Real-Time Evolution for Ultracompact Hamiltonian Eigenstates on Quantum Hardware

Katherine Klymko,[1,2,*] Carlos Mejuto-Zaera,[1,3,†] Stephen J. Cotton,[4,5] Filip Wudarski,[4,6]
Miroslav Urbanek,[1] Diptarka Hait,[3,7] Martin Head-Gordon,[3,7] K. Birgitta Whaley,[3]
Jonathan Moussa,[8] Nathan Wiebe,[9] Wibe A. de Jong,[1,‡] and Norm M. Tubman[4,§]

[1] *Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

[2] *NERSC, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

[3] *Department of Chemistry, University of California, Berkeley, California 94720, USA*

[4] *Quantum Artificial Intelligence Laboratory (QuAIL), Exploration Technology Directorate, NASA Ames Research Center, Moffett Field, California 94035, USA*

[5] *KBR, 601 Jefferson St., Houston, Texas 77002, USA*

[6] *USRA Research Institute for Advanced Computer Science, Mountain View, California 94043, USA*

[7] *Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

[8] *Molecular Sciences Software Institute, Blacksburg, Virginia 24060, USA*

[9] *Department of Computer Science, University of Toronto, Canada*

In this work we present a detailed analysis of variational quantum phase estimation (VQPE), a method based on real-time evolution for ground- and excited-state estimation on near-term hardware. We derive the theoretical ground on which the approach stands, and demonstrate that it provides one of the most compact variational expansions to date for solving strongly correlated Hamiltonians, when starting from an appropriate reference state. At the center of VQPE lies a set of equations, with a simple geometrical interpretation, which provides conditions for the time evolution grid in order to decouple eigenstates out of the set of time-evolved expansion states, and connects the method to the classical filter-diagonalization algorithm. Furthermore, we introduce what we call the unitary formulation of VQPE, in which the number of matrix elements that need to be measured scales linearly with the number of expansion states, and we provide an analysis of the effects of noise that substantially improves previous considerations. The unitary formulation allows for a direct comparison to iterative phase estimation. Our results mark VQPE as both a natural and highly efficient quantum algorithm for ground- and excited-state calculations of general many-body systems. We demonstrate a hardware implementation of VQPE for the transverse field Ising model. Furthermore, we illustrate its power on a paradigmatic example of strong correlation (Cr_2 in the def2-SVP basis set), and show that it is possible to reach chemical accuracy with as few as approximately 50 time steps.

I. INTRODUCTION

In fulfilling the promise of quantum computation [1–3] and enabling the exact solution of many-body

*kklymko@lbl.gov

†carlos_mejutozaera@berkeley.edu

‡wadejong@lbl.gov

§norman.m.tubman@nasa.gov

quantum systems, numerous algorithms of different resource requirements have been proposed [4–7], which require quantum and classical resources of different complexity. Many of these algorithms are focused on efficient eigenvalue extraction, important for solving problems in chemistry [8], physics [9], and materials science [10] and limited classically by the exponential scaling of Hilbert space with system size. Though immense progress has been made in the development of quantum algorithms for eigenvalue estimation, the resource requirements remain prohibitively high with regards to noisy intermediate-scale quantum (NISQ) hardware [10,11].

For example, quantum phase estimation (QPE) [12–14] is considered an algorithm that will need

considerable quantum resources to run, but will ultimately be a highly accurate approach to quantum simulation. Adiabatic state preparation [15,16] allows the ground state of a particular Hamiltonian to be reached by preparing an initial ground state of a simpler system and slowly changing the Hamiltonian to the desired system, requiring long coherence times and low gate errors. Particularly in the current era of NISQ quantum computers, the variational quantum eigensolver (VQE) framework [17], and its nonorthogonal variant NOVQE [18], are promising approaches for the exact solution of many-body quantum systems. However, the common formulation of this family of methods relies on the solution of a highly complex variational optimization problem on classical computers, which remains an open challenge [19,20].

The methods described above are generally used to solve the time-independent Schrödinger equation to determine the Hamiltonian eigenvalues and eigenstates. However, time evolution is a more natural operation on a quantum computer and thus simulation of the time-dependent Schrödinger equation is a more ideal framework to implement. Given the intrinsically quantum-mechanical relation between the time and energy domains [21], a different family of quantum algorithms focuses on using a time-dependent perspective to solve time-independent problems [22,23]. These methods propose a linear wave function ansatz expanded in time-evolved states and solve the thereby defined generalized eigenvalue equation classically, while the Hamiltonian and overlap matrix elements are measured on quantum hardware. Exploiting this, quantum computers hold unique potential to outperform their classical counterparts with algorithms based on real-time evolution. In this work we focus on that advantage and use real-time evolution to generate a basis of states to extract Hamiltonian eigenvalues.

Using real-time evolution to generate a basis of states to solve a Hamiltonian is not a new idea [24–26], but it is not widely used in classical simulation due to the computational limitations of simulating real-time evolution. Approximate imaginary-time evolution [27,28] and Krylov diagonalization methods [29] are far more widely used in classical simulation, and the intuition behind such approaches is simple to understand: each new state generated in these approaches has a larger overlap with the true ground state. This of course is not how real-time evolution works, as the expectation value of the energy remains constant, and one never gets closer to the ground state during the evolution. However, the states that can be generated through real-time evolution in fact do provide a basis from which one can extract ground and excited states. This is not only highly efficient, but in some cases may be faster than other quantum methods that use time evolution, such as QPE. Thus, the main goal in this work is to develop an approach for computing ground and excited states, using states generated by real-time evolution, that is, as fast and efficient as possible. With this in mind, we analyze the theoretical underpinnings of a class of algorithms we term variational quantum phase estimation (VQPE) that is based on real-time expansion methods [22,23]. We use the term VQPE because of its relationship to both QPE and VQE, as we detail below (see Secs. II A and II G).

Our VQPE algorithm and analysis go beyond previous proposals in the following ways. We reduce the number of quantum measurements needed to be *linear* instead of quadratic in the number of expansion states. We introduce the phase cancelation conditions, providing the underlying theory for why this approach works and deriving a direct link between time steps and band gaps in the spectrum. We also analyze the effects of noise on our convergence properties, providing significantly improved intuition for ill-conditioning of VQPE methods. We also provide analytical and numerical comparisons to phase estimation, providing evidence that VQPE could be a useful algorithm beyond the NISQ era of quantum computing. Furthermore, we demonstrate the method classically for several weakly and moderately correlated molecules as well as a strongly correlated transition metal dimer, Cr_2. We show that, for all the systems, regardless of the level of electronic correlation, less than 50 real-time-evolved expansion states are needed to reach agreement within chemical accuracy for ground-state energies obtained with state-of-the-art classical methods requiring $O(10^6)$ variational parameters [30]. Additionally, we implement the algorithm on quantum hardware for the transverse field Ising model, providing the first hardware demonstration of a VQPE method. Since real-time evolution is natural to implement on quantum hardware, this approach holds immense promise for NISQ implementation.

The paper is structured as follows. In Sec. II, we analyze the theoretical structure of VQPE. This starts with a brief overview and intuition of existing methods corresponding to the VQPE family (Secs. II A, II B). Subsequently, we present a novel examination of the theoretical underpinnings of the method based on the phase cancelation picture in Sec. II C and we propose a procedure for choosing optimal time step sizes. With the practical implementation of VQPE on NISQ devices in mind, we present the details of our unitary formulation of VQPE in Secs. II D and II E. We conclude the theory section with a careful analysis of the effects of noise (Sec. II F), a comparison between VQPE and QPE (Sec. II G), and a discussion of the inclusion of other time-evolved states into VQPE (Sec. II H). Section III provides details of the systems studied as well as the classical and quantum simulation methods. Section IV summarizes and discusses the results of the simulations. Concluding remarks are found in Sec. V.

II. THEORY

A. Landscape of existing variational algorithms

VQE approaches are highly relevant for the NISQ era of quantum computation [4]. They comprise fairly simple quantum circuit implementations at the price of relying on the solution of a high-dimensional, classical optimization problem in the presence of noise. Despite the optimization challenge, many of the currently existing examples of actual quantum simulations for many-body physics correspond to implementations of this algorithm. The basic premise of VQE relies on the variational approximation: a parametrized wave function ansatz $|\Phi(\vec{\alpha}_j)\rangle$, where $\vec{\alpha}_j$ are the variational parameters, is chosen such that it can be efficiently implemented on a quantum computer. The energy expectation value $E(\vec{\alpha}) = \langle\Phi(\vec{\alpha}_j)|H|\Phi(\vec{\alpha}_j)\rangle/\langle\Phi(\vec{\alpha}_j)|\Phi(\vec{\alpha}_j)\rangle$ of this ansatz is evaluated on quantum hardware, and then the optimization problem $\vec{\nabla}E(\vec{\alpha}) = 0$ is solved on a classical computer. As in any variational approach, the efficacy of the approximation depends on the flexibility of the ansatz. A way to increase this flexibility is to choose a more general ground-state estimate, namely as a linear combination of several parametrized expansion states $|\Phi_j(\vec{\alpha}_j)\rangle$. The total wave function ansatz thus becomes

$$|\Psi(\vec{c}, \{\vec{\alpha}\})\rangle = \sum_j c_j |\Phi_j(\vec{\alpha}_j)\rangle, \qquad (1)$$

where the c_j are the expansion coefficients and $\vec{\alpha}_j$ the optimization parameters of the jth expansion states. Applying the variational principle to the coefficients $\vec{c}$ alone leads to the secular equations [31]

$$\sum_j H_{i,j} c_j^I = \varepsilon_I \sum_j S_{i,j} c_j^I, \qquad (2)$$

where ε_I is an estimate for the Ith Hamiltonian eigenvalue E_I. Equation (2) is a standard generalized eigenvalue equation, which can be solved classically. The Hamiltonian, which we assume to be time independent, and overlap matrix elements are measured on quantum hardware in the "basis" of expansion states following

$$\begin{aligned} H_{i,j} &= \langle\Phi_i(\vec{\alpha}_i)|H|\Phi_j(\vec{\alpha}_j)\rangle, \\ S_{i,j} &= \langle\Phi_i(\vec{\alpha}_i)|\Phi_j(\vec{\alpha}_j)\rangle. \end{aligned} \qquad (3)$$

Thus, the expansion coefficients $\vec{c}$ can be determined by classically solving the noisy, generalized eigenvalue problem in Eq. (2). The expansion states $|\Phi_j(\vec{\alpha}_j)\rangle$ themselves can be then optimized with a classical minimization method, further improving the energy estimates. One example in which this has been used recently is with chemically motivated ansatz, such as unitary coupled cluster

expansions [32], which had nonetheless some difficulties related to the optimization of parameters.

The optimization of the $\vec{\alpha}_j$ parameters is an open field due to (i) the high-dimensional nature of the optimization problem and (ii) the presence of noise, which is necessarily part of any approach on quantum hardware. To alleviate these problems, an alternative framework has recently been explored [5,22,23,33–37], which completely bypasses the need for optimization routines. In this family of methods, one does not employ a set of parametrized expansion states $|\Phi_j(\vec{\alpha}_j)\rangle$, but instead generates a set of expansion states $|\Phi_j\rangle$ systematically from one or several reference states $|\Psi_I\rangle$. The only variational parameters left are thus the expansion coefficient $\vec{c}$, and consequently the only task to be performed by a classical computer is solving the (noisy) generalized eigenvalue problem in Eq. (2).

The way of generating these expansion states should balance ease of implementation on quantum hardware with creating a flexible expansion set, in a variational sense, to obtain accurate energies. While *a priori*, by changing from parametrized $|\Phi_j(\vec{\alpha}_j)\rangle$ to systematically generated $|\Phi_j\rangle$, we are reducing the variational flexibility of our ansatz, the expansion state generation can still be performed in such a way that it is natural to both the description of ground and excited states, as well as to the implementation on a quantum computer. One such approach is the quantum subspace expansion method [33,35,36], which, after optimizing a ground-state estimate with regular VQE, generates expansion states by applying single excitation operators on top of this VQE reference. A recent variation introduced a more general multireference ansatz, targeting ground and exited states simultaneously at the optimization step [34].

Alternatively, the expansion states can be formed by applying the time evolution operator $U(t) = e^{-iHt}$ to the reference states. This is the approach followed in the QLanczos method [5,37], where the time evolution is performed along imaginary time (i.e., $t \to i\tau$), and the quantum filter-diagonalization and quantum Krylov approaches [22,23], where the time evolution is performed along real time. In the case of the QLanczos method, it is clear that evolving to large enough imaginary times will provide an expansion set that is well suited to describe the ground state, provided that the reference state is not orthogonal to it. In the next subsections, we discuss the formalism behind using a set of the real-time states in the expansion, which we refer to as VQPE. We also analyze VQPE's robustness to noise, which is critical to assess its applicability in NISQ devices.

