

# Ground- and excited-state energies with analytic errors and short time evolution on a quantum computer

Timothy Stroschein , Davide Castaldo , and Markus Reiher <sup>\*</sup>

Department of Chemistry and Applied Biosciences, *ETH Zurich*, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland



(Received 22 July 2025; accepted 16 July 2026; published 14 August 2026)

Accurately solving the Schrödinger equation remains a central challenge in computational physics, chemistry, and materials science. Here, we propose an alternative eigenvalue problem based on a system's autocorrelation function, avoiding direct reference to a wave function. In particular, we develop a rigorous approximation framework that enables precise frequency estimation from a finite number of signal samples. Our analysis builds on results involving prolate spheroidal wave functions and yields error bounds that reveal a sharp accuracy transition governed by the observation time and spectral density of the signal. These results are very general and thus carry far. As one important example application, we consider the quantum computation for molecular systems. By combining our spectral method with a quantum subroutine for signal generation, we define quantum prolate diagonalization (QPD)—a hybrid classical-quantum algorithm. QPD simultaneously estimates ground- and excited-state energies within chemical accuracy at the Heisenberg limit. An analysis of different input states demonstrates the robustness of the method, showing that high precision can be retained even under imperfect state preparation.

DOI: [10.1103/zy63-92c4](https://doi.org/10.1103/zy63-92c4)

## I. INTRODUCTION

The efficient solution of eigenvalue problems for large matrices [1–3] is a foundational task in many areas of natural [4,5] and social sciences [6,7]. In recent years, significant efforts have been directed toward mapping an abundance of eigenvalue problems onto quantum Hamiltonians [8–11], enabling the exploitation of quantum computers for their solution. This surge of interest stems from the initial insight that a programmable quantum device could prevail over the exponential complexity that hampers classical simulations of quantum systems [12,13]. Looking more closely at problems in (bio)chemistry [14–16], catalysis [17], and condensed matter physics [18] and materials science [19,20], realizing these advantages [21] remains challenging due to limitations in current quantum hardware. In particular, despite progress in architecture design [22–25], the necessity of error correction [26] introduces substantial overheads [27–29].

These constraints are especially evident in the implementation of the quantum phase estimation (QPE) algorithm [30], which is known for achieving the Heisenberg limit in precision scaling [31]. Despite its theoretical appeal, even the most recent estimates obtained for the simulation of molecular systems show that QPE suffers from prohibitively high gate counts and qubit requirements [32]. These estimates originate

from three key factors: (1) high connectivity requirements due to the multicontrolled evolutions between two different qubit registers make the circuits deep and fragile with respect to noise [33], (2) it is necessary to prepare quantum states with sufficiently high overlap with the target state [34–37], and (3) despite their beneficial formal scaling with system size [38–40], leading time-evolution methods still demand excessively deep circuits.

These factors have motivated the exploration of alternative approaches that rely on shallower circuits. The first example of a quantum algorithm using a single-ancilla quantum circuit (i.e., the Hadamard test) was given by Kitaev [41]. Since then, many advances have been made to develop faster and more accurate algorithms, which we refer to as *post*-Kitaev approaches [42–49]. In particular, the current state-of-the-art single-ancilla phase estimation algorithms are able to satisfy the Heisenberg limit [i.e.,  $T_{\text{run-time}} = \mathcal{O}(\epsilon^{-1})$ ] while requiring shorter maximal time evolutions  $T_{\text{max}}$  compared to standard QPE. The latter exhibits a scaling of  $T_{\text{max}} = \mathcal{O}(\epsilon^{-1})$  and satisfies  $T_{\text{max}}\epsilon \approx \pi$ , whereas the very recent result of Ref. [50] demonstrates a scaling of  $T_{\text{max}} = \mathcal{O}(\epsilon^{-1.5})$  with  $T_{\text{max}}\epsilon \ll \pi$ .

The methods mentioned above [42–50] are particularly intriguing because they often rely on a formal link between QPE and the signal recovery problem. This connection is also the first key ingredient of the present work and will be discussed in depth in the next section. A second key ingredient is a thorough exploration of the relationship between signal recovery and subspace methods. A clear equivalence between the two has been established in Refs. [51,52], where filter diagonalization has been introduced. In the context of quantum algorithms development, the connection between QPE and subspace methods has been first made in Refs. [53–55] while

<sup>\*</sup>Contact author: mreiher@ethz.ch

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

Refs. [56,57] first developed a quantum subspace method inspired by filter diagonalization.

In this work, we analyze subspace methods through the lens of a post-Kitaev algorithm, characterizing  $T_{\max}$ , the number of samples and shots per sample needed to achieve a given error estimate  $\epsilon$ ; this perspective differs from previous works in the literature [54,58,59]. We make use of a framework to characterize errors in the accuracy and dimensionality of subspace methods [60,61] to derive tight error bounds for an improved version of the filter diagonalization method. Our framework builds upon and significantly expands the field of prolate Fourier theory [62–69].

With regard to filter diagonalization, we highlight Ref. [70], which was the first to employ prolates as filter functions and which made significant progress in detecting the correct number of frequencies within a band. We build upon their work by introducing truncation estimates and providing several results that formalize the approach. On the numerical side, the implementation in Ref. [70] did not account for catastrophic cancellation effects in the generation of prolate functions, a problem that was addressed and resolved in Ref. [71] and accounted for in our implementation.

All our contributions are enabled by key results in prolate Fourier theory, which we employed to provide a rigorous error analysis of the estimated eigenvalues. We note that these results are of general significance and will therefore have a broader impact, beyond their immediate application to quantum algorithms. Eventually, we introduce sampled prolate filter diagonalization (PFD) for extracting spectral information from a finite number of signal samples. By generating these samples on a quantum computer via the Hadamard test, we arrive at a hybrid classical-quantum algorithm that we call quantum prolate diagonalization (QPD).

This article is organized as follows: In Sec. II, we develop the theoretical framework underlying this work and summarize the main results. Specifically, the core idea is to shift focus from solving the time-independent Schrödinger equation to a related eigenvalue problem involving the autocorrelation function  $C(t) = \langle \Psi(t) | \Psi(0) \rangle$ . This eigenvalue problem is the conceptual bridge between post-Kitaev QPE methods and subspace techniques. We solve this spectral problem using prolate filter diagonalization (Sec. IID), which we review within the general framework of subspace protocols (Appendix A). Notably, in Sec. IID we unveil a phase-transition-like behavior of the accuracy of the protocol crucially depending on the frequency density within the spectral region of interest. Subsequently, we discretize prolate filter diagonalization to enable efficient computations that only require a finite number of signal samples (Sec. IIE). In Sec. IIE to Appendix C, we introduce improved truncation bounds for the prolate sampling formula, originally proposed by Walter and Shen [72]. Section III presents the numerical results of our QPD implementation applied to the electronic Hamiltonian. We analyze the accuracy scaling behavior of the protocol for both weakly and strongly correlated molecules. We show that our method reaches chemical accuracy (i.e., the error  $\epsilon \leq 1$  mhartree) and satisfies the Heisenberg limit (Secs. IIIB and IIIC). Finally, we investigate the impact of the initial quantum state on performance of QPD to assess the potential quantum advantage offered by our approach (Sec. IIID).

## II. THEORY

In this section, we introduce our framework for rigorous frequency approximation of signals with dense spectra requiring only a finite number of equidistant samples. Our method facilitates efficient digital computation and provides sharp guarantees on the frequency estimates. Notably, we uncover a phase-transition-like behavior in accuracy: As an *effective spectral density* approaches a critical threshold  $T/\pi$ , the error increases sharply. As this critical threshold is proportional to the observation time, this transition resembles the Heisenberg uncertainty relation.

Although the mathematical structure underlying our approach is closely related to the mathematical formulation of quantum mechanics [73–75], we stress that we shift from a traditional wave packet description to a perspective that puts system-specific signals,

$$C(t) = \int e^{i\lambda t} d\alpha(\lambda), \quad (1)$$

to the center of attention. Here,  $\alpha$  is a probability measure on  $\mathbb{R}$  [76]. In the standard formalism, such signals arise as quantum autocorrelation functions  $C(t) = \langle \Psi | e^{iHt} | \Psi \rangle$ . Furthermore, we assume that the signal is only known within the time interval  $[-2T, 2T]$ .

To determine frequencies of a system, we do not solve the time-independent Schrödinger equation. Instead, we turn to the convolution-based eigenvalue problem

$$-i\partial_\tau(C * f) = E(C * f). \quad (2)$$

Here,  $f$  is a function in  $\mathcal{L}^2(\mathbb{R})$ ,  $E \in \mathbb{R}$ , and  $C * f$  denotes the convolution of  $C$  with  $f$  evaluated in the time variable  $\tau$ . This spectral equation finds a purely mathematical reason to justify its existence: It facilitates a spectral analysis of signals without assuming an infinite observation time. In this way, it overcomes the greatest limitation of the Fourier transform. The main result of this work is a mature mathematical framework to approximate solutions of Eq. (2). This eigenvalue problem is still defined on an infinite-dimensional Hilbert space. It therefore requires approximations for practical computation, which raises fundamental questions on the optimal relationship between observation time, spectral complexity, computational dimensionality and achievable accuracy.

A seminal series of papers by Slepian, Landau, and Pollak has reached deep insights toward achieving such optimal approximation [62–69]. These authors discovered the remarkable properties of prolate spheroidal wave functions, which allowed them to prove that the space of  $W$ -band-limited and  $T$ -time concentrated functions is well approximated in essentially  $2WT/\pi$  degrees of freedom. We refer to the field that their series of papers has initiated as prolate Fourier theory (PFT).

In this work, we apply and extend the rich set of mathematical properties of prolate spheroidal wave functions to develop a method for frequency estimation that can access fundamental approximation bounds. We use the commutation relation of an integral operator and differential operator, to derive new concentration identities of prolates. The new identities enable sharp accuracy guarantees for PFD. They further allow us to give a sampling theorem that puts mathe-

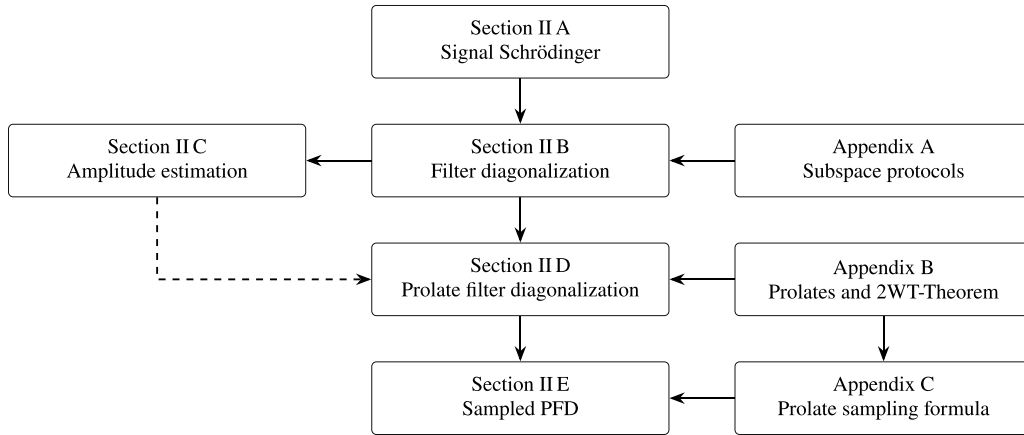


FIG. 1. Logical dependency graph of the theory underlying our approach: For the main part of the paper, we keep the essential building blocks of our approach, while we cover additional key parts of the theoretical framework in the Appendix. The relation between these parts is highlighted in this figure to aid the understanding of our work. Section II C has indirect implications for Sec. II D.

mathematical truth to the long-standing intuition that approximately  $2WT/\pi$  samples suffice to reconstruct a signal of bandwidth  $W$  and high concentration in the time interval  $[-T, T]$  with high accuracy. The prolate sampling formula is applied to achieve a discretization of PFD, which only assumes a finite number of samples of the signal and guarantees high accuracy. To clarify the structure of our theory, Figure 1 shows the logical dependencies between key results and developments within our framework.

#### A. From signals to quantum dynamics: An inverse approach

Within quantum mechanics, a system is described by a self-adjoint operator  $H$  that acts on a Hilbert space  $\mathcal{H}$ . The states of the system are described by time-dependent wave functions  $|\Psi(\cdot)\rangle : \mathbb{R} \rightarrow \mathcal{H}$ , which evolve according to the time-dependent Schrödinger equation:

$$i\partial_t |\Psi(t)\rangle = H |\Psi(t)\rangle. \quad (3)$$

For a time-independent Hamiltonian  $H$ , stationary states are solutions of the form  $|\varphi_n(t)\rangle = e^{-iE_n t} |\varphi_n\rangle$ , where  $|\varphi_n\rangle$  is an eigenvector of  $H$ ,

$$H |\varphi_n\rangle = E_n |\varphi_n\rangle. \quad (4)$$

They satisfy, in particular,

$$i\partial_t |\varphi_n(t)\rangle = E_n |\varphi_n(t)\rangle \quad (5)$$

and their probability distribution remains constant in time. The time evolution of a general wave function is described by its decomposition into stationary states  $|\Psi(t)\rangle = \sum a_n e^{-iE_n t} |\varphi_n\rangle$ . While fundamental for encoding physical laws, this formalism relies on a high-dimensional (precisely, on an infinite-dimensional) Hilbert space and the numerical representation of  $H$  and state vectors suffers from the curse of dimensionality. In contrast, experimental observations consist of low-dimensional quantities such as expectation values, low-lying eigenvalues, transition rates, and magnetization. This discrepancy in dimensionality between experimental observation and mathematical description causes a large overhead in computational applications.

To address this imbalance, we adopt an approach focusing directly on signals rather than wave packets. Specifically, we consider a continuous positive definite function  $C(t)$  with  $C(0) = 1$  and aim to determine its frequencies while only assuming a finite observation time. In this framework, we do not postulate the wave packet formalism or the Schrödinger equation; instead, these emerge as mathematical auxiliaries to analyze the signal.

We start by justifying Eq. (2) to determine the frequencies of  $C$ . By Bochner's theorem [76], any continuous positive definite function  $C$  is characterized by a unique measure  $\alpha : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$  such that Eq. (1) holds. We define the spectrum of  $C$  as  $\sigma(C) := \text{supp}(\alpha)$ . Let  $F$  denote the Fourier transform of a function  $f \in \mathcal{L}^2(\mathbb{R})$ . Inserting Eq. (1) into Eq. (2) and interchanging the integrals yields

$$\int \lambda e^{i\lambda\tau} F(\lambda) d\alpha(\lambda) \stackrel{!}{=} E \int e^{i\lambda\tau} F(\lambda) d\alpha(\lambda). \quad (6)$$

This equality holds for all  $\tau$  if and only if the complex measure  $(\lambda - E)F(\lambda)d\alpha(\lambda)$  is zero. For a nontrivial solution, this implies that  $E \in \sigma(C)$  and the  $\alpha$ -essential support of  $F$  is the singleton  $\{E\}$ , i.e.,  $\text{supp}_\alpha(F) = \{E\}$ . For a purely discrete signal  $C(t) = \sum_k |a_k|^2 e^{iE_k t}$ , this simplifies to the condition

$$F(E_n) \neq 0 \quad \text{and} \quad F(E_k) = 0 \quad \text{for all} \quad E_k \neq E_n. \quad (7)$$

Hence, the eigenvalue problem in Eq. (2) corresponds to finding a function with a Fourier transform whose roots coincide with all frequencies except one.

To perform a frequency analysis without assuming an infinite observation time, we restrict  $f$  to the subspace  $\mathcal{D}_T \subset \mathcal{L}^2(\mathbb{R})$  of functions supported on the interval  $[-T, T]$ . Consequently, the convolution in Eq. (2) only requires access to  $C(t)$  for a finite duration, allowing us to overcome a major limitation of the standard Fourier transformation. A nontrivial solution to Eq. (7) exists in  $\mathcal{D}_T$  if and only if the spectrum of  $C$  is discrete and has no accumulation point. This follows from the identity theorem and the fact that the Fourier transform of a time-limited function is entire.

While the restriction to  $\mathcal{D}_T$  addresses the finite observation time, directly identifying a function  $f$  with the requisite zeros in the Fourier transform remains a challenging continuous

search problem. In order to formulate an efficient numerical eigensolver, it is advantageous to map this signal processing task onto the linear algebraic structure of quantum mechanics.

To this end, we note that any positive definite function  $C$  with  $C(0) = 1$  can be interpreted as a correlation function arising from unitary time evolution on a Hilbert space. That is, we can construct a Hilbert space  $\mathcal{H}$ , a self-adjoint operator  $H$  acting on  $\mathcal{H}$ , and a vector  $\Psi \in \mathcal{H}$ , such that the spectra of the signal and operator  $H$  coincide,  $\sigma(C) = \sigma(H)$ , and

$$C(t) = \langle \Psi, e^{iHt} \Psi \rangle = \int_{\mathbb{R}} e^{i\lambda t} d\langle \Psi, P^{(H)}(\lambda) \Psi \rangle. \quad (8)$$

Here,  $P^{(H)}$  denotes the unique projection-valued spectral measure with  $H = \int \lambda dP^{(H)}(\lambda)$ . For a signal with a purely discrete spectrum  $C(t) = \sum_k |a_k|^2 e^{iE_k t}$ , each frequency  $E_k$  is associated with an eigenvector  $|\varphi_k\rangle$  of  $H$  and a valid construction has  $|\Psi\rangle = \sum_k a_k |\varphi_k\rangle$ . In particular,  $H$ ,  $|\Psi\rangle$ , and  $\{\varphi_k\}$  satisfy the Schrödinger equations (3)–(5) by construction. Hence, we do not postulate the wave packet formalism, but instead construct it as a mathematical auxiliary. Crucially, this construction remains implicit, effectively hiding the redundant degrees of freedom arising from unitary invariance in explicit basis expansions of  $H$ . This circumvents the curse of dimensionality of operator representations while preserving access to the relevant spectral data. In Sec. II B, this perspective provides the rigorous foundation to formulate filter diagonalization as a subspace protocol (Appendix A).

## B. Filter diagonalization

Filter diagonalization is a numerical procedure to approximate the frequencies of a signal,

$$C(t) = \sum_n |a_n|^2 e^{iE_n t}, \quad (9)$$

within a target band  $B_{f,\omega^*} = [-W_f + \omega^*, W_f + \omega^*]$ . Without loss of generality, we assume the band is centered at zero [otherwise, we shift the signal  $C(t) \mapsto C(t)e^{-i\omega^* t}$ ].

This method fits naturally into the framework of *subspace protocols* [61]. A subspace protocol approximates the eigenvalues of an operator  $H$  within a specific spectral subspace  $\mathcal{E}$  by constructing a set of trial vectors  $\{v_l\}$  that span  $\mathcal{E}$  (Appendix A reviews the structure and algorithmic operation of subspace protocols and restates the master theorem on their accuracy). We construct these trial vectors within the implicit quantum mechanical structure established in Sec. II A, where the signal is assumed to be generated by a state  $|\Psi\rangle = \sum_k a_k |\varphi_k\rangle$  evolving under a Hamiltonian  $H$ . However, the final protocol requires only the signal  $C(t)$  and no explicit access to  $H$  or  $|\Psi\rangle$  (for a comprehensive derivation, see Sec. 5.2 of Ref. [60]).

