}) \), respectively. While the linear combination of the Suzuki-Trotter decomposed time-evolution operators is treated classically, each Suzuki-Trotter decomposed time-evolution operator \( [\hat{S}^{(p)}_{2m}(\Delta_r/2)]^{n-2k} \) is evaluated on quantum computers. As illustrated in Fig. 1 (c), a quantum circuit for a single \( \hat{S}^{(p)}_{2m}(\Delta_r/2) \) has the circuit depth \( D^{(p)}_{2m} = 2p^{m+1} + 1 \), and thus the circuit depth required for \( \hat{H}^n_{ST}(\Delta_r) \) is at most \( O(n) \) with a prefactor \( D^{(p)}_{2m} \). Hence, assuming that \( \hat{H} \) consists of \( O(N) \) local terms, the number of gates required for \( \hat{H}^n_{ST}(\Delta_r) \) is \( O(Nn) \). When the \( r \)-th order Richardson extrapolation is employed, the number of gates required remains the same, but the number of terms in the linear combination of the Suzuki-Trotter decomposed time-evolution operators is \( O(nm) \). Therefore, for example, to evaluate the expectation value of \( \hat{H}^n_{ST(r)}(\Delta_r) \) with respect to a given state \( \psi \), the \( O(\Delta_r^7) \) state overlaps such as \( \langle \psi | \hat{S}^{(p)}_{2m}(\Delta_r/2)^{n-2k} | \psi \rangle \) have to be estimated. However, these quantities can be evaluated on quantum computers separately in parallel. Considering the gate count that scales as \( O(Nn) \) for approximating the Hamiltonian power \( \hat{H}^n \), the quantum power method is a potentially promising application for near-term quantum devices.

In contrast to the quantum power method, the direct evaluation of the expectation value \( \langle \psi | \hat{H}^n | \psi \rangle \) requires the average of \( O(N^n) \) operators, possibly containing long strings of Pauli
operators, provided that $\hat{H}$ consists of $O(N)$ local terms. Although the depth of the circuits for these averages is $O(1)$, the number $O(N^n)$ of averages required makes the direct evaluation of $\langle \psi | \hat{H}^n | \psi \rangle$ unfeasible as soon as the power $n$ is large.

In classical computation, the computational complexity scales as $O(N^{Dn})$ for the evaluation of $\langle \psi | \hat{H}^n | \psi \rangle$, when the Hamiltonian $\hat{H}$ is local and thus the Hamiltonian matrix is sparse. Here, $N_D$ is the dimension of the Hilbert space, i.e., $N_D = 2^N$ for the spin-1/2 Heisenberg model given in Eq. (1). This implies that the computational complexity of the classical computation scales exponentially in $N$. Therefore, the quantum power method introduced here, though it is approximate, would have a quantum advantage over the classical counterpart of the power method.

![Diagram](Fig. 2) A circuit with $N$ register qubits and $n$ ancilla qubits for probabilistically generating the state $\propto [\hat{S}^{(p)}_{2m}(\Delta_r/2) - \hat{S}^{(p)}_{2m}(\Delta_r/2)]^n|\psi\rangle$ in the register qubits for (a) $n = 1$ and (b) $n = 2$. $H$, $\hat{S}^{(p)}_{2m}$, and $\hat{S}^{(p)\dagger}_{2m}$ in the circuit denote the Hadamard gate, $\hat{S}^{(p)}_{2m}(\Delta_r/2)$, and $\hat{S}^{(p)\dagger}_{2m}(\Delta_r/2)$, 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 $P_{b_1b_2\cdots b_n}$ for finding the bit string $b_1b_2\cdots b_n = 11\cdots1$ in the ancilla qubits is given in Eq. (43).

Finally, we should note that the form of the approximated

$$K_\alpha (\hat{H}, \{|q_{k}\rangle\}_{k=1}^{M_B}) = \text{span} (|q_{1}\rangle, \ldots, |q_{M_B}\rangle, \hat{H}|q_{1}\rangle, \ldots, \hat{H}|q_{M_B}\rangle, \ldots, \hat{H}^{n-1}|q_{1}\rangle, \ldots, \hat{H}^{n-1}|q_{M_B}\rangle),$$

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 $M_B = 1$, $K_\alpha (\hat{H}, \{|q_{k}\rangle\}_{k=1}^{M_B})$ reduces to the conventional Krylov subspace. By defining

$$|u_i\rangle = \hat{H}^{l-1}|q_{k}\rangle,$$

with $i = k + (l - 1)M_B$ and $l = 1, 2, \ldots, n$, the block Krylov subspace can be written simply as $K_\alpha (\hat{H}, \{|q_{k}\rangle\}_{k=1}^{M_B}) = \text{span} (|u_{1}\rangle, \ldots, |u_{M_B}\rangle)$.

Hamiltonian power $\hat{H}_\alpha^n(\Delta_r)$ in Eq. (37) suggests a direct treatment of the linear combination of the Suzuki-Trotter decomposed time-evolution operators with a single quantum circuit [74-76] forming a simple recursive structure. Figure 2 shows such a circuit structure for probabilistically generating the state $\propto [\hat{S}^{(p)}_{2m}(\Delta_r/2) - \hat{S}^{(p)}_{2m}(\Delta_r/2)]^n|\psi\rangle$, among $2^n$ superposed states, in the $N$ register qubits using $n$ ancilla qubits. However, finding the desired state in the register qubits becomes exponentially small in general if $n$ is large. Let us define $P_{b_1b_2\cdots b_n}$ as the probability for finding a bit string $b_1b_2\cdots b_n$ 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 $11\cdots1$, which is relevant for $\hat{H}_\alpha^n(\Delta_r)$ [77], is given by

$$P_{11\cdots1} = \frac{1}{4^n} (-1)^n \langle \psi | [\hat{S}^{(p)}_{2m}(\Delta_r/2) - \hat{S}^{(p)}_{2m}(\Delta_r/2)]^n |\psi\rangle. $$

If $|\psi\rangle$ were an eigenstate of $\hat{S}^{(p)}_{2m}(\Delta_r/2)$ with an eigenvalue $e^{\pm i \lambda \Delta_r}$, it oscillates as $P_{11\cdots1} = |\sin (\lambda \Delta_r)|^n$, but otherwise is exponentially small. Therefore, the linear combination of the Suzuki-Trotter decomposed time-evolution operators is better treated with classical computers in the form of Eq. (36).

**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 [78]. 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 Appendix C and Appendix D.

**A. Block Krylov subspace**

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

$$K_\alpha (\hat{H}, \{|q_{k}\rangle\}_{k=1}^{M_B}) = \text{span} (|q_{1}\rangle, \ldots, |q_{M_B}\rangle, \hat{H}|q_{1}\rangle, \ldots, \hat{H}|q_{M_B}\rangle, \ldots, \hat{H}^{n-1}|q_{1}\rangle, \ldots, \hat{H}^{n-1}|q_{M_B}\rangle),$$

where $M_B \geq 1$ 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 $M_B = 1$, $K_\alpha (\hat{H}, \{|q_{k}\rangle\}_{k=1}^{M_B})$ reduces to the conventional Krylov subspace. By defining

$$|u_i\rangle = \hat{H}^{l-1}|q_{k}\rangle,$$

with $i = k + (l - 1)M_B$ and $l = 1, 2, \ldots, n$, the block Krylov subspace can be written simply as $K_\alpha (\hat{H}, \{|q_{k}\rangle\}_{k=1}^{M_B}) = \text{span} \left(|u_{1}\rangle, \ldots, |u_{M_B}\rangle\right)$.

**B. Rayleigh-Ritz technique**

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

$$\hat{H}|\Psi_0\rangle = E_0|\Psi_0\rangle$$

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

\[
|\Psi_0\rangle \approx |\Psi_{KS}\rangle \equiv \sum_{i=1}^{nM_B} v_i |u_i\rangle,
\]  
(47)

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 \( \langle \Psi_{KS}|\hat{H}|\Psi_{KS}\rangle \) under the constraint \( \langle \Psi_{KS}|\Psi_{KS}\rangle = 1 \). To this end, let us define the following function:

\[
\mathcal{F}(v, v^\ast) = \langle \Psi_{KS}|\hat{H}|\Psi_{KS}\rangle - \epsilon (\langle \Psi_{KS}|\Psi_{KS}\rangle - 1) = v^\dagger \hat{H} v - \epsilon (v^\dagger S v - 1) = \sum_{ij} v_i^* \left( \hat{H}_{ij} - \epsilon S_{ij} \right) v_j + \epsilon,
\]  
(48)

where \( \epsilon \) is a Lagrange multiplier, \( |v_i\rangle = v_i \),

\[ [\hat{H}]_{ij} = H_{ij} = \langle u_i|\hat{H}|u_j\rangle \]  
(49)

is the subspace Hamiltonian matrix, and

\[ [S]_{ij} = S_{ij} = \langle u_i|u_j\rangle \]  
(50)

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

\[
\hat{H} v = \epsilon S v.
\]  
(51)

Since both \( \hat{H} \) and \( S \) 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 \( \epsilon \) and the corresponding eigenvector \( v \) in Eq. (51) provide an approximation to the ground-state energy \( E_0 \) and the expansion coefficients \( \{v_i\}_{i=1}^{nM_B} \) in Eq. (47), respectively. Note that when \( M_B = 1 \), the matrices \( \hat{H} \) and \( S \) correspond to the Hankel matrices \( \mathcal{M}_n \) and \( \mathcal{L}_n \), respectively, defined in Eqs. (D6) and (D5).

Since \( S \) is a Hermitian matrix, it can be diagonalized by a unitary matrix \( V \) as

\[
V^\dagger S V = s,
\]  
(52)

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

\[
W = V s^{-1/2},
\]  
(53)

Eq. (51) can be transformed to a standard Hermitian eigenvalue problem of the form

\[
T q = \epsilon q,
\]  
(54)

where

\[
T = W^\dagger \hat{H} W
\]  
(55)

and \( q = W^{-1} v \). Thus, by solving the eigenvalue problem of Eq. (54), one can obtain \( \epsilon \) and \( v = W q \). The eigenvector \( v \) with the lowest eigenvalue \( \epsilon \) provides the coefficients in the approximate ground state \( |\Psi_{KS}\rangle \) [see Eq. (47)] with its energy \( E_{KS} \) of the Hamiltonian \( \hat{H} \) in the Krylov subspace \( \mathcal{K}_n(\hat{H}, \{|q_k\})_{k=1}^{M_B} \).

We note that if we use the Cholesky decomposition \( S = R^\dagger R \) with \( R \) being an upper-triangular matrix, instead of the eigen decomposition in Eq. (52), \( T \) reduces to the tridiagonal matrix in the Lanczos method when \( M_B = 1 \) [78].

### 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 \( \hat{H} \) in Eq. (49) and \( S \) in Eq. (50), because the states \( |u_i\rangle \) are defined on the Hilbert space of \( N_D = 2^N \) dimensions. On the other hand, the eigenvalue problem in the \( nM_B \)-dimensional block Krylov subspace given in Eq. (51) or Eq. (54) 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 [52–57].

We can now approximate the Hamiltonian power \( \hat{H}^{l-1} \) appearing in the Krylov-subspace basis \( |u_i\rangle \) given in Eq. (45) as

\[
|u_i\rangle = \tilde{u}_i + O(\Delta_t^{2+2r}) + O(\Delta_t^{2m+2r}),
\]  
(56)

where

\[
\tilde{u}_i = \hat{H}^{l-1}_{ST(r)}(\Delta_t)|q_k\rangle
\]  
(57)

with \( i = k + (l - 1)M_B \) and \( l = 1, 2, \ldots, n \). Note that the systematic errors in Eq. (56) are absent when \( l = 1 \). As described in Sec. III D, to approximate the Hamiltonian power \( \hat{H}^{l-1} \) by \( \hat{H}^{l-1}_{ST(r)}(\Delta_t) \) as in Eq. (57), the Suzuki-Trotter decomposed time-evolution operators \( \hat{S}^{(l)}_{ST(r)}(\pm \Delta_t/2) \) 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 \( \mathcal{K}_n(\hat{H}^{l}_{ST(r)}(\Delta_t), \{|q_k\})_{k=1}^{M_B} \) is at most \( O(n) \). The circuit depth does not depend on the order \( r \) of the Richardson extrapolation.