B. VQPE—basic intuition

In the VQPE approach [22,23], the expansion states $|\Phi_{j,I}\rangle$ are generated from the reference states $|\Psi_I\rangle$ through

time evolution as

$$|\Phi_{j,I}\rangle = e^{-iHt_j}|\Psi_I\rangle. \tag{4}$$

If there are N_R reference states, and N_T time steps are considered, this produces a "basis" of $N_R(N_T + 1)$ expansion states, where $|\Phi_{0,I}\rangle = |\Psi_I\rangle$. For simplicity, we consider a single reference state $|\Psi_0\rangle$ for now, and discuss the use of multiple reference states below. Throughout the paper, we set the reduced Planck constant $\hbar = 1$.

On a first glance, it seems counterintuitive that the set of expansion states in Eq. (4) would improve the ground-state estimate given by the reference state $|\Psi_0\rangle$. After all, the $|\Phi_{j,0}\rangle$ states all have the same energy expectation value. Stair *et al.* [23] suggested an interpretation based on short time evolution, in which Taylor expanding Eq. (4) shows that the expansion states span the same space as the Krylov vectors $H^j|\Psi_0\rangle$, making it equivalent to power methods such as the Lanczos algorithm [29]. This justifies why a set of expansion states concentrated in a time grid over a short time scale should produce a good variational ansatz for the ground state. Including multiple reference states should then provide for good and stable approximates for the first few excited states as well, in the spirit of the band Lanczos method [38]. This interpretation suggests that the VQPE method should work best for short time evolution. We want to complement this interpretation with a more general one, not limited to short time steps, although this still remains the most interesting regime from an implementation perspective. In particular, the VQPE approach is reminiscent of the computation of the autocorrelation function $g(t) = \langle\Psi_0|e^{-iHt}|\Psi_0\rangle$, which contains the full spectral information of H [21]. In $g(t)$, the ground-state information is encoded in the long time limit, rather than the short time one, since the ground state evolves with the slowest frequency $\omega_0 = E_0$ in units where $\hbar = 1$. Thus, there should be nothing particular about the short time evolution limit. Instead, we articulate the precise requirements for the implementation that will lead to accurate energy estimates. This is related to the linear independence of the expansion states that poses a lower bound to the optimal time step size. We also note that VQPE is closely related to the classical filter-diagonalization method [24–26,39], as pointed out by Parrish and McMahon [22].

We thus want to depart from the Krylov intuition, and provide a different, hopefully more general heuristic, which we then rigorously formalize, to understand why VQPE approaches should provide good ground- and excited-state approximations. We begin by writing out the decomposition of the reference state $|\Psi_0\rangle$ into Hamiltonian eigenstates $|N\rangle$, such that $H|N\rangle = E_N|N\rangle$. This can be written out explicitly as

$$|\Psi_0\rangle = \sum_N \psi_N^0 |N\rangle, \tag{5}$$

where $\psi_N^0 = \langle N|\Psi_0\rangle$ are the coefficients of the reference state $|\Psi_0\rangle$ in the eigenbasis of H. We refer to those Hamiltonian eigenstates $|N\rangle$ for which ψ_N^0 is beyond some nonnegligible threshold as the *support space* of state $|\Psi_0\rangle$ with respect to H. This decomposition gives, for the expansion states,

$$|\Phi_{j,0}\rangle = e^{-iHt_j}|\Psi_0\rangle = \sum_N \psi_N^0 e^{-iE_N t_j}|N\rangle. \tag{6}$$

The equation above just states the obvious: each component of $|\Psi_0\rangle$ in the support space with respect to H evolves with its own frequency. This, however, makes transparent why the VQPE method can work: choosing the time grid $\{t_j\}$ accordingly, it is possible to make linear combinations of the expansion states $|\Phi_{j,0}\rangle$ such that the different phases $e^{-iE_N t_j}$ cancel out targeted components of $|\Psi_0\rangle$ along the support space state $|N\rangle$. In this way it is possible to "extract" eigenstates in the support space of $|\Psi_0\rangle$ by including enough expansion states $|\Phi_{j,0}\rangle$. Note that this is not exclusive to the ground state, nor is this limited to short time scales t_j. The only requirement is given by the number of eigenstates of H in the support space of $|\Psi_0\rangle$, defining how many time steps are needed for perfect state extraction, and by the energy gaps (i.e., relative frequencies) of those states, which govern the phase cancellation conditions.

In essence, VQPE allows one to extract "the most out of the reference state," in the sense that if there are Q states in its support space, it should be possible to produce Q time-evolved states from which to reconstruct the corresponding Q Hamiltonian eigenstates, by solving secular equation (2). Of course, this presumes that it is possible to produce Q linearly independent time-evolved states, and that our time evolution is noiseless and performed at arbitrary numerical precision. For general reference states, the size of the support space will be too large in general to recover all eigenstates, but a modest amount of these should be enough to approximate the lowest-lying energy eigenstates in it.

C. Phase cancellation conditions and relation to filter diagonalization

We now formalize the phase cancellation heuristic on a more solid mathematical footing. To this end, we derive a set of equations, the phase cancellation conditions, which set sufficient conditions to exactly extract the Hamiltonian eigenstates from the support space. These conditions embody the intuition in terms of autocorrelation functions described before, and are effectively discrete versions of the main relations at the heart of the classical filter-diagonalization approach [24–26]. When the reference state $|\Psi_0\rangle$ is exactly confined to a support space of size Q with nondegenerate eigenstates, the necessary and

sufficient condition for exact eigenstate recovery is simply $N_T + 1 \geq Q$ for the expansion states $|\Phi_{j,0}\rangle$ to span the support space. However, this condition can be numerically unstable to small contributions to $|\Psi_0\rangle$ from outside the support space because $S_{i,j}$ in Eq. (2) can be ill-conditioned, and more robust conditions such as the phase cancellation conditions are needed in practice.

We consider the overlap matrix $S_{j,k}$ in Eq. (3) for the expansion states in Eq. (6). It is convenient to consider a spanning operator W that is contained within the subspace of the expansion states,

$$W = \sum_{j=0}^{N_T} |\Phi_{j,0}\rangle \langle \Phi_{j,0}|, \tag{7}$$

and encodes the overlap matrix in its action on individual expansion states,

$$W|\Phi_{j,0}\rangle = \sum_{k=0}^{N_T} S_{kj} |\Phi_{k,0}\rangle. \tag{8}$$

We analyze W to identify the eigenstates that are accessible to VQPE.

Substituting Eq. (6) into the spanning operator gives, after some minor reordering of terms,

$$W = \sum_{N,M}^{Q} \psi_N^0 \psi_M^{0,*} \left[\sum_{j=0}^{N_T} e^{-it_j(E_N - E_M)} \right] |N\rangle \langle M|. \tag{9}$$

In the equation above, Q is the number of Hamiltonian eigenstates in the support of $|\Psi_0\rangle$, and we can ignore Hamiltonian eigenstates outside the support space due to their small coefficients ψ_N^0. Now, we can define the phase cancellation conditions (PCCs) as

$$\frac{1}{N_T + 1} \sum_{j=0}^{N_T} e^{-it_j(E_N - E_M)} = \delta_{N,M}. \tag{10}$$

The nontrivial off-diagonal PCCs ($N \neq M$) are the $Q(Q - 1)/2$ conditions for the $N_T + 1$ time steps in the time grid. Given that the support space is spanned by just Q vectors, it seems that the PCCs impose stricter conditions on the time grid than absolutely necessary to recover the full support space. Still, they mathematically embody the phase cancelation heuristic that we have discussed above. Indeed, the condition in Eq. (10) enforces the cancelation of the time-evolved phase between all Hamiltonian eigenstates in the support of $|\Psi_0\rangle$, and can be represented graphically as a sum of phases in the unit circle. When the phase cancellation conditions are fulfilled, the spanning operator

simplifies into a weighted projector into the support of $|\Psi_0\rangle$ with respect to H, namely,

$$W \overset{\mathrm{PCC}}{=} \sum_{N}^{Q} (N_T + 1)|\psi_N^0|^2 |N\rangle \langle N|, \tag{11}$$

weighted by the coefficients of the reference state $|\Psi_0\rangle$ on the support space; cf. Eq. (5). In this case, the expansion states span exactly the same space as the Hamiltonian eigenstates in the support space, and solving secular equation (2) returns the exact eigenstates and eigenvalues of H. The PCCs in Eq. (10) can be understood as the discrete limit of the eigenstate extraction through Fourier transform of a time-evolved state exploited in the classical filter-diagonalization literature [24].

Furthermore, in the limit where long time evolutions are used the phase cancellation conditions will also be approximately satisfied with high probability as N_T tends to infinity. To see this, let $t_j = (j + 1)/\omega$, where ω^{-1} is a uniform random variable on $[0, 1/\min(E_N - E_M)] := [0, \Delta E_{\min}^{-1}]$ (where $M \neq N$). First, we have

$$\left| \mathbb{E} \sum_{j=1}^{N_T+1} e^{-ij(E_N - E_M)/\omega} \right|$$

$$= \left| \sum_{j=0}^{N_T} \Delta E_{\min} \int_0^{\Delta E_{\min}^{-1}} e^{-ij\omega^{-1}(E_N - E_M)} d\omega^{-1} \right|$$

$$= \left| \sum_{j=1}^{N_T+1} \frac{\Delta E_{\min}(1 - e^{-ij\Delta E_{\min}^{-1}(E_N - E_M)})}{ij(E_N - E_M)} \right|$$

$$\in O\left(\frac{\Delta E_{\min}}{|E_N - E_M|} \sum_{j=1}^{N_T} \frac{1}{j} \right)$$

$$\subseteq O\left(\frac{\Delta E_{\min} \log(N_T)}{|E_N - E_M|} \right). \tag{12}$$

Intuitively, it is reasonable to expect that if the mean is small then, with high probability, the PCCs should hold approximately. In order to demonstrate such a concentration for the oscillating functions that we use here, however, we need to also bound the variance:

$$\left| \mathbb{V} \sum_{j=1}^{N_T+1} e^{-ij(E_N - E_M)/\omega} \right|$$

$$\leq \mathbb{E}\left(\sum_{j \neq k} e^{-i(j-k)(E_N - E_M)\omega^{-1}} \right) + (N_T + 1)$$

$$\in O\left(N_T + \frac{\Delta E_{\min}}{|E_N - E_M|} \sum_{j \neq k} \frac{1}{|j - k|}\right)$$

$$\in O\left(\frac{\Delta E_{\min} N_T \log(N_T)}{|E_N - E_M|}\right). \tag{13}$$

Thus, from Chebyshev's inequality, it follows that, with high probability, $\sum_j e^{-ij(E_N - E_M)\omega^{-1}}$ will be within

$$O\left(\sqrt{\frac{N_T \log(N_T) \Delta E_{\min}}{|E_N - E_M|}}\right) \tag{14}$$

of the expectation value. Thus, the phase cancelation condition's error for the N, M component is in

$$O\left(\sqrt{\frac{\log(N_T) \Delta E_{\min}}{N_T |E_N - E_M|}}\right). \tag{15}$$

Thus, the value of N_T needed to ensure that the PCC holds within an error of at most ϵ (with high probability) obeys

$$N_T \in \widetilde{O}\left(\frac{\Delta E_{\min}}{|E_N - E_M|\epsilon^2}\right). \tag{16}$$

Here $\widetilde{O}(\cdot)$ denotes an asymptotic upper bound with multiplicative polylogarithmic factors neglected. Thus, an approximate solution to the phase cancelation conditions will generically hold for a gapped system.

We exemplify the previous theory on the example of a Hamiltonian of linear spectrum $E_N = N \Delta E$, akin to a harmonic oscillator, in Fig. 1. In this case, the PCCs in Eq. (10) can be fulfilled exactly by a linear time grid $t_j = j \Delta t_P$ with the perfect time step size Δt_P defined as

$$\Delta t_P = \frac{2\pi}{(N_T + 1)\Delta E}. \tag{17}$$

Indeed, it is easy to check that in the case of a linear spectrum, a linear time grid with time step size given by Eq. (17) fulfills the PCCs exactly after $N_T = Q - 1$ time steps. This can be accomplished with a single time step size since in the case of a linear spectrum the PCCs effectively reduce to a single condition. This can be seen in the rightmost panel of Fig. 1, where after exactly 15 time steps the first four eigenvalues of the secular equation match the exact eigenvalues to the maximal precision. This precision is determined by the singular value threshold, s_{SV}, introduced into the general eigenvalue problem, which truncates the singular values of the overlap matrix. In Fig. 1 this precision corresponds to $s_{SV} = 10^{-12}$, i.e., midway between double and single machine precision. This threshold also determines the support space size Q. The reference state in all examples in Fig. 1 is defined

as $|\Psi_0\rangle \propto \sum_N e^{-E_N} |N\rangle$, excited states being exponentially suppressed. The support space is then defined as the Hamiltonian eigenstates with squared coefficients in $|\Psi_0\rangle$ above s_{SV}. The inset in this panel represents the PCCs graphically, as the phases of all eigenstates in the support space perfectly span the unit circle, thus canceling each other.