To target the spectral subspace  $\mathcal{E}$  corresponding to the frequency band  $B_f$ , we apply a set of filter functions  $\{f_l(t)\}_{l=0}^{m-1}$  to the trajectory of the state  $|\Psi(t)\rangle$ . Explicitly, the  $l$ th trial vector is generated by the integral transform

$$v_l := \int_{-\infty}^{\infty} f_l(\tau) |\Psi(\tau)\rangle d\tau. \quad (10)$$

Substituting the eigen decomposition  $|\Psi(\tau)\rangle = \sum_k a_k e^{-iE_k \tau} |\varphi_k\rangle$  into Eq. (10) yields

$$v_l = \sum_k a_k \left( \int_{-\infty}^{\infty} f_l(\tau) e^{-iE_k \tau} d\tau \right) |\varphi_k\rangle = \sum_k a_k F_l(E_k) |\varphi_k\rangle, \quad (11)$$

where  $F_l(E)$  is the Fourier transform of the filter function  $f_l(t)$ . Finally, filter diagonalization can be formulated as a subspace protocol (Definition A1) with matrix elements

$$(V^\dagger H V)_{sl} = -i \int f_s(\tau) \partial_\tau (C * f_l)(\tau) d\tau, \quad (12)$$

$$(V^\dagger V)_{sl} = \int f_s(\tau) (C * f_l)(\tau) d\tau, \quad (13)$$

$$\delta A = \delta B = 0. \quad (14)$$

Given a signal  $C$  and a target band  $B_f$ , Algorithm 1 defines a numerical routine  $\mathbf{P}_{(C,B_f)}$  to extract the frequencies within  $B_f$  (cf. Appendix A). The routine  $\mathbf{P}_{(C,B_f)}$  takes the number  $M$  of trial vectors to use and the assumed number  $m$  of frequencies inside the band  $B_f$  as input. Step 2 of Algorithm 1 can be used to estimate  $m$  through the spectrum of the weight matrix  $B_M$  and a noise threshold  $\epsilon_{\text{th}}$ . Ideally, the spectrum of  $B_M$  exhibits a sharp drop, such that the number of significant frequencies inside the band  $B_f$  is detected with high confidence. Step 3 of Algorithm 1 performs a refinement of the generalized eigenvalue problem (GEP)  $[A_M, B_M]$ , to obtain a well-conditioned GEP  $[A_M^m, B_M^m]$  of dimension  $m$ . The refinement allows us to use an initial trial dimension  $M$  that exceeds  $m$ . A larger  $M$  increases the expressiveness of the trial vector space and improves the conditioning of the refined GEP. This can also enable a more reliable detection of  $m$ . However, there is a trade-off as contributions from noise typically also increase with  $M$ . So, the choice of initial trial dimension  $M$  must be balanced between improved signal representation and increased distortions from noise (see Sec. 3.1 in Ref. [61]).

After the correct number of frequencies has been detected, Theorem A1 describes the accuracy of the frequency estimates computed by  $\mathbf{P}_{(C,B_f)}(M, m)$ . The approximation error is quantified through an error measure  $\varepsilon_M : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ , which, in the case of filter diagonalization, is given by

$$\begin{aligned} \varepsilon_M(I) &= \sum_{i=0}^{M-1} \int_I |F_i(E)|^2 d\alpha(E) \\ &= \sum_{i=0}^{M-1} \sum_{E_k \in I} |F_i(E_k)|^2 |a_k|^2. \end{aligned} \quad (15)$$

## C. Amplitude estimation

Once the frequencies have been approximated, filter diagonalization can also recover the corresponding amplitudes. We introduce an approach that facilitates accuracy guarantees for its amplitude estimates. In addition, we give a bound on the minimal amplitude required for a frequency to distinguish itself from a noise threshold. Early implementations of filter diagonalization extracted amplitudes from the eigenvectors of the GEP [52,77,78]. However, this approach remained heuristic and lacked a rigorous error or stability analysis.

Our approach is based on the observation that the filter diagonalization matrices can be decomposed into matrix multiplications of alternant matrices and diagonal matrices containing the amplitudes (Secs. 5.2 and 5.3 of Ref. [60]): We let  $\vec{E} \in \mathbb{R}^m$  denote the vector of frequencies inside the target band  $B_f$  and define the *alternant matrix*  $\mathfrak{F}_M(\vec{E}) \in \mathbb{C}^{m \times m}$  by

$$\mathfrak{F}_M(\vec{E}) := (F_l(E_k))_{kl}, \quad (16)$$

where  $F_l$  is the Fourier transform of the  $l$ th filter function. Consider the refined GEP  $[A_M^m, B_M^m]$  returned by filter diagonalization through Algorithm 1. The weight matrix is found to have the form

$$B_M^m = \mathfrak{F}_M^m(\vec{E})^\dagger \Lambda(\vec{a}) \mathfrak{F}_M^m(\vec{E}) + (\mathcal{N}^{(B)})_M^m, \quad (17)$$

where  $\Lambda(\vec{a}) = \text{diag}(|a_k|^2)$  is the amplitude matrix,  $\mathfrak{F}_M^m(\vec{E}) := \mathfrak{F}_M(\vec{E})U_m \in \mathbb{C}^{m \times m}$  is the alternant matrix formed from the refined set of filter functions obtained after step 3 of Algorithm 1, and  $(\mathcal{N}^{(B)})_M^m := U_m^\dagger (N_M^\dagger N_M + \delta B_M) U_m$  is the projected noise weight matrix. This representation of the weight matrix follows from rewriting sums over the frequency contributions as a matrix multiplication.

Equation (17) provides a practical method to approximate the amplitudes that also enables accuracy guarantees. Let  $\vec{E}$  denote the vector of eigenvalues obtained by filter diagonalization. Since  $\vec{E}$  approximates the true frequency vector  $\vec{E}$ , we estimate the amplitude matrix as

$$\Lambda(\vec{a}) \approx \mathfrak{F}_M^m(\vec{E})^{-\dagger} B_M^m \mathfrak{F}_M^m(\vec{E})^{-1}. \quad (18)$$

Proposition 7 of Ref. [60] (Sec. 5.3.2 of Ref. [60]) bounds the deviation of the diagonal elements of  $\mathfrak{F}_M^m(\vec{E})^{-\dagger} B_M^m \mathfrak{F}_M^m(\vec{E})^{-1}$  to the true amplitudes. Moreover, one can assess the convergence of the method by examining whether this matrix is indeed close to diagonal.

The alternant matrix also encodes meaningful information on the stability of the method. In particular, we establish an amplitude threshold that guarantees that all frequencies can be distinguished from a given noise level:

$$|a|_{\min}^2 > \lambda_{\min}(\mathfrak{F}_M^m(\vec{E})^\dagger \mathfrak{F}_M^m(\vec{E}))^{-1} \epsilon_{\text{th}}. \quad (19)$$

If all relevant amplitudes satisfy Eq. (19), Algorithm 1 is indeed guaranteed to detect the correct number of frequencies within the band:

$$\lambda_m(B_M^m) \geq \lambda_m(\mathfrak{F}_M^m(\vec{E})^\dagger \mathfrak{F}_M^m(\vec{E})) |a|_{\min}^2 + \lambda_m((\mathcal{N}^{(B)})_M^m) \quad (20)$$

$$> \epsilon_{\text{th}}. \quad (21)$$

In the last inequality, we assume that  $(\mathcal{N}^{(B)})_M^m$  is positive definite, as is the case for purely subspace-based noise  $\mathcal{N}_M^{(B)} = N_M^\dagger N_M$ .

#### D. Prolate filter diagonalization

In this section, we review PFD and its error bounds (Chap. 5 of Ref. [60]). We highlight a phase-transition-like behavior in the accuracy of PFD that resembles the Heisenberg uncertainty relation: The accuracy drastically decreases as the spectral density of the signal approaches a critical threshold given by the observation time  $\sim T/\pi$ .

PFD uses time-limited prolates  $\mathcal{D}_T \xi_n^{(W_f)}$  as filtering functions, due to their optimal concentration in the frequency band  $B_f$  and favorable orthogonality relations. They are eigenfunctions of the time- and band-limiting integral operator  $\mathcal{D}_T \mathcal{B}_f$  (see Appendix B). In particular, there are approximately  $2W_f T/\pi$  such time-limited prolates with high concentration in the frequency band  $B_f$ . We deferred a review of the origin of prolates and their optimal approximation properties to Appendix B and note that in Appendix B 2, we review inequalities on prolates that enable rigorous accuracy guarantees for PFD.

We formalize PFD as a subspace protocol that can access the theory of Appendix A.

**Definition 1 (Prolate filter diagonalization).** Given a signal  $C(t)$  for  $t \in [-2T, 2T]$  and a frequency band  $B_f = [-W_f, W_f]$  of interest. We define PFD as the protocol  $\mathbf{P}$  that assumes infinite-dimensional matrices  $V^\dagger H V$  and  $V^\dagger V$  with matrix elements:

$$(V^\dagger H V)_{sl} = -i \int_{-T}^T \int_{-T}^T \xi_s^{(W_f)}(\tau) \partial_\tau C(\tau - t) \xi_l^{(W_f)}(\tau) d\tau d\tau, \quad (22)$$

$$(V^\dagger V)_{sl} = \int_{-T}^T \int_{-T}^T \xi_s^{(W_f)}(\tau) C(\tau - t) \xi_l^{(W_f)}(\tau) d\tau d\tau. \quad (23)$$

Here,  $\{\xi_n^{(W_f)}\}$  denotes the sequence of prolates of bandwidth  $W_f$  and concentration in the time interval  $[-T, T]$ . At a given trial dimension  $M$ , PFD returns the  $M$ th leading principal submatrices of  $V^\dagger H V$  and  $V^\dagger V$ .

PFD constitutes a subspace protocol as in Definition A1, with  $\delta A = \delta B = 0$  and noise matrix  $\mathcal{N}^{(B)} = N^\dagger N$ ,

$$(N^\dagger N)_{sl} = \int_{-T}^T \int_{-T}^T \int_{\sigma(C) \setminus B_f} \xi_s^{(W_f)}(\tau) e^{iE(\tau-t)} \xi_l^{(W_f)}(t) \times d\alpha(E) d\tau d\tau, \quad (24)$$

arising from frequency components outside the target band. Moreover, PFD can be formulated as an  $\epsilon$ -subspace protocol  $\mathbf{P}_{\{\epsilon_M\}}$  (see Definition A2) with bounds on the noise weight

$$\|\mathcal{N}^{(B)}\| \leq \epsilon_M^{\text{Prft}} \int_{\sigma(C) \setminus B_{W_f}} d\alpha(\lambda) = \epsilon_M^{\text{Prft}} \sum_{E_n \notin B_f} |a_n| e^{iE_n t} \quad (25)$$

$$\leq \epsilon_M^{\text{Prft}} C(0), \quad (26)$$

where  $\epsilon_M^{\text{Prft}}$  is defined in Eq. (29).

Theorem A1 describes the accuracy of a subspace protocol through an error measure  $\epsilon_M$ . For filter diagonalization, this error measure is determined by the Fourier transforms of the filtering functions, as seen in Eq. (15). The Fourier transform of a time-truncated prolate is

$$\mathcal{F}[\mathcal{D}_T \xi_n](\omega) = \bar{\mu}_n \tilde{\xi}_n(\omega), \quad (27)$$

where  $\tilde{\xi}_n$  denotes *dual prolates*, that is, prolates with roles of bandwidth  $W$  and time interval  $T$  exchanged. The prefactors  $\bar{\mu}_n \in \mathbb{C}$  are complex conjugates of the eigenvalues of the integral operator given in Eq. (B5).

The bound in Eq. (B17) of Theorem B4 controls the maximal value that a prolate function can attain outside its concentration region. We derived Theorem B4 in Chap. 2 of Ref. [60] to yield sharp accuracy bounds for PFD. The proof

of Theorem B4 relies on a commutation relation between an integral operator and a differential operator. This commutation relation has played an important role in the development of PFT and continues to lead to remarkable connections [69,79,80]. In Appendix B3, we highlight a spectral duality between the integral and differential operator at the critical dimension  $n = \lfloor 2WT/\pi \rfloor$ , which appears to have remained previously unnoticed.

By applying Eq. (B17) on the prolate duals  $\tilde{\xi}_n(\omega)$ , we obtain an estimate on the error measure of PFD.

*Corollary 1.* Let  $\epsilon_M$  be the error measure that is obtained from the first  $M$  time-limited prolates  $\mathcal{D}_T \xi_n^{(W_f)}$ . We have

$$\epsilon_M(I) \leq 2\pi \sum_{l=0}^{M-1} \gamma_l(1 - \gamma_l) C_{\text{extra},l}(T, W_f) \alpha(I), \quad (28)$$

where  $C_{\text{extra},l}(T, W_f)$  is defined as in Theorem B4 with arguments  $W_f$  and  $T$  exchanged.

We abbreviate the right-hand side as

$$\epsilon_M^{\text{Prlt}} := 2\pi \sum_{l=0}^{M-1} \gamma_l(1 - \gamma_l) C_{\text{extra},l}(T, W_f), \quad (29)$$

since  $\epsilon_M^{\text{Prlt}}$  characterizes the accuracy of PFD. The magnitude of  $\epsilon_M^{\text{Prlt}}$  is strongly governed by the factor  $1 - \gamma_{M-1}(c_f)$ , where  $c_f := W_f T$ . We refer to  $\tilde{c}_f := 2W_f T/\pi$  as the *essential dimensionality*, since the  $2WT$  theorem (Theorem B2) states that the space of  $W_f$ -band-limited and  $T$ -time-concentrated functions is well approximated by essentially  $2W_f T/\pi$  degrees of freedom. We can estimate

$$\epsilon_M^{\text{Prlt}} < 2\pi M T c_f (1 + \mathcal{O}(c_f^{-2})) (1 - \gamma_{M-1}(c_f)) \quad (30)$$

$$< M T \frac{\pi^{\frac{3}{2}} 2^{3M} c_f^{M+\frac{1}{2}} e^{-2c_f}}{(M-1)!} [1 + \mathcal{O}(c_f^{-1})]. \quad (31)$$

The first inequality applies bounds on  $C_{\text{extra},l}(T, W)$  from Appendix B2, and the second uses the asymptotic expansion

$$1 - \gamma_n(c) = \frac{4\sqrt{\pi} 2^{3n} c^{n+\frac{1}{2}} e^{-2c}}{n!} (1 + \mathcal{O}(c^{-1})) \quad (32)$$

according to Refs. [81,82]. In particular,  $\epsilon_M^{\text{Prlt}}$  remains very small as long as the number of used prolate filters  $M$  remains sufficiently lower than the essential dimension  $2W_f T/\pi$ . At a fixed target accuracy and trial dimension,  $\epsilon_M^{\text{Prlt}}$  and Eq. (31) can also be used to estimate the required observation duration  $T$  or phase space volume  $c_f = W_f T$ .

Corollary 1 and Theorem A1 together yield the following sharp accuracy guarantee for PFD.

*Theorem 1.* Let  $[A_M^m, B_M^m]$  be the GEP that is obtained from PFD through Algorithm 1 with a trial dimension  $M$  and a detected dimension  $m$ . Assume that the detected dimension  $m$  coincides with number of frequencies in  $B_f$ , that the GEP is well conditioned in the sense

$$\lambda_m(B_M^m) > \|\mathcal{N}_M^{(B)}\|, \quad (33)$$

and that the generalized eigenvalues lie within  $[-W_f, W_f]$ . Let  $\tilde{E}_1 > \tilde{E}_2 > \dots > \tilde{E}_m$  denote these eigenvalues, and  $E_1 > E_2 > \dots > E_m$  the frequencies of  $C$  within  $[-W_f, W_f]$ . Then,

we have error bounds on the approximated frequencies:

$$\begin{aligned} \epsilon_M^{\text{Prlt}} \frac{\int_{\lambda < -W_f} (\lambda - \tilde{E}_i) d\alpha(\lambda)}{\lambda_m(B_M^m) - \|\mathcal{N}_M^{(B)}\|} \\ \leq \tilde{E}_i - E_i \leq \epsilon_M^{\text{Prlt}} \frac{\int_{\lambda > W_f} (\lambda - \tilde{E}_i) d\alpha(\lambda)}{\lambda_m(B_M^m) - \|\mathcal{N}_M^{(B)}\|}. \end{aligned} \quad (34)$$

Theorem 1 provides sharp error bounds for the estimated frequencies. These bounds are governed by the error parameter  $\epsilon_M^{\text{Prlt}}$ , which essentially behaves like  $1 - \gamma_M(c_f)$ . This leads to a *phase-transition*-like behavior in the accuracy of the method: Prolate filter diagonalization achieves high accuracy within a spectral region of interest, as long as the effective spectral density  $\delta_{\text{eff}}$  of the signal has

$$\delta_{\text{eff}} \lesssim \frac{T}{\pi}. \quad (35)$$

As  $\delta_{\text{eff}}$  approaches  $T/\pi$ , the accuracy of PFD sharply decreases.

This statement follows from the number of trial vectors necessary to represent a signal within  $B_f$  and  $\epsilon_M^{\text{Prlt}} \sim 1 - \gamma_M(c_f)$ . The effective spectral density is slightly more locally defined than the spectral density, as we explain below.

The first  $\sim 2c/\pi$  eigenvalues  $\gamma_n(c)$  are very close to one, followed by a transition region that is logarithmically sharp in  $c$ . Beyond the transition region, all remaining eigenvalues rapidly decay toward zero (Theorem B3). The characterizing accuracy parameter  $\epsilon_M^{\text{Prlt}}$  of PFD inherits this behavior; as  $M$  approaches  $2W_f T/\pi$ , there is a sharp increase in  $\epsilon_M^{\text{Prlt}}$ .

The number of trial vectors necessary to use depends on the effective spectral density of the signal. A necessary condition to stay out of the transition region in terms of the spectral density is

$$\frac{\#\{E_k \in B_{f,\omega^*}\}}{2W_f} \lesssim \frac{T}{\pi}, \quad (36)$$

where  $\#$  denotes the number of elements of a set. The condition is not sufficient, as the error bounds of Theorem 1 also depend on the conditioning of  $B_M^m$ . In Sec. II C, we derived that the conditioning of the weight matrix does not only depend on the amplitudes, but also on the conditioning of the filter matrix evaluated at the in-band frequencies:

$$\lambda_m(B_M^m) \geq \lambda_m(\mathfrak{F}_M^m(\vec{E})^\dagger \mathfrak{F}_M^m(\vec{E})) |a|_{\min}^2. \quad (37)$$

Hence, we require, for  $M$  large enough but lower than  $2W_f T/\pi$ , that a good conditioning of  $B_M^m$  can be ensured. Despite the necessary condition in Eq. (36) being satisfied, the contrary can occur if two frequencies are very close with respect to the analyzed frequency band  $B_f$ . This can cause near-linear dependencies in the rows of  $\mathfrak{F}_M^m(\vec{E})$ . In this case, it is appropriate to “close-up” on the approximate degeneracy through a more informed choice of  $B_f$ . This also suggests an intuitive definition of the *effective spectral density*: It is defined through  $\#\{E_k \in B_{f,\omega^*}\}/2W_f$  but with a choice of  $B_{f,\omega^*}$  such that the spectral density within the band remains relatively homogeneous.

The accuracy transition we just uncovered through PFT is reminiscent of a generalization of the Heisenberg uncertainty principle. As long as the observation time  $T$  satisfies Eq. (35),

PFD achieves high accuracy despite dense spectra with many frequencies outside the target band  $B_f$ .

### E. Sampled prolate filter diagonalization

In this section, we discretize PFD through the *prolate sampling formula* to obtain a method particularly suited for digital computation. This discretization requires an additional assumption on the bandwidth of the signal, but in return, only relies on a finite number of equidistant samples.

We emphasize that the sampling-induced error exhibits a phase-transition behavior, which is simultaneous with the phase transition of the subspace-based error derived in Sec. IID.

The prolate sampling formula was introduced in Ref. [72], and it improves upon the Whittaker-Shannon sampling formula because it truncates better. However, the truncation estimates in Ref. [72] were still relatively loose. Using the inequalities of Theorem B4, we can derive sharper estimates. In Appendix C, we formulate an approximation theorem that resembles the engineering folk theorem [69,83]: Approximately  $2WT/\pi$  equidistant samples suffice to reconstruct a signal of bandwidth  $W$  and essential duration  $2T$  with high accuracy.

This theorem was a long-standing intuition that preceded the  $2WT$  theorem (see Appendix C and references therein). Let  $\mathcal{B}_s \subset \mathcal{L}^2(\mathbb{R})$  denote the subspace of functions with a Fourier transform that has support in  $[-W_s, W_s]$ . Given a band-limited function  $f \in \mathcal{B}_s$  that is highly concentrated in the interval  $[-T, T]$ , we approximate  $f$  on  $[-T, T]$  through

$$f(t) \approx \frac{\pi}{W_s} \sum_{n=0}^{\lfloor 2W_s T/\pi \rfloor} \sum_{k=-\lfloor W_s T/\pi \rfloor}^{\lfloor W_s T/\pi \rfloor} f\left(\frac{k\pi}{W}\right) \xi_n^{(W_s)}\left(\frac{k\pi}{W}\right) \xi_n^{(W_s)}(t), \quad (38)$$

where  $\{\xi_n^{(W_s)}\}$  denotes the sequence of prolates of bandwidth  $W_s$  and concentration in  $[-T, T]$ . The accuracy of the approximation is described by Theorem C1.

To apply this truncated sampling formula to PFD, we require the additional assumption that the signal  $C$  is of finite bandwidth  $W_C$ . Then, the integrands appearing in the matrix elements of PFD,

$$f_l^{(\tau)}(t) := C(\tau - t) \xi_l^{(W_s)}(t), \quad (39)$$

have bandwidth at most  $W_C + W_f$ . And note that  $f_l^{(\tau)}(t)$  is concentrated in  $[-T, T]$  provided  $l \lesssim 2W_f T/\pi$ . Therefore, Theorem C1 applies for any sampling rate  $W_s \geq W_C + W_f$ . Applying the interpolation formula to each matrix element (twice) yields *sampled prolate filter diagonalization*.

