Title

Quantum Multiple Eigenvalue Gaussian filtered Search: an efficient and versatile quantum phase estimation method

Permalink

https://escholarship.org/uc/item/2sc8j70m

Authors

Ding, Zhiyan
Li, Haoya
Lin, Lin
et al.

Publication Date

2024

DOI

10.48550/arXiv.2402.01013

Peer reviewed

Quantum Multiple Eigenvalue Gaussian filtered Search: an efficient and versatile quantum phase estimation method

Zhiyan Ding[1], Haoya Li[2], Lin Lin[1,3,4], Hongkang Ni[5], Lexing Ying[2,5], and Ruizhe Zhang[6]

[1]Department of Mathematics, University of California, Berkeley, CA 94720, USA

[2]Department of Mathematics, Stanford University, CA 94305, USA

[3]Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[4]Challenge Institute for Quantum Computation, University of California, Berkeley, CA 94720, USA

[5]Institute for Computational and Mathematical Engineering, Stanford University, CA 94305, USA

[6]Simons Institute for the Theory of Computing, University of California, Berkeley, CA 94720, USA

Quantum phase estimation is one of the most powerful quantum primitives. This work proposes a new approach for the problem of multiple eigenvalue estimation: Quantum Multiple Eigenvalue Gaussian filtered Search (QMEGS). QMEGS leverages the Hadamard test circuit structure and only requires simple classical postprocessing. QMEGS is the first algorithm to simultaneously satisfy the following two properties: (1) It can achieve the Heisenberg-limited scaling without relying on any spectral gap assumption. (2) With a positive energy gap and additional assumptions on the initial state, QMEGS can estimate all dominant eigenvalues to ϵ accuracy utilizing a significantly reduced circuit depth compared to the standard quantum phase estimation algorithm. In the most favorable scenario, the maximal runtime can be reduced to as low as $\log(1/\epsilon)$. This implies that QMEGS serves as an efficient and versatile approach, achieving the best-known results for both gapped and gapless systems. Numerical results validate the efficiency of our proposed algorithm in various regimes.

1 Introduction

Phase estimation is among the most powerful quantum primitives, offering eigenvalue estimates of a Hamiltonian H when given quantum access to Hamiltonian simulation $\exp(-iHt)$ or block encoding of H. While initial algorithms, such as the textbook version of phase estimation [1], depend on techniques such as the Quantum Fourier Transform

Zhiyan Ding: zding.m@berkeley.edu

Haoya Li: lihaoya@stanford.edu

Lin Lin: linlin@berkeley.edu

Hongkang Ni: hongkang@stanford.edu

Lexing Ying: lexing@stanford.edu

Ruizhe Zhang: rzzhang@berkeley.edu

(QFT) and multiple ancilla qubits, recent advances demonstrate that comparable results can be achieved with as few as one ancilla qubit. Despite the much simpler quantum circuit, these "modern" approaches often match or exceed the performance of their predecessors due to enhanced postprocessing techniques. This improvement has been demonstrated in applications such as estimating the ground-state energy, which involves determining the smallest eigenvalue of H and is an important application of phase estimation [2–12]. More recently, this enhancement has also been evident in the broader and more general problem of Multiple Eigenvalue Estimation (MEE) [13–20]. A typical example of MEE is estimating the low-lying energies of Hamiltonian H, which has many applications, such as determining the electronic and optical properties of materials.