The smaller panels on the left of Fig. 1 show the outcome of choosing a time step size differing from Δt_p. Small time step sizes are shown in the upper two subfigures, and would be the natural choice from the Krylov interpretation of VQPE [23]. These clearly show a significantly slower convergence than the perfect time step derived from the PCCs, which is easy to explain from the phase distribution on the unit circle in the insets. Only once we cover the unit circle close to homogeneously, thus approximately fulfilling the PCCs, can we extract all eigenstates essentially exactly after the minimal number of time steps (see the lower left panel in Fig. 1). Because of the periodic nature of the complex phases in Eq. (10), large time steps can result in as poor approximations as short ones, as shown in the lower right panel in Fig. 1. The worst such longer time step sizes correspond to particular integer multiples of Δt_P, namely $zQ\Delta t_P$, where z is an integer. For these time step sizes, the PCCs in Eq. (10) cannot be fulfilled, even approximately. These are very particular time steps, and thus, for the linear spectrum, a randomly chosen time step $\Delta t \geq \Delta t_p$ is still likely to provide good results for the Hamiltonian eigenstates.

For general spectra, a single time step size Δt in a linear grid is unlikely to fulfill all $Q(Q - 1)/2$ PCCs exactly. From the above analysis, a valid strategy would be to choose a time step size Δt and number of time steps N_T such that we sample a full period of the slowest oscillation in the support space. This is given, e.g., by the minimal energy gap $\Delta E_{\min}$ if we are interested in all excited states contained in the support space, and the ground-state gap $\Delta E_{1,2} = E_2 - E_1$ if we only need an estimate of the ground state. However, in practical implementations it is advantageous, and sometimes necessary, to limit the total simulation time in order to minimize the error. At the same time, we have to choose a time step size large enough such that each new state is linearly independent from the previous ones. Otherwise, no new information is added and the variational ansatz is not improved (see Fig. 1, upper left-hand panel, where the energy decreases in a steplike fashion). Reconciling these two notions, we propose the following systematic approach.

1. Choose a small enough time step size such that the energy convergence is steplike. Steplike convergence refers to the situation in which adding a new expansion state, i.e., propagating for an additional time step, does not improve the variational ansatz, resulting in the same (or slightly worse) energy
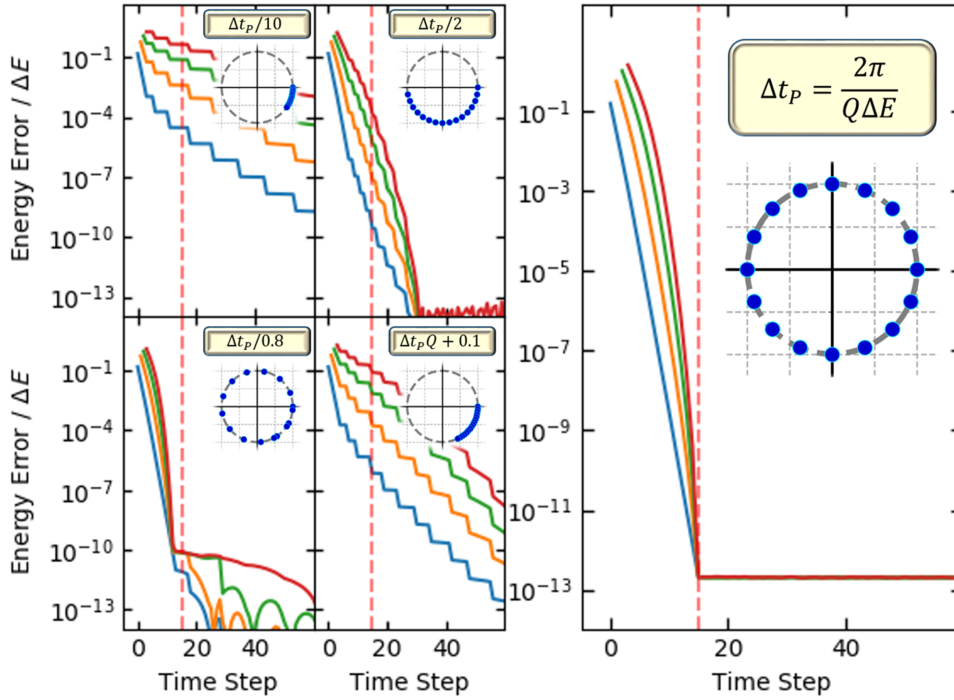
FIG. 1. Relative error of the first four eigenvalues from the real-time NOVQE secular equation for a Hamiltonian of linear spectrum $E_N = N \Delta E$ ($\Delta E = 0.75$ here) and different time steps Δt as a function of the number of time steps. The reference state follows $|\Psi_0\rangle \propto \sum_N e^{-E_N} |N\rangle$, such that only the 16 Hamiltonian eigenstates of lowest energy are part of the support space, choosing a coefficient threshold of 10^{-12}. When solving the secular equation, we choose the same threshold for the SVD decomposition of the overlap matrix. The vertical dashed line marks the 15th time step, after which we have as many expansion states as vectors in the support space. The large subplot corresponds to the smallest time step that fulfills the phase cancelation condition, Eq. (10), which reduces to a single condition for this Hamiltonian. The insets in each subfigure correspond to a geometric representation of the phase cancelation condition, with each phase $e^{-it_j \Delta E}$ a point in the unit circle on the complex plane. See the text for details.

estimates as before, including the new step (cf. upper left panel in Fig. 1). This happens when the inclusion of the new expansion state produces an overlap matrix that has no additional singular value over the threshold s_{SV}.

2. Perform the VQPE algorithm using the previously identified small time step size, until the first expansion state resulting in an improvement of the energy estimates is produced.

3. Plotting the lowest eigenvalue of the previous VQPE simulation as a function of the propagation time will result in a nearly horizontal line. This plateau ends after the addition of the final expansion state, which does improve the energies. The length of this plateau defines a new, larger time step size, which can be used in a new VQPE simulation.

4. Repeat the simulation with the new time step size. If there is still steplike convergence, go back to step 3. Otherwise, use this as your simulation time step.

This procedure underlies the fact that, for devising practical implementations of VQPE, the guiding principle should be to generate linearly independent expansion states

rather than to exactly fulfill the PCCs, which nonetheless are a useful perspective for the theoretical analysis of the algorithm. This is the strategy we adopt in the results section below.

D. Towards an optimal implementation—Toeplitz structure of S

Besides being natural to implement on quantum hardware, and presenting the interesting phase cancelation structure described above, VQPE approaches show a further theoretical advantage: when using a linear time grid $t_j = j \Delta t$, the Hamiltonian and overlap matrices in Eq. (3) have a restrictive structure, which formally reduces the number of measurements that should be needed to solve the generalized eigenvalue problem. Indeed, as pointed out by Parrish and McMahon [22], using the real-time expansion set, these matrices become Toeplitz, meaning that, e.g., $S_{j,k} = S_{j+1,k+1}$. In particular, the concrete expressions read

$$H_{j,k} = \langle \Phi_{j,0}|H|\Phi_{k,0}\rangle = \langle \Psi_0|He^{-iH\Delta t(k-j)}|\Psi_0\rangle,$$
$$S_{j,k} = \langle \Phi_{j,0}|\Phi_{k,0}\rangle = \langle \Psi_0|e^{-iH\Delta t(k-j)}|\Psi_0\rangle, \quad (18)$$

where we have only used the fact that a time-independent Hamiltonian commutes with itself at all times, thus relying exclusively on time translational symmetry. From Eq. (18), it follows that we can reconstruct the two $(N_T + 1) \times (N_T + 1)$ matrices by measuring only $2(N_T + 1)$ overlaps in total.

Unfortunately, as pointed out in Ref. [22], the Toeplitz property of the Hamiltonian matrix is lost in actual quantum hardware implementations, if the time evolution operator is Trotterized. In those cases, the commutativity of $U(t_j)$ with H is lost, and thus we either need to evaluate all $(N_T + 1)^2$ Hamiltonian matrix elements separately or transition to a higher-order Trotter formula that better approximates the commutation relations. Nevertheless, the same is not the case for the overlap matrix. As long as the expansion states $|\Phi_{j,0}\rangle$ are constructed using a linear grid with a unitary time evolution approximation

$$U(\Delta t) = e^{-iH\Delta t} \approx U_a(\Delta t), \quad (19)$$

the Toeplitz condition will prevail. Here, $U_a(\Delta t)$ is the Trotterized time evolution operator and Δt is the common time step size of the linear time grid, such that the approximate expansion states obey

$$|\Phi_{j,0}^a\rangle = [U_a(\Delta t)]^j |\Psi_0\rangle, \quad (20)$$

which in the limit of an exact time evolution operator recovers Eq. (6). Now, if $U_a(\Delta t)$ is unitary, the overlap matrix of the approximated expansion states $|\Phi^a(\Delta t)\rangle$ will clearly be Toeplitz, since

$$S_{j,k}^a = \langle \Phi_{0,j}^a | \Phi_{0,k}^a \rangle = \langle \Psi_0 | [U_a(\Delta t)]^{k-j} |\Psi_0\rangle. \quad (21)$$

Given that gate operations in quantum hardware are naturally unitary, this means that it is always possible to guarantee the Toeplitz condition of the overlap matrix, simply by choosing a linear time grid. This is true of course for the often invoked first-order Trotterization [8] approximation to the time step evolution $U(\Delta t)$, which is indeed unitary.

E. Towards an optimal implementation—unitary formulation

It is possible to rewrite this generalized eigenvalue problem in a simpler form, exploiting the particular relationship between the Hamiltonian and overlap matrices in Eq. (18), essentially formulating it equivalently to the classical filter-diagonalization problem found in signal processing [26]. This proves to be the ideal formulation of VQPE for quantum computation.

The main insight relies on substituting the Hamiltonian in secular equation (2) by the time evolution operator $U(\Delta t) = e^{-iH\Delta t}$. This operator is effectively isospectral

with the Hamiltonian; indeed the eigenstates $|N\rangle$ of H fulfill

$$U(\Delta t) |N\rangle = e^{-iE_N \Delta t} |N\rangle. \quad (22)$$

It is important to note that, unlike the Hamiltonian, the time evolution operator is not Hermitian, but unitary, thus having complex eigenvalues of unit modulus. We can therefore write a secular equation for the time evolution operator $U(\Delta t)$ as

$$\sum_k U(\Delta t)_{j,k} c_k^I = e^{-i\varepsilon_I \Delta t} \sum_k S_{j,k} c_k^I, \quad (23)$$

where the overlap matrix, eigenvalues ε_I, and expansion coefficients c_j^I are the same as in Eq. (2), and the time evolution matrix elements follow, in the single reference implementation,

$$U(\Delta t)_{j,k} = \langle \Phi_{j,0} | U(\Delta t) | \Phi_{k,0} \rangle = \langle \Psi_0 | e^{-iH(\Delta t + t_k - t_j)} | \Psi_0 \rangle. \quad (24)$$

To transform from eigenvalues of Eq. (22) to Eq. (2), Δt must also be small enough that we can distinguish a physical E_N value from its unphysical periodic images $E_N \pm 2\pi/\Delta t$.

From Eq. (2) to Eq. (23), we have simply reformulated the VQPE problem into an equivalent generalized eigenvalue problem with a unitary matrix. The key simplification for the implementation on quantum hardware relies on the realization that the time evolution matrix elements in Eq. (24) have the same structure as the overlap matrix elements in Eq. (18). Thus, again choosing the time grid $\{t_j\}$ to be linear, i.e., $t_j = j \Delta t$, the time evolution matrix elements coincide with the overlap matrix elements as

$$U(\Delta t)_{j,k} = \langle \Psi_0 | e^{-iH\Delta t(1+k-j)} | \Psi_0 \rangle = S_{j,k+1} = S_{j-1,k}. \quad (25)$$

The last equality is a manifestation of the Toeplitz structure. Thus, according to Eq. (25), for linear time grids, there is no need to measure the time evolution matrix explicitly, since it can be recovered from the measurements for the overlap matrix plus an additional measurement involving an extra expansion state $|\Phi_{N_T+1,0}\rangle$. In this way, exploiting Eq. (25) and the Toeplitz structure of the overlap matrix, the number of measurements reduces from $2(N_T + 1)^2$ to just $N_T + 2$. Furthermore, as shown in the previous subsection, the Toeplitz structure prevails when implementing the time evolution operators with a unitary approximation, such as first-order Trotterization, making the reduction in the number of measurements applicable for real implementation on quantum hardware.

Intuitively, the unitary formulation of VQPE is the quantum algorithm equivalent to measuring the autocorrelation function $g(t) = \langle\Psi_0|e^{-iHt}|\Psi_0\rangle$ and analyzing its Fourier spectrum. The overlap matrix elements are essentially sampling $g(t)$ at different points, and one can approximate the underlying spectrum once enough samples are obtained. Thus, we are fundamentally expressing the VQPE algorithm in its most natural language, that of autocorrelation functions. In this work, we implement both the traditional and unitary formulations of the VQPE secular equation.

F. Diminishing the effect of noise through singular value decomposition

In the previous subsections, we briefly reviewed the theoretical formalism of VQPE approaches, in the new light of the phase cancelation interpretation, but without considering the effects of noise. We turn our attention now to how the presence of noise, comprising both finite numerical precision on the classical computer and measurement uncertainty from the quantum hardware, limits the final accuracy of the VQPE results. We consider systematic noise, due to *a priori* uncontrollable or unavoidable sources, plus any remaining statistical uncertainty after repeated measurements. To this end, the phase cancelation formalism will simplify the analysis. We restrict ourselves for simplicity to the single reference implementation, but generalizing our conclusions to the multireference case is straightforward.