However, for the purpose of solving the generalized eigenvalue problem in Eq. (51) or the corresponding standard eigenvalue problem in Eq. (54), we can evaluate the matrix elements in Eqs. (49) and (50) more directly as

\[
H_{ij} = \tilde{H}_{ij} + O(\Delta_t^{2+2r}) + O(\Delta_t^{2m+2r})
\]  
(58)

and

\[
S_{ij} = \tilde{S}_{ij} + O(\Delta_t^{2+2r}) + O(\Delta_t^{2m+2r}),
\]  
(59)

where

\[
\tilde{H}_{ij} = \langle q_k|\hat{H}^{l-1}_{ST(r)}(\Delta_t)|q_k\rangle
\]  
(60)

and

\[
\tilde{S}_{ij} = \langle q_k|\hat{S}^{(l-1)}_{ST(r)}(\Delta_t)|q_k\rangle
\]  
(61)
with $i = k + (l - 1)M_n$ and $j = k' + (l' - 1)M_n$ for $1 \leq k, k' \leq M_n$ and $1 \leq l, l' \leq n$ in the block Krylov subspace $\mathcal{K}_n(\mathcal{H}_{\text{STr}}(\Delta_r), \{|q_k\}_k)_{k=1}^{M_n}$.

To be more specific, the matrix elements of $\hat{H}$ and $\hat{S}$ for $r = 0$, i.e., without the Richardson extrapolation, are given as

$$\hat{H}_{ij} = \sum_{x=0}^{l-1} c_{l-1,x} \langle q_l | \hat{S}_{2m}^{(p)}(\Delta_r/2) | q_{l'} \rangle$$

and

$$\hat{S}_{ij} = \sum_{x=0}^{l'-2} c_{l'-2,x} \langle q_l | \hat{S}_{2m}^{(p)}(\Delta_r/2) | q_{l'} \rangle.$$  \hfill (62)

The number of state overlaps required for constructing all matrix elements of both $\hat{H}$ and $\hat{S}$ is thus $O(nM_n^2)$. If the $r$-th order Richardson extrapolation is employed, the number of state overlaps to be evaluated is increased by a factor of $(r+1)$. The state overlaps in Eqs. (62) and (63) can be evaluated with a Hadamard-test like circuit, for example [47, 79–81].

**V. RESULTS**

In this section, we use numerical simulations to first show how the quantum power method can control the systematic errors in approximating the Hamiltonian power $\hat{H}^n$ and then demonstrate the application of the quantum power method combined with the multireference Krylov-subspace diagonalization for the spin-1/2 Heisenberg model.

**A. Degree of approximation**

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

$$d(\hat{A}, \hat{B}) = \sqrt{1 - \frac{|\langle \hat{A}, \hat{B} \rangle_F|^2}{||\hat{A}||_F ||\hat{B}||_F}}.$$  \hfill (64)

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

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

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

$$||\hat{A}||_F = \sqrt{\text{Tr}[\hat{A}^\dagger \hat{A}]}.$$  \hfill (65)

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 $\hat{A} = \hat{H}^r$ and $\hat{B} = \hat{H}_{\text{STr}}^{(r)}(\Delta_r)$ given in Eq. (36) for $r = 0$ and Eq. (41) for $r \geq 1$. The Hamiltonian $\hat{H}$ is for the spin-1/2 Heisenberg model on an $N$-qubit ring given in Eq. (1).

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 [82–86]

$$\text{Tr}[\hat{X}] = \lim_{R \to \infty} \frac{1}{R} \sum_{i=1}^{R} \langle \phi_i | \hat{X} | \phi_i \rangle,$$  \hfill (67)

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

$$|\phi_i \rangle = \sum_x \frac{e^{i\delta_x(i)}}{\sqrt{\theta_x}} |x \rangle.$$  \hfill (68)

is a random-phase state with $|\langle x | \rangle$ being a complete orthonormal basis set such that $\langle \langle x | y \rangle \rangle = \delta_{xy}$ and $\phi_x(x)$ being a random variable drawn uniformly from $[0, 2\pi)$. Note that $\langle \phi_i | \phi_i \rangle = 2^N$, i.e., the dimension $N_D$ of the Hilbert space. We choose $|\langle x | \rangle$ as the orthonormal basis set that diagonalizes the local Pauli $Z$ operators. The stochastic evaluation of the trace in Eq. (67) requires only sparse matrix-vector multiplications and a single inner-product calculation for each $\xi$, if $\hat{X}$ is represented as a product of sparse matrices, which is indeed the case here. Instead of taking the limit $R \to \infty$, we fix $R = 16$ for $N \geq 12$ and $R = 256$ for $N = 10$ and estimate error bars. Since $\langle \phi_i | \hat{A}^\dagger \hat{A} | \phi_i \rangle$, $\langle \phi_i | \hat{B}^\dagger \hat{B} | \phi_i \rangle$, and $\langle \phi_i | \hat{A} \hat{B} | \phi_i \rangle$ for $\hat{A} = \hat{H}^n$ and $\hat{B} = \hat{H}_{\text{STr}}^{(r)}(\Delta_r)$ are highly correlated to each other, error bars of $d(\hat{A}, \hat{B})$ must be estimated using the corresponding $3 \times 3$ covariance matrix.

Figure 3 shows the distance as a function of $\Delta_r$ 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 $\hat{S}_3^{(3)}$. Figures 3(a)–3(c) show the results without the Richardson extrapolation ($r = 0$). Since the leading systematic error in $\hat{H}_{\text{STr}}^{(0)}(\Delta_r)$ is $O(\Delta_r^2)$, the distance scales almost linearly in $\Delta_r^2$ 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 $\hat{H}_{\text{STr}}^{(r=1)}(\Delta_r)$ is $O(\Delta_r^4)$, and the distance indeed scales almost linearly in $\Delta_r^4$. As expected from Eq. (40), essentially no difference can be found between the results with $\hat{S}_2$ and $\hat{S}_3^{(3)}$, indicated respectively by empty and filled symbols in Fig. 3. These results clearly demonstrate that the systematic errors in approximating the Hamiltonian power $\hat{H}^n$ are well controlled.

Figure 4(a) shows the $n$ dependence of the distance for $N = 24$ with various values of $\Delta_r$ 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 dependence of the distance on $\Delta_r^2$ remains in a wide range of $\Delta_r (\Delta_r J \leq 0.1)$.
and the distance is smoothly extrapolated to zero in the limit of $\Delta \tau \to 0$, 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 $\hat{H}^n$ are expected to be $O(\Delta \tau^2)$. Indeed, our numerical simulations find the linear dependence of distance on $\Delta \tau^2$ for at least $\Delta \tau J \leq 0.05$ when $n = 100$ [see the inset in Fig. 4(d)]. Notice also that the distance itself becomes smaller by the factor of approximately 5 even for large $n$ when the first-order Richardson extrapolation is employed.

### B. Ground-state energy and fidelity

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 spin-1/2 Heisenberg model described by the Hamiltonian $\hat{H}$ in Eq. (1) 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 Sec. III D and Sec. IV C, the depth of the circuit required for constructing the block Krylov subspace $\mathcal{K}_{n}(\hat{H}_{\text{ST}}(\Delta \tau), \{|\psi_k\rangle\}_{k=1}^{M_n})$ scales as $O(n)$ with a prefactor $D_n^{(p)}$. Since $m$ and $p$ in the symmetric Suzuki-Trotter decomposition $\hat{S}_n^{(p)}$ 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 $\mathcal{K}_{n}(\hat{H}_{\text{ST}}(\Delta \tau), \{|\psi_k\rangle\}_{k=1}^{M_n})$ 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 $\hat{H}$. Then we show by numerical simulations how the selection of the reference states affects the convergence to the ground state with $n$.  

the following product states for the subspace diagonalization:

\[ |q_1 \rangle = |\Phi_A \rangle = \otimes_{i=1}^{N/2} |s_{2i-1},2i \rangle, \]
\[ |q_2 \rangle = |\Phi_B \rangle = \otimes_{i=1}^{N/2} |s_{2i-2},2i \rangle, \]
\[ |q_3 \rangle = |X_{AFM1} \rangle = \otimes_{i=1}^{N/2} |+ \rangle_{2i-1} | - \rangle_{2i}, \]
\[ |q_4 \rangle = |X_{AFM2} \rangle = \otimes_{i=1}^{N/2} |+ \rangle_{2i-1} | - \rangle_{2i+1}, \]
\[ |q_5 \rangle = |Y_{AFM1} \rangle = \otimes_{i=1}^{N/2} |R \rangle_{2i-1} |L \rangle_{2i}, \]
\[ |q_6 \rangle = |Y_{AFM2} \rangle = \otimes_{i=1}^{N/2} |R \rangle_{2i-1} |L \rangle_{2i+1}, \]
\[ |q_7 \rangle = |Z_{AFM1} \rangle = \otimes_{i=1}^{N/2} |0 \rangle_{2i-1} |1 \rangle_{2i}, \]
\[ |q_8 \rangle = |Z_{AFM2} \rangle = \otimes_{i=1}^{N/2} |0 \rangle_{2i-1} |1 \rangle_{2i+1}, \]

where \(|s_{ij}\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{i} |1\rangle_{j} - |1\rangle_{i} |0\rangle_{j})\) is the spin-singlet state which is an eigenstate of the swap operator \(\hat{P}_{ij}\) with eigenvalue \(-1\) and is also known as one of the Bell states, \(|+\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle_i + |1\rangle_i\rangle\) and \(|-\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle_i - |1\rangle_i\rangle\) are the eigenstates of \(\hat{X}_i\) with eigenvalues \( \pm 1\), \(|R\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle_i + i |1\rangle_i\rangle\) and \(|L\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle_i - i |1\rangle_i\rangle\) are the eigenstates of \(\hat{Y}_i\) with eigenvalues \( \pm 1\), and \(|0\rangle_i\) and \(|1\rangle_i\) are the eigenstates of \(\hat{Z}_i\) with eigenvalues \( \pm 1\).

|\(\Phi_A\rangle\) and \(|\Phi_B\rangle\) are the ground states of \(\hat{H}_A\) and \(\hat{H}_B\), respectively, while others are the 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 Eq. 6) and, moreover, 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

\[ |q_9 \rangle = |\Psi_{\text{VQE}} \rangle \]  

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. [87]. While the exact ground-state energy is 
\(E_0/NJ = -0.196393522\), our variational state \(|\Psi_{\text{VQE}} \rangle\) has the variational energy \(\langle \Psi_{\text{VQE}} | \hat{H} | \Psi_{\text{VQE}} \rangle / NJ = -0.1885\) (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 [87], 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 \(|\tilde{q}_k\rangle \rangle_{k=1}^{N}\) with

\[ |\tilde{q}_k \rangle = \hat{F}_{k-1} |\Psi_{\text{VQE}} \rangle, \]  

where \(\hat{F}_k\) is a unitary operator representing the one-dimensional \(k\)-lattice-space translation with \(\hat{F}_0 = I\), and \(|\Psi_{\text{VQE}} \rangle\) is the same state given in Eq. (77). 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 $|\tilde{q}_k\rangle_{k=1}^N$, i.e., $\sum_{k=1}^N |\tilde{q}_k\rangle$, is translationally symmetric with momentum zero.

The reference states $|\Phi_A\rangle$, $|\Phi_B\rangle$, $|\Psi_{\text{VQE}}\rangle$, and $|\tilde{q}_k\rangle_{k=1}^N$ 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 zero for the reference states $|X_{\text{AFM}}\rangle$, $|Y_{\text{AFM}}\rangle$, and $|Z_{\text{AFM}}\rangle$, respectively. Because the Hamiltonian $\hat{H}$ considered here is spin singlet [88], 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 [88].

2. Ground-state energy and fidelity

Figures 5 and 6 show the estimated ground-state energy $E_{\text{KS}}$ and the ground-state fidelity $F = |\langle \Psi_0 | \Psi_{\text{KS}} \rangle|^2$, obtained by solving Eq. (54), as a function of $n = \text{dim} \mathcal{K}_n/M_B$, i.e., the dimension of the Krylov subspace $\mathcal{K}_n$ per block size $M_B$. Note that $\hat{H}^{(\text{ST})}_{\text{ST}}(\Delta_r)$ is the maximum approximated Hamiltonian power multiplied to the reference states when the Krylov subspace $\mathcal{K}_n(\hat{H}^{(\text{ST})}_{\text{ST}}(\Delta_r), |q_k\rangle_{k=1}^N)$ is constructed in Eq. (44).