*Definition 2 (Sampled prolate filter diagonalization).*

Consider a signal  $C(t)$  that is of bandwidth  $W_C$ , a filter width  $W_f$ , a sample rate  $W_s$  with  $W_s \geq W_C + W_f$ , and a duration  $T$ . Assume sample points  $C(t_k)$  of the signal at  $t_k = \pi k/W_s$  for  $k = -N_s, -N_s + 1, \dots, 0, \dots, N_s - 1, N_s$ , where  $N_s = 2\lceil W_s T/\pi \rceil$ .

*Sampled prolate filter diagonalization* is the protocol that generates infinite-dimensional matrices  $A$  and  $B$  with matrix elements:

$$\begin{aligned} A_{sl} = & -i\pi^2 \frac{T}{W_s^3} \sum_{n_1, n_2=0}^{\lfloor 2W_s T/\pi \rfloor} \mu_{n_2}(c_s) \mu_{n_1}(c_s) \xi_{n_1}^{(W_s)}(0) \xi_{n_2}^{(W_s)}(0) \\ & \times \sum_{k_1, k_2=-\lceil W_s T/\pi \rceil}^{\lceil W_s T/\pi \rceil} \xi_{n_2}^{(W_s)}\left(\frac{\pi k_2}{W_s}\right) \xi_s^{(W_f)}\left(\frac{\pi k_2}{W_s}\right) \\ & \times C\left(\frac{\pi k_2}{W_s} - \frac{\pi k_1}{W_s}\right) \xi_l^{(W_f)}\left(\frac{\pi k_1}{W_s}\right) \xi_{n_1}^{(W_s)}\left(\frac{\pi k_1}{W_s}\right) \\ & + i\pi \sqrt{\frac{T}{W_s^3}} \sum_{n=0}^{\lfloor 2W_s T/\pi \rfloor} \mu_n(c_s) \xi_n^{(W_s)}(0) \sum_{k=-\lceil W_s T/\pi \rceil}^{\lceil W_s T/\pi \rceil} \\ & \times \xi_n^{(W_s)}\left(\frac{\pi k}{W_s}\right) \left[ C\left(\frac{\pi k}{W_s} - T\right) \xi_l^{(W_f)}(T) \right. \\ & \left. - C\left(\frac{\pi k}{W_s} + T\right) \xi_l^{(W_f)}(-T) \right], \quad (40) \end{aligned}$$

$$\begin{aligned} B_{sl} = & \pi^2 \frac{T}{W_s^3} \sum_{n_1, n_2=0}^{\lfloor 2W_s T/\pi \rfloor} \mu_{n_2}(c_s) \mu_{n_1}(c_s) \xi_{n_1}^{(W_s)}(0) \xi_{n_1}^{(W_s)}(0) \\ & \times \sum_{k_1, k_2=-\lceil W_s T/\pi \rceil}^{\lceil W_s T/\pi \rceil} \xi_{n_2}^{(W_s)}\left(\frac{\pi k_2}{W_s}\right) \xi_s^{(W_f)}\left(\frac{\pi k_2}{W_s}\right) \\ & \times C\left(\frac{\pi k_2}{W_s} - \frac{\pi k_1}{W_s}\right) \xi_l^{(W_f)}\left(\frac{\pi k_1}{W_s}\right) \xi_{n_1}^{(W_s)}\left(\frac{\pi k_1}{W_s}\right). \quad (41) \end{aligned}$$

Here,  $c_s = W_s T$  and  $\{\xi_n^{(W_s)}\}$  denotes the sequence of prolates of bandwidth  $W_s$  and concentration in the time interval  $[-T, T]$ . At a given trial dimension  $M$ , sampled PFD returns the  $M$ th leading principal submatrices of  $A$  and  $B$ .

Sampled PFD is a subspace protocol as in Definition A1 with

$$\delta A = A - V^\dagger H V, \quad \delta B = B - V^\dagger V, \quad \mathcal{N}^{(B)} = N^\dagger N + \delta B. \quad (42)$$

Here,  $V^\dagger H V$  and  $V^\dagger V$  refer to matrices as in PFD (Definition 1). Then, the Master Theorem A1 yields for sampled PFD the accuracy guarantee.

*Theorem 2.* Let  $[A_M^m, B_M^m]$  be the GEP that is obtained from sampled PFD through Algorithm 1 with a trial dimension  $M$  and a detected dimension  $m$ . Assume that the detected dimension  $m$  coincides with number of frequencies in  $B_f$ , that the GEP is well conditioned in the sense

$$\lambda_m(B_M^m) > \|\mathcal{N}_M^{(B)}\|, \quad (43)$$

and that the generalized eigenvalues lie within  $[-W_f, W_f]$ . Let  $\tilde{E}_1 > \tilde{E}_2 > \dots > \tilde{E}_m$  denote these eigenvalues, and  $E_1 > E_2 > \dots > E_m$  the frequencies of  $C$  within  $[-W_f, W_f]$ . Then,



general task of multiple eigenvalue estimation. We anticipate that, in this case as well, the protocol achieves Heisenberg scaling for both weakly (LiH) and strongly (H<sub>8</sub>) correlated systems, with total run-time scaling as  $T_{\text{run-time}} = \mathcal{O}(\epsilon^{-1})$ , and with maximal evolution times shorter than those required by standard QPE, i.e.,  $T_{\text{max}} \ll \frac{\pi}{\epsilon}$ . Finally, we discuss another important aspect of interference-based quantum algorithms: the effect of the initial input state.

## A. Computational details

### 1. Efficient computation of prolates

Prolates are eigenfunctions of the time- and band-limiting operator  $\mathcal{B}_W \mathcal{D}_T$ . However, the direct discretization of the integral equation is ill conditioned, as the eigenvalues are highly clustered around one and zero. Instead, we utilize the property—famously termed the “lucky accident” [69]—that prolates are also eigenfunctions of the prolate spheroidal wave equation. Specifically,  $\xi_n^{(W)}(t)$  corresponds to the angular prolate spheroidal wave function  $S_{0n}(c, \eta)$ , with  $c = WT$  and scaled time coordinate  $\eta = t/T \in [-1, 1]$ . The corresponding integral operator eigenvalues  $\gamma_n$  (required for the error bounds  $\epsilon_M^{\text{Prlt}}$ ) are obtained from the radial prolate functions  $R_{0n}^{(1)}(c, 1)$  using the relation  $\gamma_n = \frac{2c}{\pi} [R_{0n}^{(1)}(c, 1)]^2$  [62].

We employ the open-source library `prolate_swf` [86], which implements the efficient numerical scheme detailed by Van Buren and Boisvert [71]. Their approach overcomes catastrophic cancellation effects in the computation of the radial functions  $R_{0n}^{(1)}(c, \zeta)$  that limited previous methods [87]. The angular wave functions are calculated through an expansion in Legendre polynomials:

$$S_{0n}(c, \eta) = \sum_{k=0,1}^{\infty} d_k(c|0n)P_k(\eta), \quad (55)$$

where the summation runs over even  $k$  for even  $n$  and odd  $k$  for odd  $n$ . The same coefficients  $d_k$  determine the radial functions through a series of spherical Bessel functions  $j_k(c\zeta)$ , which remains numerically stable even for large  $c$  where traditional power series fail. These coefficients satisfy a three-term recurrence relation. The library determines them with Bouwkamp’s algorithm [88] and fixes the overall scaling through the normalization scheme of Meixner and Schäfer [89] to prevent subtraction errors. As the prolate generation is part of the classical postprocessing routine (sampled PFD), it is carried out entirely on a classical computer.

### 2. Hamiltonian representation and state preparation

The Jordan-Wigner mapping was employed in all calculations, as implemented in the PennyLane library [90]. All calculations were performed employing the Slater-Type Orbitals approximated by 3 Gaussian functions (STO-3G) basis set [91]. While we acknowledge that this basis set does not yield highly accurate results, we adopt it for the sake of simplicity, as the primary goal of the following numerical experiments is to evaluate the features of the QPD. The initial input wave function was prepared and uploaded to the quantum register using a custom modification of the Overlapper library [92], which called PySCF [93] (see below). All initial

approximate wave functions applied in our calculations had the following structure:

$$|\Psi_{\text{input}}\rangle = \frac{1}{\sqrt{n_{\text{ts}}}} \sum_i^{n_{\text{ts}}} |i\rangle, \quad (56)$$

where  $n_{\text{ts}}$  is the number of target states of our calculation and  $|i\rangle$  is a normalized approximation of the true eigenstate  $|i_{\text{FCI}}\rangle$  expressed as a linear combination of  $K$  Slater determinants  $|\Phi_j\rangle$ ,

$$|i\rangle = \mathcal{N} \sum_j^K c_j |\Phi_j\rangle, \quad (57)$$

with normalization constant  $\mathcal{N}$ .

### 3. Emulation of QPD

We emulate the Hadamard test circuit through a custom implementation based on JAX [94]. This quantum subroutine of QPD generates signal samples  $C(t_k = \pi k/W_s)$  for  $k = 1, \dots, N_s$ . We denote the maximal evolution time as  $T_{\text{max}} = \pi N_s/W_s$  and set  $T = T_{\text{max}}/2$ . Exploiting the symmetry  $C(-t) = C(t)^*$  gives  $2N_s + 1$  signal samples in  $[-2T, 2T]$ . Then, sampled PFD can be applied to compute the frequencies of the signal within the target frequency band  $B_{f,\omega^*} = [\omega^* - W_f, \omega^* + W_f]$ . Note that we require a sampling rate large enough such that  $W_s \geq W_f + W_C$ , where  $W_C$  is the signal bandwidth. To obtain a pretrial for the filter center  $\omega^*$  when approximating the lowest-lying energies, we used a configuration interaction singles and doubles [95] calculation with the PySCF program [93] (note that alternatives are also possible, such as Ref. [96]). All reference wave functions of both full configuration interaction (FCI, representing an exact diagonalization result) and complete active space configuration interaction (CASCI, equivalent to FCI in a restricted orbital space) quality have been obtained with PySCF [93].

As discussed in Sec. II F, finite circuit measurements introduce statistical noise to the signal samples  $C(t_k)$ . This stochastic noise constitutes an error source distinct from the deterministic subspace and discretization errors, and is not captured by the approximation theorem in Sec. II E. To evaluate robustness of QPD under realistic conditions, we incorporate this shot noise into our simulations and analyze the average performance as a function of the total quantum run-time  $T_{\text{run-time}}$  [46]:

$$T_{\text{run-time}} = \#_{\text{shots}} \sum_{k=1}^{N_s} t_k. \quad (58)$$

Evaluating Eq. (58) with the Gauss formula yields

$$T_{\text{run-time}} \approx \#_{\text{shots}} \frac{\pi}{W_s} \frac{N_s^2}{2} = \#_{\text{shots}} \mathcal{O}(W_s T^2). \quad (59)$$

Furthermore, we characterize the error scaling with respect to the maximum evolution time  $T_{\text{max}}$ . While our simulations employ exact time evolution, practical implementations on real quantum hardware incur evolution errors (such as Trotterization) that accumulate with  $T_{\text{max}}$ . Consequently,  $T_{\text{max}}$  serves as a direct proxy for the required circuit depth and hardware coherence time. Demonstrating chemical accuracy at low  $T_{\text{max}}$

TABLE I. Energies (E) and excitation energies (EE) of the benzene electronic states included in the initial state used to obtain the results of Fig. 2 and Table II. Values are reported in atomic units (hartree) and shifted by the filter center  $\omega^*$  and frozen core energy  $E_{\text{core}} = -221.50532$  hartree.

|    | $E_{ \text{GS}\rangle}$ | $E_{ \text{S}_1\rangle}$ | $E_{ \text{S}_6\rangle}$ | $E_{ \text{S}_7\rangle}$ | $E_{ \text{S}_8\rangle}$ | $E_{ \text{S}_9\rangle}$ |
|----|-------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| E  | -0.130 08               | 0.130 08                 | 0.299 95                 | 0.309 97                 | 0.310 04                 | 0.411 83                 |
| EE | —                       | 0.260 15                 | 0.430 03                 | 0.440 05                 | 0.440 12                 | 0.541 91                 |

is therefore necessary for the algorithm's applicability on early fault-tolerant devices.

For our scans in  $T_{\text{max}} \sim N_s$ , we set the number of shots per sample to  $\#_{\text{shots}} = F \sqrt{N_s \log(N_s)}$ , where  $F$  is some constant independent of  $N_s$ . We found that an appropriate choice of  $F$  depends on the amplitude of the targeted frequencies. For weakly correlated systems, where the initial trial with few determinants already has a large overlap with the target state, a value of  $F = 2$  was sufficient. For strongly correlated systems—characterized by a low overlap (i.e., small amplitudes for the target frequencies) with the target eigenstates and a high number of frequencies in the autocorrelation signal (see Sec. III D)—we set  $F = 10$ . We will rationalize this behavior when discussing the results of Fig. 5.

All numerical results were obtained with the QUANTUM-PROLATE-DIAGONALIZATION package, which we developed for the purpose of this paper and which we made available in the GitHub repository [97].

### B. Multiple eigenvalue estimation at chemical accuracy

We propose an algorithm for the simultaneous estimation of multiple eigenvalues of a quantum Hamiltonian  $H$ . As an example, we focus on the electronic Hamiltonian  $H_{\text{el}}$ ,

$$H_{\text{el}} = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} g_{pqrs} a_p^\dagger a_r^\dagger a_s a_q, \quad (60)$$

where  $a_p$  ( $a_p^\dagger$ ) is the annihilation (creation) operator acting on the spin orbital  $p$ , and  $h_{pq}$  and  $g_{pqrs}$  are the usual one- and two-electron integrals [98] accounting for the kinetic and electron-nuclei and for the electron-electron interaction terms, respectively.

We demonstrate in the following that QPD can resolve multiple energy levels at chemical accuracy (that is, an error  $\epsilon \leq 1$  mhartree in the total electronic energy is obtained). As a prototypical test case, we estimate the excitation energy between the two lowest singlet states of benzene, corresponding to a  $\pi \rightarrow \pi^*$  transition. We restricted the calculation to an active space of six electrons in six orbitals, denoted as (6e, 6o), which corresponds to the  $\pi$  system of benzene.

First, we quantify the subspace error estimates provided by Theorem 1. For this purpose, we do not include shot noise on the signal samples. We prepare an initial state with multiple excitations outside of the band of interest. In more detail, our initial state contains the ground state and the first excited state—representing the in-band frequencies—and the sixth, seventh, eighth, and ninth excited states as out-of-band frequencies. All the eigenstates have the same squared overlap with the initial state  $|\langle \Psi_{\text{input}} | i \rangle|^2 \approx 0.166$ . We obtained the excited states  $|i\rangle$  using the 35 most relevant configurations obtained from the corresponding CASCI wave function [see

Eq. (57)]. Table I lists all excitation energies of all the states included in the signal with respect to the ground-state energy. These frequencies have been shifted by  $\omega^* = (E'_{\text{GS}} - E'_1)/2 = 6.31271$  with respect to the energies  $E'_i$  of the CASCI calculation. By construction, the signal bandwidth is  $W_C = 0.54191$ .

We set for our QPD calculations  $W_f = E_{|\text{S}_3\rangle} - \omega^* = 0.29170$ ,  $T_{\text{max}} = 62.832$ , and  $W_s = 5$ . Once the signal is sampled, PFD generates a refined GEP according to Algorithm 1. The choice of the trial dimension  $M$  can be optimized to minimize the subspace error. By Theorem 1, this subspace error is bounded by

$$\tilde{E}_i - E_i \leq \epsilon_M^{\text{PrIt}} \frac{\int_{\lambda > W_f} (\lambda - \tilde{E}_i) d\alpha(\lambda)}{\lambda_m(B_M^m) - \|\mathcal{N}_M^{(B)}\|} \quad (61)$$

$$\leq \epsilon_M^{\text{PrIt}} \frac{\sum_{E_k > W_f} (E_k - \tilde{E}_i) |a_k|^2}{\lambda_m(B_M^m) - \epsilon_M^{\text{PrIt}}}, \quad (62)$$

where, in the last inequality, we exploited Eq. (26). Hence, we neglect the discretization error in our analysis. Estimate (62) is characterized by three different contributions: (1) an error factor  $\epsilon_M^{\text{PrIt}}$ , (2) conditioning  $\lambda_m(B_M)^{-1}$ , and (3) integral terms that are specific to a given eigenvalue. As shown in Fig. 2, increasing the guess dimension leads to an increase in the subspace error due to the growth of  $\epsilon_M^{\text{PrIt}}$ . Conversely, the conditioning of the GEP improves as the refinement is performed with a larger number of vectors.

However, Fig. 2 shows that the decrease of  $\lambda_m(B_M)^{-1}$  in  $M$  is not significant. Therefore, we choose  $M = 2$  in our

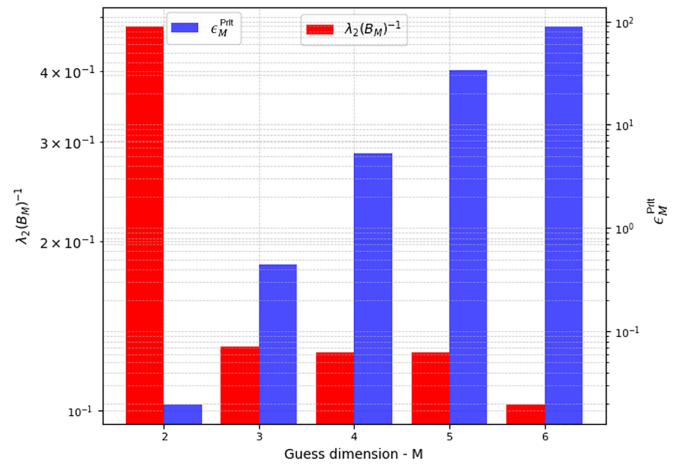


FIG. 2. Effect of the guess dimension  $M$  on the stability and accuracy of QPD. Conditioning  $\lambda_2(B_M)^{-1}$  of the refined GEP (red bars) and the subspace-based error parameter  $\epsilon_M^{\text{PrIt}}$  (blue bars) at different trial dimensions  $M$ . The error parameter is shown on a logarithmic scale.

TABLE II. Error quantification of QPD. We report the error obtained numerically,  $\epsilon$ , and the one estimated by Eq. (62),  $\epsilon_{\text{bound}}$  (given in hartree atomic units).

|                           | $E_{ \text{GS}\rangle}$ | $E_{ \text{S}_1\rangle}$ |
|---------------------------|-------------------------|--------------------------|
| $\epsilon$                | 0.000 924               | 0.001 077                |
| $\epsilon_{\text{bound}}$ | 0.002 931               | 0.001 286                |

computations to suppress out-of-band frequencies. Table II lists the errors of the energy approximations and the corresponding bounds. The results confirm that Eq. (62) provides sharp estimates of the actual errors of QPD.

We now move to our second experiment where we include statistical noise due to finite measurements on the signal samples and demonstrate that QPD, nevertheless, resolves multiple energy estimates at chemical accuracy. In this example, we construct the input state from the Hartree-Fock determinant for the ground state and the five dominant configurations of the first excited singlet, including the HOMO  $-1 \rightarrow$  LUMO and HOMO  $\rightarrow$  LUMO  $+1$  transitions. Owing to the weakly correlated character of the system, this initial state suffices to generate an autocorrelation function dominated by the ground and first excited state.