As shown in Eq. (4), VQPE generates a series of N_T states from a reference $|\Psi_0\rangle$ defined by a time grid $\{t_j\}$. In a noiseless simulation, any given time grid is more likely than not to produce a set of N_T linearly independent vectors. For them to be linearly dependent requires the following determinant to vanish exactly:

$$\begin{vmatrix} c_0 & c_0 e^{-iE_0 t_1} & c_0 e^{-iE_0 t_2} & \cdots \\ c_1 & c_1 e^{-iE_1 t_1} & c_1 e^{-iE_1 t_2} & \cdots \\ c_2 & c_2 e^{-iE_2 t_1} & c_2 e^{-iE_2 t_2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ c_{N_T} & c_{N_T} e^{-iE_{N_T} t_1} & c_{N_T} e^{-iE_{N_T} t_2} & \cdots \end{vmatrix} = 0. \quad (26)$$

This equation is one constraint on N_T unknowns $\{t_j\}$, which is generically satisfied by an $(N_T - 1)$-dimensional manifold of $\{t_j\}$ embedded in $\mathbb{R}^{N_T}$. For a linear grid $t_j = j\,\Delta t$, the choice of time step size Δt will generically cause linear dependencies on a subset of $\mathbb{R}$ with measure zero. For example, $\Delta t = 2\pi n/(E_2 - E_1)$ causes a linear dependency in the case of $N_T = 2$. Thus, with the exception of Hamiltonians with a restricted spectrum such as $E_j = j\,\Delta E$, it seems safe to assume that in almost any time grid chosen, a noiseless simulation will generate N_T linearly independent vectors. Since all the expansion states share the same support space [40], a noiseless simulation with Q steps,

Q being the size of the support space, should recover all eigenstates exactly. We henceforth assume a linear time grid with time step size Δt.

This ideal notion stops holding the moment we consider noise, both from numerical and measurement origins. Noise can for example make states close to linearly dependent, and thus introduce errors in the eigenvalues ε_I of the secular equation. We quantify noise by introducing the parameter ϵ. Two measured or computed values α, β are only distinguishable if $|\alpha - \beta| > \epsilon$. Noise becomes important, for example, in the small time step size limit. When $\Delta t \Delta E_{\min}$ is small, where $\Delta E_{\min}$ is the minimal spectral gap in the support space, the first expansion steps will produce states that are only marginally different to the reference $|\Psi_0\rangle$. These will not improve the variational ansatz if

$$\frac{\Delta T_\epsilon}{\hbar} \Delta E_{\min} < \epsilon, \quad (27)$$

where we have recovered Planck's constant to make the units clear. If Eq. (27) is fulfilled, the magnitude of the difference between the expansion state and the reference will fall below the noise threshold, making the new expansion state $U(\Delta T_\epsilon)|\Psi_0\rangle$ useless from a variational perspective. This is the reason behind the steplike decreasing behavior in the small time step panels of Fig. 1. A finite ϵ thus determines a minimal time step $\Delta T_\epsilon > \hbar\epsilon/\Delta E_{\min}$.

There will be cases where it is hard to generate precise expansion states with the minimal time step size ΔT_ϵ required to offset a given noise level. It thus becomes important to prune the Hamiltonian and overlap matrices of numerical and measurement noise. This can be done by means of a singular value decomposition (SVD) of the overlap matrix: neglecting all singular values below some threshold, which should be larger than the magnitude of the noise. In the case of measurement error, this noise scales as $1/\sqrt{M}$, where M is the number of samples. In the majority of this work, we have thus conservatively chosen a threshold of 10^{-1}, corresponding to > 100 samples. We already used such a truncation in the results shown in Fig. 1. This singular value truncation effectively produces a new but smaller expansion basis of N_{SVD} elements. As a consequence, the number N_T is not the significant measure of how much information is collected in the expansion set, and instead $N_{\text{SVD}} \leq N_T$ becomes the measure to follow. Only when $N_{\text{SVD}} = Q$ will the secular equations recover the exact support space spectrum.

We exemplify this on the harmonic spectrum in Fig. 2, where in the upper panels we show the relative noise error for the first four eigenstates as a function of number of expansion states, introducing Gaussian noise $\mathcal{N}(0, \epsilon)$ of standard deviation ϵ on the Hamiltonian and overlap matrix elements. We choose the singular value truncation threshold s_{SV} to be between 10^{-1} and 1, and in the
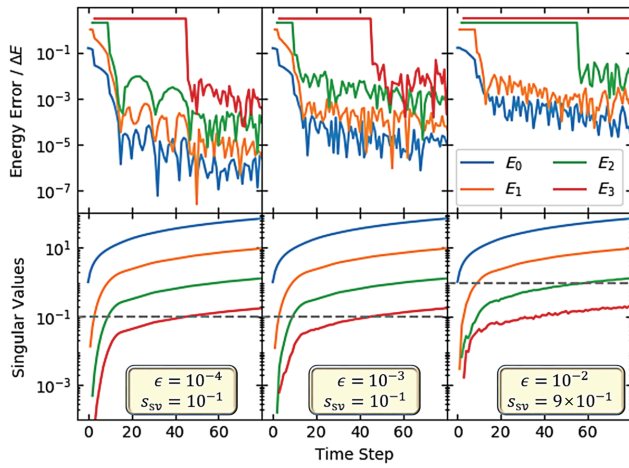
FIG. 2. Relative error and corresponding singular values of the overlap matrix for a Hamiltonian of linear spectrum with different noise values. Upper panels: relative error of the first four eigenvalues from the VQPE secular equation for a Hamiltonian of linear spectrum $E_N = N\Delta E$ ($\Delta E = 0.75$ here) and perfect time step Δt_P, including Gaussian noise $\mathcal{N}(0, \epsilon)$ to the Hamiltonian and overlap matrix elements. We choose the singular value truncation threshold s_{SV} to be at least 2 orders of magnitude larger than the noise standard deviation ϵ. The reference state follows $|\Psi_0\rangle \propto \sum_N e^{-E_N} |N\rangle$, the effective support space size determined using the singular value truncation threshold $N_{SVD}^{max} = -\ln(s_{SV})/(2\Delta E) - 1$. Lower panels: corresponding singular values of the overlap matrix as a function of time. The horizontal dashed line represents the threshold s_{SV}. Note that, until a given singular value is larger than s_{SV}, we cannot extract the corresponding eigenvalue from the generalized eigenvalue problem. This is represented by the horizontal lines in the upper panels. See the text for details.

lower panels we plot the singular values of $S_{j,k}$ as a function of the number of expansion states, marking s_{SV} as a dashed line. In each simulation, we choose as time step size Δt the optimal time step in Eq. (17), considering a possible support space of 16 elements, regardless of s_{SV} [41]. In a noiseless simulation, this choice of time step would result in an optimally compact number of effective expansion vectors N_{SVD}, which equals the number of actual expansion vectors until the maximal number $N_{SVD}^{max} = Q$ is reached, after which all eigenstates would be resolved accurately. The presence of statistical noise has two consequences. On the one hand, the asymptotic accuracy decreases with increasing noise variance. As mentioned above, this type of statistical noise can be reduced by sampling.

As a second effect of statistical noise, not all Q eigenstates in the support space are resolved after exactly Q steps, since the corresponding singular values of the overlap matrix fall below the truncation threshold s_{SV} (see the lower panels of Fig. 2). As long as a given singular value falls below s_{SV}, the corresponding eigenvalue cannot be extracted from the generalized eigenvalue equation, which

is represented in the upper panels of Fig. 2 by horizontal straight lines. Thus, the noise limits what states can be extracted from the reference state $|\Psi_0\rangle$, by setting the minimal singular value truncation threshold s_{SV}. Those Hamiltonian eigenstates with smaller absolute coefficient squared than s_{SV} cannot be resolved. However, examining Eq. (11), we observe that the singular values of the overlap matrix are enhanced linearly with an increasing number of expansion states N_T, i.e., with an increased number of time steps in the VQPE approach, once the PCCs are reasonably fulfilled. Thus, it should be possible to extract eigenstates with reference state components below the error threshold by increasing the number of time steps. We exemplify this in Fig. 3, again on the harmonic spectrum example with exponentially suppressed initial state. Because of this choice of starting state, it takes an exponentially large number of extra time steps to resolve every new eigenstate, but it is in principle possible. The ideal strategy is of course to propose a reference state with large overlap with the eigenstate of interest, but this discussion shows that it is possible to extract states beyond the dominant one accurately. Once enough time steps have been produced, any singular value of the overlap matrix can be made to increase above s_{SV}, the horizontal dashed line in the lower
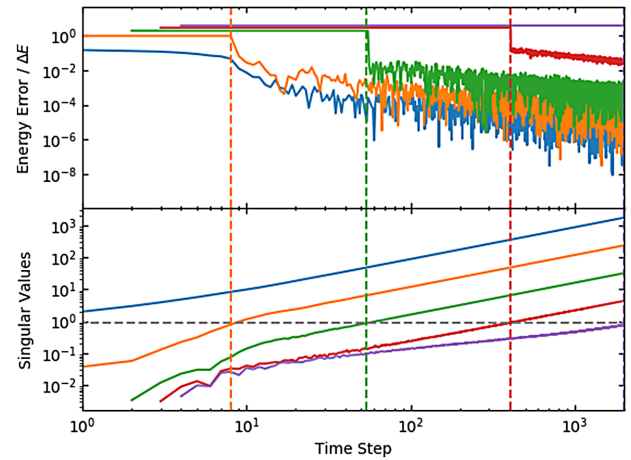

FIG. 3. Relative error and corresponding singular values of the overlap matrix for a Hamiltonian of linear spectrum for a large enough number of time steps to extract eigenstates below the error threshold. Upper panel: relative error of the first five eigenstates from the VQPE secular equation for a Hamiltonian of linear spectrum $E_N = N\Delta E$ ($\Delta E = 0.75$ here) and perfect time step Δt_P, including Gaussian noise $\mathcal{N}(0, 10^{-2})$ to the Hamiltonian and overlap matrix elements. We choose the singular value truncation threshold $s_{SV} = 9 \times 10^{-1}$. The reference state follows $|\Psi_0\rangle \propto \sum_N e^{-E_N} |N\rangle$. Lower panel: corresponding singular values of the overlap matrix as a function of time. The horizontal dashed line represents the singular value threshold s_{SV}. The time step at which each singular value becomes larger than s_{SV} is marked with a vertical dashed line, connecting upper and lower panels. See the text for details.

panel of the figure. Of course, this is not unexpected, since the initial starting signal on each excited state is exponentially small by construction. When a uniform overlap is used as a starting state, then all the SVD values are the same.

We want to address another notion that has been brought up with regards to the effect of noise in VQPE simulations: the overlap matrix condition number $n(S)$ [23]. As shown in Ref. [23], typical eigenvalue problems arising in VQPE have extremely large condition numbers, which may suggest high sensitivity to noise. This is expected for Hartree-Fock starting states, which ideally have exponentially small overlap with most of the Hilbert space, leading to very large condition numbers. Still, this can be dealt with by performing a SVD of the overlap matrix and truncating the singular values below s_{SV}. For our application, this reduces the number of linearly independent states we can resolve within our error threshold, thus reducing the best possible accuracy of the results. This can be remedied by including additional time-evolved states, as discussed above.

In the Supplemental Material [42], we investigate the role of $n(S)$ for the eigenvalue accuracy in the VQPE approach. Our results show that, upon pruning the expansion space from the singular values of the overlap matrix below the threshold s_{SV}, we consistently obtain accurate eigenvalues in the presence of noise even with matrices of large condition number. This is in itself not surprising, since the singular value truncation is effectively proposing an auxiliary generalized eigenvalue problem with smaller condition number.

At this point, we can return our attention to the notion of multireference implementations of the VQPE, such as the multireference quantum Krylov method of Stair *et al.* [23]. In this work the use of several reference states is proposed in order to reduce the condition number of the Hamiltonian and overlap matrices in the expansion space, at the cost of requiring a larger number of expansion states for the same ground-state accuracy. Using the phase cancelation picture, we argue that the worsened ground-state energy convergence is due to two distinct, cooperating factors: the increased size of the total support space, and the smaller number of expansion vectors in each individual support space. In the multireference formulation of VQPE, each state $|\Psi_I\rangle$ has its own support space with respect to H, which we refer to as individual support spaces, the union of these forming the total support space of the implementation. The individual support spaces will in general be distinct from each other. Clearly, the larger total support space allows for a more flexible variational ansatz, from which it is possible to extract more Hamiltonian eigenstates than in the single reference case. However, this comes at the price of requiring more expansion states to perform the phase cancelation procedure to purify individual eigenstates. From our results, performed in classical simulations with noise, and on actual noisy quantum hardware, the larger condition numbers do not result in large errors in the eigenvalue estimates ε_I, and thus we conclude that the condition number alone should not be a reason to employ multireference VQPE implementations. However, in some excited state simulations a multireference approach might accelerate convergence, in the same way that band Lanczos improves normal Lanczos in this regard [29,38].