Here, the Krylov-subspace Hamiltonian matrix $[\hat{H}]_{ij}$ and the overlap matrix $[S]_{ij}$ are computed as $\langle \tilde{q}_i | \hat{H} | \tilde{q}_j \rangle$ and $\langle \tilde{q}_i | \tilde{q}_j \rangle$, respectively. The first-order Richardson extrapolation ($r = 1$) and the lowest-order symmetric Suzuki-Trotter decomposition $\hat{S}_2$ are used for $[\hat{H}^{(\text{ST})}_{\text{ST}}(\Delta_r)]_{ij}^{(\text{ST})}$ with $\Delta_r = 0.05$, in which the systematic errors are practically negligible for our purpose (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 Krylov subspace 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 indicated at $n = 1$ are merely the expectation values of $\hat{H}$ with respect to the corresponding reference state, e.g., $\langle \Phi_{\text{K}} | \hat{H} | \Phi_{\text{K}} \rangle/NJ = -0.125$ and $\langle \Psi_{\text{VQE}} | \hat{H} | \Psi_{\text{VQE}} \rangle/NJ = -0.1885$. 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 the power $n$, the energy decreases monotonically and the fidelity keeps increasing towards one, implying that the ground state estimation can be improved system-

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

atically over a chosen set of reference states without any parameter optimization. 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 [10]. Notice also that the energy as well as the fidelity for $M_B = 16$ is consistently better than those for $M_B < 9$ 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 $M_B = 1$, 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 $E_{\text{KS}}$. Therefore, the multireference scheme with $M_B > 1$ works effectively for reducing the power $n$ 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 I summarizes the minimum dimension $n$ per block size of the Krylov subspace and the corresponding circuit depth required for converging the ground-state energy $E_{\text{KS}}$ with an accuracy $(E_{\text{KS}} - E_0)/NJ \leq 10^{-3}$ for $N = 16$. Note here that the commuting exponentials in $[\hat{S}_2(\pm \Delta_r/2)]^{n-1}$ are contracted
which are not included in the maximum circuit depth. The key ingredients of the quantum power method are the central-finite-difference scheme for the time derivatives and the symmetric Suzuki-Trotter decomposition that decomposes each time-evolution operator, both of which guarantee that the approximated Hamiltonian power \( \hat{\mathcal{H}}^n_{\text{ST}}(\Delta_t) \) retains the Hermiticity and the even parity in \( \Delta_t \), i.e., \( \hat{\mathcal{H}}^n_{\text{ST}}(\Delta_t) = \hat{\mathcal{H}}^n_{\text{ST}}(\Delta_t)\), with the controlled accuracy of the finite-difference error \( \epsilon_{\text{FD}} \sim O(\Delta_t^2) \) and the Suzuki-Trotter error \( \epsilon_{\text{ST}} \sim O(\Delta_t^{2m}) \). The number of gates required for approximating the Hamiltonian power \( \hat{\mathcal{H}}^n_{\text{ST}}(\Delta_t) \) is \( O(Nn) \), where \( N \) is the number of qubits and the Hamiltonian \( \hat{\mathcal{H}} \) is assumed to be local. This should be contrasted to the classical power method which scales exponentially in \( N \).

The \( r \)th order Richardson extrapolation can be adopted to systematically improve the systematic errors as \( \epsilon_{\text{FD}} \sim O(\Delta_t^{2+2r}) \) and \( \epsilon_{\text{ST}} \sim O(\Delta_t^{2m+2r}) \) in the approximated Hamiltonian power \( \hat{\mathcal{H}}^n_{\text{ST}}(\Delta_t) \), 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 \( \epsilon_{\text{FD}} \) and \( \epsilon_{\text{ST}} \) 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 as compared with the algorithm that requires the higher-order Suzuki-Trotter decomposition with increasing the power \( n \) (see Appendix B). Therefore, the quantum power method proposed here is potentially promising 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 \( \hat{\mathcal{H}}^n_{\text{ST}}(\Delta_t) \) 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 \( N_D = 2^{24} \approx 10^7 \).

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 with simple parameter-free product states 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 [78], 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 [17], but here one can treat higher powers of the Hamiltonian on quantum computers though approximately yet with a controlled accuracy.

### 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

![Graph](image_url)
The quantum power method proposed here can be easily generalized to higher spatial dimensions or a more complicated system described by a Hamiltonian that is decomposed into \( \hat{H} = \hat{H}_A + \hat{H}_B + \hat{H}_C + \cdots \), where generally \( [\hat{H}_A, \hat{H}_B] \neq 0 \) if \( \Gamma \neq \Gamma' \) but terms within each \( \hat{H}_A \) commute to each other (here, \( \Gamma = A, B, C, \cdots \)). Even in this case, the number of gates required in approximating the Hamiltonian power \( \hat{H}^n \) scales similarly, except for the additional prefactor of \( O(N_{\Gamma}) \), where \( N_{\Gamma} \) is the number of sub-divided Hamiltonians in \( \hat{H} \). This is because when the lowest-order symmetric Suzuki-Trotter decomposition is employed, e.g., \( e^{-i\Delta t(\hat{H}_A+\hat{H}_B+\hat{H}_C)} = e^{-i\frac{\Delta t}{2}\hat{H}_A}e^{-\frac{i\Delta t}{2}\hat{H}_B}e^{\frac{i\Delta t}{2}\hat{H}_C}e^{-i\frac{\Delta t}{2}\hat{H}_A} + O(\Delta^3) \), the time-evolution operator \( e^{-i\Delta t\hat{H}} \) is approximated generally as a product of \( 2N_{\Gamma} - 1 \) time-evolution operators of the sub-divided systems. In the multireference Krylov-subspace diagonalization, one can choose, for example, the ground states \( |\Phi_{\Delta_A}\rangle, |\Phi_{\Delta_B}\rangle, |\Phi_{\Delta_C}\rangle, \cdots \) of the sub-divided Hamiltonians \( \hat{H}_A, \hat{H}_B, \hat{H}_C, \cdots \) for the reference states. The mean-field ground states can also be used for the reference states.

Although we have simulated only the ground-state energy, the expectation value of other observables that commute with the Hamiltonian \( \hat{H} \) can be evaluated similarly. When an observable \( \hat{O} \) does not commute with the Hamiltonian \( \hat{H} \), the expectation value with respect to the approximate ground state \( \langle \Psi_0 \rangle \approx \langle \Psi_{KS} \rangle \equiv \sum_{i=1}^{M_B} \langle \tilde{u}_i | \tilde{u}_i \rangle \), with the coefficients \( v_i \) already determined by solving Eq. (54) in the block Krylov subspace \( \mathcal{K}_c(\mathcal{H}_{\Sigma}(\Delta_r), \{|q_k\}_k=1)^{M_B} \rangle \), can also be evaluated as

\[
\langle \Psi_0 | \hat{O} | \Psi_0 \rangle \approx \sum_{i=1}^{M_B} \sum_{j=1}^{M_B} v_i^* v_j |\tilde{u}_i \rangle \langle \tilde{u}_i | \hat{O} | \tilde{u}_j \rangle \\
= \sum_{i=1}^{M_B} \sum_{j=1}^{M_B} v_i^* v_j \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} C_{l-i-1,j-i-1}^{2l-i-j-1} \langle \hat{S}_{2m}^{(p)}(\Delta_r) \rangle^{l-1-2i} \langle \hat{S}_{2m}^{(p)}(\Delta_r) \rangle^{l-1-2j} \langle \Phi | \hat{O} | \Phi \rangle, 
\]

(79)

where \( |\tilde{u}_i \rangle = \hat{H}_{\Sigma}^{(i-l)}(\Delta_r) |q_k \rangle \), as given in Eq. (57), and the explicit form of \( \hat{H}_{\Sigma}^{(i-l)}(\Delta_r) \) with \( r = 0 \) is used in the second line. Here, \( i = k + (l-1)M_B \) and \( j = k' + (l' - 1)M_B \).

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 to near-term quantum computers. Other applications of the quantum power method include various moment based methods, which are briefly outlined in Appendix C and Appendix D. In these appendixes, we show that the power method can evaluate \( \langle \Psi | \hat{H}^{2} | \Psi \rangle \) with exactly the same amount of resource that is required for \( \langle \Psi | \hat{H}^{2} | \Psi \rangle \), and therefore, for example, the energy variance \( \sigma^2 = \langle \Psi | \hat{H}^{2} | \Psi \rangle - \langle \Psi | \hat{H} | \Psi \rangle^2 \) 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 [89].

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

\[
\hat{A}^n = \int_0^\tau dt e^{i\hat{A}t/n} = \int_0^\tau dt e^{i\hat{A}t/n} 
\]

with the generating function \( \hat{V}(t) = e^{-i\hat{A}t/\tau} \). 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 have been done on the HOKUSAI supercomputer at RIKEN (Project ID: G200015). This work was supported by Grant-in-Aid for Research Activity start-up (No. JP19K23433) and Grant-in-Aid for Scientific Research (B) (No. JP18H01183) from MEXT, Japan.

**Appendix A: Higher-order symmetric Suzuki-Trotter decompositions \( \hat{S}_{2m}^{(p)}(\Delta_r) \)**

In this appendix, we provide a \texttt{Python} program that generates coefficients required for the higher-order symmetric Suzuki-Trotter decompositions \( \hat{S}_{2m}^{(p)}(\Delta_r) \) introduced in Sec. III C 2, and examine numerically the systematic errors due to the Suzuki-Trotter decompositions \( \hat{S}_{2m}^{(p)}(\Delta_r) \) 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 \texttt{Python} program that generates the coefficients \( \Delta s_i \) for a given set of parameters \( m \) and \( p \) in the symmetric Suzuki-Trotter decompositions \( \hat{S}_{2m}^{(p)}(\Delta_r) \):

\[
\hat{S}_{2m}^{(p)}(\Delta_r) = e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A} \times \cdots \times e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A}, \quad (A1)
\]

where \( x = -i\Delta_r \) and \( D^{(p)}_{2m} = 2p^{m-1} + 1 \) derived in Eq. (34). The program includes an example for \( m = 2 \) and \( p = 5 \). In this case, the symmetric Suzuki-Trotter decomposition has a form

\[
\hat{S}_{4}^{(5)}(\Delta_r) = e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_B} \times e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A} \times e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_A} e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_B} e^{i\Delta_r \hat{H}_A}, \quad (A2)
\]
and the output of the program gives the 11 coefficients

\[
\begin{align*}
  s_1 &= 0.20724538589718786, \\
  s_2 &= 0.4144907717943757, \\
  s_3 &= 0.4144907717943757, \\
  s_4 &= 0.4144907717943757, \\
  s_5 &= -0.12173615769156357, \\
  s_6 &= -0.6579630871775028, \\
  s_7 &= -0.12173615769156357, \\
  s_8 &= 0.4144907717943757, \\
  s_9 &= 0.4144907717943757, \\
  s_{10} &= 0.4144907717943757, \\
  s_{11} &= 0.20724538589718786.
\end{align*}
\]

By modifying lines 20-21 in the program, one can obtain \( \{s_i\}_{i=1}^{D_{2m}^{(p)}} \) for other values of \( m \) and \( p \).

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

\[ s_i = s_{D_{2m}^{(p)} - i + 1} \quad (A3) \]

and satisfy the following sum rule:

\[ \sum_{i=1}^{D_{2m}^{(p)}} s_i = 2 \quad (A4) \]

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

\[ T_i = \sum_{k=1}^{i} s_k. \quad (A5) \]

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 [58, 72].

import numpy

def Suzuki_Trotter(m, p):
    s = numpy.array([0.5, 1.0, 0.5])
    pl = int((p-1)/2)
    for mm in range(m-1):
        k0 = 1.0 / (p-1) ** (1.0 / (2*mm+3))
        s1 = s[-1] * k0
        sl = numpy.concatenate([s1 for _ in range(pl)])
        s = numpy.concatenate([s, sl[::-1]])
    s[-1] = s[0]
    s[0] = s[0] + s[1]
    s = s * k1
    s[0] = s[0] + s[1]
    s[-1] = s[0]
    return s

# example
m=2
p=5
s=Suzuki_Trotter(m, p)
T=numpy.cumsum(s)
print('order =', 2*m)
print('depth =', len(s))
print('sum(s) =', numpy.sum(s))
for i in range(len(s)):
    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 \( \hat{S}_{2m}^{(p)} \).

2. Numerical examination of a Suzuki-Trotter error

Here we numerically examine the systematic errors due to the Suzuki-Trotter decompositions \( \hat{S}_{2m}^{(p)}(\Delta t) \) with different parameters \( m \) and \( p \). The Trotter formula [90–92] combined with \( \hat{S}_{2m}^{(p)}(\Delta t) \) yields

\[ \hat{U}(t) = \left[ \hat{S}_{2m}^{(p)}(\Delta t) \right]^M + O(t\Delta^2 m), \quad (A6) \]

where \( M \) is an integer such that \( t = MA \). Figure 7 shows the real part of the difference between the exact propagator

\[ K(t) = \langle \Psi_0 | \hat{U}(t) | \Psi_0 \rangle \quad (A7) \]

and the approximated propagator

\[ \hat{K}(t) = \langle \Psi_0 | \left[ \hat{S}_{2m}^{(p)}(\Delta t) \right]^M | \Psi_0 \rangle, \quad (A8) \]

i.e.,

\[ \text{Re} \delta K(t) = \text{Re} \hat{K}(t) - \text{Re} K(t), \quad (A9) \]

with \( \Delta, J = 0.07 \) and 0.1 for the spin-1/2 Heisenberg model on an \( N = 16 \) qubit ring described by the Hamiltonian \( \tilde{H} \) in Eq. (1). Here, \( | \Psi_0 \rangle \) is the exact ground state. The exact propagator is simply given by \( K(t) = e^{-i \hat{H} t} \), where \( \hat{E}_0 \) 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 Re$\delta K(t)$, the imaginary part of the difference, Im$\delta K(t)$, behaves similarly.