For this calculation, we set  $\omega^* = 6.3$ ,  $W_f = 1$ ,  $W_s = 3$ , and  $T_{\text{max}} = 104.72$  (in hartree atomic units, a.u.). The autocorrelation samples  $C(t_k)$  were estimated using only  $\#_{\text{shots}} = 13$  circuit shots per signal sample, yielding particularly noisy

TABLE III. Absolute energies of the ground (GS) and first excited ( $\text{S}_1$ ) states (middle columns) and the lowest  $\pi \rightarrow \pi^*$  singlet energy gap (last column) of the benzene molecule. Results obtained with QPD and with CASCI are reported in hartree.

|                | $E_{ \text{GS}\rangle}$ | $E_{ \text{S}_1\rangle}$ | $\Delta E_{ \text{GS}\rangle \rightarrow  \text{S}_1\rangle}$ |
|----------------|-------------------------|--------------------------|---------------------------------------------------------------|
| QPD (6e, 6o)   | -227.948 210            | -227.688 626             | 0.259 584                                                     |
| CASCI (6e, 6o) | -227.948 115            | -227.687 960             | 0.260 154                                                     |
| $\epsilon$     | 0.000 095               | 0.000 666                | 0.000 570                                                     |

input data  $\tilde{C}(t_k)$ . The results following the workflow of PFD are presented in Fig. 3. The spectra of the weight matrices  $B_M$  confirm that the signal has  $m = 2$  dominant frequencies in  $B_{f,\omega^*}$  [Fig. 3(a)]. We set the trial dimension to half of the critical dimension  $M = \lfloor W_f T / \pi \rfloor = 16$ . Algorithm 1 generates a refined GEP  $[A_M^m, B_M^m]$  whose eigenvalues approximate the in-band frequencies. The corresponding amplitudes are estimated through Eq. (18). Figures 3(c) and 3(d) depict the reconstruction of the filtered signal (blue solid line) from the noisy samples (red stars). Note that we actually depict the shifted autocorrelation  $C(t)e^{-i\omega^*t}$  such that relevant energies appear in the low-frequency range. The energy estimates of our computation and their accuracy with respect to a CASCI reference calculation are listed in Table III. In particular, we find that QPD surpasses chemical accuracy.

Our choice to set the trial dimension  $M$  to half the essential dimension is not necessarily optimal. However, we

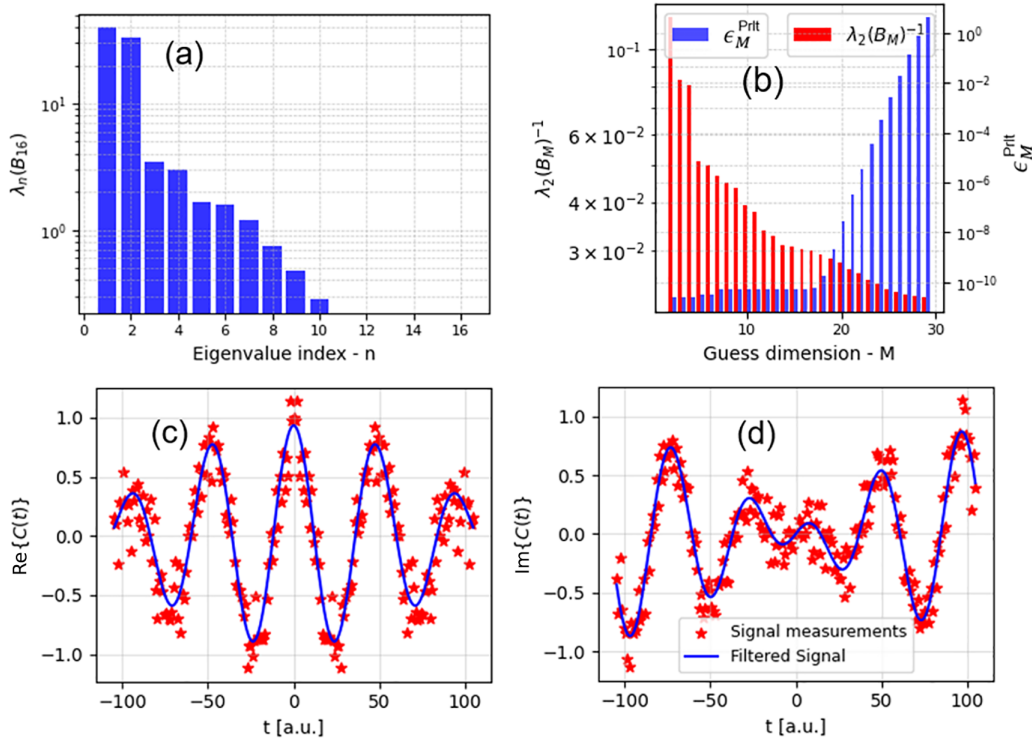


FIG. 3. QPD applied to the benzene molecule. (a) Noisy samples of the electronic autocorrelation function were applied to generate the weight matrix  $B_{16}$  [Eq. (41)] at trial dimension  $M = 16$ . Its spectrum was exploited to detect the number of dominant frequencies  $m$  within  $B_{f,\omega^*}$ . (c), (d) Red stars denote noisy samples of the shifted signal  $\tilde{C}(t_k)e^{-i\omega^*t}$ ; the blue solid line shows the reconstructed filtered signal. Panel (c) shows the real part, and panel (d) shows the imaginary part. (b) Conditioning  $\lambda_2(B_M)^{-1}$  of the refined GEP (red bars) and the subspace-based error parameter  $\epsilon_M^{\text{Prit}}$  (blue bars) at different trial dimensions  $M$ . The error parameter is shown on a logarithmic scale.

found that this setting consistently achieves low errors due to off-band frequencies and prolate sampling, while maintaining good conditioning of the weight matrix  $B_M^m$ . For noisy input data  $\tilde{C}(t_k)$ , it is especially beneficial to choose a larger trial dimension  $M \gtrsim m$ , as it improves the conditioning  $\lambda_m(B_M^m)^{-1}$  (see Appendix A and Sec. 4 of Ref. [61]). Better conditioning, in turn, improves the stability of the method with respect to additional error sources (Theorem A1).

Regarding this aspect, the stability achieved by QPD is quite remarkable, as we are able to estimate the two eigenvalues within chemical accuracy using only 13 shots per signal sample. We compare this result with other examples in the literature [53,99,100].

We begin with the in-depth analysis presented in Ref. [99], where the authors prove that, in order to maintain a constant conditioning error, the number of measurements per matrix element of a GEP must scale as  $\mathcal{O}(n^3)$ , with  $n$  denoting the dimension of the GEP. For comparison, all numerical experiments reported in their work require a number of shots on the order of at least  $10^6$  measurements per sample. References [53,100] propose an approach very similar to the one considered here, in which a subspace problem constructed by sampling snapshots of the autocorrelation function is used to determine a subset of the eigenvalues of a Hamiltonian  $H$ . In both cases, to mitigate the ill conditioning of the noisy GEP, the authors adopt a thresholding strategy [101]. Their results are obtained by adding Gaussian noise  $\mathcal{N}(0, 10^{-2})$ , which approximately corresponds to a number of measurements on the order of  $10^4$ . In contrast, our results demonstrate that our Algorithm 1 is able to reduce the required number of measurements by several orders of magnitude with respect to the results of Refs. [53,99,100]. Given the importance of this aspect for the entire class of quantum subspace algorithms [102], we will provide a more detailed discussion of this topic in future work.

We recall that sampled PFD requires a sufficiently large sampling rate  $W_s \geq W_f + W_C$ , where  $W_C$  is the bandwidth of the shifted signal  $C(t)e^{-i\omega^*t}$ . A simple way to rigorously estimate  $W_C$  is provided by spectral bounds on the electronic Hamiltonian as shown in Ref. [103]. As this probably overestimates the resources needed, a possible alternative might be a cheap classical calculation to obtain a more informed trial of  $W_C$ .

Finally, we conclude this section by remarking that, as stated in Eq. (36), the frequency density is a critical parameter at fixed  $T_{\max}$ . Of course, in routine applications the spectral density of the signal is not known *a priori*, but we note that cheap and quick classical calculations [104] can provide a useful guide to not overestimate the minimum  $T_{\max}$  needed.

### C. Numerical evidence for Heisenberg-limited scaling

In this section, we provide numerical evidence that QPD achieves Heisenberg scaling in both the weak and strong correlation limits for ground-state energy estimation and for multiple eigenvalue estimation. This has been previously achieved by only a few other algorithms relying on a single-ancilla phase estimation circuit [43,50]. We recall that Heisenberg scaling is the optimal bound achievable according to quantum information theory [31]. From an algorithmic

point of view, it implies that the quantum computational resources employed are inversely proportional to the error estimate of the protocol, that is,  $T_{\text{run-time}} = \mathcal{O}(\epsilon^{-1})$ . Opposed to this regime, we have the shot noise limit, where the accuracy of our eigenvalue estimate is bounded by the accuracy of the single sample estimate (see Sec. IIF), for which we have  $T_{\text{run-time}} = \mathcal{O}(\epsilon^{-2})$ . Moreover, we also monitored how the error estimate depends on  $T_{\max}$ , which is another important parameter for evaluating the performance of QPD. Algorithms with a better accuracy scaling in  $T_{\max}$  are particularly valuable for quantum circuits, because they lead to shallower circuits and shorter run times [see Eq. (59)].

#### 1. Ground-state energy estimation

We consider ground-state energy estimation for the linear,  $D_{\text{och}}$ -symmetric  $H_8$  system with internuclear distance  $r_{\text{HH}} = 2\text{\AA}$ . Although experimentally unstable [105], hydrogen chains have been extensively characterized as the simplest systems revealing strong electronic correlation phenomena [106,107]. Besides this system having a strong multireference character, it also features a dense spectrum in the low-energy section.

Although dense spectra are generally challenging to describe, our results demonstrate that QPD is capable of efficiently and reliably addressing them. The initial input state used for these calculations was obtained by uploading the first five most relevant determinants from the exact FCI wave function. Of course, in an actual application where QPD is applied to larger systems (specifically, too large orbital bases), no FCI reference (i.e., reference from exact diagonalization) wave function will be available. Instead, one may exploit approximate CI methods or approximate matrix product state calculations. The latter have been successfully exploited for the automated construction of active orbital spaces [104] and for the analysis and preparation of initial states for QPE [108]. With such an input state, we obtain an initial overlap  $|\langle \Psi | \text{GS}_{\text{FCI}} \rangle|^2 \approx 0.33$ . We set the sampling rate to  $W_s = 3$  a.u. Figure 4 collects the results of multiple calculations obtained by scanning the total length of the sampled signal. Final results were computed by binning the run-time axis (or the  $T_{\max}$  axis) into 20 bins and calculating the mean and standard deviation within each bin.

In Figure 4(a), which reports the error on the ground-state energy estimate (in hartree) as a function of the run-time  $T_{\text{run-time}}$  [Eq. (58)], we notice a very good agreement with the inverse scaling law dictated by the Heisenberg limit. Figure 4(b) shows that QPD exhibits a scaling of  $\epsilon = \mathcal{O}(T_{\max}^{-3})$ .

This result represents a cubic speedup compared to the standard QPE scaling,  $\epsilon = \mathcal{O}(T_{\max}^{-1})$ , and a quadratic improvement with respect to other state-of-the-art approaches [50,109]. We point out that Eq. (59) and our choice of  $\#_{\text{shots}}$  imply a scaling  $T_{\text{run-time}} = \mathcal{O}(T_{\max}^{2.5} \sqrt{\log(T_{\max})})$ , which is almost cubical and in agreement with our numerical results. Finally, we note that a cubic improvement in circuit depth (maximal evolution time) has already been shown in the literature when a cosine window was inserted into the standard QPE circuit [110].

#### 2. Multiple eigenvalue estimation

We now consider the more general task of multiple eigenvalue estimation. Thus, we aim to determine more than one

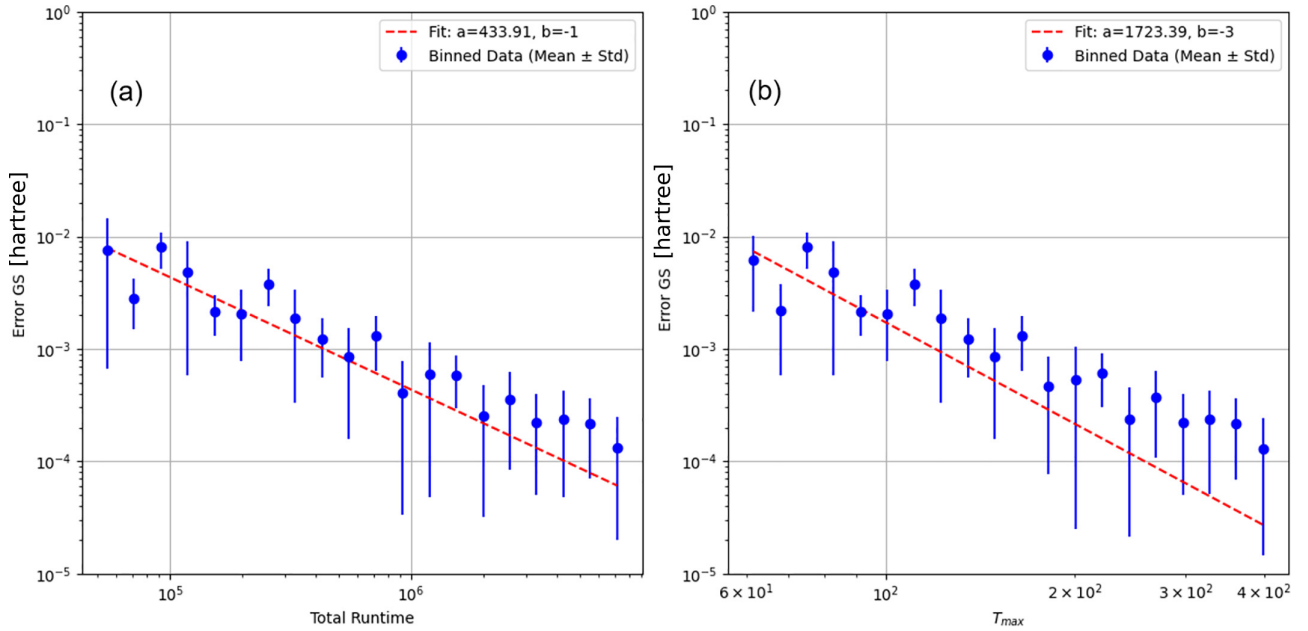


FIG. 4. Empirical scaling of QPD for the task of ground-state energy estimation for  $H_8$ . Left (a) Average error on the ground-state energy as a function of the total run-time [see Eq. (58)]. Right (b) Average error on the ground-state energy as a function of the maximal evolution time  $T_{\max}$ . Blue dots and corresponding error bars are obtained by averaging different runs of QPD where we scan over the maximal evolution time  $T_{\max}$ . All other parameters have been set according to Sec. III A. Red dashed lines correspond to the fit of the function  $y = a e^b$ ;  $b$  values are kept fixed and set to  $b = -1$  ( $-3$ ) in the left (right) plot.

frequency within a given band  $B_{f,\omega^*}$ . We will examine two systems representing opposite limits of weak and strong correlation. As an example for the weak correlation case, we study the LiH molecule at its equilibrium internuclear distance; for the strong correlation regime, we again exploit the  $H_8$  system. In the first case, all frequencies present in the signal also lie in band  $B_{f,\omega^*}$ . Therefore, there are no out-of-band contributions. In contrast, the second example corresponds to signals with many frequencies outside of  $B_{f,\omega^*}$ . In this case, we only recover a small fraction of all the frequencies present in  $C(t)$ .

Although quantum computing approaches are typically considered to offer the greatest advantages for strongly correlated systems [111], we believe that a utility-scale quantum algorithm should also be able to address efficiently, that is, with provable advantages over classical computations, a wider span of the electron-correlation (and hence, chemical) space. While electronic correlation has been thoroughly classified [108,112,113], we here focus specifically on its impact on correlation functions  $C(t)$ . In Ref. [114], Zurek showed that correlation function decay rates and fluctuation amplitudes depend on the number of dynamically active states. The higher the number of states involved in the dynamics, the faster the decay of autocorrelations. This is consistent with the orthogonality catastrophe [115]: The system's evolution tends toward states with exponentially shrinking overlaps with a given eigenstate. Increasing the number of states also suppresses long-time fluctuations. In weakly correlated systems,  $C(t)$  preserves long-time oscillations, even with few initial determinants in the initial input state. In contrast, for strong correlations, we observe rapid damping and suppression of recurrence beatings. Noise sensitivity increases as oscillations fade, masking signal information.

Analogously to the previous section, Fig. 5 collects the results of multiple calculations obtained by scanning the total length of the sampled signal. Final results were computed by binning the run-time axis (or the  $T_{\max}$  axis) into 20 bins and calculating the mean and standard deviation within each bin. For the weak correlation example, we estimated simultaneously the eigenvalues of the four lowest eigenstates of LiH at equilibrium distance ( $r_{\text{LiH}} = 1.6 \text{ \AA}$ ). These are  $|\text{GS}\rangle$ ,  $|\text{S}_1\rangle$ ,  $|\text{S}_2\rangle$ , and  $|\text{S}_4\rangle$ . We built the input state by selecting, for each eigenstate, the four most important determinants. Again, we point out that scaling up the initial state preparation problem to larger systems can be done via approximate matrix product states calculations [104,108]. Because of the weakly correlated nature of the system, the overlap of this few-determinant approximation with the target state is about 1.0. In all LiH calculations, we set a value of  $W_s = 3 \text{ a.u.}$  As we can understand from the results in Fig. 5, all considerations and observations made for the ground-state energy estimation problem in the previous section also apply to this case, with the average error on the excitation energies being inversely proportional to the quantum run-time and scaling as  $\mathcal{O}(T_{\max}^{-3})$  with the maximal evolution time [see Figs. 5(a) and 5(b)]. Notably, the multiplicative prefactor [that is, the  $a$  parameter in the legend of Figs. 5(a) and 5(b)] resulting from the power-law fit (red dashed line) is slightly smaller than the ones obtained in the previous section (the  $a$  parameter in Fig. 4). This observation may be related to the facts that (1) the overall input wave function shows a high fidelity with the target state and that (2) the energy spectrum is much less dense for LiH than for  $H_8$ . For the multiple eigenvalue estimation on  $H_8$  [Figs. 5(c) and 5(d)], the choice of the initial input state requires more attention. Figure 6 presents a characterization of

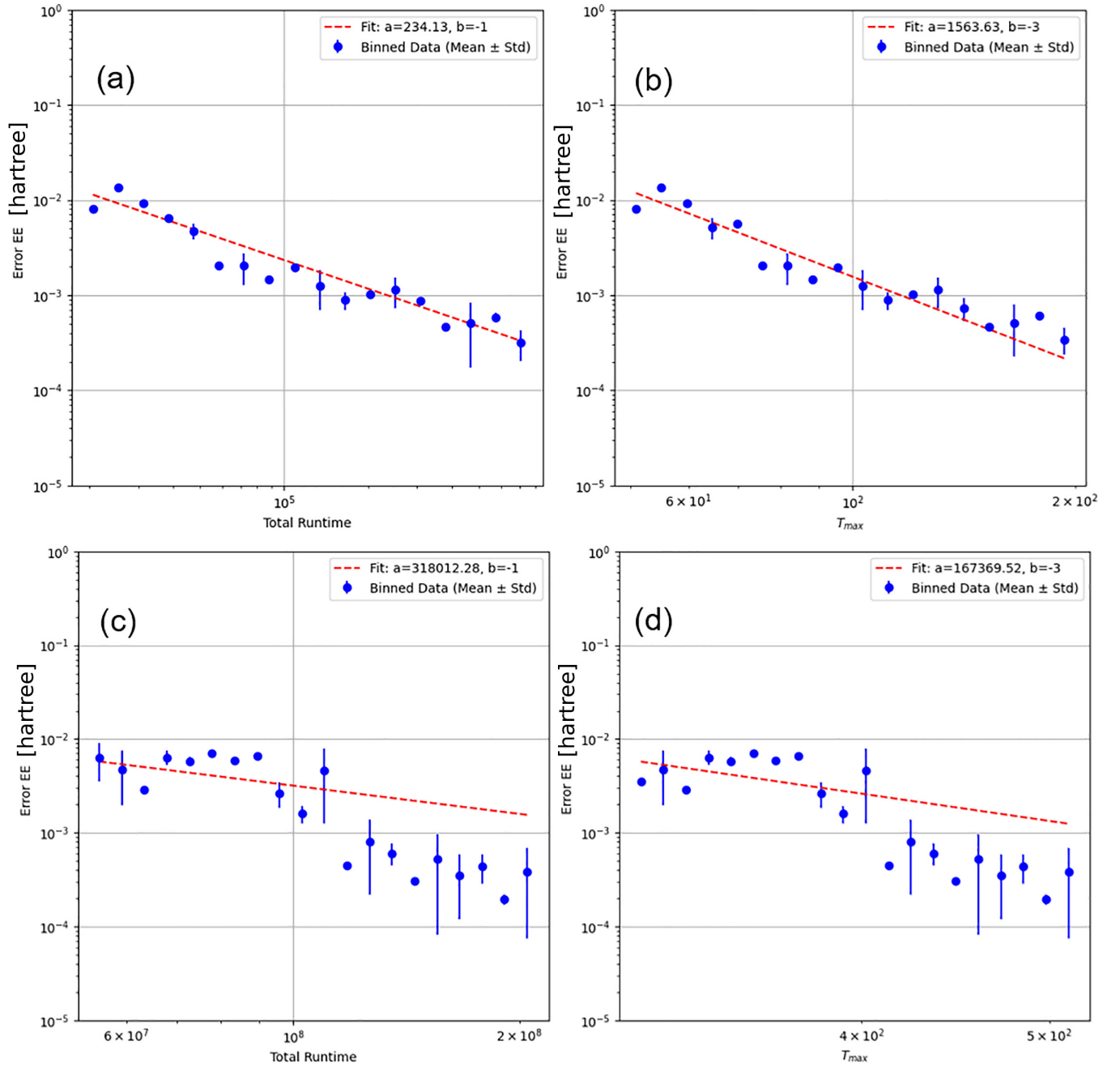


FIG. 5. Empirical scaling of QPD in the weak [LiH molecule, panels (a) and (b)] and strong [ $H_8$  system, panels (c) and (d)] correlation limit for the task of simultaneous multiple eigenvalue estimation. Left: Average error on the excitation energies as a function of the total run-time [see Eq. (58)]. Right: Average error on the excitation energies as a function of the maximal evolution time  $T_{\max}$ . Blue dots and corresponding error bars were obtained by averaging different runs of QPD where we scan over the maximal evolution time  $T_{\max}$ . All other parameters have been set according to Sec. III A. Red dashed lines correspond to the fit of the function  $y = a e^b$ ;  $b$  values were kept fixed and set to  $b = -1$  ( $-3$ ) in the left (right) plot.

the input state used to obtain the results of Figs. 5(c) and 5(d). The initial input state is defined as noted above in Eq. (56) and thus reads

$$|\Psi\rangle = \frac{1}{2}[|GS\rangle + |S_1\rangle + |S_2\rangle + |S_4\rangle]. \quad (63)$$

Figure 6(a) presents the total squared overlap  $|\langle\Psi|\Psi_{\text{FCI}}\rangle|^2 = \frac{1}{4} \sum_i |\langle\Psi|i_{\text{FCI}}\rangle|^2$  of the state  $|\Psi\rangle$  and each corresponding FCI reference eigenstate wave function as a function of the num-

ber of determinants included in each approximation  $|i\rangle$ . With increasing number of determinants, the total squared overlap improves monotonically, approaching unity, which shows the systematic convergence of the prepared state toward the target state. After having characterized the total squared overlap as a function of the number of determinants included in our approximation for the target eigenstates, we selected for the calculations shown in Figs. 5(c) and 5(d), the 90 most relevant determinants per eigenstate.