The previous considerations hold for statistical errors, but a more careful analysis needs to be performed for systematic errors in the implementation of the Hamiltonian dynamics. For example, in the case of a Trotterized Hamiltonian, the reference states are created under the evolution of a Hamiltonian different than the one of interest, which will result in errors in the eigenvalue estimation that cannot be reduced through sampling.

G. Comparison between unitary VQPE and QPE

Here we compare the unitary formulation of VQPE in Eq. (22) to conventional QPE in the general case of a multidimensional support space. For the special case of a one-dimensional support space, there are adaptive variants of QPE that use one ancilla qubit and achieve Heisenberg-limit measurement [43]. Recently, developed methods [44, 45] devise QPE variants that use one ancilla qubit and are suitable for larger support spaces (which is also the case for VQPE). In an adaptive approach where $|\Psi_0\rangle$ is the ground state, a different t would be chosen for each measurement to maximize the extraction of information about E_0 rather than performing multiple measurements to estimate the expectation value of $g(t) = \langle\Psi_0|e^{-iHt}|\Psi_0\rangle = e^{-iE_0t}$ for a single choice of t.

QPE is a natural algorithm to compare VQPE against; however, there are a wealth of different phase estimation algorithms known in the literature and further some applications of phase estimation can even be used in concert with VQPE. Our aim in this section is to compare and contrast different flavors of phase estimation to VQPE and also show how QPE can be used to accelerate learning the expectation values of the VQPE circuit through amplitude estimation.

There are broadly two categories of phase estimation algorithms, iterative phase estimation and Fourier-based phase estimation. Fourier-based phase estimation is perhaps the best understood approach to performing phase estimation. An advantage of this approach is that it is known to precisely achieve optimal scaling of the uncertainty with the number of applications of the underlying unitary (i.e., it saturates the Heisenberg limit [46,47]). The optimal approach to Fourier-based phase estimation deviates slightly from traditional approaches by using an optimized initial state that deviates from the Fourier state typically used in older approaches. Specifically, the input

state is taken to be an m-qubit state of the form

$$\chi |0\rangle = \sqrt{\frac{2}{2^m + 1}} \sum_{n=0}^{2^m - 1} \sin\left(\frac{\pi(n+1)}{2^m + 1}\right) |n\rangle. \qquad (28)$$

This state is chosen to minimize an estimate of the circular variance, known as the Holevo variance, of the eigenphases of the unitary e^{-iHt} that results from the phase estimation protocol. Next let us define the following notation for a controlled directional evolution:
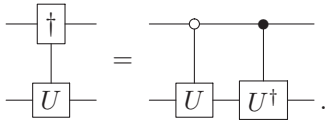
This notation is useful here because in quantum phase estimation for the simulation of electronic structure, the cost of performing a controlled-directional evolution as seen above is approximately the same as performing an ordinary controlled evolution. Applying phase estimation with a circuit of this form (see Figure 4) provides twice as much phase per application of the unitary as a controlled-U application would experience. The uncertainty from this phase estimation circuit, as quantified by the Holevo variance, is approximately $\epsilon_{\text{PE}} = \pi/2^{m+1}$, where m is the number of qubits used in the QPE register. This is precisely the Heisenberg limit.

Of course, error in the simulation dynamics needs to be considered in any such simulation. For simplicity, we consider using the lowest-order Trotter-Suzuki formula within the phase estimation protocol. In the limit where the uncertainty in the eigenphase is small relative to π, the Holevo variance approximately corresponds to the variance. Then, since the Trotter-Suzuki error is uncorrelated with the phase estimation error, the total error in the phase
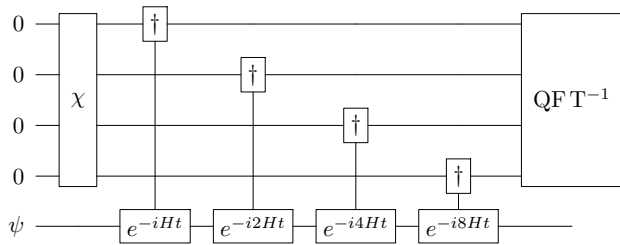

FIG. 4. Four-qubit version of the asymptotically optimal phase estimation circuit.

estimate that arises from using a time step of duration t is

$$\epsilon \leq \frac{\sqrt{\epsilon_{\text{PE}}^2 + \pi^2 \epsilon_{\text{TS}}^2(t)}}{t}$$

$$\approx \frac{\pi \sqrt{1/2^{2m+2} + \epsilon_{\text{TS}}^2(t)}}{t}. \qquad (29)$$

This suggests that if we choose both error sources to be $\epsilon t/(2\sqrt{2}\pi)$ then the overall error will be ϵ. This corresponds to

$$m = \left\lceil -\tfrac{1}{2} + \log_2(2\pi/\epsilon t) \right\rceil \leq \log_2(2\sqrt{2}\pi/\epsilon t). \qquad (30)$$

The error in the Trotter-Suzuki formula is difficult to estimate *a priori* and, for the purposes of simplifying the comparison, we choose the symmetric Trotter formula. Using the results of Proposition 16 of Ref. [48], if $H = \sum_j \alpha_j H_j$ then the error in the symmetric Trotter formula is upper bounded by

$$\epsilon_{\text{TS}}(t) \leq \frac{t^3}{12} \sum_j \sum_{p>j} \sum_{q>j} \alpha_p \alpha_q \alpha_j \, \|[H_q, [H_p, H_j]]\|$$

$$+ \frac{t^3}{24} \sum_j \sum_{p>j} \alpha_p \alpha_j^2 \, \|[H_j, [H_j, H_p]]\|. \qquad (31)$$

This bound can be further simplified using the fact that each of the Hamiltonian terms H_j in quantum chemistry is of norm 1. Therefore, let us denote S to be the set of all tuples (p, q, j) such that $[H_q, [H_p, H_j]] \neq 0$. Then we can further bound

$$\epsilon_{\text{TS}}(t) \leq \frac{2t^3}{3} \sum_j \sum_{p>j} \sum_{q>j} |\alpha_p \alpha_q \alpha_j| \delta_{(q,p,j) \in S}$$

$$+ \frac{t^3}{3} \sum_j \sum_{p>j} |\alpha_p \alpha_j^2| \delta_{(j,p,j) \in S}$$

$$:= \Gamma^3 t^3. \qquad (32)$$

Under the assumption that the error from QPE is equal to the Trotter error, we then find that a sufficient choice for the Trotter step is

$$t = \Gamma^{-3/2} \sqrt{\frac{\epsilon}{2\sqrt{2}\pi}}. \qquad (33)$$

Therefore, the value of m needed in the phase estimation step is

$$m \leq \frac{3}{2} \log_2\left(\frac{2\sqrt{2}\pi \Gamma}{\epsilon}\right). \qquad (34)$$

Thus, if M is the number of terms in the Hamiltonian then the maximum number of operator exponentials that need

to be simulated in the circuit to achieve Holevo variance ϵ is

$$N_{\exp} \leq 2M\left(\frac{2\sqrt{2}\pi\Gamma}{\epsilon}\right)^{3/2}. \tag{35}$$

This process needs to be repeated a number of times to find an eigenstate j that occurs with a probability of $|\psi_j^0|^2$. This implies that the total number of repetitions needed to reconstruct the entire spectrum with high probability is in $\widetilde{O}(1/\min_j |\psi_j|^2)$ and, thus,

$$N_{\exp,QPE} \in \widetilde{O}\left[\frac{M}{\min_j |\psi_j^0|^2}\left(\frac{\Gamma}{\epsilon}\right)^{3/2}\right]. \tag{36}$$

Note that the performance of this algorithm can be improved using optimized simulation methods and the use of amplitude amplification, but we forgo these optimizations here since they are less appropriate for experiments in the NISQ era.

Now let us consider the analogous problem for estimating the corresponding eigenvalue using VQPE. The first step involves learning the matrices $H_{j,k}$ and $S_{j,k}$. In order to do so, we need to apply e^{-iHt_j} for all j considered. There are of course a host of quantum simulation algorithms that can be employed to perform this simulation. For simplicity, we consider the second-order Trotter-Suzuki formula. Consider Eq. (32). From this expression, we can perform e^{-iHt_j} using $O[M(\Gamma t_{\max})^{3/2}/\sqrt{\epsilon_{TS}}]$ operator exponentials, where $|t_j| \leq t_{\max}$ for all t_j. Thus we can use this circuit and the Hadamard test to compute the real and imaginary components of $\langle\phi_j|\phi_k\rangle$ within an error of at most ϵ with high probability (using the Chernoff bound to justify the high-probability statement) using

$$N_{\exp,S_{ij}} \in \widetilde{O}\left(\frac{M(\Gamma t_{\max})^{3/2}}{\epsilon^{5/2}}\right). \tag{37}$$

Similarly, using the arguments of [49] we have that the number of operator exponentials needed to approximate the mean energy within the same error requirements is in

$$N_{\exp,H_{ij}} \in \widetilde{O}\left(\frac{M(\sum_j |\alpha_j|)^2(\Gamma t_{\max})^{3/2}}{\epsilon^{5/2}}\right). \tag{38}$$

Next, using the fact that the induced 2-norm is at most N_T times the max-norm, then if $\widetilde{S}$ and $\widetilde{H}$ are the approximated matrices, we need to take $\epsilon \to \epsilon/N_T$ in order to ensure that the total error in the reconstructed matrix (as measured by the operator norm) is at most ϵ. This implies that the number of exponentials needed to reconstruct both

matrices within the required error budget is in

$$N_{\exp,S} \in \widetilde{O}\left(\frac{N_T^{9/2}M(\Gamma t_{\max})^{3/2}}{\epsilon^{5/2}}\right), \tag{39}$$

$$N_{\exp,H} \in \widetilde{O}\left(\frac{N_T^{9/2}M(\sum_j |\alpha_j|)^2(\Gamma t_{\max})^{3/2}}{\epsilon^{5/2}}\right). \tag{40}$$

Then, from the Hellman-Feynman theorem, it follows that the eigenvalues of both matrices lie within an ϵ neighborhood of the original eigenvalues. Next, using standard matrix inequalities for the error in the matrix inverse [50] (under the assumption that ϵ is less than the minimum eigenvalue gap), it follows that the error in any eigenvalue reconstructed from Eq. (2) has error in $O(\|H\|\|S^{-1}\|^2\epsilon)$. Thus, if we wish the error in the reconstructed eigenvalues to be at most ϵ then we finally need to take $\epsilon \mapsto \epsilon/\|H\|\|S^{-1}\|^2$. This, combined with the bound that $\|H\| \leq \sum_j |\alpha_j|$, leads to a final complexity scaling for the number of exponentials needed to reconstruct the eigenvalues within an error of at most ϵ with high probability:

$$N_{\exp} \in \widetilde{O}\left(\frac{M(N_T \sum_j |\alpha_j|)^{9/2}\|S^{-1}\|^5(\Gamma t_{\max})^{3/2}}{\epsilon^{5/2}}\right). \tag{41}$$

Asymptotically, this analysis suggests that VQPE may be favorable in cases where

$$\min_j |\psi_j^0|^2 \ll \frac{\epsilon}{(N_T \sum_j |\alpha_j|)^{9/2}\|S^{-1}\|^5}, \tag{42}$$

due to the potential benefits that could arise from inferring having to sample low probability events using phase estimation. However, such asymptotic analysis does not conclusively show which method will be superior as both are upper bounds that are potentially loose. For this reason, numerical studies such as the one undertaken in this paper are essential for gauging the actual performance of these algorithms.

For Toeplitz matrices such as in Eqs. (18) and (25) to be diagonal in a Fourier basis, they must have additional circulant matrix structure satisfying

$$S_{j,N_T} = S_{j+1,0}, \tag{43}$$

which corresponds to an aliasing condition whereby all energy eigenvalues E_N are an integer multiple of a base energy,

$$E_N/\omega_1 \in \mathbb{Z}. \tag{44}$$

An accurate approximation of this condition requires a very small value of Δt and thus a large number of ancilla

qubits in the QPE circuit and large N_T. There is a formal equivalence between QPE and unitary VQPE when this condition is exactly satisfied, although this occurs beyond the intended small-N_T operating regime of any VQPE variant. Practically, VQPE and QPE remain distinct in that VQPE uses measurements of one ancilla qubit to statistically estimate overlap matrix elements, while QPE uses measurements of many ancilla qubits to sample from approximate eigenpairs.

Besides requiring fewer ancilla qubits, the main benefit of VQPE over QPE in the small-N_T regime is its exact diagonalization of a generalized eigenvalue problem rather than relying on approximate diagonalization within a Fourier basis. The overlap matrix of this eigenvalue problem has a numerical rank corresponding to the size of the support space, and this reduced rank is not reliably exposed by the diagonal elements of the overlap matrix in the Fourier basis. As a result, VQPE achieves a higher effective energy resolution than QPE for small N_T values and an even smaller support space by exploiting a SVD-based projection of the eigenvalue problem into a numerically relevant subspace.