We should emphasize here that while the deviation of the approximated propagator $K(t)$ 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 evolution operators $\hat{U}(t)$.

It is also found that the error decreases with increasing $p$. As expected from Eq. (A6), the values of $\delta$ are approximated by the symmetric Suzuki-Trotter decomposition described in Sec. III D, the time-evolution operators in Eq. (14) being larger in the long time limit ($\Delta_0, \Delta_1, \Delta_2$). Although we only show $\delta_k(\Delta)$, the exact time-evolution operators satisfy the multiplicity law $\hat{U}(\Delta_0)\hat{U}(\Delta_1) = \hat{U}(\Delta_0 + \Delta_1)$. Note also that $\hat{H}_{\text{ST}}^{2n}(\Delta) = \hat{H}_{\text{ST}}^{2n+1}(\Delta)$.

We can readily confirm that $\hat{H}_{\text{ST}}^{2n}(\Delta)$ is Hermitian and an even function of $\Delta$, i.e.,

$$\hat{H}_{\text{ST}}^{2n}(\Delta) = [\hat{H}_{\text{ST}}^{2n}(\Delta)]^\dagger = \hat{H}_{\text{ST}}^{2n}(-\Delta),$$

as in the case of $\hat{H}_{\text{ST}}^{2n}(\Delta)$ given in Eq. (39) and hence the systematic error $E_{\text{ST}}$ as well as the systematic error $E_{\text{TD}}$, see Sec. III.A) in odd powers of $\Delta$ is absent in Eq. (B1). We can also show that $\hat{H}_{\text{ST}}^{2n}(\Delta)$ does not satisfy the law of exponents, i.e.,

$$\hat{H}_{\text{ST}}^{2n}(\Delta) \neq \left[\hat{H}_{\text{ST}}^{2n}(\Delta)\right]^n$$

for $n \geq 2$, simply because of Eq. (B3), but only satisfies it approximately within the systematic errors. This is in sharp contrast to the case of $\hat{H}_{\text{ST}}^{2n}(\Delta)$, which satisfies exactly the law of exponents in Eq. (38).

At first glance, one would tend to conclude that $\hat{H}_{\text{ST}}^{2n}(\Delta)$ in Eq. (B2) is more suitable to approximate the Hamiltonian power $\hat{H}^n$ than $\hat{H}_{\text{ST}}^{n}(\Delta)$ in Eq. (36), because each team in $\hat{H}_{\text{ST}}^{2n}(\Delta)$ contains a single $\hat{S}^{(p)}$ not a product of multiple $\hat{S}^{(p)}$ as in $\hat{H}_{\text{ST}}^{2n}(\Delta)$, thus expecting the less number of gates in the circuit. However, the disadvantage of $\hat{H}_{\text{ST}}^{2n}(\Delta)$ in Eq. (B2) is that the higher-order Suzuki-Trotter decompositions are required for approximating the Hamiltonian power $\hat{H}^n$ with larger $n$.

This can be understood by recalling that $\hat{S}^{(p)}(t)$ has a form of Eq. (29):

$$\hat{S}^{(p)}_{2m}(t) = \exp \left[-i\hat{H} + (i)^{m+1}\hat{R}_{n+1} + \cdots\right].$$

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

$$i^n \frac{d^n \hat{S}^{(p)}_{2m}(t)}{dt^n} \bigg|_{t=0} = \hat{H}^n$$

for $n \leq 2m$ but

$$i^n \frac{d^n \hat{S}^{(p)}_{2m}(t)}{dt^n} \bigg|_{t=0} \neq \hat{H}^n,$$

where

$$\hat{H}_{\text{ST}}^{2n}(\Delta) = \sum_{k=0}^{n} c_n k \hat{S}^{(p)}_{2m} \left(\frac{n}{2} - k\right) \Delta.$$

$O(\Delta)$ represents the systematic error $E_{\text{TD}}$ due to the finite-difference scheme for the time derivatives, and $E_{\text{ST}}$ denotes the systematic error due to the Suzuki-Trotter decomposition of the time-evolution operators, the order of $E_{\text{ST}}$ being discussed below. We should emphasize here that $\hat{H}_{\text{ST}}^{2n}(\Delta) \neq \hat{H}_{\text{ST}}^{2n+1}(\Delta)$ for $n \geq 2$ because

$$\hat{S}^{(p)}_{2m}(\Delta) \hat{S}^{(p)}_{2m}(\Delta') \neq \hat{S}^{(p)}_{2m}(\Delta + \Delta')$$

for $\Delta \neq -\Delta'$, although the exact time-evolution operators satisfy the multiplication law $\hat{U}(\Delta)\hat{U}(\Delta') = \hat{U}(\Delta + \Delta')$. Note also that $\hat{H}_{\text{ST}}^{2n}(\Delta) = \hat{H}_{\text{ST}}^{2n+1}(\Delta)$.

Finally, we note that several exponential-product formulas, not limited to those found by Suzuki, up to the depth $\leq 11$ with an error analysis can be found in Ref. [93]. Other error analysis of the Suzuki-Trotter decomposition devoted for quantum computing can be found in Refs. [94–96].
for $n > 2m$. For example, if $n = 2m + 1$, the derivative reads

$$i^{2m+1} \frac{d^{2m+1} S^{(p)}_{2m}(t)}{d\tau^{2m+1}} \bigg|_{\tau=0} = \hat{H}^{2m+1} + (2m + 1)! \hat{R}_{2m+1}. \quad (B9)$$

It is now important to notice that the right-hand side of Eq. (B2) corresponds to the central finite-difference approximation of $i^n \frac{d^n S^{(p)}_{2m}(t)}{d\tau^n}$, i.e.,

$$i^n \frac{d^n S^{(p)}_{2m}(t)}{d\tau^n} \bigg|_{\tau=0} = \sum_{k=0}^{n} c_{n,k} S^{(p)}_{2m} \left( \frac{n-k}{2} \Delta \tau \right) + O(\Delta \tau^2). \quad (B10)$$

In other words, the approximated Hamiltonian power $\hat{H}^{\text{ST}}_{\lambda_n}(\Delta \tau)$ in Eq. (B2) is given by the higher-order derivative of $S^{(p)}_{2m}(t)$ at $t = 0$ as

$$\hat{H}^{\text{ST}}_{\lambda_n}(\Delta \tau) = i^n \frac{d^n S^{(p)}_{2m}(t)}{d\tau^n} \bigg|_{\tau=0} + O(\Delta \tau^2). \quad (B11)$$

It is now obvious that the formalism in Eq. (B2) breaks down if $n > 2m$ because in this case, according to Eq. (B8), $\lim_{\Delta \tau \to 0} \hat{H}^{\text{ST}}_{\lambda_n}(\Delta \tau) \neq \hat{H}^n$, which contradicts to Eq. (B1). Therefore,

$$2m \gg n \quad (B12)$$

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

$$E_{\text{ST}} \sim O(\Delta \tau^2). \quad (B13)$$

This is the most important difference from the algorithm described in Sec. III D, 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 $\hat{H}^{\text{ST}}_{\lambda_n}(\Delta \tau)$ in Eq. (37) can be considered as

$$\hat{H}^{\text{ST}}_{\lambda_n}(\Delta \tau) = i^n \frac{d^n S^{(p)}_{2m}(t)}{d\tau^n} \bigg|_{\tau=0} + O(\Delta \tau^2) \bigg|^{(n)}. \quad (B14)$$

Therefore, the lowest-order symmetric Suzuki-Trotter decomposition with $m = 1$ is adequate to satisfy Eq. (B12) and indeed, as discussed in Sec. III D, it approximates the Hamiltonian power $\hat{H}^n$ with the controlled accuracy. Second, al-
though we have emphasized that the violation of the multiplication law $S_{2m}(\frac{\Delta r}{2})S_{2m}(\frac{\Delta r}{2}) \neq S_{2m}(\Delta r)$ is the essential point that distinguishes the two algorithms described here and in Sec. III D, this equation is satisfied within the systematic error. i.e.,

$$S_{2m}(\frac{\Delta r}{2})S_{2m}(\frac{\Delta r}{2}) = S_{2m}(\Delta r) + O(\Delta r^{2m+1}).$$ \hfill (B15)

Accordingly, the two algorithms described here and in Sec. III D should be the same within the systematic error. In fact, the approximated Hamiltonian powers $\hat{H}^m_{ST}(\Delta r)$ and $\hat{H}^m_{ST}(\Delta r)$ in Eqs. (36) and (B2), respectively, are equivalent within the systematic error because

$$\hat{H}^m_{ST}(\Delta r) = \sum_{k=0}^{n} C_n k \left[ S_{2m}^{(p)} \frac{\Delta r}{2} \right]^{n-2k}$$ \hfill (B16)

$$= \sum_{k=0}^{n} C_n k \left[ S_{2m}^{(p)} \left( \frac{n}{2} - k \right) \Delta r \right] + O(\Delta r^{2m+1})$$ \hfill (B17)

$$= \hat{H}^m_{ST}(\Delta r) + O(\Delta r^{2m+1-n}),$$ \hfill (B18)

provided that $2m + 1 > n$, which is consistent with Eq. (B12).

As an example, we show in Fig. 9 the expectation values $\langle \hat{H}^m_{ST}(\Delta r) \rangle$ and $\langle \hat{H}^m_{VQE}(\Delta r) \rangle$ 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 $n = 3$. Here, a simplified notation of the expectation value

$$\langle \cdots \rangle \equiv \langle \Psi | \cdots | \Psi \rangle$$ \hfill (B19)

with $|\Psi\rangle \in \{|\Phi_A\rangle, |\Psi_{\text{VQE}}\rangle\}$ is introduced. According to Eqs. (B7)–(B9), $\langle \hat{H}^3_{ST}(\Delta r) \rangle$ in the limit of $\Delta r \to 0$ should converge as

$$\lim_{\Delta r \to 0} \langle \hat{H}^3_{ST}(\Delta r) \rangle = \langle \hat{H}^3 \rangle$$ \hfill (B20)

for $m \geq 2$, but

$$\lim_{\Delta r \to 0} \langle \hat{H}^2_{ST}(\Delta r) \rangle = \langle \hat{H}^3 + 3! \hat{R}_3 \rangle$$ \hfill (B21)

for $m = 1$. Here, the explicit form of the residual term $\hat{R}_3$ in Eq. (B21) can be derived by using the Baker-Campbell-Hausdorff formula for $\hat{S}_2$ as \cite{59, 93, 97}

$$\hat{R}_3 = -\frac{1}{24} \left[ \hat{H}_A, [\hat{H}_A, \hat{H}_B] \right] + \frac{1}{12} \left[ \hat{H}_B, [\hat{H}_B, \hat{H}_A] \right].$$ \hfill (B22)

The numerical results in Fig. 9 confirm Eqs. (B20) and (B21), as well as the expected behavior $\lim_{\Delta r \to 0} \langle \hat{H}^3_{ST}(\Delta r) \rangle = \langle \hat{H}^3 \rangle$. 

---

**FIG. 8.** Same as Fig. 7 but the deviation $\text{Re} \hat{S}(t)$ divided by $(\Delta r)^{2m}$ for several values of $\Delta r$ as indicated in the figures.
for \(m = 1\). Note also that the linear convergence of these quantities to the exact values as a function of \(N^2\) shown in Fig. 9 corroborates the systematic errors expected for \(\langle \hat{H}_{ST}^n(\Delta_r) \rangle\) in Eqs. (B10) and \(\langle \hat{H}_{ST}^n(\Delta_r) \rangle\) in Eq. (35).

Now we discuss the gate count for approximating \(\hat{H}^n\) with \(\hat{H}_{ST}^n(\Delta_r)\). As described above, Eq. (B12) 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 \(\hat{H}^n\) being \(m = [n/2]\), where \([\cdot]\) is the ceiling function that returns the minimum integer larger than or equal to the argument. Therefore, assuming that the Hamiltonian \(\hat{H}\) is local, the number of gates required for approximating \(\hat{H}^n\) with \(\hat{H}_{ST}^n(\Delta_r)\) is \(O(Np^{n/2})\) because the circuit depth \(D_{2m}^{(p)}\) for the single Suzuki-Trotter decomposed time-evolution operator \(\hat{S}_{2m}^{(p)}\) is given by Eq. (33), and thus increases exponentially in the power \(n\). In contrast, as described in Sec. III D, the number of gates required for approximating \(\hat{H}^n\) with \(\hat{H}_{ST}^n(\Delta_r)\) is \(O(Nn)\) with a prefactor \(D_{2m}^{(p)} \sim O(1)\), i.e., increasing polynomially in \(N\) and \(n\).