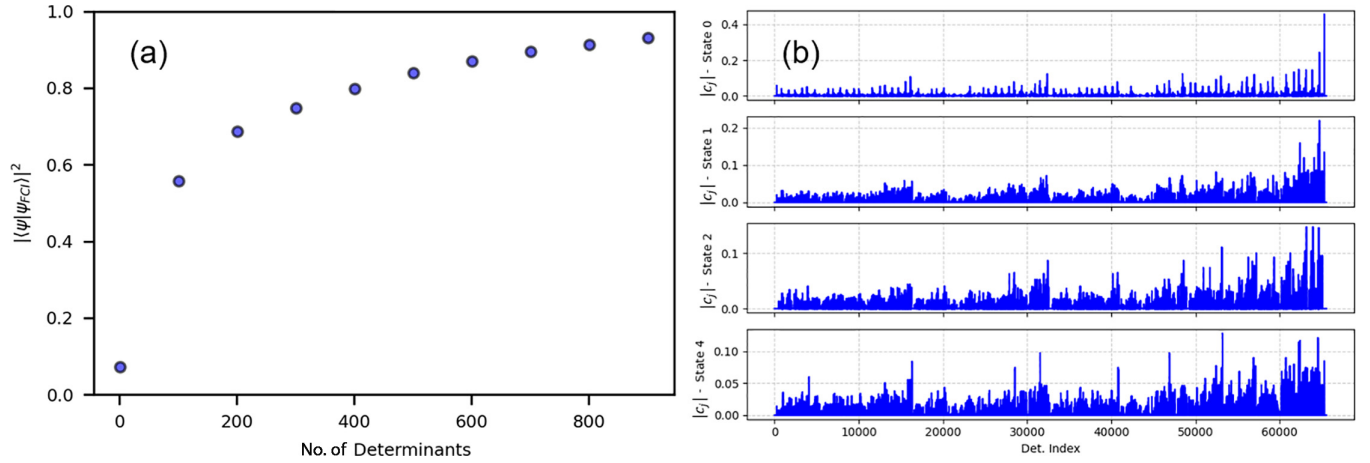


FIG. 6. Initial input state characterization for the calculations on  $H_8$  reported in Figs. 5(c) and 5(d). (a) Initial input state total squared overlap as a function of the number of Slater determinants considered. (b) Absolute values of the expansion coefficients  $c_j$  for the targeted states:  $|GS\rangle$ ,  $|S_1\rangle$ ,  $|S_2\rangle$ , and  $|S_4\rangle$  [see Eq. (57)].

With this setting, the total squared overlap between the initial input state and the target eigenstates is approximately 0.55. This value implies for the excited states (whose wave function is less peaked than the ground-state wave function; Fig. 6) a squared overlap  $|\langle\Psi|\psi_{\text{FCI}}\rangle|^2 \approx 0.1$ . In all calculations on  $H_8$ , we set  $W_s = 6$  a.u. and  $W_f \approx 0.47$ , and the number of frequencies within  $B_{f,\omega^*}$  was set to  $m = 7$ .

For the sake of simplicity, we chose to include the same number of electronic configurations for each target eigenstate. This choice is not optimal and implies that the frequencies we aim to recover are not equally represented in  $C(t)$ . The reason for this can be seen in Fig. 6(b), which shows the distribution of the absolute value of the expansion coefficients,  $|c_j|$ , [cf. Eq. (57)] for each determinant contributing to the target eigenstates. While the ground state still shows a slight predominance of the  $|\text{HF}\rangle$  determinant, we observe that the higher the energy of a state is, the less compact its representation becomes. This is not only due to the intrinsic electronic structure of the system, but also reflects the nature of the CI expansion built upon the mean-field determinant  $|\text{HF}\rangle$ , with orbitals optimized for the ground state by virtue of the variational principle.

We also observe that the average error in the excitation energies (plotted as a function of both the total run-time and  $T_{\text{max}}$ ) follows two distinct regimes, with a transition occurring around  $T_{\text{run-time}} \approx 10^8$  and  $T_{\text{max}} \approx 400$  a.u. This behavior can be explained in the light of Eq. (36): Data points at lower values of  $T_{\text{run-time}}$  (and  $T_{\text{max}}$ ) do not satisfy this condition. This is due to the high density of the spectrum of  $H_8$ , where the eigenvalues of interest are very close in energy. As a result, the accuracy exhibits a marked change in behavior once  $T_{\text{max}}$  (and therefore also the total run-time) exceeds a certain threshold. If we, however, focus on the region  $T_{\text{run-time}} \geq 10^8$  and  $T_{\text{max}} \geq 400$  a.u., we find again that the average error on the excitation energies behaves as per Figs. 4, 5(a), and 5(b).

Finally, to conclude this section, we emphasize that the spectral density relevant to the accuracy of the sampled PFD is dictated by the choice of  $B_{f,\omega^*}$ , since QPD is tasked with recovering all frequencies within this bandwidth. For this reason, we study the effect of the initial input state overlap

in the next section, which can deeply affect the frequency distribution determining  $C(t)$ .

#### D. Effect of the initial input state

Selecting and initializing well-suited quantum states is deeply intertwined with the computational difficulty of solving electronic structure problems. In the general case, computing ground-state energies of local Hamiltonians is proven to be QMA-hard (Quantum-Merlin-Arthur, the quantum analogue of NP) [116,117]. Nonetheless, from a total resource perspective, encoding a classically obtained state on a quantum device is generally far less resource intensive than executing a complete quantum energy estimation algorithm, regardless of the specific technique employed [35,118]. In more detail, it has been shown [35] that preparing an initial input state consisting of about a thousand determinants [e.g., expansion terms in Eq. (57)] for a system with one hundred orbitals would require  $10^5$ – $10^6$  Toffoli gates, while the total circuit cost for a standard QPE circuit for such a system would be on the order of  $10^{10}$  Toffoli gates. This means that the cost of preparing an initial input state with sufficient support to run efficiently a phase estimation procedure should be considered as relatively negligible with respect to the gains in terms of total run-time as compared to a poor initialization. Furthermore, if one considers embedding methods [34] or recent results regarding MPS preparation [119,120], this cost is reduced even further.

This fact, along with the consequences of the orthogonality catastrophe [115], emphasizes the need to understand how the initial state selection affects the efficiency and accuracy of a quantum algorithm. Assessing the utility of a quantum algorithm hinges on the capacity of the latter to extract eigenvalues at chemically relevant accuracy despite an initial state that can be generated with limited amount of classical resources. In Fig. 7, we report the results of the effect of the initial input state quality on the accuracy of the energy estimate of QPD. All calculations were carried out with a sampling rate equal to  $W_s = 6$  a.u. For each number of determinants included in the initial input state (ranging from 1 up to 901 determinants),

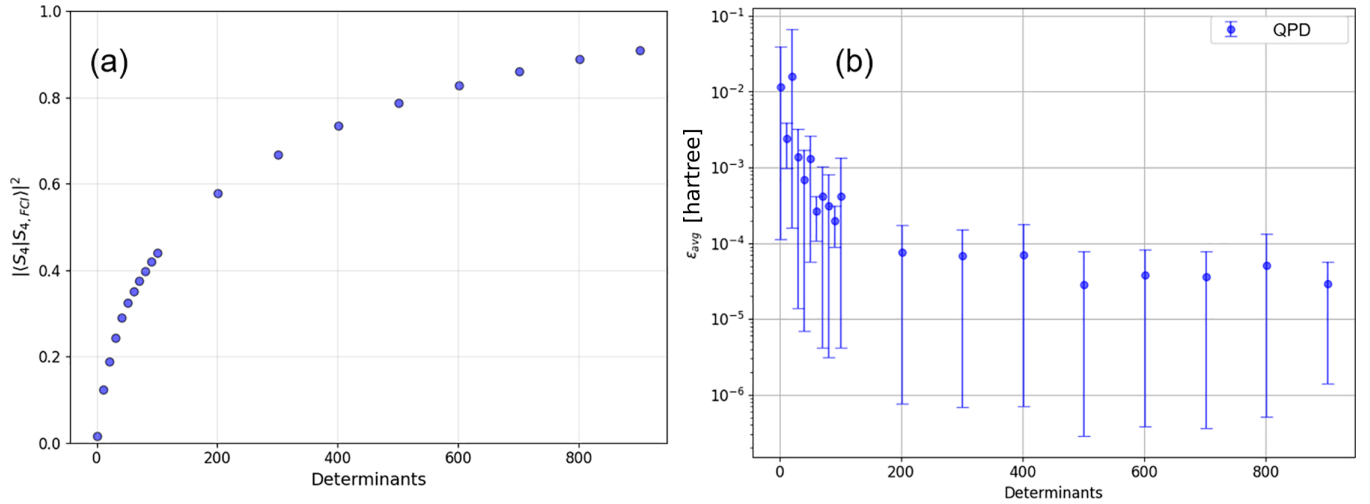


FIG. 7. Effect of the initial input state on the QPD algorithm for the  $H_8$  system. (a) Initial input state squared overlap as a function of the number of Slater determinants included in the initial input state for the  $|S_4\rangle$  state. (b) Average error on the energy estimate as a function of the number of Slater determinants included in the initial input state. Averages were taken over different QPD runs spanning a range of  $T_{\text{max}}$  for the sampled autocorrelation function.

we ran 60 QPD calculations scanning the value of  $T_{\text{max}}$  from  $T_{\text{max}} = 400$  a.u. to  $T_{\text{max}} = 1000$  a.u. In all linear combinations constructed for the different initial input states, we selected the determinants with the highest weight in the exact FCI reference wave function.

In Fig. 7, we report the results of in total 1140 different calculations with average results per each fixed number of determinants included in the initial input state (dots) and the corresponding standard deviation (bars).

Figure 7(a) shows the squared overlap of an approximate state and the exact FCI state as a function of the number of determinants employed in the approximation. Figure 7(b) shows the average error  $\epsilon_{\text{avg}}$  (over the multiple runs) for the fourth excited singlet state  $|S_4\rangle$  of the  $H_8$  system, again as a function of the number of determinants. A logarithmic scale is used on the vertical axis to better visualize the wide range of errors.

In Fig. 7(b), we observe three distinct regimes in the error scaling with respect to the initial state quality. First, the error decreases rapidly until approximately 60 determinants are included, corresponding to an initial squared overlap of  $|\langle \Psi | S_{4,FCI} \rangle|^2 \approx 0.35$ . In the second regime, adding further determinants continues to reduce the error below the chemical accuracy threshold ( $\epsilon_{\text{avg}} \leq 10^{-3}$  hartree), although the rate of improvement is reduced. Finally, beyond 200 determinants, the average error plateaus in the range of  $10^{-4}$  to  $10^{-5}$  hartree.

Our analytical framework (particularly Sec. II C) provides a rigorous explanation for this behavior. In the first regime, the initial state reaches a sufficient overlap such that the in-band amplitude  $|a|_{\text{min}} = |\langle \Psi | S_{4,FCI} \rangle|$  satisfies the detectability condition of Eq. (20). Below this threshold, the signal is indistinguishable from the noise background. Beyond this threshold (second regime), increasing the overlap further improves the conditioning of the weight matrix,  $\lambda_m(B_M^m)^{-1}$ . As established in the general Theorem A1, a better-conditioned weight matrix linearly suppresses perturbations from arbitrary error sources—including stochastic shot noise—leading to the observed accuracy gains. However, in this regime, Eq. (20)

implies  $\lambda_m(B_M^m)^{-1} \sim |a|_{\text{min}}^{-2}$ , and the advantage of increasing  $|a|_{\text{min}}$  eventually saturates (third regime). Beyond sufficiently large overlaps ( $|a|_{\text{min}}^2 \gtrsim 0.5$ ), further increases yield diminishing returns, as QPD already efficiently suppresses out-of-band frequencies. Consequently, the overall error plateaus at the dominant error level set by the shot noise.

These results demonstrate the robustness of QPD to imperfect state preparation: The method achieves chemical accuracy with initial states accessible via classical approximations, provided they satisfy a modest overlap threshold. This threshold depends on the magnitude of the other error sources.

#### IV. CONCLUSIONS

We presented an eigenvalue problem that makes the duality between spectral analysis, as in eigen decompositions of operators, and frequency decompositions of functions explicit. To solve this spectral equation, we developed a general subspace-based approximation framework that enables a nonasymptotic error analysis of multiple error sources. We rigorously studied the constraint imposed by finite observation/simulation time (deriving a nonasymptotic error bound) and the effect of a finite number of signal samples. As a result, we obtained a method that extracts the frequencies of an autocorrelation function  $C(t)$  with high accuracy. Additional error sources, such as shot noise, can be incorporated seamlessly into the same framework by quantifying their effect on the relevant matrices.

By starting from such a nonasymptotic error framework, our work deviates from previous approaches to rigorous error analysis, such as those in Refs. [42–50]. The framework makes it possible to study multiple error sources in a nonasymptotic manner, which is necessary for error balancing. This paves the way toward optimal approximation that connects phase estimation with subspace-based approximation and draws on insights from prolate Fourier theory.

Our analysis yields a quantitative relationship between observation time, spectral density, and achievable precision.

In particular, we derived a sharp *accuracy transition* that resembles the Heisenberg uncertainty relation (Secs. II D and II E). Our accuracy guarantees distinguish between subspace-based error and discretization-based error. Remarkably, the transition appears simultaneously in both, which underscores the maturity of our protocol.

To provide a self-contained and complete picture of our theory, we deferred some of its key parts to the Appendixes in favor of a streamlined presentation in the main article. In particular, our theory is built upon the framework of subspace protocols (Appendix A) and insights on optimal time and band concentration (Appendix B). We derived new identities on prolates to give error bounds for our routines (Appendix B 2). These new identities also enabled a refined error bound for the truncated prolate sampling formula that gives a rigorous grounding for the long-standing engineering folk theorem (Appendix C). Much of the mathematical depth in prolate Fourier theory lies in a commutation relation between an integral and a differential operator. In this work, we highlighted a spectral identity for the differential operator that mirrors Landau's classical threshold result for the integral operator (Appendix B 3 [68]). We provided a simplified proof of this inequality, which strengthens a previously known bound [121].

For the study of molecular systems, we combined our classical signal processing protocol with the Hadamard test circuit to generate samples of the autocorrelation function. This algorithm we call quantum prolate diagonalization. Our numerical experiments show that QPD exhibits Heisenberg scaling, with the estimation error decreasing proportionally to  $1/T_{\text{run-time}}$  over several orders of magnitude in  $T_{\text{run-time}}$ , for both single- and multiple-eigenvalue estimation. Furthermore, our numerical experiments indicate that the average error of the eigenvalues decreases quadratically faster with the maximal evolution time compared to most recent results in the literature [50]. Our finding is similar to what is known in the literature for the standard QPE algorithm with a cosine filter applied on the readout register [110].

Unlike most quantum subspace methods [102], QPD is resilient to shot noise and does not suffer from conditioning issues, as demonstrated in our simulations. The theory of subspace protocols effectively resolves the conditioning issue through a trial space refinement and accurate dimension detection (Appendix A). This has been a long-standing issue for quantum subspace methods. Previous approaches addressed it with a thresholding strategy applied to the overlap matrix [53,99,100], but the number of measurements required to achieve chemical accuracy was on the order of 10 000 measurements per evaluated matrix element. In stark contrast, QPD achieves the same level of chemical accuracy with only 10–100 shots per sample of the autocorrelation function.

We validated the performance of QPD on realistic quantum chemistry problems, with special attention paid to the role of the input state for the observed accuracy. These results served also as a test for QPD, because, as the recent work by Kiss *et al.* [118] highlights, practical implementations of early fault-tolerant algorithms can expose details that may be overlooked from a purely theoretical standpoint.

We have demonstrated that the concentration properties of prolates allow us to accurately estimate multiple eigenvalues with very short evolution times. As a future perspective to

further reduce the computational cost of QPD, we plan to implement sparse sampling approaches [122,123], which may decrease the number of necessary signal samples.

From a general perspective, not solely focusing on quantum computation, the results presented in this work mark progress toward a deeper understanding of optimal computation. Developing approximation routines to a given accuracy with minimal resources require a rigorous understanding of all contributing error sources. This, in turn, allows us to allocate computational resources across different subroutines to balance their respective error contributions. Here, we have taken a first step in this direction by providing a rigorous description of both subspace-based and discretization errors.

## ACKNOWLEDGMENTS

The authors acknowledge financial support from the Swiss National Science Foundation through Grant No. 200021\_-219616 and from the Novo Nordisk Foundation (Grant No. NNF20OC0059939 “Quantum for Life”). We thank the organizers of WATOC2025 for the opportunity to present this work in Oslo in June 2025.

## DATA AVAILABILITY

The source code that supports the findings of this article is openly available in a GitHub repository [97], and all data underlying the figures can be reproduced with it.

## APPENDIX A: SUBSPACE PROTOCOLS

To facilitate the derivation of accuracy guarantees for such subspace-based methods, we established Theorem A1 as a master theorem. In fact, the accuracy guarantees of PFD (Theorem 1) and sampled PFD (Theorem 2) follow as corollaries of Theorem A1. In this Appendix, we briefly review subspace-based methods and present key results adapted from Ref. [61].

We assume a self-adjoint operator  $H$  acting on a Hilbert space  $\mathcal{H}$  and consider the eigenvalue problem

$$H\varphi_i = \lambda_i\varphi_i, \quad (\text{A1})$$

with  $\varphi_i \in \mathcal{H}$  and  $\lambda_i \in \mathbb{R}$ . Subspace-based methods aim to approximate the eigenvalues of  $H$  within a spectral subspace  $\mathcal{E}[a, b]$  using a GEP. A numerical routine  $\mathbf{P}$  generates a set of trial vectors  $v_1, \dots, v_m \in \mathcal{H}$  such that they *approximately* span the spectral subspace  $\mathcal{E}[a, b]$ . Let  $V : \mathbb{C}^m \rightarrow \mathcal{H}$  be the linear map defined by  $b \mapsto \sum_i v_i b_i$  and consider the subspace eigenvalue problem

$$V^\dagger H V b_i = \tilde{\lambda}_i V^\dagger V b_k. \quad (\text{A2})$$

Here,  $b_k \in \mathbb{C}^m$  and  $\tilde{\lambda}_i \in \mathbb{R}$ . The generalized eigenvalues  $\tilde{\lambda}_i$  serve as approximations to the eigenvalues of  $H$  contained in the interval  $[a, b]$ . Reference [61] develops a mathematical framework to quantify the accuracy of such subspace-based methods. To enable a rigorous analysis, the conceptual abstraction of *subspace protocols* is introduced as a unifying formalism.