In Fig. 5, we plot two different resource (for LiH discussed further in Sec. III) estimates that exemplify the near-term benefits of VQPE and the long-term benefits of QPE, here shown as the Heisenberg limit. VQPE is able to achieve substantially higher accuracy than idealized QPE for the same number of time steps per circuit; thus, it is better suited for utilizing near-term hardware with shorter circuit depths and fewer qubits available for use as ancillae. However, QPE utilizes its deeper circuits to achieve Heisenberg-limited energy resolution, which is more efficient in overall run time for achieving high accuracy. But, as with modern parallel algorithms, it is clear that each matrix element can be calculated independently and can thus be run as a parallel algorithm over multiple quantum computers. The expected QPE run time is also amplified by $|\langle \Psi_0 | 0 \rangle|^{-2} \approx 1.02$ to account for the probability of collapsing into a state other than the ground state, which is a negligible overhead in this example. Other than finite-sampling costs and errors, this example does not consider other sources of error such as Trotterization or quantum circuit noise, which can further degrade the accuracy of both QPE and VQPE. For example, Fig. 6 depicts a Δt-dependent error floor induced by Trotterization errors (for the transverse field Ising model defined in Eq. (45) and discussed further in Sec. III) that remains overwhelmed by sampling errors in VQPE. With shorter circuit depths and fewer ancilla qubits, VQPE is less susceptible than QPE to these extrinsic sources of error.

H. Inclusion of other time-evolved states

In this work we focus on real-time evolution. However, in practice any time-evolved state shown in Fig. 7,
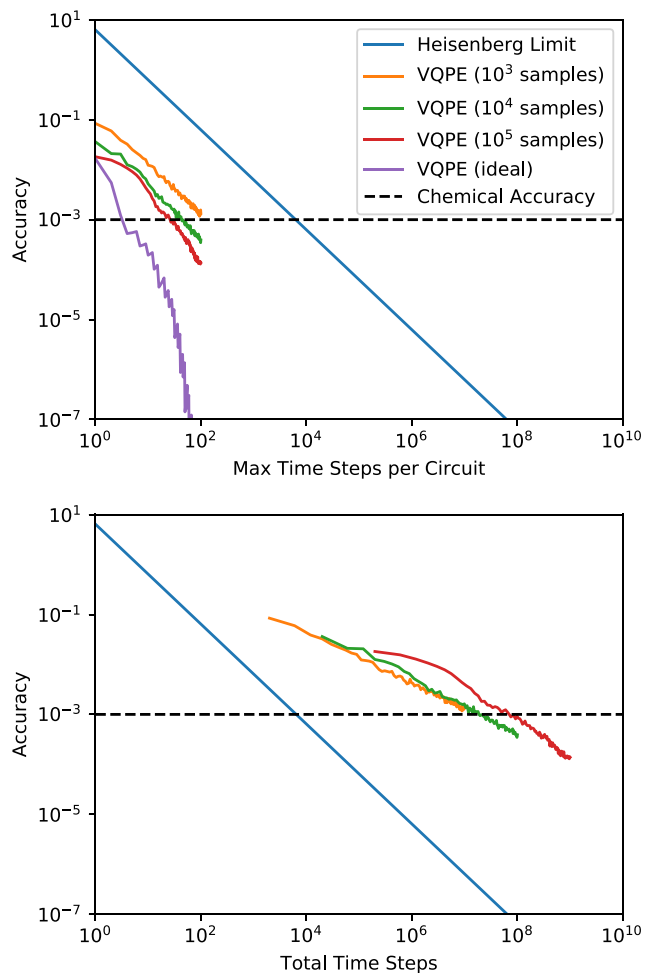

FIG. 5. Accuracy comparison between VQPE and the Heisenberg limit for the ground-state energy of LiH. Varying numbers of samples per expectation value are shown for VQPE alongside the idealization of VQPE without finite-sampling errors. We use SVD cutoffs of 2, 0.9, and 0.3 respectively corresponding to 10^3, 10^4, and 10^5 samples and 10^{-6} for the ideal case. All methods use $\Delta t = 0.5$. The top and bottom plots show the maximal evolution time (i.e., how many time steps need to be included for energy convergence), related to the circuit depth, and the total evolution time, i.e., the sum of all time segments needed.

real or imaginary, could be included in the expansion set for solving the generalized eigenvalue equation. Moreover, any unitaries $f(H;j)$ that commute with the Hamiltonian would be viable candidates to produce an expansion set with the following important property: all the states in the expansion set share an identical support space with respect to the Hamiltonian. Thus, any expansion $|\Phi_{j,I}\rangle = f(H;j)|\Psi_I\rangle$ formed by any such unitary will eventually cover the whole support space without introducing more eigenstates. Recent algorithms such as QLanczos [5, 37], quantum filter diagonalization [22], quantum Krylov approaches [23], and quantum power methods [52,53] have all proposed using real- or imaginary-time evolution
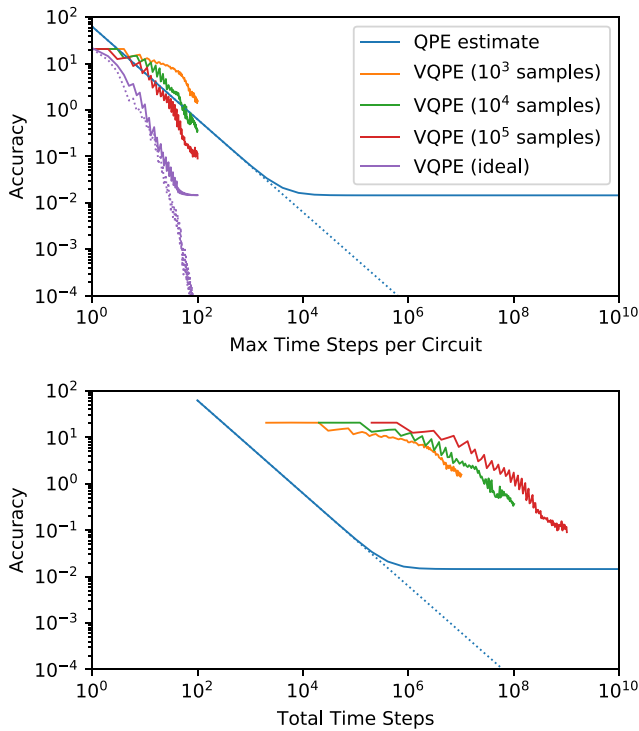
FIG. 6. Accuracy comparison between VQPE and QPE for the ground-state energy of the ten-qubit transverse field Ising model with first-order Trotterization and $\Delta t = 0.05$. We use SVD cut-offs of 1, 0.6, and 0.2 respectively corresponding to 10^3, 10^4, and 10^5 samples and 10^{-5} for the ideal case. Dotted lines show QPE and idealized VQPE results without Trotter errors. The reference state, $|\Phi\rangle = |\phi\rangle^{\otimes 10}$, is a product state the maximizes overlap with the exact ground state, $|\langle \Psi_0 | \Phi \rangle|^2 \approx 0.014$ for $|\phi\rangle \approx 0.979|0\rangle + 0.205|1\rangle$.

in various ways to generate states for creating a wave function ansatz. The real- or imaginary-time evolution plane is illustrated in Fig. 7, along with the states used in these various algorithms. As long as the support space is retained in the time evolution and the states are linearly independent, efficient eigenvalue extraction will be possible. There are several benefits derived from using only the real-time axis that include only a linear number of measurements for setting up the generalized eigenvalue problem, almost no optimization of parameters for determining the time evolution, and a well-conditioned problem for extracting information about the ground and low-lying excited states.

III. METHODOLOGICAL DETAILS

We turn now to presenting concrete numerical data for Hamiltonians from condensed matter physics and quantum chemistry. For practical applications, the most important quality of the VQPE approach is doubtlessly its compactness. By this, we refer to the number of variational parameters required to obtain accurate eigenvalues. The number of variational parameters is $(N_T + 1)N_R$, where N_R
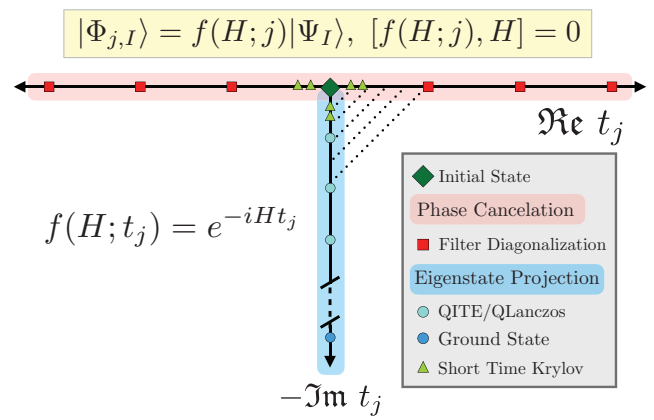

FIG. 7. A schematic picture of recent eigenvalue algorithms relying on time evolution proposed for quantum computers. The plane corresponds to the complex time plane, with the horizontal axis being real time and the vertical axis imaginary time. The colored dots denote expansion states $|\phi_{j,I}\rangle$. The operator $f(H; j)$ denotes any unitary that commutes with the Hamiltonian, and thus the expansion states formed by applying this unitary will remain in the support space and can be used to extract eigenstates. The real-time axis is where the phase cancelation picture applies. On the imaginary-time axis, the coefficients of the excited states are suppressed as time increases, leading to the ground state. The dotted lines indicate that short time imaginary- and real-time states will be very similar, as argued in the short time Krylov method [51]. We emphasize that combinations of states generated on the entire real- or imaginary-time plane have essentially the same support space and can be used as a basis to solve the generalized eigenvalue problem.

is the number of reference states (for all examples shown below, we use a single reference $N_R = 1$), and we choose chemical accuracy (approximately 1.6 mHa) as our target accuracy. Thus, we apply VQPE to several many-body systems of different complexity, taken from the fields of quantum chemistry and condensed matter physics: several small molecules, including weakly correlated ones, LiH (3-21g basis set [54], 11 orbitals, and 4 electrons), H_2O (cc-pVDZ basis set [55], 24 orbitals, and 10 electrons), N_2 (cc-pVDZ basis set [55], 28 orbitals, and 14 electrons), two moderately correlated, linear H_6 (STO-6g basis [56], bond length 1.5 Å, six orbitals, and six electrons) and C_2 (cc-pVDZ basis set [55], 28 orbitals, and 12 electrons), and one strongly correlated Cr_2 (def2-SVP basis set [57], 30 orbitals, and 24 electrons), and the transverse field Ising model (two site on hardware, ten site on simulator). We simulate the VQPE calculation for the molecular systems classically, while running the Ising model calculations on IBM's quantum hardware and simulator. Further details, such as molecular geometries, can be found in the Supplemental Material [42].

For the classical simulations, we perform time evolution using exact diagonalization dynamics based on the Lanczos algorithm [58–60]. In the case of all molecular

systems except for Cr_2, the small basis sets allow us to perform the time evolution including all electrons and all orbitals explicitly, while for Cr_2 we restrict the simulation to the widely studied 30 orbital active space. Still, only for LiH and H_6 is it computationally feasible to perform exact dynamics. For all other systems, we must truncate the Hilbert space, considering only a subset of all Slater determinants in the corresponding active spaces. This presents an approximation, and we perform a finite-size effect study by comparing the dynamics of progressively larger truncations, from one thousand to one million determinants, which we show for Cr_2, the most complicated of the systems considered. The determinants of each truncation are chosen using the adaptive sampling configuration interaction (ASCI) algorithm [30,61,62]. This is an iterative selected configuration interaction approach, which explores the Hilbert space and identifies the most important determinants for a ground-state approximation, providing highly accurate yet moderately sized truncations. Thus, by optimizing the truncated spaces with respect to the ground-state description, we can reliably infer the full Hilbert space limit from our finite size calculations. In practice, we determined a one million determinant Hilbert space truncation with ASCI for Cr_2 (10^5 determinants for H_2O, C_2, and N_2), and built all smaller truncations of size N from this one, by picking the N determinants with the largest coefficients in the one million (10^5 for H_2O, C_2, and N_2) determinant ground-state wave function. For all molecular systems, we determined the optimal time step size Δt using the strategy described in the theory section (see Fig. 9), and use the Hartree-Fock state as reference state $|\Psi_0\rangle$. For these classical simulations, we did not exploit the unitary formulation of VQPE, instead solving the generalized eigenvalue equation (2).

For the simulations on IBM's quantum hardware and simulator, we consider the transverse field Ising model (TFIM) with open boundary conditions, defined by the Hamiltonian

$$H = -J\left(\sum_i Z_i Z_{i+1} + h\sum_i X_i\right), \qquad (45)$$

where X, Y, and Z above and henceforth denote the Pauli operators, $J = 1$ corresponds to the spin coupling and $h = 2$ the external field. In these simulations, we choose a time step size $\Delta t = 0.05$ and approximate the time evolution operator via first-order Trotterization [8]. Each run is performed with 8192 shots. For all hardware data shown, $|\Psi_0\rangle$ readout error mitigation is performed. We emphasize that other time evolution methods could be substituted for Trotterization, such as variational approaches with constant or adaptive circuits [63–66], without changing the nature of the algorithm, and may result in substantially reduced circuit depth.
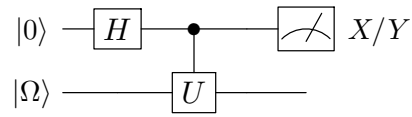

FIG. 8. Hadamard test circuit, which computes the real and imaginary parts of the overlap matrix, where $|\Omega\rangle$ denotes the system and the top qubit is the ancilla. Here $U = e^{-itH}$.