This indicates that the algorithm based on \(\hat{H}_{ST}^n(\Delta_r)\) suffers from the exponential increase of the number of gates for large \(n\). However, the algorithm based on \(\hat{H}_{ST}^n(\Delta_r)\) can be more favorable than that based on \(\hat{H}_{ST}^n(\Delta_r)\) when the power \(n\) is small. To be more specific, let us consider the case of \(p = 3\). Then the circuit depth for \(\hat{H}_{ST}^n(\Delta_r)\) is given by \(D_{2m}^{(3)}(2^{n/2}) = 3, 3, 7, 19, 19, 55, 55, 163, \ldots\), while the largest circuit depth for \(\hat{H}_{ST}^n(\Delta_r)\) involving \([\hat{S}^{(p)}_2(\pm \Delta/2)n]\) is \(n(D_{2m}^{(3)} - 1) + 1 = 2n + 1 = 3, 5, 7, 9, 11, 13, 15, 17, 19, \ldots\), for the Hamiltonian power \(n\) is \(1, 2, 3, 4, 5, 6, 7, 8, 9, \ldots\). Here, for the latter, the depth is counted by assuming that the commuting exponentials in \([\hat{S}^{(p)}_2(\pm \Delta/2)n]\) are contracted. Therefore, in this case with \(p = 3\), the algorithm based on \(\hat{H}_{ST}^n(\Delta_r)\) is more preferable than that based on \(\hat{H}_{ST}^n(\Delta_r)\) as long as the power \(n \ll 4\). As shown in Appendix C 4, the algorithm based on \(\hat{H}_{ST}^n(\Delta_r)\) 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. By defining

\[
|\tilde{u}\rangle = |\hat{H}^{\frac{n-1}{2}}_{ST}(\Delta_r)q_k\rangle
\]

for the basis set generated in the block Krylov subspace \(\mathcal{K}_n(\hat{H}_{ST}(\Delta_r), \{ |\varphi_k\rangle \}_{k=1}^M)\), the matrix elements \(\hat{H}\) and \(\hat{S}\) in Eqs. (49) and (50) are now approximated by replacing \(|u_i\rangle\) with \(|\tilde{u}_i\rangle\) as

\[
\hat{R}_{ij} = \langle \tilde{u}_i | \hat{H} | \tilde{u}_j \rangle = \langle q_k | \hat{H}^{\frac{n-1}{2}}_{ST}(\Delta_r) \hat{H} \hat{H}^{\frac{n-1}{2}}_{ST}(\Delta_r) | q_{k'} \rangle
\]

and

\[
\hat{S}_{ij} = \langle \tilde{u}_i | \hat{S} | \tilde{u}_j \rangle = \langle q_k | \hat{H}^{\frac{n-1}{2}}_{ST}(\Delta_r) \hat{H} \hat{H}^{\frac{n-1}{2}}_{ST}(\Delta_r) | q_{k'} \rangle
\]

where \(i = k + (l - 1)M_B\) and \(j = k' + (l' - 1)M_B\) for \(1 \leq k, k' \leq M_B\) and \(1 \leq l, l' \leq n\). As compared with Eqs. (60) and (61), the power exponents are now distributed to the left and the right basis states.

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

\[
\hat{R}_{ij} = \frac{1}{2} \sum_{v=0}^{n} \sum_{r=0}^{n-1} \sum_{r'=0}^{n-1} c_{r-1,v}^{(r')} \times \langle q_k | \hat{S}_{2m}^{(p)} (-i^{(r')}) | \hat{P}_{l,v+1}^{(p)} \hat{S}_{2m}^{(p)} (i^{(r')}) | q_{k'} \rangle
\]

and

\[
\hat{S}_{ij} = \sum_{v=0}^{n} \sum_{r=0}^{n-1} c_{r-1,v}^{(r')} \times \langle q_k | \hat{S}_{2m}^{(p)} (-i^{(r')}) \hat{S}_{2m}^{(p)} (i^{(r')}) | q_{k'} \rangle,
\]

where \(i^{(r')} = \left( \frac{r}{2} + v - 1 \right) \Delta_r\). The swap operator \(\hat{P}_{l,v+1}\) is the local term of the Hamiltonian in Eq. (3) and can be further simplified to a product of two Pauli operators as in Eq. (2). The number of terms in Eqs. (B26) and (B27) is \(O(Nn)\) and \(O(ll')\), respectively. In total, \(O(M_B^2 n^2)\) and \(O(M_B^2 n^2)\) state overlaps are required to be evaluated for constructing all matrix elements of the \(nM_B \times nM_B\) matrices \(\hat{R}\) and \(\hat{S}\), respectively.
Finally, we note that the systematic errors \( E_{FD} \) and \( E_{ST} \) in Eq. (B1) can be improved systematically, without increasing the gate count of each circuit, by adopting the Richardson extrapolation as

\[
\hat{H}'^n = \hat{H}^n_{ST(r)}(\Delta_r) + O(\Delta_r^{2r}) \tag{B28}
\]

where \( \hat{H}^n_{ST(r)}(\Delta_r) \) is the \( r \)th order Richardson extrapolation of the approximate Hamiltonian power, i.e.,

\[
\hat{H}^n_{ST(r)}(\Delta_r) = \frac{h^{2r} \hat{H}^n_{ST(r-1)}(\Delta_r/h) - \hat{H}^n_{ST(r-1)}(\Delta_r)}{h^{2r} - 1}, \tag{B29}
\]

with \( \hat{H}^n_{ST(0)}(\Delta_r) \equiv \hat{H}^n_{ST}(\Delta_r) \). Since \( \hat{H}^n_{ST(0)}(\Delta_r) \) is a linear combination of \( n + 1 \) unitaries \( S_{2m}^{(n)} \), \( \hat{H}^n_{ST(r)}(\Delta_r) \) is a linear combination of \( (r + 1)(n + 1) \) unitaries \( S_{2m}^{(n)} \).

### Appendix C: 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 connected-cluster expansion (CMX) for a short-time 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 B 2. 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

\[
K(t) = \langle \hat{U}(t) \rangle = \sum_{n=0}^{\infty} \frac{(-i)^n t^n}{n!} \mu_n, \tag{C1}
\]

where \( \hat{U}(t) \) is the time-evolution operator given in Eq. (7) and

\[
\mu_n = \langle \hat{H}^n \rangle, \tag{C2}
\]

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

\[
\Phi(t) \equiv \ln K(t) = \ln(e^{-i\hat{H}t}) \equiv \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} \kappa_n. \tag{C3}
\]

Thus, the \( n \)th moment \( \mu_n \) and cumulant \( \kappa_n \) are given by the \( n \)th time derivative of generating functions \( K(t) \) and \( \Phi(t) \), respectively, as

\[
\mu_n = i^n \frac{d^n K(t)}{dt^n} \bigg|_{t=0}, \tag{C4}
\]

and

\[
\kappa_n = i^n \frac{d^n \Phi(t)}{dt^n} \bigg|_{t=0}. \tag{C5}
\]

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 \([98]\).

It should be noticed \([99]\) that the \( n \)th moment \( \mu_n \) can be expressed as

\[
\mu_n = \kappa_0 + \sum_{k=1}^{n-1} \binom{n-1}{k-1} \kappa_k \mu_{n-k} \tag{C6}
\]

and, equivalently, the \( n \)th cumulant \( \kappa_n \) can be expressed as

\[
\kappa_n = \mu_n - \sum_{k=1}^{n-1} \binom{n-1}{k-1} \kappa_k \mu_{n-k}. \tag{C7}
\]

Therefore, from the moments \( \{\mu_n\}_{l=0}^{\infty} \), one can obtain the cumulants \( \{\kappa_l\}_{l=0}^{\infty} \) 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 \( \kappa_n \sim O(N) \) \([100]\).

#### 2. Finite-difference approximation

By approximating the derivative in Eq. (C4) with the central-finite-difference method, we obtain that

\[
\mu_n = \mu_n(\Delta_r) + O(\Delta_r^2), \tag{C8}
\]

where

\[
\mu_n(\Delta_r) = \sum_{i=0}^{n} c_{n,i} K \left( \left( \frac{n}{2} - i \right) \Delta_r \right). \tag{C9}
\]

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

\[
\mu_{2m+1}(\Delta_r) = 2i \sum_{i=0}^{m} c_{2m+1,i} \text{Im} K \left( \left( m + \frac{1}{2} - i \right) \Delta_r \right) \tag{C10}
\]

and

\[
\mu_{2m}(\Delta_r) = c_{2m,m} + 2 \sum_{i=0}^{m-1} c_{2m,i} \text{Re} K \left( \left( m - i \right) \Delta_r \right), \tag{C11}
\]

where \( K(0) = 1 \) and \( c_{2m,m} = \frac{1}{\Delta_r^2} \binom{2m}{m} \) are used in Eq. (C11). Thus, for obtaining \( \mu_n(\Delta_r) \), it suffices to evaluate \( K(t) \) at equally spaced \( \lfloor n/2 \rfloor \) different points, where \( \lfloor . \rfloor \) denotes the ceiling function defined previously. Moreover, if \( \{K(l\Delta_r/2)\}_{l=1}^{n-1} \) used for evaluating the moments \( \{\mu_l(\Delta_r)\}_{l=1}^{n-1} \) are all stored, only \( K(n\Delta_r/2) \) to be evaluated for \( \mu_n(\Delta_r) \). Therefore, for obtaining all \( n \) moments \( \{\mu_l(\Delta_r)\}_{l=1}^{n} \), it is sufficient to evaluate the propagator at \( n \) different points, i.e., \( 
\{K(l\Delta_r/2)\}_{l=1}^{n} \) only once.
We can apply the same argument for the cumulants. Using the central-finite-difference method, the nth cumulant is evaluated as

\[ \kappa_n = \kappa_n(\Delta r) + O(\Delta r^2), \]  

(C12)

where

\[ \kappa_n(\Delta r) = \sum_{i=0}^{n} c_{n,i} \Phi \left( \left( \frac{n}{2} - i \right) \Delta r \right). \]  

(C13)

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

\[ \kappa_{2m+1}(\Delta r) = 2i \sum_{i=0}^{m} c_{2m+1,i} \text{Im} \Phi \left( \left( m + \frac{1}{2} - i \right) \Delta r \right), \]  

(C14)

and

\[ \kappa_{2m}(\Delta r) = 2 \sum_{i=0}^{m-1} c_{2m,i} \text{Re} \Phi \left( \left( m - i \right) \Delta r \right), \]  

(C15)

where \( \Phi(0) = \ln K(0) = 0 \) is used in Eq. (C15). Note that, if we write the propagator as \( K(t) \) with \( \Phi(t) \) real, then \( \text{Re} \Phi(t) = \ln \langle a(t) \rangle \) 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 [101]. 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 nth moment \( \mu_n \) 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 \( t_i^{(n)} = \left( \frac{n}{2} - i \right) \Delta r \) for \( i = 0, 1, \ldots, n \). Similarly, the nth cumulant \( \kappa_n \) can be approximated as a linear combination of the propagator \( \Phi(t) \), 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. (27). However, at this point, it is crucially important to recall the argument given in Sec. III D and Appendix B. Although the time evolution operator \( \hat{U}(t) \) evaluated at time \( t_i^{(n)} = \left( \frac{n}{2} - i \right) \Delta r \) satisfies that

\[ \hat{U} \left( \left( \frac{n}{2} - i \right) \Delta r \right) = \left[ \hat{U} \left( \Delta r \right) \right]^{n-2i}, \]  

(C16)

and thus the approximated nth moment \( \mu_n(\Delta r) \) in Eq. (C9) is equivalent to

\[ \mu_n(\Delta r) = \sum_{i=0}^{n} c_{n,i} \left[ \hat{U} \left( \frac{\Delta r}{2} \right) \right]^{n-2i}. \]  

(C17)

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

\[ \hat{S}_{2m}^{(p)} \left( \left( \frac{n}{2} - i \right) \Delta r \right) \neq \left[ \hat{S}_{2m}^{(p)} \left( \frac{\Delta r}{2} \right) \right]^{n-2i}. \]  

Therefore, the Feynman propagator \( K(t_i^{(n)}) \) in Eq. (C9) can be approximated either as

\[ K(t_i^{(n)}) = \left[ \hat{S}_{2m}^{(p)} \left( \left( \frac{n}{2} - i \right) \Delta r \right) \right] + O(\Delta r^{2m+1}) \]  

(C19)

or

\[ K(t_i^{(n)}) = \left[ \hat{S}_{2m}^{(p)} \left( \frac{\Delta r}{2} \right) \right]^{n-2i} + O(\Delta r^{2m+1}). \]  

(C20)

If the Feynman propagator \( K(t_i^{(n)}) \) is approximated as in Eq. (C20), the nth moment \( \mu_n \) is given by

\[ \mu_n = \sum_{i=0}^{n} c_{n,i} \left[ \hat{S}_{2m}^{(p)} \left( \left( \frac{n}{2} - i \right) \Delta r \right) \right] + O(\Delta r^{2m}) \]  

(C21)

and thus the lowest-order symmetric Suzuki-Trotter decomposition \( \hat{S}_{2m}^{(p)} \) with \( m = 1 \) can be adopted (see Sec. III D). 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. (C19), the higher-order symmetric Suzuki-Trotter decomposition \( \hat{S}_{2m}^{(p)} \) is required. As discussed in Appendix B, in order to evaluate the nth moment \( \mu_n \) with the controlled accuracy, the order of the symmetric Suzuki-Trotter decomposition \( \hat{S}_{2m}^{(p)} \) must be \( 2m \geq n \) [see Eq. (B12)]. In this case, the systematic error is \( O(\Delta r^2) \), i.e.,

\[ \mu_n = \sum_{i=0}^{n} c_{n,i} \left[ \hat{S}_{2m}^{(p)} \left( \left( \frac{n}{2} - i \right) \Delta r \right) \right] + O(\Delta r^2). \]  

(C22)

Therefore, this approach is not suitable for large \( n \) but is more preferable than the other approach when \( n \leq 4 \). The same argument is applied for the cumulant \( \kappa_n \).