*Definition A1 (Subspace protocols).* A numerical routine  $\mathbf{P}$  is called a *subspace protocol* if, for any self-adjoint operator  $H$  on any Hilbert space  $\mathcal{H}$  and any spectral subspace of  $\mathcal{E}$

ALGORITHM 1. Operation of a subspace protocol  $\mathbf{P}_{(H,\mathcal{E})}(M, \epsilon_{\text{th}})$ .

- 
1. Use  $\mathbf{P}$  to compute the matrices  $A_M, B_M \in \mathbb{C}^{M \times M}$ , where  $M$  is the current trial dimension. Diagonalize  $B_M$  and obtain its eigenvalues.
  2. Determine the *detected dimensionality*  $m$  as the number of eigenvalues of  $B_M$  exceeding the threshold  $\epsilon_{\text{th}}$ . If  $m = M$ , increase the trial dimension  $M$  and repeat the process.
  3. Let  $U_m \in \mathbb{C}^{M \times m}$  be the matrix whose columns are the  $m$  leading eigenvectors of  $B_M$ . Form the projected matrices  $A_M^m := U_m^\dagger A_M U_m$  and  $B_M^m := U_m^\dagger B_M U_m$ . The resulting GEP  $[A_M^m, B_M^m]$  is of dimension  $m$  and satisfies  $\lambda_m(B_M^m) > \epsilon_{\text{th}}$ , ensuring it is well conditioned.
  4. Compute the generalized eigenvalues of  $[A_M^m, B_M^m]$ . These serve as approximations to the eigenvalues of  $H$  in the subspace  $\mathcal{E}$ .
- 

of  $H$ , the protocol  $\mathbf{P}$  generates for any given trial dimension  $M \in \mathbb{N}$  a generalized eigenvalue problem of the form

$$A_M b_k = \tilde{\lambda}_k B_M b_k. \quad (\text{A3})$$

Here,  $A_M, B_M \in \mathbb{C}^{M \times M}$  are self-adjoint matrices,  $b_k \in \mathbb{C}^M$  and  $\tilde{\lambda}_k \in \mathbb{R}$ . We denote the GEPs as Eq. (A3) as  $[A_M, B_M]$  and refer to  $B_M$  as the *weight matrix*. For a fixed pair  $(H, \mathcal{E})$ , we write  $\mathbf{P}_{(H,\mathcal{E})}(M)$  for the GEP generated by  $\mathbf{P}$  at a trial dimension  $M$ . Moreover, we assume the GEPs of a subspace protocol admit decompositions

$$A_M = V_M^\dagger H V_M + \delta A_M, \quad (\text{A4})$$

$$B_M = V_M^\dagger V_M + \delta B_M, \quad (\text{A5})$$

where  $V_M : \mathbb{C}^M \rightarrow \mathcal{H}$  is a linear map, called the *trial vector map*, and  $\delta A_M, \delta B_M \in \mathbb{C}^{M \times M}$  are self-adjoint matrices modeling numerical error. Let  $P_{\mathcal{E}}$  denote the orthogonal projection operator onto  $\mathcal{E}$  and define  $N_M := P_{\mathcal{E}}^\perp V_M$ . Then, the matrix

$$\mathcal{N}_M^{(B)} := N_M^\dagger N_M + \delta B_M \quad (\text{A6})$$

quantifies the deviation from the ideal projection onto  $\mathcal{E}$  and is referred to as the *noise weight*.

The matrices  $\delta A_M, \delta B_M$  can capture deviations in the computation of the matrix elements of  $V_M^\dagger H V_M$  and  $V_M^\dagger V_M$ . For example, in sampled PFD of Sec. II E, these matrix elements are approximated through the prolate sampling formula. Noise in processed data may introduce additional deviations, which can also be modeled by  $(\delta A_M, \delta B_M)$ . It is advantageous to have analytical bounds on the noise weights—particularly for a reliable detection of spectral subspace dimensions. This motivates a stricter class of subspace protocols.

**Definition A2 ( $\epsilon$ -subspace protocols).** Let  $\mathbf{P}$  be a subspace protocol as in Definition A1, and let  $\{\epsilon_M\}_{M=1}^\infty$  be a sequence of non-negative real numbers that bound the noise weights:

$$\|\mathcal{N}_M^{(B)}\| \leq \epsilon_M \quad \text{for all } M \in \mathbb{N}. \quad (\text{A7})$$

Then, the pair  $(\mathbf{P}, \{\epsilon_M\}_{M=1}^\infty)$  is called an  $\epsilon$ -subspace protocol and we denote it by  $\mathbf{P}^{\{\epsilon_M\}}$ .

The purpose of a subspace protocol is to approximate the eigenvalues of  $H$  in the spectral subspace  $\mathcal{E}$ . However, the dimension of  $\mathcal{E}$  is often unknown *a priori*. In such cases, the dimension can be inferred either from a significant drop in the spectrum of the weight matrix  $B_M$  or by applying a noise threshold  $\epsilon_{\text{th}} > 0$ .

During the numerical operation, the protocol  $\mathbf{P}$ , the operator  $H$ , and the target subspace  $\mathcal{E}$  remain fixed, while the trial dimension  $M$  and the threshold  $\epsilon_{\text{th}}$  are variable parameters. Accordingly, we define

$$\mathbf{P}_{(H,\mathcal{E})}(M, \epsilon_{\text{th}}) \quad (\text{A8})$$

as an instance of Algorithm 1, specified by fixed choices of  $\mathbf{P}, H$ , and  $\mathcal{E}$ . The algorithm takes  $(M, \epsilon_{\text{th}})$  as input, detects the dimension of  $\mathcal{E}$ , and returns approximations of the eigenvalues of  $H$  in  $\mathcal{E}$ .

If, in step 2, all eigenvalues of  $B_M$  are smaller than  $\epsilon_{\text{th}}$ , a zero subspace is detected, and we set  $m = 0$ . This can occur either when the target spectral subspace  $\mathcal{E}$  is trivial (i.e.,  $\mathcal{E} = 0$ ), or when the weight matrix  $B_M$  is too poor to distinguish signal from noise. The refinement in step 3 restricts the GEP to the dominant subspace of  $B_M$ . This projection improves numerical stability and allows to use an initial trial dimension  $M$  that exceeds  $\dim(\mathcal{E})$ .

If  $\mathbf{P}$  is an  $\epsilon$ -subspace protocol, the noise threshold  $\epsilon_{\text{th}}$  is determined by the analytic bounds  $\epsilon_M$  satisfying Eq. (A7). In particular, since the noise level generally increases with the number of trial vectors, an appropriate threshold  $\epsilon_{\text{th}}$  is typically dimension dependent:

$$\epsilon_{\text{th}}(M) = \epsilon_M. \quad (\text{A9})$$

For  $\epsilon$ -subspace protocols, Theorem 1 of Ref. [61] guarantees that dimension detected by  $\mathbf{P}_{(H,\mathcal{E})}^{\{\epsilon_M\}}(M, \epsilon_M)$  does not exceed  $\dim \mathcal{E}$ .

Increasing  $M$  beyond  $\dim(\mathcal{E})$  enhances the expressiveness of the trial vector space, making it more likely to capture the signal subspace and to reveal a sharp spectral gap in  $B_M$ . This improves both the reliability of dimension detection and the conditioning of the refined GEP (Sec. 4.1 of Ref. [61]). However, this benefit comes at a cost: Higher  $M$  typically also increases the noise level  $\epsilon_M$ . Therefore, the choice of  $M$  presents a trade-off between improved signal representation and increased contributions from noise.

If the dimension of  $\mathcal{E}$  is known in advance or has been reliably detected using  $\mathbf{P}_{(H,\mathcal{E})}(M, \epsilon_{\text{th}})$ , we define a related protocol  $\mathbf{P}_{(H,\mathcal{E})}(M, m)$ , which takes the dimension  $m$  directly as input instead of a noise threshold. The algorithm  $\mathbf{P}_{(H,\mathcal{E})}(M, m)$  is identical to Algorithm 1, except that step 2 is omitted.

Once  $\mathbf{P}_{(H,\mathcal{E})}(M, \epsilon_{\text{th}})$  has correctly identified the dimension  $m$ , Theorem A1 provides an accuracy guarantee for the eigenvalues returned by  $\mathbf{P}_{(H,\mathcal{E})}(M, m)$ . The theorem distinguishes between two sources of error: A *subspace-based error*, arising from  $\text{Range } V_M \neq \mathcal{E}$  and arbitrary other errors, modeled by perturbations  $\delta A_M$  and  $\delta B_M$  in the matrix elements of the GEP. The subspace-based error is quantified by the measure  $\varepsilon_M^m : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ , defined as

$$\varepsilon(I) = \text{Tr}[V^\dagger P_{\mathcal{E}^\perp} P^{(H)}(I) V], \quad (\text{A10})$$

where  $P^{(H)}$  is the unique projection-valued measure associated with  $H$ , such that  $H = \int \lambda dP^{(H)}(\lambda)$ . We denote by  $\tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_m$  the generalized eigenvalues of the GEP  $[A_M^m, B_M^m]$ ,

and by  $\lambda_1 \geq \dots \geq \lambda_m$  the eigenvalues of  $H$  restricted to  $\mathcal{E}$ , where  $m = \dim \mathcal{E}$ . To compactly handle indefinite behavior in self-adjoint error matrices  $E$  (such as  $\delta A_M^m - \tilde{\lambda}_i \delta B_M^m$ ), we adopt the notation  $\lambda_{\min}^*(E) := \min(\lambda_{\min}(E), 0)$  and  $\lambda_{\max}^*(E) := \max(\lambda_{\max}(E), 0)$ .

**Theorem A1.** Let  $\mathcal{P}$  be a subspace protocol as in Definition A1 and  $m = \dim \mathcal{E}$ . Consider a trial dimension  $M \geq m$  and let  $[A_M^m, B_M^m]$  be the GEP returned by  $\mathcal{P}_{(H, \mathcal{E})}(M, m)$ . Assume that the GEP is well conditioned in the sense

$$\lambda_m(B_M) > \|\mathcal{N}_M^{(B)}\| \quad (\text{A11})$$

and  $\tilde{\lambda}_i \in [a, b]$  for all  $i$ . Then, for each  $i = 1, \dots, m$ , the eigenvalues of  $[A_M^m, B_M^m]$  satisfy

$$\begin{aligned} \frac{\int_{-\infty}^a (\lambda - \tilde{\lambda}_i) d\varepsilon(\lambda) + \lambda_{\min}^*(\delta A_M^m - \tilde{\lambda}_i \delta B_M^m)}{\lambda_m(B_M) - \|\mathcal{N}_M^{(B)}\|} &\leq \tilde{\lambda}_i - \lambda_i \\ &\leq \frac{\int_b^\infty (\lambda - \tilde{\lambda}_i) d\varepsilon(\lambda) + \lambda_{\max}^*(\delta A_M^m - \tilde{\lambda}_i \delta B_M^m)}{\lambda_m(B_M) - \|\mathcal{N}_M^{(B)}\|}. \end{aligned} \quad (\text{A12})$$

## APPENDIX B: PROLATE SPHEROIDAL WAVE FUNCTIONS AND ESSENTIAL DIMENSIONALITY

In this Appendix, we provide a review of PFT that aims to promote the development of approximation routines with optimal accuracy guarantees. A core result of PFT is the 2WT theorem on the *essential dimension* of the space of functions of bandwidth  $W$  and approximate duration  $2T$ . The 2WT theorem delineates an intriguing path to fundamental approximation guarantees rooted in spectral and dimensional optimality.

While our exposition builds on this theory, we make some adaptations and introduce extensions. We present a form of the 2WT theorem that highlights the eigenvalues of integral operator of time-and-band-limiting  $\mathcal{B}_W \mathcal{D}_T$  as fundamental accuracy bounds and improves on a constant prefactor with respect to the original theorem in Ref. [64]. The statement on the essential dimensionality of the 2WT theorem is tied to a sharp transition region the spectrum of  $\mathcal{B}_W \mathcal{D}_T$ . The transition region is precisely located through a sequence of inequalities proven by Landau in Ref. [68]. We refer to such a sequence of inequalities as *spectral signature* and note that the spectral signature of the integral operator has a dual in the spectrum of differential operator which commutes with  $\mathcal{B}_W \mathcal{D}_T$ .

We consider PFT to be an essential corner stone of approximation theory, which can guide the formulation of optimal approximation methods. To facilitate the applicability of PFT we introduce concentration inequalities, on the derivatives of prolates as well as a bound on their maximal magnitude outside of the concentration inequalities (Appendix B 2). The inequalities enable sharp accuracy guarantees for PFD (Sec. IID) and a reformulation of the prolate sampling formula [72] that highlights its relation to the 2WT theorem (Appendix C). The proof also relies on the differential equation, which commutes with the integral operator.

### 1. Prolates as optimal filtering system and the 2WT theorem

We motivate the significance of prolates and their properties by sketching their derivation as optimal filter systems. An

error measure that integrates over the energy regions outside of  $[-W, W]$  describes the accuracy of a subspace protocol, as seen in Theorem A1. For filter diagonalization, this suggests that an optimal filtering system is a sequence of time-limited functions, which is maximally confined inside the frequency band  $\mathcal{B}_W$ . We are led to the following optimization problem:

$$\sup_{f \in \mathcal{D}_T} \frac{\|\mathcal{B}_W f\|^2}{\|f\|^2} = \sup_{f \in \mathcal{L}^2(\mathbb{R})} \frac{\|\mathcal{B}_W \mathcal{D}_T f\|^2}{\|\mathcal{D}_T f\|^2}, \quad (\text{B1})$$

where  $\mathcal{B}_W$  is the projection operator onto the space of band-limited functions  $\mathcal{B}_W$  and  $\mathcal{D}_T$  onto the space of time-limited functions  $\mathcal{D}_T$ . Functions in  $\mathcal{B}_W$  have Fourier transforms with support in  $[-W, W]$ . References [62,63] show how optimization problems such as Eq. (B1) lead to the discovery of prolate spheroidal wave functions as eigenfunctions to the integral operator  $\mathcal{B}_W \mathcal{D}_T$ ,

$$\mathcal{B}_W \mathcal{D}_T f(\tau) = \int_{-T}^T \frac{\sin W(\tau - t)}{\pi(\tau - t)} f(t) dt. \quad (\text{B2})$$

The remarkable properties of prolates have been studied in a seminal series of papers [62–64,67,68,81,82,124]. Theorem B1 summarizes a selection of key properties adopted from Ref. [62].

**Theorem B1.** There exists a sequence of functions  $\{\xi_n\}_{n=0}^\infty$  in  $\mathcal{L}_\infty^2$  with the following properties:

(1) For all  $t \in \mathbb{C}$ ,

$$\mathcal{B}_W \mathcal{D}_T \xi_n(t) = \gamma_n \xi_n(t), \quad (\text{B3})$$

where  $1 > \gamma_0 > \gamma_1 > \dots$  and  $\gamma_n$  converges to 0 for  $n \rightarrow \infty$ .

(2) The system  $\{\xi_n\}_{n=0}^\infty$  is complete and orthogonal in  $\mathcal{B}_W$  and  $\mathcal{L}_{[-T, T]}^2$ .

(3) The energy of  $\xi_n$  in the time interval  $[-T, T]$  is given by  $\gamma_n$ . In particular,

$$\int_{-\infty}^\infty \xi_n(t) \xi_m(t) dt = \delta_{nm} \quad \text{and} \quad \int_{-T}^T \xi_n(t) \xi_m(t) dt = \gamma_n \delta_{nm}. \quad (\text{B4})$$

(4) The functions  $\xi_n(t)$  are real valued for  $t \in \mathbb{R}$ , even for  $n$  even and odd for  $n$  odd.

(5) For all  $t \in \mathbb{C}$ :

$$\int_{-T}^T e^{i \frac{Wt}{T}} \xi_n(t) dt = \mu_n \sqrt{\frac{T}{W}} \xi_n(\tau), \quad (\text{B5})$$

where  $\mu_n \in \mathbb{C}$  and  $\frac{|\mu_n|^2}{2\pi} = \gamma_n$ .

The eigenvalues of  $\mathcal{B}_W \mathcal{D}_T$  are functions of the product  $WT$ . To make this explicit, we write  $\gamma_n(c)$ , where  $c = WT$ . The integral operator  $\mathcal{B}_W \mathcal{D}_T$  is Hilbert-Schmidt in  $\mathcal{L}_\infty^2$ , and the supremum in Eq. (B1) is attained by  $\mathcal{D}_T \xi_0$ .

The optimal concentration of prolates as well as their peculiar double completeness relation enabled the deep result that is the 2WT theorem [69]. The theorem states that the space of functions of *essential* bandwidth  $W$  and *essential* duration  $2T$ , is approximately  $2WT/\pi$  dimensional. The 2WT theorem was first proved in Ref. [64]. First, we give the formal definition of the space of band-limited functions, which is up to an error  $\epsilon_T^2$  of duration  $2T$ .

Let  $E(\epsilon_T) \subset \mathcal{L}_\infty^2$  be the set of functions of total energy 1, bandwidth  $W$ , and energy outside of  $[-T, T]$  at most  $\epsilon_T^2$ , that is,

$$\int_{|t|>T} |f(t)|^2 dt \leq \epsilon_T^2 \quad (\text{B6})$$

for all  $f \in E(\epsilon_T)$ . The extended  $2WT$  theorem, as we write it here, encompasses that the space  $E(\epsilon_T)$  is optimally approximated by the prolate sequence. To make this statement precise, the concept of deflection is used.

Let  $\{\psi_n\}$  be a sequence of functions in  $\mathcal{L}_\infty^2$  and  $N \in \mathbb{N}$ . We call

$$\delta(\{\psi_n\}, N) := \sup_{f \in E(\epsilon_T)} \inf_{a_n} \int_{-T}^T \left| f(t) - \sum_{n=0}^{N-1} a_n \psi_n(t) \right|^2 dt \quad (\text{B7})$$

the deflection of  $E(\epsilon_T)$  from  $\{\psi_n\}$  at dimension  $N$ . We are now in a position to state the  $2WT$  theorem.

**Theorem B2 (2WT theorem).** Prolates optimally approximate the space of band-limited and time-concentrated functions  $E(\epsilon_T)$  in the sense that  $\{\xi_n\}$  minimizes the deflection  $\delta(\cdot, N)$  for all  $N$ . In particular,

$$\sup_{f \in E(\epsilon_T)} \inf_{a_n} \int_{-T}^T \left| f(t) - \sum_{n=0}^{N-1} a_n \xi_n(t) \right|^2 dt < \frac{\epsilon_T^2}{1 - \gamma_N(c)}. \quad (\text{B8})$$

For  $N - 1 = \lfloor 2WT/\pi \rfloor$ , we have  $(1 - \gamma_N(c))^{-1} \leq 2$ .

The original version of the  $2WT$  theorem was stated as

$$\sup_{f \in E(\epsilon_T)} \inf_{a_n} \int_{-T}^T \left| f - \sum_{n=0}^{\lfloor 2WT/\pi \rfloor} a_n \xi_n \right|^2 dt < C \epsilon_T^2 \quad (\text{B9})$$

and it was shown that  $C < 12$ . We prefer Eq. (B8) as it underlines the significance of  $1 - \gamma_n(c)$  in the description of the optimal relation between Fourier volume  $c = WT$ , dimensionality  $n$ , and accuracy. In particular, we consider  $1 - \gamma_n(c)$  as fundamental approximation parameters of broader significance in the derivation of more extensive approximation routines. Moreover, Eq. (B8) and the spectral signature in Eq. (B11) immediately allow us to improve the bound on the constant to  $C < 2$ . The optimality of prolates in approximation  $E(\epsilon_T)$  was established in Theorem 1 of Ref. [64].

Motivated by the  $2WT$  Theorem, we call  $\tilde{c} = 2WT/\pi$  the essential dimensionality of  $W$ -band-limited and  $T$ -concentrated functions. The  $2WT$  theorem is a consequence of the eigenvalue distribution of the integral operator  $\mathcal{B}_W \mathcal{D}_T$ , which is summarized in Theorem B3.