In order to solve MEE for a given quantum Hamiltonian $H \in \mathbb{C}^{M \times M}$, we assume the availability of an initial state $|\psi\rangle$ that contains several dominant modes. Specifically, let $\{(\lambda_m, |\psi_m\rangle)\}_{m=1}^M$ represent pairs of eigenvalues and eigenvectors of H. We define $p_m = |\langle \psi_m | \psi \rangle|^2$ as the overlap between the initial state and the m-th eigenvector. Our primary assumption in this paper is the *Sufficiently Dominant Condition*: there exists a set of indices $\mathcal{D} \subset 1, 2, \cdots, M$ such that $p_{\min} =: \min_{i \in \mathcal{D}} p_i > p_{\text{tail}} =: \sum_{i \in \mathcal{D}^c} p_i$, where $\mathcal{D}^c = \{1, 2, \cdots, M\} \backslash \mathcal{D}$. The eigenvalues $\{\lambda_m\}_{m \in \mathcal{D}}$ are then called the dominant eigenvalues of H with respect to the initial state $|\psi\rangle$ (or simply the dominant eigenvalues), and the associated eigenvectors are referred to as the dominant eigenvectors. For simplicity, we assume $\|H\| \leq \pi$, which also implies $\{\lambda_m\}_{m \in \mathcal{D}} \subset [-\pi, \pi]$. Our objective is to estimate the dominant eigenvalues $\{\lambda_m\}_{m \in \mathcal{D}}$. From a signal processing perspective, the *Sufficiently Dominant Condition* allows us to differentiate signal and noise, providing a natural basis for analysis. Different versions of the condition have appeared in previous works [10, 11, 14, 18, 19]. A notable instance is in ground-state energy estimation, where the condition is equivalent to the initial overlap p_1 between the initial state and the ground state being greater than 0.5. The $\mathcal{O}(1)$ error allowed by this condition is the key to the robustness of many algorithms based on phase estimation.

To estimate the dominant eigenvalues, we assume an oracle access to the Hamiltonian simulation $\exp(-itH)$ for any $t \in \mathbb{R}$. Specifically, given any $t \in \mathbb{R}$, we assume the ability to implement the Hadamard test circuit (see Section 2 for details) to obtain an unbiased estimation to $\langle \psi | \exp(-itH) | \psi \rangle^{\text{i}}$. Several quantum phase estimation algorithms [5, 9–11, 14, 18, 19] have been developed assuming access to the Hadamard test circuit. In general, these algorithms, including the one proposed in this paper, involve three steps (refer to Fig. 1 for the flowchart): 1. Generate a proper set of $\{t_n\}_{n=1}^N \subset \mathbb{R}$; 2. Execute the Hadamard test circuits with t_n and obtain the dataset $\{(t_n, Z_n)\}_{n=1}^N$, where Z_n is an approximation of $\langle \psi | \exp(-it_n H) | \psi \rangle$; 3. Classically post-process Z_n to derive the estimation for $\{\lambda_m\}_{m \in \mathcal{D}}$. The efficiency of a quantum phase estimation algorithm is then quantified by two metrics: the maximal runtime denoted by $T_{\max} = \max_{1 \leq n \leq N} |t_n|$, and the total runtime $T_{\text{total}} = \sum_{n=1}^N |t_n|$. Here, $T_{\max}$ and T_{total} approximately measure the depth of the circuit and the total cost of the algorithm, respectively. Although the access to the Hamiltonian simulation $\exp(-itH)$ is assumed to be exact in this work for the sake of simplicity, we anticipate that a certain level of simulation error can be allowed as well by treating the simulation error together with the statistical errors and the effect of non-dominant eigenvectors (see Section 5 for details).

We further define two types of spectral gaps. The first is the spectral gap between the dominant eigenvalues, denoted by $\Delta_{\text{dom}} := \min_{i,j \in \mathcal{D}, j \neq i} |\lambda_i - \lambda_j|$. The second is the

---

[i] Throughout this paper, we assume access to $\exp(-itH)$ for any $t \in \mathbb{R}$ for simplicity of presentation. It is straightforward to extend our algorithm to the case with integer powers, where access only to $\exp(-inH)$ for $n \in \mathbb{N}$ is assumed. Refer to Section 5 and Appendix C for details.

spectral gap between the dominant eigenvalues and the remaining eigenvalues, denoted by $\Delta := \min_{i \in \mathcal{D}, j \neq i} |\lambda_i - \lambda_j|$. In this paper, we use the notations $\mathcal{O}_{[a]}$, $\Omega_{[a]}$, and $\Theta_{[a]}$ to indicate that the quantity polynomially depends on the parameters in $[a]$ if a is a number greater than one, or inversely proportional to $[a]$ if a is a number smaller than one.