4. First and second moments

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

From Eqs. (C10) and (C11), we can approximate the first and second moments \( \mu_1 \) and \( \mu_2 \) as

\[ \mu_1(\Delta r) = -2 \Delta r \text{Im} \left( \hat{U} \left( \frac{\Delta r}{2} \right) \right), \]  

(C23)
and

\[ \mu_2(\Delta_r) = \frac{2}{\Delta_r^2} \left[ 1 - \text{Re} \left\langle \hat{S}_2(\Delta_r) \right\rangle \right], \quad (C24) \]

respectively. This is already remarkable because the second moment \( \mu_2 \) 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 \( \hat{U}(\Delta_r) \) is approximated by the lowest-order symmetric Suzuki-Trotter decomposition \( \hat{S}_2(\Delta_r) \) (see Appendix C 3). Therefore, in the quantum power method, the first and second moments \( \mu_1 \) and \( \mu_2 \) are estimated simply by evaluating \( \text{Im} \left\langle \hat{S}_2(\frac{\Delta_r}{2}) \right\rangle \) and \( \text{Re} \left\langle \hat{S}_2(\Delta_r) \right\rangle \), i.e.,

\[ \mu_1(\Delta_r) = -\frac{2}{\Delta_r} \text{Im} \left\langle \hat{S}_2\left(\frac{\Delta_r}{2}\right) \right\rangle \quad (C25) \]

and

\[ \mu_2(\Delta_r) \approx \frac{2}{\Delta_r^2} \left[ 1 - \text{Re} \left\langle \hat{S}_2(\Delta_r) \right\rangle \right], \quad (C26) \]

respectively. Although we have to introduce an ancilla qubit (see Fig. 10), \( \mu_1 = \langle \hat{H} \rangle \) and \( \mu_2 = \langle \hat{H}^2 \rangle \) can be thus 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 \( \hat{H} \) and \( \hat{H}^2 \).

Figure 11 shows the numerical results of \( \mu_1 \) and \( \mu_2 \) evaluated from Eqs. (C25) and (C26) for the spin-1/2 Heisenberg model with two different quantum states. We also show the results obtained by employing the first-order Richardson extrapolation, i.e.,

\[ \mu_{m(1)}(\Delta_r) = \frac{\hbar^2 \mu_n(\Delta_r)}{\hbar^2} - \mu_n(\Delta_r) \quad (C27) \]

for \( n = 1 \) and \( 2 \), which expects that the systematic error scales as \( O(\Delta_r^2) \), in stead of \( O(\Delta_r) \) 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 \( \Delta_r \to 0 \). 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 [102].

FIG. 10. Quantum circuit to evaluate \( \text{Re} \left\langle \Psi | \hat{S}_2(\Delta_r) | \Psi \right\rangle \) or \( \text{Im} \left\langle \Psi | \hat{S}_2(\Delta_r) | \Psi \right\rangle \). \( \theta \) in the circuit denotes the phase gates such that \( \theta = 0 \) and \( \theta(1) = e^{\pi} |1 \rangle \). Since \( P_0 - P_1 = \text{Re} \left\langle e^{\pi} | \hat{S}_2(\Delta_r) | \Psi \right\rangle \), one can evaluate \( \text{Re} \left\langle \Psi | \hat{S}_2(\Delta_r) | \Psi \right\rangle \) if \( \theta = 0 \) and \( \text{Im} \left\langle \Psi | \hat{S}_2(\Delta_r) | \Psi \right\rangle \) if \( \theta = -\pi/2 \) from the difference of the probabilities \( P_0 \) and \( P_1 \). Here, \( P_0 \) is the probability for finding a bit \( b = 0 \) by measuring the ancilla qubit.

Figure 11 shows the numerical results of \( \mu_1 \) and \( \mu_2 \) 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. (69) and (b,d) the VQE state \( |\Psi_{\text{VQE}} \rangle \) in Eq. (77). 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 \( \Delta_r^4 \). The exact values are indicated at \( \Delta_r = 0 \) with the filled symbols.
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

\[ \mu_3(\Delta_r) = \frac{2}{\Delta_r^3} \left[ \text{Im} \left( \hat{U} \left( \frac{3\Delta_r}{2} \right) \right) - 3\text{Im} \left( \hat{U} \left( \frac{\Delta_r}{2} \right) \right) \right] \] (C28)

and

\[ \mu_4(\Delta_r) = \frac{2}{\Delta_r^4} \left[ \text{Re} \left( \hat{U} \left( 2\Delta_r \right) \right) - 4\text{Re} \left( \hat{U} \left( \Delta_r \right) \right) + 3 \right]. \] (C29)

respectively. To implement these on quantum computers, the Trotter decomposition \( \hat{S}^p_m(\Delta_r) \) with \( m = 2 \) (also see Fig. 9), which is still affordable.

5. Imaginary-time evolution

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

\[ |\Psi(\tau)\rangle = \frac{e^{-\tau \hat{H}/2} |\Psi\rangle}{\sqrt{\langle \Psi | e^{-\tau \hat{H}} |\Psi\rangle}} \] (C30)

for \( \tau \) 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

\[ E(\tau) = \langle \hat{H} \rangle_\tau = \frac{\langle \Psi | e^{-\tau \hat{H}} \hat{H} e^{-\tau \hat{H}/2} |\Psi\rangle}{\langle \Psi | e^{-\tau \hat{H}} |\Psi\rangle} = \langle \hat{H} e^{-\tau \hat{H}} \rangle_{e^{-\tau \hat{H}}} . \] (C31)

Observing that \( E(\tau) = -\frac{d}{d\tau} \ln(e^{-\tau \hat{H}}) = -\frac{d}{d\tau} \sum_{n=0}^{\infty} (-\tau)^n n! \kappa_n \), the CMX of the energy is given as \([99]\)

\[ E(\tau) = \sum_{n=0}^{\infty} \frac{(-\tau)^n}{n!} \kappa_{n+1} . \] (C32)

Figure 12 shows the exact \( E(\tau) \) and the CMX of the energy truncated at the \( n_{\text{max}} \)th cumulant

\[ E_{n_{\text{max}}}(\tau) = \sum_{n=0}^{n_{\text{max}}-1} \frac{(-\tau)^n}{n!} \kappa_{n+1} \] (C33)

for the spin-1/2 Heisenberg model, where the VQE state \( |\Psi_{\text{VQE}}\rangle \) in Eq. (77) is selected for the quantum state \( |\Psi\rangle \) in Eq. (C30). The energy \( E(\tau) \) with the exact ITE decreases monotonically in \( \tau \), because the first derivative of \( E(\tau) \) is minus of the energy fluctuation \([99]\)

\[ \frac{dE(\tau)}{d\tau} = -\left( \langle \hat{H}^2 \rangle_\tau - \langle \hat{H} \rangle_\tau^2 \right) \leq 0, \] (C34)

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 \( \tau \) diverges to \( -\infty \) for even \( n_{\text{max}} \geq 2 \) or to \( +\infty \) for odd \( n_{\text{max}} \geq 3 \). Note that \( E_2(\tau) = k_1 - k_2 \tau \) 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 \( n_{\text{max}} \) in the cumulants required is rather slower, as compared with the Krylov-subspace diagonalization with either \( M_B = 1 \) or \( M_B = 9 \) discussed in Sec. V B. This is not quite surprising because the form of \( E(\tau) \) in Eq. (C33) is an expansion around \( \tau = 0 \), which is analogous to the high-temperature expansion.

Appendix D: Lanczos method

In this appendix, we briefly outline the Lanczos method with an emphasis on its aspect as a moment method \([103, 104]\), 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

\[ \mathcal{H} |q_i\rangle = \beta_i |q_i\rangle + \alpha_i |q_{i+1}\rangle, \] (D1)

with \( \alpha_i = \langle q_i | \mathcal{H} |q_i\rangle, \beta_i = \langle q_i | \mathcal{H} |q_{i+1}\rangle, \beta_0 \equiv 0, \) and \( \beta_0 \equiv 0 \). After obtaining \( |q_i\rangle \)\(^{n_{\text{max}}}_{\text{max}} \), the Hamiltonian \( \mathcal{H} \) can be represented as a tridiagonal matrix \( |T_n| \) as

\[ T_n = \begin{bmatrix} \alpha_1 & \beta_1 & 0 & \cdots & 0 \\ \beta_1 & \alpha_2 & \beta_2 & \cdots & 0 \\ 0 & \beta_2 & \alpha_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & \beta_{n-1} \end{bmatrix}. \] (D2)

The matrix elements \( \{\alpha_i\} \) and \( \{\beta_i\} \) can also be constructed recursively using the Hamiltonian moments \([103, 104]\). Following Ref. [103], \( \{\alpha_i\} \) and \( \{\beta_i\} \) are given in terms of \( \{\mu_n\} \) recursively as

\[ \alpha_i = \left( \frac{L_i}{L_{i-1}} \right) \left( M_{i-2} M_{i-3} \right)^{-1} + \left( M_{i-1} M_{i-2} \right) \left( L_{i-1} L_{i-2} \right)^{-1} \] (D3)

and

\[ \beta_i^2 = \left( \frac{L_i}{L_{i-1}} \right) \left( L_{i-1} L_{i-2} \right)^{-1}, \] (D4)
where $L_n \equiv \det L_n$ and $M_n \equiv \det M_n$ are determinants of $(n+1) \times (n+1)$ Hankel matrices defined respectively as $[L_n]_{ij} = \mu_{i+j-2}$ and $[M_n]_{ij} = \mu_{i+j-1}$, or more explicitly

$$
L_n = 
\begin{bmatrix}
\mu_0 & \mu_1 & \mu_2 & \cdots & \mu_{n-1} & \mu_n \\
\mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{n} & \mu_{n+1} \\
\mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{n+1} & \mu_{n+2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\mu_{n-1} & \mu_n & \mu_{n+1} & \cdots & \mu_{2n-2} & \mu_{2n-1} \\
\mu_n & \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n-1} & \mu_{2n}
\end{bmatrix}
$$

(D5)

and

$$
M_n = 
\begin{bmatrix}
\mu_1 & \mu_2 & \mu_3 & \cdots & \mu_n & \mu_{n+1} \\
\mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{n+1} & \mu_{n+2} \\
\mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{n+2} & \mu_{n+3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\mu_n & \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n-1} & \mu_{2n} \\
\mu_{n+1} & \mu_{n+2} & \mu_{n+3} & \cdots & \mu_{2n-2} & \mu_{2n}
\end{bmatrix}
$$

(D6)

The Hankel matrices $L_{n-1}$ and $M_{n-1}$ are identical respectively to $S$ in Eq. (50) and $H$ in Eq. (49) if $M_0 = 1$. It is noticed in Eqs. (D3) and (D4) that the Lanczos matrix elements $\alpha_i$ and $\beta_i$ 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 $L_n$ and $M_n$ allows us to evaluate the ratios of the determinants appearing in Eqs. (D3) and (D4) recursively, as described in Appendix D2.