**Theorem B3 (Spectrum of  $\mathcal{B}_W \mathcal{D}_T$ ).** The integral operator  $\mathcal{B}_W \mathcal{D}_T$  is Hilbert-Schmidt in  $\mathcal{L}_\infty^2$  and

$$\sum_{n=0}^{\infty} \gamma_n(c) = \frac{2WT}{\pi}. \quad (\text{B10})$$

The sequence of eigenvalues crosses a threshold around the essential dimensionality  $\tilde{c} = 2WT/\pi$ :

$$\gamma_{\lfloor \tilde{c} - 1 \rfloor}(c) \geq 1/2 \geq \gamma_{\lceil \tilde{c} \rceil}(c). \quad (\text{B11})$$

There exists a constant  $A$  independent of  $c$ , such that for all  $\gamma \in (0, 1)$  the number of eigenvalues  $\gamma_n(c)$  in between  $\gamma$  and

$1 - \gamma$  is bounded as

$$\#\{k \mid \gamma \leq \gamma_k(c) \leq 1 - \gamma\} \leq \frac{A}{\gamma(1 - \gamma)} \log \frac{2WT}{\pi}. \quad (\text{B12})$$

The asymptotic expansion of  $1 - \gamma_n(c)$  for  $n$  fixed and large  $c$  is

$$1 - \gamma_n(c) = \frac{4\sqrt{\pi} 2^{3n} c^{n+\frac{1}{2}} e^{-2c}}{n!} (1 + \mathcal{O}(c^{-1})). \quad (\text{B13})$$

Equation (B11) was established in Ref. [68]. We refer to such a crossing in the eigenvalues as a *spectral signature*. In Sec. B3, we point on a dual of Eq. (B11) appearing in the spectrum of the prolate spheroidal wave equation. We adopted Eq. (B12) from Refs. [64,68]. The asymptotic expansion of Eq. (B13) was derived in Refs. [81,82].

Theorem B3 shows that the spectrum of  $\mathcal{B}_W \mathcal{D}_T$  is highly clustered around two values: 1 and 0, with a logarithmically sharp transition region between. Equation (B11) precisely locates the transition region, and it follows that the number of prolates, which are highly concentrated in the time interval, is approximately  $\tilde{c} = 2WT/\pi$ , up to a logarithmic correction. Hence, it was the connection to the spectrum of  $\mathcal{B}_W \mathcal{D}_T$  that allowed us to describe the essential dimension of the space of signals. The significance of the eigenvalues  $\gamma_n(c)$  as optimal approximation bounds is underlined by their definition through a Fourier space optimization problem [Eq. (B1)]. In Theorem B2,  $1 - \gamma_n(c)$  gives an optimal relation between the duration-bandwidth product  $c = WT$ , the degrees of freedom used in approximation  $n$ , and the achieved accuracy. Therefore, PFT delineates an intriguing path toward an approximation theory that offers fundamental accuracy guarantees rooted in spectral and dimensional optimality. Inspired by the  $2WT$  theorem, we aim to develop numerical routines with accuracy guarantees essentially described by  $1 - \gamma_n(c)$ . For PFD, this endeavor requires additions to the set of known concentration properties on prolate functions, which we present in the next section (Appendix B2).

## 2. New concentration identities

To derive accuracy guarantees for larger information processing protocols using prolate Fourier theory, it was necessary to establish how the  $\ell_2$  concentration of prolates extends to their derivatives, as well as to a supremum bound outside their concentration region. These new identities are presented in Theorem B4 and, in particular, yield inequalities of the form  $\sim 1 - \gamma_n(c)$ . Consequently, the sharp spectral transition described in Theorem B3 carries over to these new identities. Thus, Theorem B4 provides additional tools for deriving fundamental approximation schemes that exhibit a sharp accuracy transition around a critical dimension  $\tilde{c} = 2WT/\pi$ .

However, the new concentration identities also rely on a commutation relation of the integral operator  $\mathcal{B}_W \mathcal{D}_T$  with a second-order differential operator. Specifically, prolates are also eigenfunctions to the generalized prolate spheroidal wave equation:

$$(T^2 - t^2)\xi_n''(t) - 2t\xi_n'(t) + W^2(T^2 - t^2)\xi_n(t) = -\lambda_n \xi_n(t). \quad (\text{B14})$$

The eigenvalues of the differential operator are functions of  $c = WT$  and have

$$-c^2 < \lambda_0(c) < \lambda_1(c) < \dots < \dots$$

were  $\lambda_n(c) \rightarrow \infty$  for  $n \rightarrow \infty$  and  $c$  fixed. The peculiar commutation relation has played a significant role in the discovery and development of prolate Fourier theory [62,65–67]. During the derivation of Theorem B4, we noticed that the spectrum of the differential operator exhibits signature similar to the one of the integral operator Eq. (B11). In Appendix B3, we further remark on the commutation relation and the more recently discovered dual property.

Theorem B4 presents our inequalities, which are proven in Chap. 2 of Ref. [60].

*Theorem B4.* The energy concentration of prolates transfers to their derivatives as

$$\|\xi'_n\|_{>T}^2 = (1 - \gamma_n)C_{\text{extra},n}^2, \quad (\text{B15})$$

$$\|\xi'_n\|_T^2 = \gamma_n C_{\text{intra},n}^2. \quad (\text{B16})$$

Prolates have bounds within and outside of the concentration region

$$\sup_{|t| \geq T} \xi_n(t)^2 \leq (1 - \gamma_n)C_{\text{extra},n}, \quad (\text{B17})$$

$$\sup_{|t| \leq T} \xi_n(t)^2 \leq \gamma_n \tilde{C}_{\text{intra},n}. \quad (\text{B18})$$

Here,

$$C_{\text{extra},n} = \sqrt{\|\xi'_n\|_\infty^2 - \frac{\lambda_n \xi_n(T)^2}{T(1 - \gamma_n)}},$$

$$C_{\text{intra},n} = \sqrt{\|\xi'_n\|_\infty^2 + \frac{\lambda_n \xi_n(T)^2}{T\gamma_n}}, \quad (\text{B19})$$

$$\text{and } \tilde{C}_{\text{intra},n} = C_{\text{intra},n} + \delta_{n0} \frac{\xi_0(T)^2}{\gamma_0}.$$

The prefactors satisfy the following inequalities:

(1) For  $n \leq \lfloor 2WT/\pi \rfloor - 1$ :

$$C_{\text{extra},n} \leq \left(C_n - \frac{\lambda_n}{2T}\right) \quad \text{and} \quad C_{\text{intra},n} < W. \quad (\text{B20})$$

(2) For  $n \geq \lceil 2WT/\pi \rceil$ :

$$C_{\text{extra},n} < W \quad \text{and} \quad C_{\text{intra},n} \leq \left(C_n + \frac{\lambda_n}{2T}\right), \quad (\text{B21})$$

where  $C_n = \sqrt{\|\xi'_n\|_\infty^2 + \frac{\lambda_n^2}{4T^2}}$  and we have the bound  $\|\xi'_n\|_\infty < W$ .  $C_{\text{extra},n}$  is a function of  $W$  and  $T$ . Sometimes we make this explicit by writing  $C_{\text{extra},n} \equiv C_{\text{extra},n}(W, T)$ . For  $n \leq \lfloor 2WT/\pi \rfloor - 1$ , we can estimate

$$C_{\text{extra},n} < W \left( \sqrt{1 + \frac{c^2}{4}} + \frac{c}{2} \right) \quad (\text{B22})$$

$$= W \left( c + \frac{1}{c} + \mathcal{O}\left(\frac{1}{c^3}\right) \right). \quad (\text{B23})$$

In particular,  $C_{\text{extra},n}$  grows much more slowly with  $c$  compared to the exponential decay of  $1 - \gamma_n(c)$  described in Eq. (B13). Therefore, the bounds  $(1 - \gamma_n)C_{\text{extra},n}$  and  $(1 -$

$\gamma_n)C_{\text{extra},n}^2$  inherit from  $1 - \gamma_n$  the characterization as an approximation parameter exhibiting a sharp transition as  $n$  approaches  $2WT/\pi$ .

### 3. A dual in the spectrum of an integral and a differential operator

A surprising commutation relation between an integral operator and a differential operator played a significant role in the development of PFT, its generalizations, and its numerical applications [62,65,66]. Slepian famously described this commutation relation as a “lucky accident” [69], and it later inspired the development of bispectral theory [79,125–127]. The bispectrality of the prolate spheroidal wave functions has indeed led to unexpected developments in recent years [80]. However, the geometry behind the commutation relation still appears mysterious. On this note, we point to a dualism in the spectrum of the integral operator and differential operator, which seems to have been unnoticed.

The spectral signature due to Landau is of conceptual depth for PFT:

$$\gamma_{\lfloor \tilde{c} \rfloor - 1}(c) \geq 1/2 \geq \gamma_{\lceil \tilde{c} \rceil}(c). \quad (\text{B24})$$

Here, we notice that the spectrum of the prolate spheroidal integral operator exhibits a similar signature:

$$\lambda_{\lfloor \tilde{c} \rfloor - 1}(c) \leq -1 \quad \text{and} \quad 0 \leq \lambda_{\lceil \tilde{c} \rceil}(c). \quad (\text{B25})$$

A similar statement has, in fact, already been proved in Theorem 3.1 of Ref. [121]:

$$\lambda_{\lfloor \tilde{c} \rfloor - 1}(c) \leq 0 \leq \lambda_{\lceil \tilde{c} \rceil}(c). \quad (\text{B26})$$

However, the duality between this spectral signature and that of the integral operator appears to have remained unnoticed in Ref. [121]. The proof in Ref. [121] is based on the Prüfer transformation and rather technical. We include here a simpler proof for the slightly stronger statement given in the lower inequality of Eq. (B25).

For the sake of simplicity, we set  $T = 1; W = c$  and consider the angular part of the prolate spheroidal wave equation as a differential operator:

$$H_c = -\frac{d}{dx} \left[ (1 - x^2) \frac{d}{dx} \right] - c^2(1 - x^2). \quad (\text{B27})$$

For  $\varphi \in Q(H_c)$ , where  $Q(H_c)$  denotes the form domain, we have

$$(\varphi, H_c \varphi) = \int_{-1}^1 (1 - x^2)(|\varphi'(x)|^2 - c^2|\varphi(x)|^2) dx. \quad (\text{B28})$$

The form domain of  $H_c$  is given by  $C^1[-1, 1]$ . We prove the left inequality of Eq. (B25) through the variational principle (Theorem XIII.2, page 76 of Ref. [128]):

*Lemma B1.* Let  $H$  be a self-adjoint operator on a Hilbert space, bounded from below and with a purely discrete spectrum. Assume its eigenvalues are indexed in increasing order:

$$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots \quad (\text{B29})$$

Then, its  $n$ th eigenvalue satisfies

$$\lambda_n = \inf_{S_{n+1} \subset Q(H)} \sup_{\psi \in S_n} \frac{(\psi, H\psi)}{(\psi, \psi)},$$

where  $Q(H)$  is the form domain of  $H$ , and the infimum is taken over all  $(n+1)$ -dimensional subspaces of  $Q(H)$ .

*Proof of  $\lambda_{\lfloor \frac{2c}{\pi} \rfloor - 1}(c) \leq -1$ .* Assume that  $\lfloor \frac{2c}{\pi} \rfloor$  is odd. For  $k$  odd, we have

$$\int_{k\pi/2c}^{(k+2)\pi/2c} (1-x^2)(\sin^2(cx) - \cos^2(cx))dx = -\frac{\pi}{2c^3}.$$

Define

$$\varphi_k(x) = \begin{cases} \sqrt{\frac{2c}{\pi}} \cos(cx), & \text{for } x \in \left[\frac{k\pi}{2c}, \frac{(k+2)\pi}{2c}\right], \\ 0, & \text{else.} \end{cases}$$

The index  $k$  runs through all odd integers. There are  $2\lfloor \frac{c}{\pi} - \frac{1}{2} \rfloor + 1$  such functions  $\varphi_k$  with support contained in  $[-1, 1]$ . For  $\lfloor \frac{2c}{\pi} \rfloor$ , odd holds  $\lfloor \frac{2c}{\pi} \rfloor = 2\lfloor \frac{c}{\pi} - \frac{1}{2} \rfloor + 1$ . By construction,  $\{\varphi_k\}$  forms an orthonormal system in  $\mathcal{L}^2[-1, 1]$ . Let  $\psi = \sum_k a_k \varphi_k$  with  $\sum_k |a_k|^2 = 1$ . Evidently,  $\psi \in C^1[-1, 1] = Q(H_c)$ . In particular,

$$\begin{aligned} (\Psi, H_c \Psi) &= \sum_k |a_k|^2 \frac{2c^3}{\pi} \int_{k\pi/2c}^{(k+2)\pi/2c} (1-x^2) \\ &\quad \times (\sin^2(cx) - \cos^2(cx))dx = -1. \end{aligned} \quad (\text{B30})$$

By Lemma B1, we have therefore shown that

$$\lambda_{\lfloor \frac{2c}{\pi} \rfloor - 1}(c) \leq -1. \quad (\text{B31})$$

For the case where  $\lfloor \frac{2c}{\pi} \rfloor$  is even, we have  $\lfloor \frac{2c}{\pi} \rfloor = 2\lfloor \frac{c}{\pi} \rfloor$ . For  $k$  even,

$$\int_{k\pi/2c}^{(k+2)\pi/2c} (1-x^2)(\cos^2(cx) - \sin^2(cx))dx = -\frac{\pi}{2c^3}, \quad (\text{B32})$$

and we can construct analogously a subspace spanned by sine functions supported on mutually disjoint intervals. Hence, inequality (B31) holds for all  $c > 0$ . ■

### APPENDIX C: PROLATE SAMPLING FORMULA

The  $2WT$  theorem is a milestone in mathematical electrical engineering, as it provides a rigorous foundation for the long-standing *engineering folk theorem*, which originated in the early development of communication theory. This folk theorem on the dimensionality of signals emerged from sampling principles such as the Whittaker-Shannon interpolation formula [69,83,129].

The Whittaker-Shannon formula reconstructs a signal of bandwidth  $W$  from uniform samples taken at rate  $W/\pi$ . In particular, it assumes  $2WT/\pi$  samples within  $[-T, T]$ . This has led to the intuition that  $2WT/\pi$  independent measurements should suffice to accurately reconstruct a signal

of bandwidth and approximate duration  $2T$ . However, the Whittaker-Shannon interpolation formula does not succeed in proving this statement, as it converges very slowly and also requires a lot of samples outside of the interval of interest.

Only the  $2WT/\pi$  theorem ultimately established that approximately  $2WT/\pi$  degrees of freedom suffice to accurately represent a signal. However, strictly speaking, the  $2WT$  theorem does not fully reflect the original context of the engineering folk theorem, which is rooted in representations using equidistant samples. In practice, such equidistant sampling remains particularly relevant due to its simplicity and implementational convenience. We highlight that PFT also enables us to give a sampling-based formulation of the engineering folk theorem. This result is formalized in Theorem C1.

*Theorem C1.* Consider a band limited function  $f \in \mathcal{B}_W$  in the time domain  $[-T, T]$ . Let  $\{\xi_n\}$  be the sequence of  $WT$  prolates and  $c = TW/\pi$ . We denote the truncated sampling series  $f_N(t)$  defined as

$$f_N(t) = \frac{\pi}{W} \sum_{n=0}^{N-1} \sum_{k=-\lfloor TW/\pi \rfloor}^{\lfloor TW/\pi \rfloor} f\left(\frac{k\pi}{W}\right) \xi_n\left(\frac{k\pi}{W}\right) \xi_n(t). \quad (\text{C1})$$

The sampling series  $f_N(t)$  approximates  $f$  within the interval  $[-T, T]$  with a precision guarantee of

$$\begin{aligned} \|f - f_N\|_T^2 &\leq \left( \|f\|_{>T}^2 + 2\frac{\pi}{W} \|f\|_{>T} \|f'\|_{>T} \right) \sum_{n=0}^{N-1} \gamma_n (1 - \gamma_n) C_n \\ &\quad + \frac{\|f\|_{>T}^2}{1 - \gamma_N}. \end{aligned} \quad (\text{C2})$$

Here,  $C_n = (1 + 2\frac{\pi}{W} C_{\text{extra},n})$  with  $C_{\text{extra},n}$  as in Theorem B4.

Theorem C1 establishes that  $2WT/\pi$  equidistant sample points suffice to reconstruct a signal of bandwidth  $W$  with high accuracy, provided the signal is sufficiently concentrated within the sampled interval. A proof is given in Chap. 3 of Ref. [60].

Theorem C1 builds upon the prolate sampling formula that was introduced in Ref. [72]. However, the truncation estimate of Theorem C1 is sharper than in Ref. [72], as it accesses the bounds of Theorem B4.

To formulate a sampling result more closely aligned with the original engineering folk theorem, we also made a modification in the assumptions. Reference [72] and Chap. 6 of Ref. [130] consider functions  $f$  lying in the subspace spanned by leading prolates,  $\text{PSWF}_N = \{\xi_0, \dots, \xi_{N-1}\}$ . Theorem C1 focuses instead on the space of band-limited and time-concentrated functions  $E(\epsilon_T)$ , which we regard as a more natural model for real-world signals. A direct application of the  $2WT$  theorem, as stated in Theorem B2, allows us to bound the error in the sampling error due to contributions outside of  $\text{PSWF}_N$ .

- [1] S. R. Langhoff and E. R. Davidson, Configuration interaction calculations on the nitrogen molecule, *Int. J. Quantum Chem.* **8**, 61 (1974).
- [2] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*, Cambridge Monographs on Applied and Computational Mathematics, Vol. 13 (Cambridge University Press, Cambridge, UK, 2003).
- [3] E. J. Heller, L. Kaplan, and F. Pollmann, Inflationary dynamics for matrix eigenvalue problems, *Proc. Natl. Acad. Sci. USA* **105**, 7631 (2008).
- [4] K.-J. Bathe and E. L. Wilson, Solution methods for eigenvalue problems in structural mechanics, *Int. J. Numer. Methods Eng.* **6**, 213 (1973).
- [5] R. Pederson, J. Kozłowski, R. Song, J. Beall, M. Ganahl, M. Hauru, A. G. Lewis, Y. Yao, S. B. Mallick, V. Blum, *et al.*, Large scale quantum chemistry with tensor processing units, *J. Chem. Theory Comput.* **19**, 25 (2023).
- [6] S. Brin and L. Page, The anatomy of a large-scale hypertextual web search engine, *Comput. Networks ISDN Syst.* **30**, 107 (1998).
- [7] S. Ramachandran, O. Deshpande, C. C. Roseman, N. A. Rosenberg, M. W. Feldman, and L. L. Cavalli-Sforza, Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa, *Proc. Natl. Acad. Sci. USA* **102**, 15942 (2005).
- [8] N. Stamatopoulos and W. J. Zeng, Derivative pricing using quantum signal processing, *Quantum* **8**, 1322 (2024).
- [9] R. Babbush, D. W. Berry, R. Kothari, R. D. Somma, and N. Wiebe, Exponential quantum speedup in simulating coupled classical oscillators, *Phys. Rev. X* **13**, 041041 (2023).
- [10] R. Y. Li, R. Di Felice, R. Rohs, and D. A. Lidar, Quantum annealing versus classical machine learning applied to a simplified computational biology problem, *npj Quantum Inf.* **4**, 14 (2018).
- [11] P. Pravatto, D. Castaldo, F. Gallina, B. Fresch, S. Corni, and G. J. Moro, Quantum computing for classical problems: Variational quantum eigensolver for activated processes, *New J. Phys.* **23**, 123045 (2021).
- [12] P. Benioff, The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines, *J. Stat. Phys.* **22**, 563 (1980).
- [13] R. P. Feynman, Quantum mechanical computers, *Found. Phys.* **16**, 507 (1986).
- [14] A. Baiardi, M. Christandl, and M. Reiher, Quantum computing for molecular biology, *ChemBioChem* **24**, e202300120 (2023).
- [15] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [16] J. J. Goings, A. White, J. Lee, C. S. Tautermann, M. Degroote, C. Gidney, T. Shiozaki, R. Babbush, and N. C. Rubin, Reliably assessing the electronic structure of cytochrome P450 on today's classical computers and tomorrow's quantum computers, *Proc. Natl. Acad. Sci. USA* **119**, e2203533119 (2022).
- [17] V. von Burg, G. H. Low, T. Häner, D. S. Steiger, M. Reiher, M. Roetteler, and M. Troyer, Quantum computing enhanced computational catalysis, *Phys. Rev. Res.* **3**, 033055 (2021).
- [18] T. Farajollahpour, Quantum algorithm software for condensed matter physics, [arXiv:2506.09308](https://arxiv.org/abs/2506.09308).
- [19] H. Liu, G. H. Low, D. S. Steiger, T. Häner, M. Reiher, and M. Troyer, Prospects of quantum computing for molecular sciences, *Mater. Theory* **6**, 11 (2022).
- [20] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science, *Chem. Rev.* **120**, 12685 (2020).
- [21] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, Elucidating reaction mechanisms on quantum computers, *Proc. Natl. Acad. Sci. USA* **114**, 7555 (2017).
- [22] F. Vigneau, S. Majumder, A. Rath, P. Parrado-Rodríguez, F. R. F. Pereira, S. Pogorzalek, T. Jones, N. Wurz, M. Renger, J. Verjauw, *et al.*, Quantum error detection in qubit-resonator star architecture, *PRX Quantum* **6**, 040369 (2025).
- [23] S. Bartolucci, P. Birchall, H. Bombin, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, *et al.*, Fusion-based quantum computation, *Nat. Commun.* **14**, 912 (2023).
- [24] H. Aghaee Rad, T. Ainsworth, R. Alexander, B. Altieri, M. Askarani, R. Baby, L. Banchi, B. Baragiola, J. Bourassa, R. Chadwick, *et al.*, Scaling and networking a modular photonic quantum computer, *Nature (London)* **638**, 912 (2025).
- [25] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. Allman, C. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, *et al.*, Demonstration of the trapped-ion quantum CCD computer architecture, *Nature (London)* **592**, 209 (2021).
- [26] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, Cambridge, 2013).
- [27] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Phys. Rev. A* **86**, 032324 (2012).
- [28] S.-H. Lee, F. Thomsen, N. Fazio, B. J. Brown, and S. D. Bartlett, Low-overhead magic state distillation with color codes, *PRX Quantum* **6**, 030317 (2025).
- [29] C. Gidney, N. Shutty, and C. Jones, Magic state cultivation: Growing T states as cheap as CNOT gates, [arXiv:2409.17595](https://arxiv.org/abs/2409.17595).
- [30] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, *Proc. R. Soc. A* **454**, 339 (1998).
- [31] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [32] A. M. Dalzell, S. McArdle, M. Berta, P. Bienias, C.-F. Chen, A. Gilyén, C. T. Hann, M. J. Kastoryano, E. T. Khabiboulline, A. Kubica, *et al.*, *Quantum Algorithms: A Survey of Applications and End-to-End Complexities* (Cambridge University Press, Cambridge, 2025).
- [33] J. S. Nelson and A. D. Baczewski, Assessment of quantum phase estimation protocols for early fault-tolerant quantum computers, *Phys. Rev. A* **110**, 042420 (2024).
- [34] M. Erakovic, F. Witteveen, D. Harley, J. Günther, M. Bensberg, O. R. Meitei, M. Cho, T. Van Voorhis, M. Reiher, and M. Christandl, High ground state overlap via quantum embedding methods, *PRX Life* **3**, 013003 (2025).
- [35] S. Fomichev, K. Hejazi, M. S. Zini, M. Kiser, J. Fraxanet, P. A. M. Casares, A. Delgado, J. Huh, A.-C. Voigt, J. E. Mueller, *et al.*, Initial state preparation for quantum chemistry on quantum computers, *PRX Quantum* **5**, 040339 (2024).
- [36] P. J. Ollitrault, C. L. Cortes, J. F. Gonthier, R. M. Parrish, D. Rocca, G.-L. Anselmetti, M. Degroote, N. Moll, R. Santagati, and M. Streif, Enhancing initial state overlap through orbital optimization for faster molecular electronic ground-state energy estimation, *Phys. Rev. Lett.* **133**, 250601 (2024).