As discussed in the theory section, the Toeplitz property of the overlap matrix means that the number of measurements is linear with the number of time steps $N_T + 2$. Furthermore, we take advantage of the unitary formulation of VQPE to measure only the real and imaginary components of the overlap matrix explicitly, not the Hamiltonian matrix elements. To measure the overlap matrix on quantum hardware, we follow the Hadamard test approach described in Ref. [22], where one additional ancilla qubit is required, as shown in Fig. 8. The price of this simplicity is the requirement for an efficient construction of a controlled time evolution operator. The general construction of controlled unitaries is challenging. Other methods have been developed to measure the required overlap matrix elements in potentially more NISQ-appropriate ways, such as the use of direct measurement [67], shadow tomography [68–71], as well as machine learning approaches [72]. For our small demonstrations, we use the Hadamard test and are able to generate short depth controlled unitary circuits with the use of QSearch [73] for circuit synthesis, where a unitary, U, is read in and a circuit with minimal controlled-NOT (CNOT) depth is constructed while keeping the similarity with the synthesized unitary (U_s) below some threshold σ (here $\sigma = 10^{-10}$ is the Hilbert-Schmidt inner product between the conjugate transpose of U and U_s). For the two-qubit transverse field Ising model with one ancilla, a total of six CNOT gates are needed, and for the three-qubit system, between 14–16 CNOT gates are needed, regardless of the time step [74]. The overlap matrix elements are obtained by measuring $S_{0j} = \langle X\rangle_j + i\langle Y\rangle_j$ on the ancilla qubit for each time step, where j denotes the number of time steps after which we make the measurement. As noted in the theory section, because this has the structure of a time correlation function, the other rows of the matrix can be filled out accordingly, with one extra measurement needed for the unitary formulation of the generalized eigenvalue equation.

IV. RESULTS AND DISCUSSION

A. Finding the optimal time step size

Here we exemplify the strategy for finding optimal time step sizes on LiH. Figure 9 shows the convergence of the ground state (solid lines) and first excited state included in the support space (dashed lines) for several time step
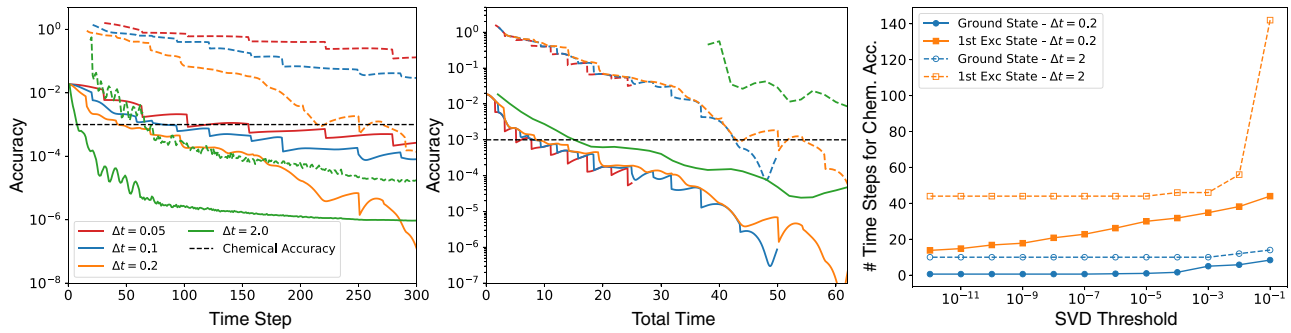
FIG. 9. Ground- (solid) and excited-state (dashed) energies of LiH as a function of the time step (left) and total simulation time (middle). The initial reference vector is the Hartree-Fock state. We use the 3-21g basis set that has 4 electrons and 11 orbitals. The SVD cutoff is 10^{-1} for both the left and middle plots and the legend is the same for both. The right figure shows the total time to reach chemical accuracy as a function of SVD.

sizes Δt, as a function of the number of expansion states (left) and the total simulation time (middle). The left panel in Fig. 9 shows how the steplike behavior appearing for the smallest time step size ($\Delta t = 0.05$) decreases with increasing Δt. The convergence as a function of simulation time (middle) clearly shows the implicit balance to be made between simulations with short total time and a small number of expansion states. The $\Delta t = 2.0$ simulation (green curves) achieves chemical accuracy with the smallest number of expansion states, i.e., with the most compact variational ansatz. However, this comes at the

cost of a significantly larger total simulation time, at least twice as large as the simulations with smaller time step sizes. For practical implementations on NISQ devices, the marginal decrease in the number of expansion states is not worth the drastic increase in total simulation time. This underlies the importance of finding a time step size that is long enough, but not too long.

We further note that the excited-state convergence is worse than the ground state in general, requiring for LiH about twice as many expansion states and total simulation time. This is likely related to the magnitude of the
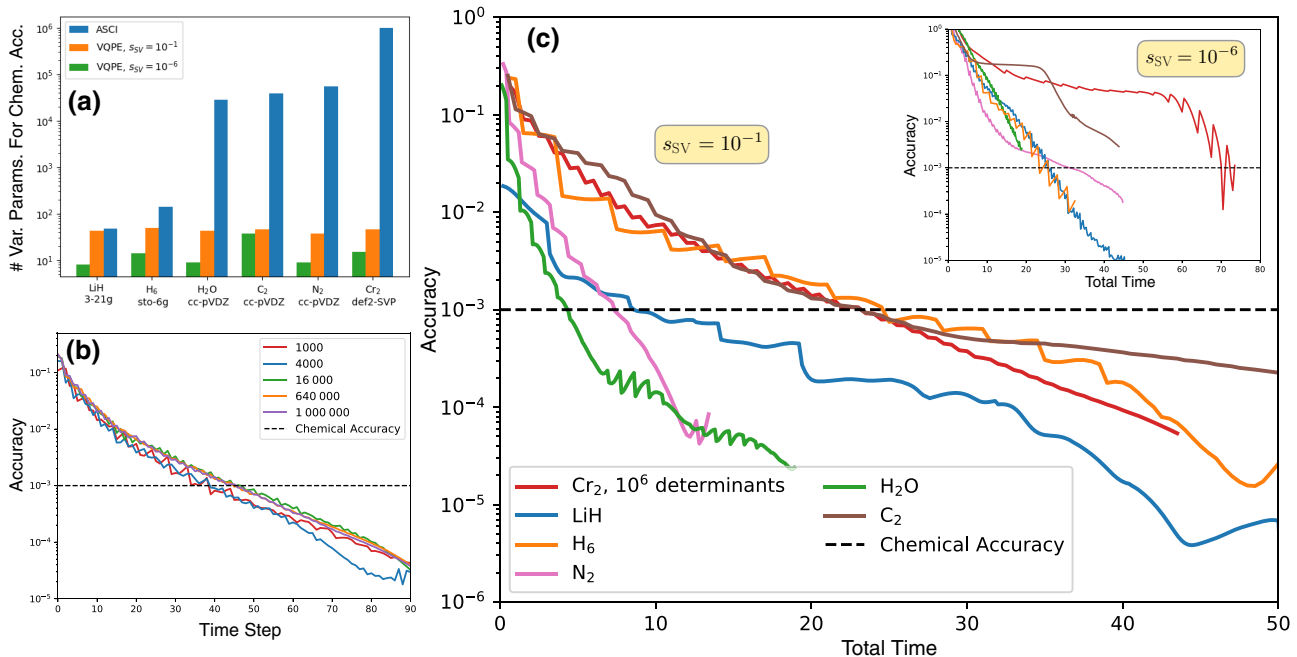

FIG. 10. (a) Number of variational parameters to reach chemical accuracy using ASCI (a classical selected configuration interaction approach), and the VQPE algorithm with different s_{SV} values. These parameters correspond to determinants in ASCI, and to expansion states in VQPE. (b) Comparison of convergence for all different Hilbert space sizes of Cr_2. (c) Convergence for all the molecules as a function of total time ($s_{SV} = 10^{-1}$) with the inset showing convergence of the first excited state that has nonzero overlap with the reference state ($s_{SV} = 10^{-6}$).

expansion coefficients of the Hartree-Fock reference vector in the basis of Hamiltonian eigenstates.

The rightmost plot of Fig. 9 shows the total simulation time for ground-state convergence to chemical accuracy using $\Delta t = 0.2$ for different singular value cutoffs s_{SV}. Clearly, the larger the value of s_{SV}, the longer the simulation time becomes, which relates to the notions of linear independence, which we elaborated upon in the theory section [see Eq. (27)].

B. Classical simulation of molecular systems

Figure 10 shows the power of the VQPE method on molecular systems with varying degrees of electronic correlation, using a singular value cutoff of 10^{-1}. As discussed in Sec. III, for all systems except for H_6 and LiH, we use a truncated Hamiltonian obtained with ASCI. In panel (b) we show that this approximation has no effect on our conclusions, since vastly different truncation sizes present the same convergence even though the actual ground-state energies are very different. Panel (a) in this figure shows the number of variational parameters to reach chemical accuracy using ASCI (a classical selected configuration interaction approach), and the VQPE algorithm with different s_{SV} values. These parameters correspond to determinants in ASCI, and to expansion states in VQPE. Given the differences in complexity with respect to the classical algorithm, it is remarkable how many of these systems converge to chemical accuracy at essentially identical rates, as shown in panel (c). Note that the differences in convergence times are likely connected to the different levels of correlation in these systems: the existence of low-lying excited states (i.e., more strongly correlated) means longer convergence time because of a smaller energy gap. Even with these differences, for all systems studied, we reach chemical accuracy with as little as 50 variational parameters and about 30 a.u. of total simulation time. This exceptionally low number of variational parameters is particularly striking when considering that classically, e.g., with ASCI, it is necessary to include 10^5–10^6 parameters (determinants) to reach the same energies. We emphasize that part of the dramatic efficiency of VQPE demonstrated here comes from having a good reference state (for all molecular systems, the Hartree-Fock state) that has small overlap with the majority of Hilbert space.

The inset in panel (c) shows the convergence of the first excited state in the support space. Here it is necessary to use a significantly smaller SVD cutoff ($s_{SV} = 10^{-6}$) to converge in reasonable simulation times. Even with this drastic restriction in the allowable noise, the excited-state energies converge significantly slower than the ground-state energies. Still, they should be possible to access efficiently with reference states tailored to describe excited states [75–79].
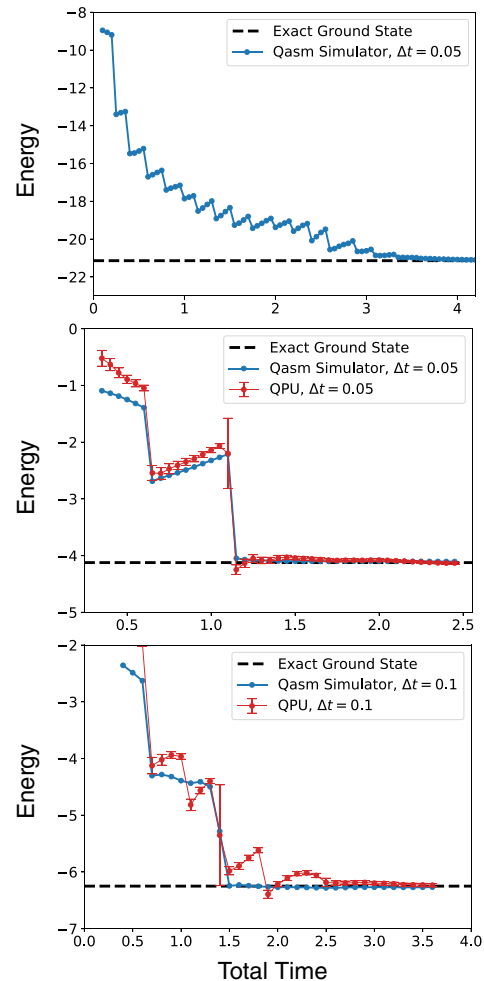

FIG. 11. The top figure shows the ground-state energy as a function of time step for the ten-qubit transverse field Ising model run on IBM's qasm simulator with $s_{SV} = 10^{-1}$. The middle figure shows results for hardware runs (IBM's Paris machine) for a two-qubit transverse field Ising model. For both the top and middle plots, VQPE and the Trotter step size are the same, $\Delta t = 0.05$. Classical processing in the middle plot (for both QPU and simulator data) used SVD cutoff $s_{SV} = 2$. The bottom figure shows results for hardware runs (IBM's Montreal machine) for a three-qubit transverse field Ising model, VQPE $\Delta t = 0.1$, Trotter step size $\Delta t = 0.05$, and $s_{SV} = 1$.