It is instructive to give the explicit forms of the first few matrix elements of $T_n$. The first three matrix elements required for constructing the $2 \times 2$ matrix $T_2$ are given by

$$
\alpha_1 = \langle \hat{H} \rangle, \quad \beta_1 = \sqrt{\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2}, \quad \alpha_2 = \frac{\langle \hat{H}^3 \rangle - 2\langle \hat{H} \rangle \langle \hat{H}^2 \rangle + \langle \hat{H}^4 \rangle}{\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2}
$$

(D7) (D8) (D9)

where $(\cdots) = \langle q_1 | \cdots | q_1 \rangle$. Therefore, $\alpha_1$ and $\beta_1^2$ 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. (D3) and (D4). Let us first review the determinant and the matrix-inversion formulas for general matrices. Let $A_n$ be an $(n \times n)$ matrix, $b$ be an $(n \times 1)$ matrix, $e$ be an $(n \times 1)$ matrix, and $b$ be a $(1 \times 1)$ matrix, and let us consider an $(n+1) \times (n+1)$ matrix $A_{n+1}$ of the form

$$
A_{n+1} = \begin{bmatrix} A_n & b \\ e^T & d \end{bmatrix}.
$$

(D10)

If we define

$$
r = d - e^T A_n^{-1} b,
$$

(D11)
the determinant of $A_{n+1}$ is given by
\[
\det A_{n+1} = \det \begin{bmatrix} A_n & b \\ c^T & d \end{bmatrix} = r \det A_n,
\] (D12)
and the inverse $A_{n+1}^{-1}$ is given by
\[
A_{n+1}^{-1} = \begin{bmatrix} A_n & b \\ c^T & d \end{bmatrix}^{-1} = \begin{bmatrix} 1/r & -A_n^{-1}b/r \\ -c^T A_n^{-1} \end{bmatrix}.
\] (D13)

Now we apply the above formulas to recursively evaluate the ratios of the determinants of $L_n$ and $L_{n-1}$. Due to its particular structure, $L_n$ can be expressed in terms of $L_{n-1}$ as
\[
L_n = \begin{bmatrix} L_{n-1} & m_n \\ m_n^T & \mu_{2n} \end{bmatrix}
\] (D14)
with the following $n$-dimensional vector:
\[
m_n^T = (\mu_n, \mu_{n+1}, \ldots, \mu_{2n-1}).
\] (D15)
From the formula in Eq. (D12), the ratio of the determinants is given by
\[
\frac{\det L_n}{\det L_{n-1}} = r_n
\] (D16)
with
\[
r_n = \mu_{2n} - m_n^T L_{n-1}^{-1} m_n,
\] (D17)
which involves the inverse $L_{n-1}^{-1}$ whose dimension is less than that of $L_n^{-1}$ by 1.

The inverse matrix $L_n^{-1}$ can be calculated using Eq. (D13). Starting with
\[
L_0^{-1} = \mu_0^{-1},
\] (D18)
$L_n^{-1}$ for $n \geq 1$ can be constructed from $L_{n-1}^{-1}$ and $m_n$ recursively as
\[
L_n^{-1} = \begin{bmatrix} L_{n-1}^{-1} m_n \\ m_n^T \mu_{2n} \end{bmatrix}^{-1} = \begin{bmatrix} 1/r_n & -L_{n-1}^{-1} m_n/r_n \\ (L_{n-1}^{-1} m_n)^T/r_n & -L_{n-1}^{-1} m_n/r_n \end{bmatrix},
\] (D19)
where $(L_n^{-1})^T = L_n^{-1}$ is used. Thus, starting with the known $L_0^{-1}$ and using Eqs. (D17) and (D19), one can obtain $\{r_n\}$ recursively as $L_0^{-1} \rightarrow r_1 \rightarrow L_1^{-1} \rightarrow r_2 \rightarrow L_2^{-1} \rightarrow r_3 \rightarrow \cdots$.

It should be noted that Eq. (D17) involves a matrix-vector multiplication and, in addition, Eq. (D19) involves a rank-1 update. Therefore, the complexity for computing the ratio of determinants in Eq. (D16) 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 $[L_n]_{ij} = \mu_{i+j-2}$ while $[M_n]_{ij} = \mu_{i+j-1}$, the similar recursive formula for $M_n$ can be readily derived simply by replacing the indexes for the moments in the above as $\{\mu_{i+j-2}\}_{i+j-2}$.

[1] J. P. F. LeBlanc, Andrey E. Antipov, Federico Becca, Ireneusz W. Bulik, Garnet Kin-Lic Chan, Chia-Min Chung, Youjin Deng, Michel Ferrari, Thomas M. Henderson, Carlos A. Jiménez-Hoyos, E. Koziuk, Xuan-Wen Liu, Andrew J. Millis, N. V. Prokof’ev, Mingpu Qin, Gustavo E. Scuseria, Hao Shi, B. V. Svitulov, Luca F. Tocchio, I. S. Tupitsyn, Steven R. White, Shiwei Zhang, Bo-Xiao Zheng, Zhenyue Zhu, and Emanuel Gull (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] Mario Motta, David M. Ceperley, Garnet Kin-Lic Chan, John A. Gomez, Emanuel Gull, Sheng Guo, Carlos A. Jiménez-Hoyos, Tran Nguyen Lan, Jia Li, Fengjie Ma, Andrew J. Millis, Nikolay V. Prokof’ev, Ushnish Ray, Enrico Ronca, Hao Shi, Sandro Sorella, Edwin M. Stoudenmire, Steven R. White, and Shiwei Zhang, “Ground-state properties of the hydrogen chain: insulator-to-metal transition, dimerization, and magnetic phases,” (2019), arXiv:1911.06178 [cond-mat.str-el].

[3] Mario Motta, Claudio Genovese, Fengjie Ma, Zhi-Hao Cui, Randy Sawaya, Garnet Kin-Lic Chan, Natalia Chepiga, Phillip Helms, Carlos Jimenez-Hoyos, Andrew J. Millis, Ushnish Ray, Enrico Ronca, Hao Shi, Sandro Sorella, Edwin M. Stoudenmire, Steven R. White, and Shiwei Zhang, “Ground-state properties of the hydrogen chain: insulator-to-metal transition, dimerization, and magnetic phases,” (2019), arXiv:1911.06178 [cond-mat.str-el].

[4] J. Liesen and Z. Strakoš, Krylov Subspace Methods (Oxford University Press, Oxford, 2013) Chap. 3.

[5] Elbio Dagotto, “Correlated electrons in high-temperature superconductors,” Rev. Mod. Phys. 66, 763–840 (1994).

[6] J. Jaklič and P. Prelošek, “Lanczos method for the calculation of finite-temperature quantities in correlated systems,” Phys. Rev. B 49, 5065–5068 (1994).

[7] J. Jaklič and P. Prelošek, “Finite-temperature properties of doped antiferromagnets,” Adv. Phys. 49, 1 (2000).

[8] Alexander Weiße and Holger Fehske, “Exact diagonalization techniques,” in Computational Many-Particle Physics, Vol. 739, edited by H. Fehske, R. Schneider, and A. Weiße (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008) pp. 529–544.

[9] P. Prelošek and J. Bonča, “Ground state and finite temperature lanczos methods,” in Strongly Correlated Systems: Numerical Methods, edited by Adolfo Avella and Ferdinando Mancini (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013) pp. 1–30.

[10] Erik Koch, “Exact diagonalization and lanczos method,”
Many-Body Methods for Real Materials, Schriften des Forschungszentrums Jülich Modeling and Simulation, Vol. 9, edited by Eva Pavarini and Schwei Zhang (Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag, Jülich, 2019) Chap. 7.

A. Weiße and H. Fehske, Computational Many-Particle Physics, edited by H. Fehske, R. Schneider, and A. Weiße, Lect. Notes Phys. 739 (Springer, Berlin Heidelberg, 2008) Chap. 19, pp. 545–577.

H. Tal-Ezer and R. Kosloff, “An accurate and efficient scheme for propagating the time dependent schrödinger equation,” The Journal of Chemical Physics 81, 3967–3971 (1984).

Tae Jun Park and J. C. Light, “Unitary quantum time evolution by iterative lanczos reduction,” The Journal of Chemical Physics 85, 5870–5876 (1986).

Amrendra Vijay and Horia Metiu, “A polynomial expansion of the quantum propagator, the greens function, and the spectral density operator,” The Journal of Chemical Physics 116, 60–68 (2002).

Toshiaki Iitaka and Toshikazu 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).

N. Mohankumar and Scott M. Auerbach, “On time-step bounds in unitary quantum evolution using the lanczos method,” Computer Physics Communications 175, 473 – 481 (2006).

Sandro Sorella, “Generalized lanczos algorithm for variational quantum monte carlo,” Phys. Rev. B 64, 024512 (2001).

Richard P. Feynman, “Simulating physics with computers,” International Journal of Theoretical Physics 21, 467–488 (1982).

Dave Wecker, Matthew B. Hastings, Nathan Wiebe, Bryan K. Clark, Chetan Nayak, and Matthias Troyer, “Solving strongly correlated electron models on a quantum computer,” Phys. Rev. A 92, 062318 (2015).

Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan, “Quantum computational chemistry,” Rev. Mod. Phys. 92, 015003 (2020).

Bela Bauer, Sergey Bravyi, Mario Motta, and Garnet Kin-Lic Chan, “Quantum algorithms for quantum chemistry and quantum materials science,” (2020), arXiv:2003.03685 [quant-ph].

Y. Nakamura, Yu A. Pashkin, and J. S. Tsai, “Coherent control of macroscopic states in a single-cooper-pair box,” Nature 398, 786–788 (1999).

Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn, “Linear optical quantum computing with photonic qubits,” Rev. Mod. Phys. 79, 135–174 (2007).

T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien, “Quantum computers,” Nature 464, 45–53 (2010).

Ze-Liang Xiang, Sahel Ashhab, J. Q. You, and Franco Nori, “Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems,” Rev. Mod. Phys. 85, 623–653 (2013).

R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O’Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis, “Superconducting quantum circuits at the surface code threshold for fault tolerance,” Nature 508, 500 EP – (2014).

Jerry M. Chow, Jay M. Gambetta, Easwar Magesan, David W. Abraham, Andrew W. Cross, B. R. Johnson, Nicholas A. Masluk, Colm A. Ryan, John A. Smolin, Srikanth J. Srinivasan, and M. Steffen, “Implementing a strand of a scalable fault-tolerant quantum computing fabric,” Nature Communications 5, 4015 (2014).

J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Yu Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O’Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and John M. Martinis, “State preservation by repetitive error detection in a superconducting quantum circuit,” Nature 519, 66 EP – (2015).

D. Risté, 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,” Nature Communications 6, 6983 (2015).

Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtayne, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerrin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hofmann, Trent Huang, Travis S. Humble, Sergey V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Kysh, Alexander Korotkov, Fedor Kostriitsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Niman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sang, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis, “Quantum supremacy using a programmable superconducting processor,” Nature 574, 505–510 (2019).

Warit Asavanant, Yu Shiozawa, Shota Yokoyama, Baramee Charoensombutamon, Hiroki Emura, Rafael N. Alexander, Shuntaro Takeda, Jun-ichi Yoshikawa, Nicolas C. Menicucci, Hidehiro Yonezawa, and Akira Furusawa, “Generation of time-domain-multiplexed two-dimensional cluster state,” Science 366, 373–376 (2019).

John Preskill, “Quantum Computing in the NISQ era and beyond,” Quantum 2, 79 (2018).

Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” Nature Communications 5, 4213 (2014).

Dave Wecker, Matthew B. Hastings, and Matthias Troyer, “Progress towards practical quantum variational algorithms,” Phys. Rev. A 92, 042303 (2015).

P. J. J. O’Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, “Scalable quantum simulation of molecular energies,” Phys. Rev. X 6, 031007 (2016).
[36] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik, “The theory of variational hybrid quantum-classical algorithms,” New Journal of Physics 18, 023023 (2016).

[37] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta, “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets,” Nature 549, 242 (2017).

[38] Jun Li, Xiaodong Yang, Xinhua Peng, and Chang-Pu Sun, “Hybrid quantum-classical approach to quantum optimal control,” Phys. Rev. Lett. 118, 150503 (2017).

[39] Guglielmo Mazzola, Pauline J. Olitrault, Panagiotis Kl. Bartsousos, and Ivano Tavernelli, “Nonunitary operations for ground-state calculations in near-term quantum computers,” Phys. Rev. Lett. 123, 130501 (2019).