- [37] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, *et al.*, Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry, *Nat. Commun.* **14**, 1952 (2023).
- [38] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Exponential improvement in precision for simulating sparse Hamiltonians, in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing* (ACM, New York, 2014), pp. 283–292.
- [39] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, *Quantum* **3**, 163 (2019).
- [40] R. D. Somma, R. King, R. Kothari, T. E. O’Brien, and R. Babbush, Shadow Hamiltonian simulation, *Nat. Commun.* **16**, 2690 (2025).
- [41] A. Y. Kitaev, Quantum computations: Algorithms and error correction, *Russ. Math. Surv.* **52**, 1191 (1997).
- [42] Z. Ding and L. Lin, Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation, *PRX Quantum* **4**, 020331 (2023).
- [43] Z. Ding and L. Lin, Simultaneous estimation of multiple eigenvalues with short-depth quantum circuit on early fault-tolerant quantum computers, *Quantum* **7**, 1136 (2023).
- [44] H. Li, H. Ni, and L. Ying, Adaptive low-depth quantum algorithms for robust multiple-phase estimation, *Phys. Rev. A* **108**, 062408 (2023).
- [45] M. Stroeks, J. Helsen, and B. Terhal, Spectral estimation for Hamiltonians: A comparison between classical imaginary-time evolution and quantum real-time evolution, *New J. Phys.* **24**, 103024 (2022).
- [46] C. Yi, C. Zhou, and J. Takahashi, Quantum phase estimation by compressed sensing, *Quantum* **8**, 1579 (2024).
- [47] J. Günther, F. Witteveen, A. Schmidhuber, M. Miller, M. Christandl, and A. Harrow, Phase estimation with partially randomized time evolution, *PRX Quantum* **7**, 020332 (2026).
- [48] A. Dutkiewicz, B. M. Terhal, and T. E. O’Brien, Heisenberg-limited quantum phase estimation of multiple eigenvalues with few control qubits, *Quantum* **6**, 830 (2022).
- [49] A. Dutkiewicz, S. Polla, M. Scheurer, C. Gogolin, W. J. Huggins, and T. E. O’Brien, Error mitigation and circuit division for early fault-tolerant quantum phase estimation, *PRX Quantum* **6**, 040318 (2025).
- [50] Z. Ding, H. Li, L. Lin, H. Ni, L. Ying, and R. Zhang, Quantum multiple eigenvalue Gaussian filtered search: An efficient and versatile quantum phase estimation method, *Quantum* **8**, 1487 (2024).
- [51] D. Neuhauser, Bound state eigenfunctions from wave packets: Time→energy resolution, *The J. Chem. Phys.* **93**, 2611 (1990).
- [52] V. Mandelshtam, FDM: The filter diagonalization method for data processing in NMR experiments, *Prog. Nucl. Magn. Reson. Spectrosc.* **38**, 159 (2001).
- [53] K. Klymko, C. Mejuto-Zaera, S. J. Cotton, F. Wudarski, M. Urbanek, D. Hait, M. Head-Gordon, K. B. Whaley, J. Moussa, N. Wiebe, *et al.*, Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware, *PRX Quantum* **3**, 020323 (2022).
- [54] Y. Shen, K. Klymko, J. Sud, D. B. Williams-Young, W. A. de Jong, and N. M. Tubman, Real-time Krylov theory for quantum computing algorithms, *Quantum* **7**, 1066 (2023).
- [55] C. L. Cortes and S. K. Gray, Quantum Krylov subspace algorithms for ground- and excited-state energy estimation, *Phys. Rev. A* **105**, 022417 (2022).
- [56] R. M. Parrish and P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation, [arXiv:1909.08925](https://arxiv.org/abs/1909.08925).
- [57] J. Cohn, M. Motta, and R. M. Parrish, Quantum filter diagonalization with compressed double-factorized Hamiltonians, *PRX Quantum* **2**, 040352 (2021).
- [58] E. N. Epperly, L. Lin, and Y. Nakatsukasa, A theory of quantum subspace diagonalization, *SIAM J. Matrix Anal. Appl.* **43**, 1263 (2022).
- [59] W. Kirby, Analysis of quantum Krylov algorithms with errors, *Quantum* **8**, 1457 (2024).
- [60] T. Stroschein, Prolate spheroidal wave functions and the accuracy and dimensionality of spectral analysis, [arXiv:2409.16584](https://arxiv.org/abs/2409.16584).
- [61] T. Stroschein, Local spectral approximation of unbounded operators: Non-asymptotic and unified error quantification for subspace methods, [arXiv:2505.07513](https://arxiv.org/abs/2505.07513).
- [62] D. Slepian and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty—I, *Bell Syst. Tech. J.* **40**, 43 (1961).
- [63] H. J. Landau and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty—II, *Bell Syst. Tech. J.* **40**, 65 (1961).
- [64] H. J. Landau and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty—III: The dimension of the space of essentially time- and band-limited signals, *Bell Syst. Tech. J.* **41**, 1295 (1962).
- [65] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty—IV: Extensions to many dimensions; generalized prolate spheroidal functions, *Bell Syst. Tech. J.* **43**, 3009 (1964).
- [66] D. Slepian, Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case, *Bell Syst. Tech. J.* **57**, 1371 (1978).
- [67] D. Slepian, On bandwidth, *Proc. IEEE* **64**, 292 (1976).
- [68] H. Landau, On the density of phase-space expansions, *IEEE Trans. Inf. Theory* **39**, 1152 (1993).
- [69] D. Slepian, Some comments on Fourier analysis, uncertainty and modeling, *SIAM Rev.* **25**, 379 (1983).
- [70] T. Levitina and E. Brändas, Filter diagonalization: Filtering and postprocessing with prolates, *Comput. Phys. Commun.* **180**, 1448 (2009).
- [71] A. L. Van Buren and J. E. Boisvert, Accurate calculation of prolate spheroidal radial functions of the first kind and their first derivatives, *Q. Appl. Math.* **60**, 589 (2002).
- [72] G. G. Walter and X. A. Shen, Sampling With Prolate Spheroidal Wave Functions, *Sampl. Theory Signal Image Process.* **2**, 25 (2003).
- [73] J. von Neumann, Mathematische Begründung der Quantenmechanik, *Nachr. Ges. Wiss. Göttingen* **1** (1927).
- [74] J. von Neumann, Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik, *Nachr. Ges. Wiss. Göttingen* **245** (1927).
- [75] J. von Neumann, Allgemeine Eigenwerttheorie Hermiteischer Funktionaloperatoren, *Math. Ann.* **102**, 49 (1930).
- [76] S. Bochner, *Lectures on Fourier Integrals* (Princeton University Press, Princeton, NJ, 1932).

- [77] 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).
- [78] V. A. Mandelshtam and H. S. Taylor, Harmonic inversion of time signals and its applications, *J. Chem. Phys.* **107**, 6756 (1997).
- [79] A. Grunbaum, B. Vasquez Campos, and J. Zubelli, Matrix bispectrality and noncommutative algebras: Beyond the prolate spheroidals, *Enseign. Math.* **69**, 335 (2023).
- [80] A. Connes and H. Moscovici, The UV prolate spectrum matches the zeros of zeta, *Proc. Natl. Acad. Sci. USA* **119**, e2123174119 (2022).
- [81] W. Fuchs, On the eigenvalues of an integral equation arising in the theory of band-limited signals, *J. Math. Anal. Appl.* **9**, 317 (1964).
- [82] D. Slepian, Some asymptotic expansions for prolate spheroidal wave functions, *J. Math. Phys.* **44**, 99 (1965).
- [83] C. Shannon, Communication in the presence of noise, *Proc. IRE* **37**, 10 (1949).
- [84] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [85] S. Lloyd, Universal quantum simulators, *Science* **273**, 1073 (1996).
- [86] A. L. Van Buren and J. E. Boisvert, Mathieu and Spheroidal Wave Functions: prolate\_swf, GitHub, 2021, [https://github.com/MathieuandSpheroidalWaveFunctions/prolate\\_swf](https://github.com/MathieuandSpheroidalWaveFunctions/prolate_swf).
- [87] We denote the radial coordinate as  $\zeta$  to distinguish it from the prolate functions  $\xi_n$ . Standard literature [71] uses  $\xi$  for this coordinate.
- [88] C. Bouwkamp, Theoretical and numerical treatment of diffraction through a circular aperture, *IEEE Trans. Antennas Propag.* **18**, 152 (1970).
- [89] J. Meixner and F. W. Schäfke, *Mathieusche Funktionen und Sphäroidfunktionen* (Springer, Berlin, 1954).
- [90] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. AkashNarayanan, A. Asadi, *et al.*, PennyLane: Automatic differentiation of hybrid quantum-classical computations, [arXiv:1811.04968](https://arxiv.org/abs/1811.04968).
- [91] 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).
- [92] F. Stepan, H. Kasra, F. Joana, and J. M. Arrazola, Overlapper, GitHub, 2024, <https://github.com/XanaduAI/Overlapper>.
- [93] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, *et al.*, Recent developments in the PySCF program package, *J. Chem. Phys.* **153**, 024109 (2020).
- [94] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: Composable transformations of Python + NumPy programs, GitHub, 2018, <https://github.com/google/jax>.
- [95] C. D. Sherrill and H. F. Schaefer III, The configuration interaction method: Advances in highly correlated approaches, in *Advances in Quantum Chemistry* (Elsevier, New York, 1999), Vol. 34, pp. 143–269.
- [96] I. Vasiliev, S. Ögüt, and J. R. Chelikowsky, *Ab initio* excitation spectra and collective electronic response in atoms and clusters, *Phys. Rev. Lett.* **82**, 1919 (1999).
- [97] T. Stroschein, D. Castaldo, and M. Reiher, Quantum-prolate-diagonalization, GitHub, 2025, <https://github.com/Timothy-St/Quantum-Prolate-Diagonalization>.
- [98] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, New York, 2013).
- [99] G. Lee, D. Lee, and J. Huh, Sampling error analysis in quantum Krylov subspace diagonalization, *Quantum* **8**, 1477 (2024).
- [100] Y. Shen, D. Camps, A. Szasz, S. Darbha, K. Klymko, D. B. Williams-Young, N. M. Tubman, and R. Van Beeumen, Estimating eigenenergies from quantum dynamics: A unified noise-resilient measurement-driven approach, *Quantum* **9**, 1836 (2025).
- [101] When using singular value thresholding, the GEP is regularized, truncating the subspace with a threshold parameter  $\delta$ . The original GEP  $[A, B]$  is then transformed into  $[\tilde{A}, \tilde{B}]$ , where  $\tilde{B} = \sum_{l, \sigma_l \geq \delta} \sigma_l \mathbf{u}_l \mathbf{v}_l^\dagger$  and  $\tilde{A} = \tilde{\mathbf{U}}^\dagger A \tilde{\mathbf{V}}$ , where  $\mathbf{U}, \mathbf{V}$  are the matrices of the left  $\mathbf{u}$  and right  $\mathbf{v}$  singular vectors after truncation (e.g.,  $\tilde{\mathbf{V}}_{ij} = \mathbf{V}_{ij}, \sigma_j \geq \delta$ ).
- [102] G. Lee, S. Choi, J. Huh, and A. F. Izmaylov, Efficient strategies for reducing sampling error in quantum Krylov subspace diagonalization, *Digital Discovery* **4**, 954 (2025).
- [103] C. L. Cortes, D. Rocca, J. F. Gonthier, P. J. Ollitrault, R. M. Parrish, G.-L. R. Anselmetti, M. Degroote, N. Moll, R. Santagati, and M. Streif, Assessing the query complexity limits of quantum phase estimation using symmetry-aware spectral bounds, *Phys. Rev. A* **110**, 022420 (2024).
- [104] C. J. Stein and M. Reiher, Automated selection of active orbital spaces, *J. Chem. Theory Comput.* **12**, 1760 (2016).
- [105] L. Stella, C. Attaccalite, S. Sorella, and A. Rubio, Strong electronic correlation in the hydrogen chain: A variational Monte Carlo study, *Phys. Rev. B* **84**, 245117 (2011).
- [106] M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma, *et al.*, Towards the solution of the many-electron problem in real materials: Equation of state of the hydrogen chain with state-of-the-art many-body methods, *Phys. Rev. X* **7**, 031059 (2017).
- [107] M. Motta, C. Genovese, F. Ma, Z.-H. Cui, R. Sawaya, G. K.-L. Chan, N. Chepiga, P. Helms, C. Jiménez-Hoyos, A. J. Millis, *et al.*, Ground-state properties of the hydrogen chain: Dimerization, insulator-to-metal transition, and magnetic phases, *Phys. Rev. X* **10**, 031058 (2020).
- [108] M. Mörchén, G. H. Low, T. Weymuth, H. Liu, M. Troyer, and M. Reiher, Classification of electronic structures and state preparation for quantum computation of reaction chemistry, [arXiv:2409.08910](https://arxiv.org/abs/2409.08910).
- [109] Z. Ding, E. N. Epperly, L. Lin, and R. Zhang, The ESPRIT algorithm under high noise: Optimal error scaling and noisy super-resolution, in *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2024), pp. 2344–2366.
- [110] G. Rendon, T. Izubuchi, and Y. Kikuchi, Effects of cosine tapering window on quantum phase estimation, *Phys. Rev. D* **106**, 034503 (2022).
- [111] M. Motta and J. E. Rice, Emerging quantum computing algorithms for quantum chemistry, *WIREs Comput. Mol. Sci.* **12**, e1580 (2022).

- [112] J. Cioslowski, Robust interpolation between weak-and strong-correlation regimes of quantum systems, *J. Chem. Phys.* **136**, 044109 (2012).
- [113] D. Materia, L. Ratini, C. Angeli, and L. Guidoni, Quantum information reveals that orbital-wise correlation is essentially classical in natural orbitals, *J. Chem. Phys.* **161**, 244108 (2024).
- [114] W. H. Zurek, Environment-induced superselection rules, *Phys. Rev. D* **26**, 1862 (1982).
- [115] P. W. Anderson, Infrared catastrophe in Fermi gases with local scattering potentials, *Phys. Rev. Lett.* **18**, 1049 (1967).
- [116] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, *SIAM J. Comput.* **35**, 1070 (2006).
- [117] B. O’Gorman, S. Irani, J. Whitfield, and B. Fefferman, Intractability of electronic structure in a fixed basis, *PRX Quantum* **3**, 020322 (2022).
- [118] O. Kiss, U. Azad, B. Requena, A. Roggero, D. Wakeham, and J. M. Arrazola, Early fault-tolerant quantum algorithms in practice: Application to ground-state energy estimation, *Quantum* **9**, 1682 (2025).
- [119] D. W. Berry, Y. Tong, T. Khattar, A. White, T. I. Kim, G. H. Low, S. Boixo, Z. Ding, L. Lin, S. Lee, G. K.-L. Chan, R. Babbush, and N. C. Rubin, Rapid initial-state preparation for the quantum simulation of strongly correlated molecules, *PRX Quantum* **6**, 020327 (2025).
- [120] K. C. Smith, A. Khan, B. K. Clark, S. M. Girvin, and T.-C. Wei, Constant-depth preparation of matrix product states with adaptive quantum circuits, *PRX Quantum* **5**, 030344 (2024).
- [121] A. Osipov, V. Rokhlin, and H. Xiao, Analysis of a differential operator, in *Prolate Spheroidal Wave Functions of Order Zero: Mathematical Tools for Bandlimited Approximation* (Springer, Boston, MA, 2013), pp. 73–133.
- [122] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, Compressed sensing off the grid, *IEEE Trans. Inf. Theory* **59**, 7465 (2013).
- [123] D. Castaldo and S. Corni, Heisenberg limited multiple eigenvalue estimation via off-the-grid compressed sensing, [arXiv:2507.12438](https://arxiv.org/abs/2507.12438).
- [124] H. Landau and H. Widom, Eigenvalue distribution of time and frequency limiting, *J. Math. Anal. Appl.* **77**, 469 (1980).
- [125] F. A. Grünbaum and M. Yakimov, The prolate spheroidal phenomenon as a consequence of bispectrality, in *Superintegrability in Classical and Quantum Systems*, CRM Proceedings and Lecture Notes (American Mathematical Society, Providence, RI, 2004), Vol. 37, pp. 301–312.
- [126] W. R. Casper, F. A. Grünbaum, M. Yakimov, and I. Zurrián, Algebras of commuting differential operators for kernels of Airy type, in *Toeplitz Operators and Random Matrices: In Memory of Harold Widom*, edited by E. Basor, A. Böttcher, T. Ehrhardt, and C. A. Tracy (Springer International Publishing, Cham, 2022), pp. 229–256.
- [127] W. R. Casper, F. A. Grünbaum, M. Yakimov, and I. Zurrián, Matrix valued discrete–continuous functions with the prolate spheroidal property and bispectrality, *Commun. Math. Phys.* **405**, 69 (2024).
- [128] M. Reed and B. Simon, *IV: Analysis of Operators, Methods of Modern Mathematical Physics* (Elsevier Science, New York, 1978).
- [129] H. Nyquist, Certain topics in telegraph transmission theory, *Trans. AIEE* **47**, 617 (1928).
- [130] J. A. Hogan and J. D. Lakey, *Duration and Bandwidth Limiting: Prolate Functions, Sampling, and Applications, Applied and Numerical Harmonic Analysis* (Birkhäuser, Boston, 2012).