C. Quantum simulation of the transverse field Ising model

Finally, Fig. 11 shows our results using IBM's QPUs (two- and three-qubit TFIM) and qasm simulator (ten-qubit TFIM), compared to the exact ground-state energy. Our initial state corresponds to a ferromagnetically ordered chain where all spins point down in the Z direction, which corresponds to the $|000 \cdots\rangle$ state in the computational basis. For processing the hardware data, we use a larger s_{SV} cutoff ($s_{SV} = 2$ and $s_{SV} = 1$ for the two- and three-qubit systems, respectively) to account for larger noise. Error

bars on the hardware data correspond to the variance of the average ground-state energy at each time point over ten individual hardware runs on IBM's Paris machine, using qubits 0, 1, and 2 for the $2 + 1$ (system + ancilla) experiment and using qubits 0, 1, 2, and 3 on IBM's Montreal machine for the $3 + 1$ experiment. Each hardware run used 8192 shots per matrix element. Figure 11 shows convergence to the exact ground-state energy for both simulator and hardware data. The saw-tooth behavior is indicative of linear dependence in the expansion states, as discussed and shown in Secs. II C and IV A.

V. CONCLUSIONS

Here we analyze the theoretical underpinnings of VQPE algorithms as well as the effects of statistical noise on its performance. We have presented evidence for VQPE as a particularly compact and natural algorithm for the NISQ era, showing that with the unitary formulation of the generalized eigenvalue equation and the prevailing Toeplitz structure of the overlap matrix, only a linear number of measurements are needed. We provided a heuristic for choosing an optimal time step size, which balances two opposing factors: on the one hand, generating linearly independent expansion states with each new time evolution, such that the variational ansatz becomes compact, which sets a lower bound to the time step size, and on the other hand, minimizing the total simulation time, as required by NISQ hardware. We have also demonstrated the effect of statistical noise on the optimal time step and final accuracy of the energies, showing that simple regularization techniques suffice to mitigate the effects of noise.

We have exemplified the power of the VQPE approach on a wide range of molecules of different complexities, simulating the algorithm classically, as well as the transverse field Ising model on IBM's quantum simulator and hardware. On a traditional example of strong correlation, Cr_2, our results suggest that VQPE achieves comparable accuracy to state-of-the-art classical simulations with orders of magnitude fewer variational parameters (approximately 50 versus 10^6). This compactness, together with their NISQ compatibility (with respect to total simulation time, number of measurements, and noise resilience), marks VQPE approaches as some of the most promising platforms to realize the long sought quantum computation goal of performing many-body simulations beyond the reach of classical computations.

ACKNOWLEDGMENTS

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

[2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, and D. A. Buell *et al.*, arXiv preprint arXiv:2004.04174 (2020).

[3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, A. Bengtsson, S. Boixo, M. Broughton, and B. B. Buckley *et al.*, arXiv preprint arXiv:2010.07965 (2020).

[4] M. Cerezo, A. Arrasmith, R. Babbush, S. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, arXiv preprint arXiv:2012.09265 (2020).

[5] M. Motta, C. Sun, A. T. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. Brandao, and G. K. L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

[6] W. M. Kirby and P. J. Love, arXiv preprint arXiv:2012.07171 (2020).

[7] M. Nielson and I. Chuang, (2006).

[8] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[9] A. Smith, M. Kim, F. Pollmann, and J. Knolle, Simulating quantum many-body dynamics on a current digital quantum computer, Npj Quantum Inf. **5**, 1 (2019).

[10] B. Bauer, S. Bravyi, M. Motta, and G. Kin-Lic Chan, quantum algorithms for quantum chemistry and Quantum

materials Science, Chem. Rev. **120**, 12685 (2020), pMID: 33090772.

[11] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, quantum chemistry in the age of Quantum Computing, Chem. Rev. **119**, 10856 (2019), pMID: 31469277.

[12] D. S. Abrams and S. Lloyd, Simulation of Many-Body Fermi Systems on a Universal Quantum Computer, Phys. Rev. Lett. **79**, 2586 (1997).

[13] D. S. Abrams and S. Lloyd, Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors, Phys. Rev. Lett. **83**, 5162 (1999).

[14] A. Y. Kitaev, Quantum computations: Algorithms and error correction, Russ. Math. Surveys **52**, 1191 (1997).

[15] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular Energies, Science **309**, 1704 (2005).

[16] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: Methods and implementations, Rep. Prog. Phys. **83**, 054401 (2020).

[17] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New J. Phys. **18**, 023023 (2016).

[18] W. J. Huggins, J. Lee, U. Baek, B. O'Gorman, and K. B. Whaley, arXiv preprint arXiv:1909.09114 (2019).

[19] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 1 (2018).

[20] D. Wierichs, C. Gogolin, and M. Kastoryano, Avoiding local minima in variational quantum eigensolvers with the natural gradient optimizer, Phys. Rev. Res. **2**, 043246 (2020).

[21] D. J. Tannor, *Introduction to Quantum Mechanics: a Time-Dependent Perspective* (University Science Books, Melville NY, USA, 2007).

[22] R. M. Parrish and P. L. McMahon, arXiv preprint arXiv:1909.08925 (2019).

[23] N. H. Stair, R. Huang, and F. A. Evangelista, A multireference quantum krylov algorithm for strongly correlated Electrons, J. Chem. Theory Comput. **16**, 2236 (2020).

[24] D. Neuhauser, Bound state eigenfunctions from wave packets: Time-energy resolution, J. Chem. Phys. **93**, 2611 (1990).

[25] D. Neuhauser, Circumventing the heisenberg principle: A rigorous demonstration of filter-diagonalization on a liCN model, J. Chem. Phys. **100**, 5076 (1994).

[26] V. A. Mandelshtam, FDM: The filter diagonalization method for data processing in NMR experiments, Prog. Nucl. Magn. Reson. Spectrosc. **38**, 159 (2001).

[27] D. M. Ceperley, Path integrals in the theory of condensed helium, Rev. Mod. Phys. **67**, 279 (1995).

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

[29] E. Koch, 'The LDA+DMFT approach to strongly correlated materials, (Verlag des Forschungszentrum Jülich, 2011) Chap. 8, p. 235.

[30] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, A deterministic alternative to the full configuration interaction quantum monte carlo method, J. Chem. Phys. **145**, 044112 (2016).

[31] P. W. Atkins and R. S. Friedman, *Molecular Quantum Mechanics, 5th Edition* (Oxford University Press, Oxford, UK, 2011).

[32] W. J. Huggins, J. Lee, U. Baek, B. O'Gorman, and K. B. Whaley, New J. Phys. (2020).

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

[34] R. M. Parrish, E. G. Hohenstein, P. L. McMahon, and T. J. Martínez, Quantum Computation of Electronic Transitions Using a Variational Quantum Eigensolver, Phys. Rev. Lett. **122**, 230401 (2019).

[35] T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, and J. R. McClean, Increasing the Representation Accuracy of Quantum Simulations of Chemistry without Extra Quantum Resources, Phys. Rev. X **10**, 011004 (2020).

[36] M. Urbanek, D. Camps, R. Van Beeumen, and W. A. de Jong, Chemistry on quantum computers with virtual quantum subspace Expansion, J. Chem. Theory Comput. **16**, 5425 (2020).

[37] K. Yeter-Aydeniz, R. C. Pooser, and G. Siopsis, Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and lanczos algorithms, Npj Quantum Inf. **6**, 1 (2020).

[38] H. D. Meyer and S. Pal, A band-lanczos method for computing matrix elements of a resolvent, J. Chem. Phys. **91**, 6195 (1989).

[39] M. R. Wall and D. Neuhauser, Extraction, through filter diagonalization, of general quantum eigenvalues or classical normal mode frequencies from a small number of residues or a short-time segment of a signal. I. Theory and application to a quantum-dynamics model, J. Chem. Phys. **102**, 8011 (1995).

[40] After all, time evolution does not change the absolute values of the expansion coefficients of $|\Psi_0\rangle$ in the Hamiltonian eigenbasis.

[41] One could consider changing the perfect time step size for this harmonic Hamiltonian according to the singular value truncation threshold s_{SV}, effectively changing the size of the support space. This would preclude, however, resolving the eigenstate energies to better accuracy than s_{SV}, which, as shown in Fig. 2, is not a true lower bound for the possible accuracy, even in the presence of noise.

[42] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PRXQuantum.3.020323 for additional information about the calculations and the used molecular geometries.

[43] N. Wiebe and C. Granade, Efficient Bayesian Phase Estimation, Phys. Rev. Lett. **117**, 010503 (2016).

[44] R. D. Somma, Quantum eigenvalue estimation via time series analysis, New J. Phys. **21**, 123025 (2019).

[45] L. Lin and Y. Tong, arXiv preprint arXiv:2102.11340 (2021).

[46] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, Phys. Rev. Lett. **96**, 010401 (2006).

[47] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).

[48] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, arXiv preprint arXiv:1912.08854 (2019).

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

[50] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, UK, 2012).

[51] N. H. Stair, R. Huang, and F. A. Evangelista, A multireference quantum krylov algorithm for strongly correlated Electrons, J. Chem. Theory Comput. **16**, 2236 (2020).

[52] O. Kyriienko, Quantum inverse iteration algorithm for programmable quantum simulators, Npj Quantum Inf. **6**, 1 (2020).

[53] K. Seki and S. Yunoki, Quantum power method by a superposition of time-evolved States, PRX Quantum **2**, 010333 (2021).

[54] J. S. Binkley, J. A. Pople, and W. J. Hehre, Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements, J. Am. Chem. Soc. **102**, 939 (1980).

[55] T. H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. **90**, 1007 (1989).

[56] W. J. Hehre, R. F. Stewart, and J. A. Pople, Self-Consistent molecular-Orbital methods. I. Use of Gaussian expansions of slater-Type atomic Orbitals, J. Chem. Phys. **51**, 2657 (1969).

[57] F. Weigend and R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to rn: Design and assessment of accuracy, Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[58] T. J. Park and J. Light, Unitary quantum time evolution by iterative lanczos reduction, J. Chem. Phys. **85**, 5870 (1986).

[59] S. R. Manmana, A. Muramatsu, and R. M. Noack, in *AIP Conference Proceedings*, (American Institute of Physics, 2005), Vol. 789, p. 269.

[60] C. Kollath, A. M. Läuchli, and E. Altman, Quench Dynamics and Nonequilibrium Phase Diagram of the Bose-Hubbard Model, Phys. Rev. Lett. **98**, 180601 (2007).

[61] N. M. Tubman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, arXiv preprint arXiv:1808.02049v1 (2018).

[62] N. M. Tubman, C. D. Freeman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI Method, J. Chem. Theory Comput. **16**, 2139 (2020).

[63] Y. Li and S. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, Phys. Rev. X **7**, 021050 (2017).

[64] Y.-X. Yao, N. Gomes, F. Zhang, C.-Z. Wang, K.-M. Ho, T. Iadecola, and P. P. Orth, Adaptive variational quantum dynamics Simulations, PRX Quantum **2**, 030307 (2021).

[65] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum **3**, 191 (2021).

[66] Z.-J. Zhang, J. Jinzhao, X. Yuan, and M.-H. Yung, arXiv:2011.05283 (2020).

[67] K. Mitarai and K. Fujii, Methodology for replacing indirect measurements with direct measurements, Phys. Rev. Res. **1**, 013006 (2019).

[68] S. Aaronson, Shadow tomography of quantum states, SIAM J. Comput. **49**, STOC18 (2019).

[69] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[70] W. J. Huggins, B. A. O'Gorman, C. Neil, N. C. Rubin, P. Roushan, D. R. Reichman, R. Babbush, and J. Lee, arXiv preprint arXiv:2106.16235 (2021).

[71] C. L. Cortes and S. K. Gray, arXiv:2109.06868 (2021).

[72] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, Learning the quantum algorithm for state overlap, New J. Phys. **20**, 113022 (2018).

[73] M. G. Davis, E. Smith, A. Tudor, K. Sen, I. Siddiqi, and C. Iancu, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2020), p. 223.

[74] L. Bassman, R. Van Beeumen, E. Smith, E. Younis, C. Iancu, and W. De Jong, Constant-depth circuits for dynamic simulations of materials on quantum computers, Mater. Theory **6**, 1 (2022).

[75] P. Phillipson and R. Mulliken, J. Chem. Phys. **28**, 1248 (1958).

[76] A. T. Gilbert, N. A. Besley, and P. M. W. Gill, Self-consistent field calculations of excited states using the maximum overlap method (MOM), J. Phys. Chem. A **112**, 13164 (2008).

[77] J. A. Shea and E. Neuscamman, Communication: A mean field platform for excited state quantum chemistry, J. Chem. Phys. **149**, 081101 (2018).

[78] J. A. Shea, E. Gwin, and E. Neuscamman, A generalized variational principle with applications to excited state mean field Theory, J. Chem. Theory Comput. **16**, 1526 (2020).

[79] D. Hait and M. Head-Gordon, arXiv preprint arXiv:2103.04573 (2021).

[80] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, and C. Chen *et al.*, Accessed on: Mar **16** (2019).