[40] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Sergio Boixo, Michael Broughton, Bob B. Buckley, David A. Buell, Brian Burkett, Nicholas Bushnell, Yu Chen, Zijun Chen, Benjamin Chiaro, Roberto Collins, William Courtney, Sean Demura, Andrew Dunsworth, Edward Farhi, Austin Fowler, Brooks Foxen, Craig Gidney, Marissa Giustina, Rob Graff, Steve Habegger, Matthew P. Harrigan, Alan Ho, Sabrina Hong, Trent Huang, William J. Huggins, Lev Ioffe, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Cody Jones, Divr Kafri, Kostyantyn Kechedzhi, Julian Kelly, Seon Kim, Paul V. Klimov, Alexander Korostov, Fedor Kostritsa, David Landhuis, Pavel Laptev, Mike Lindmark, Erik Lucero, Orinon Martin, John M. Martinis, Jarrod R. McClean, Matt McEwen, Anthony Mengran, Xiao Mi, Masoud Mohseni, Wojciech Mruczkiewicz, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Hartmut Neven, Murphy Yuezheng Niu, Thomas E. O’Brien, Eric Osby, Andre Petukhov, Harald Putterman, Chris Quintana, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. SATzinger, Vadim Smelyanskiy, Doug Strain, Kevin J. Sung, Marco Szalay, Tyler Y. Takeshita, Amit Vainsencher, Theodore White, Nathan Wiebe, Z. Jamie Yao, Ping Yeh, and Adam Zalcman, “Hartree-fock on a superconducting qubit quantum computer,” (2020), arXiv:2004.04174 [quant-ph].

[41] Jin-Guo Liu, Yi-Hong Zhang, Yuan Wei, and Lei Wang, “Variational quantum eigensolver with fewer qubits,” Phys. Rev. Research 1, 023025 (2019).

[42] Michael Foss-Feig, David Hayes, Joan M. Dreiling, Caroline Figgatt, John P. Gaebler, Steven A. Moses, Juan M. Pino, and Andrew C. Potter, “Holographic quantum algorithms for simulating correlated spin systems,” (2020), arXiv:2005.03023 [quant-ph].

[43] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O’Rourke, Erika Ye, Austin J. Mimchik, Fernando G. S. L. Brandão, and Garnet Kin-Lic Chan, “Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution,” Nature Physics (2019). 10.1038/s41567-019-0704-4.

[44] Kübra Yeter-Aydeniz, Raphael C. Pooser, and George Siopsis, “Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and lanczos algorithms,” npj Quantum Information 6, 63 (2020).

[45] Hirofumi Nishi, Taichi Kosugi, and Yuichiro Matsushita, “Implementation of quantum imaginary-time evolution method on nisq devices: Nonlocal approximation,” (2020), arXiv:2005.12715 [quant-ph].

[46] Niladri Gomes, Feng Zhang, Noah F. Berthusen, Cai-Zhuang Wang, Kai-Ming Ho, Peter P. Orth, and Yongxin Yao, “Efficient step-merged quantum imaginary time evolution algorithm for quantum chemistry,” (2020), arXiv:2006.15371 [physics.comp-ph].

[47] Nicholas H. Stair, Renke Huang, and Francesco A. Evangelista, “A multireference quantum krylov algorithm for strongly correlated electrons,” Journal of Chemical Theory and Computation 16, 2236–2245 (2020).

[48] Robert M. Parrish and Peter L. McMahon, “Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation,” (2019), arXiv:1909.08925 [quant-ph].

[49] Oleksandr Kyriienko, “Quantum inverse iteration algorithm for programmable quantum simulators,” npj Quantum Information 6 (2020), 10.1038/s41534-019-0239-7.

[50] Andrew M. Childs, Robin Kothari, and Rolando D. Somma, “Quantum algorithm for systems of linear equations with exponentially improved dependence on precision,” SIAM Journal on Computing 46, 1920–1950 (2017).

[51] Ting Liu, Jin-Guo Liu, Lei Wang, and Heng Fan, “Probabilistic nonunitary gate in imaginary time evolution,” (2020), arXiv:2006.09726 [quant-ph].

[52] Jarrod R. McClean, Mollie E. Kimchi-Schwartz, Jonathan Carter, and Wibe A. de Jong, “Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states,” Phys. Rev. A 95, 042306 (2017).

[53] J. I. Collies, 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).

[54] Robert M. Parrish, Edward G. Hohenstein, Peter L. McManus, and Todd J. Martinez, “Quantum computation of electronic transitions using a variational quantum eigensolver,” Phys. Rev. Lett. 122, 230401 (2019).

[55] Ken M. Nakanishi, Kosuke Mitarai, and Keisuke Fujii, “Subspace-search variational quantum eigensolver for excited states,” Phys. Rev. Research 1, 033062 (2019).

[56] Kentaro Heya, Ken M Nakanishi, Kosuke Mitarai, and Keisuke Fujii, “Subspace variational quantum simulator,” (2019), arXiv:1904.08566 [quant-ph].

[57] William J. Huggins, Joonho Lee, Unpil Baek, Bryan O’Gorman, and K. Birgitta Whaley, “A non-orthogonal variational quantum eigensolver,” (2019), arXiv:1909.09114 [quant-ph].

[58] Masuo Suzuki, “Fractal decomposition of exponential operators with applications to many-body theories and monte carlo simulations,” Physics Letters A 146, 319 – 323 (1990).

[59] Haruo Yoshida, “Construction of higher order symplectic integrators,” Physics Letters A 150, 262–268 (1990).

[60] Masuo Suzuki, “General theory of fractal path integrals with applications to many-body theories and statistical physics,” Journal of Mathematical Physics 32, 400–407 (1991).

[61] Daniel Loss and David P. DiVincenzo, “Quantum computation with quantum dots,” Phys. Rev. A 57, 120–126 (1998).

[62] D. P. DiVincenzo, D. Bacon, J. Kempe, G. Burkard, and K. B. Whaley, “Universal quantum computation with the exchange interaction,” Nature 408, 339–342 (2000).

[63] 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).

[64] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost, “Quantum principal component analysis,” Nature Physics 10, 631–
633 (2014).

[65] Hoi-Kwan Lau and Martin B. Plenio, “Universal quantum computing with arbitrary continuous-variable encoding,” Phys. Rev. Lett. 117, 100501 (2016).

[66] Heng Fan, Vvani Roychowdhury, and Thomas Szkopec, “Optimal two-qubit quantum circuits using exchange interactions,” Phys. Rev. A 72, 052323 (2005).

[67] S. Balakrishnan and R. Sankaranarayanan, “Entangling characteristic of swap and controlled unitary gates,” Phys. Rev. A 78, 052305 (2008).

[68] Bryan T. Gard, Linghua Zhu, George S. Barron, Nicholas J. Mayhall, Sophia E. Economou, and Edwin Barnes, “Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm,” npj Quantum Information 6 (2020).

[69] G. Vidal and C. M. Dawson, “Universal quantum circuit for two-qubit transformations with three controlled-not gates,” Phys. Rev. A 69, 010301 (2004).

[70] A. Chiesa, F. Tacchino, M. Grossi, P. Santini, I. Tavvernelli, D. Gerace, and S. Carretta, “Quantum hardware simulating four-dimensional inelastic neutron scattering,” Nature Physics 15, 455 (2019).

[71] Our decomposition of the eSWAP gate in Ref. [87] was not optimal with respect to the number of CNOT gates.

[72] Naomichi Hatano and Masuo Suzuki, “Finding exponential product formulas of higher orders,” Lecture Notes in Physics , 37–68 (2005).

[73] Masuo Suzuki, “General nonsymmetric higher-order decomposition of exponential operators and symplectic integrators,” Journal of the Physical Society of Japan 61, 3015–3019 (1992), https://doi.org/10.1143/JPSJ.61.3015.

[74] Andrew M. Childs and Nathan Wiebe, Quantum Information and Computation 12 (2012).

[75] Taichi Kosugi and Yuichiro Matsushita, “Construction of green’s functions on a quantum computer: Quasiparticle spectra of molecules,” Phys. Rev. A 101, 012330 (2020).

[76] Taichi Kosugi and Yu-ichiro Matsushita, “Linear-response functions of molecules on a quantum computer: Charge and spin responses and optical absorption,” Phys. Rev. Research 2, 033043 (2020).

[77] The expectation value of $\hat{H}_{\text{eff}}(\Delta)$ with an odd power $n$ might be evaluated with an Hadamard-test like circuit by introducing an additional ancilla qubit.

[78] F. Chatelin, Eigenvalues of Matrices (SIAM, Philadelphia, 2012).

[79] Jonathan Romero, Ryan Babbush, Jarrod R McClean, Corneliu Hempel, Peter J Love, and Alán Aspuru-Guzik, “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz,” Quantum Science and Technology 4, 014008 (2018).

[80] Pierre-Luc Dallaire-Demers, Jonathan Romero, Libor Veis, Sukin Sim, and Alán Aspuru-Guzik, “Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer,” Quantum Science and Technology 4, 045005 (2019).

[81] Sam McArdle, Tyson Jones, Suguru Endo, Ying Li, Simon C. Benjamin, and Xiao Yuan, “Variational ansatz-based quantum simulation of imaginary time evolution,” npj Quantum Information 5, 75 (2019).

[82] David A. Drabold and Otto F. Sankey, “Maximum entropy approach for linear scaling in the electronic structure problem,” Phys. Rev. Lett. 70, 3631–3634 (1993).

[83] Anthony Hams and Hans De Raedt, “Fast algorithm for finding the eigenvalue distribution of very large matrices,” Phys. Rev. E 62, 4365–4377 (2000).

[84] Toshiaki Iitaka and Toshikazu Ebisuzaki, “Random phase vector for calculating the trace of a large matrix,” Phys. Rev. E 69, 057701 (2004).

[85] Alexander Weiße, Gerhard Wellein, Andreas Alvermann, and Holger Fehske, “The kernel polynomial method,” Rev. Mod. Phys. 78, 275–306 (2006).

[86] Kazuhiro Seki and Seiji Yunoki, “Thermodynamic properties of an $s = \frac{1}{2}$ ring-exchange model on the triangular lattice,” Phys. Rev. B 101, 235115 (2020).

[87] Kazuhiro Seki, Tomonori Shirakawa, and Seiji Yunoki, “Symmetry-adapted variational quantum eigensolver,” Phys. Rev. A 101, 052340 (2020).

[88] Elliott Lieb and Daniel Mattis, “Ordering energy levels of interacting spin systems,” Journal of Mathematical Physics 3, 749–751 (1962), https://doi.org/10.1063/1.1724276.

[89] J. Oitmaa, C. Hamer, and W. Zheng, Series Expansion Methods for Strongly Interacting Lattice Models (Cambridge, New York, 2006).

[90] H. F. Trotter, “On the product of semi-groups of operators,” Proc. Am. Math. Soc. 10, 545 (1959).

[91] Masuo Suzuki, “Relationship between $d$-Dimensional Quantal Spin Systems and $(d+1)$-Dimensional Ising Systems: Equivalence, Critical Exponents and Systematic Approximants of the Partition Function and Spin Correlations,” Progress of Theoretical Physics 56, 1454–1469 (1976).

[92] Masuo Suzuki, “Generalized trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems,” Comm. Math. Phys. 51, 183–190 (1976).

[93] L.P. Omelyan, I.M. Myrglod, and R. Folk, “Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations,” Computer Physics Communications 151, 272 – 314 (2003).

[94] Anargyros Papageorgiou and Chi Zhang, “On the efficiency of quantum algorithms for hamiltonian simulation,” Quantum Information Processing 11, 541–561 (2012).

[95] Markus Heyl, Philipp Hauke, and Peter Zoller, “Quantum localization bounds trotter errors in digital quantum simulation,” Science Advances 5 (2019), 10.1126/sciadv.aau8342.

[96] Andrew M. Childs, Yuan Su, Ming C. Tran, Nathan Wiebe, and Shuchen Zhu, “A theory of trotter error,” (2019), arXiv:1912.08854 [quant-ph].

[97] F. Becca and S. Sorella, Quantum Monte Carlo Approaches for Correlated Systems (Cambridge University Press, Cambridge, 2017).

[98] Rolando D. Somma, “Quantum eigenvalue estimation via time series analysis,” (2019), arXiv:1907.11748 [quant-ph].

[99] D. Horn and M. Weinstein, “The $t$ expansion: A nonperturbative analytic tool for hamiltonian systems,” Phys. Rev. D 30, 1256–1270 (1984).

[100] R. Kubo, “Generalized Cumulant Expansion Method,” J. Phys. Soc. Jpn. 17, 1100 (1962).

[101] M. Fanzizza, 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).

[102] Dan-Bo Zhang, Zhan-Hao Yuan, and Tao Yin, “Variational quantum eigensolvers by variance minimization,” (2020), arXiv:2006.15781 [quant-ph].

[103] N. S. Witte and Lloyd C. L. Hollenberg, “Plaquette expansion and controlled unitary gates,” Phys. Rev. A 71, 052323 (2005).
for nuclear response surfaces,” Phys. Rev. C 72, 065501 (2005).