The phase estimation algorithm proposed in this paper exhibits the following properties:

(1) Allow imperfect initial state: $p_{\text{tail}} > 0$.

(2) Maintain Heisenberg-limited scaling: Assuming all other parameters remain constant, the algorithm can achieve ϵ-accuracy with $T_{\text{total}} = \widetilde{\mathcal{O}}(1/\epsilon)$.

(3) No gap requirement: The algorithm can achieve ϵ-accuracy with $T_{\text{total}} = \text{poly}(1/\epsilon)$ for any $\epsilon > 0$, where the polynomial and constants are independent of $\Delta, \Delta_{\text{dom}}$.

(4) "Short" depth: When the spectral gap between the dominant eigenvalues $\Delta_{\text{dom}} > 0$ and the precision is small enough, that is, $\epsilon = \widetilde{\mathcal{O}}_{p_{\text{tail}}, p_{\min}, |\mathcal{D}|}(\Delta_{\text{dom}})$, the maximal runtime $T_{\max}$ can be as small as $\widetilde{\mathcal{O}}_{p_{\min}, |\mathcal{D}|}(p_{\text{tail}}/\epsilon)$. Here, the constant before $1/\epsilon$ approaches zero when $p_{\text{tail}} \to 0$. In addition, the total runtime still achieves the Heisenberg-limited scaling. More specifically, $T_{\text{total}} = \widetilde{\mathcal{O}}_{p_{\text{tail}}, p_{\min}, |\mathcal{D}|}(1/\epsilon)$.

Although previous algorithms may fulfill some of the mentioned properties (see Section 2.1 for a detailed discussion), to our knowledge, our algorithm stands out as **the first algorithm** that can be **rigorously proven** to simultaneously achieve **all** four properties, which matches the best available results in the literature. In particular, the "short" circuit depth property is considered important for applications on early fault-tolerant quantum computers [10, 21].

The algorithm presented in this paper also fulfills the following two additional properties, enhancing its efficiency compared to others:

(5) "Constant" depth: When the spectral gap between dominant eigenvalues and all other eigenvalues $\Delta > 0$ and $\epsilon = \widetilde{\mathcal{O}}_{p_{\min}}(\Delta)$, we can set $T_{\max} = \widetilde{\mathcal{O}}_{p_{\min}}(\delta/\epsilon \log(1/\delta))$ and $T_{\text{total}} = \widetilde{\mathcal{O}}_{p_{\min}}(1/(\delta\epsilon))$ for any $\delta = \Omega_{p_{\min}}(\epsilon/\Delta)$. In particular, setting $\delta = \Theta_{p_{\min}}(\epsilon/\Delta)$ gives the constant depth $T_{\max} = \widetilde{\mathcal{O}}(\Delta^{-1}\log(1/\epsilon))$, and the total cost is $T_{\text{total}} = \widetilde{\mathcal{O}}_{p_{\min}}(\Delta\epsilon^{-2})$.

(6) The quantum cost of the algorithm depends *logarithmically* on the number of dominant eigenvalues $|\mathcal{D}|$.

Property (5) can be satisfied in some of the prior work, such as [9, 10] for ground state energy estimation and [18] for MEE (see the detailed discussion in Section 3). An application of Property (5) is to ensure the algorithm's efficiency in the presence of global depolarizing noise. In this scenario, it is important to maintain a short depth $T_{\max} = \mathcal{O}(\log(1/\epsilon))$ (setting $\delta = \Theta(\epsilon/\Delta)$) to ensure that T_{total} polynomially depends on $1/\epsilon$ [22].

We find that our algorithm does not require meticulous tuning of the simulation parameters. Specifically, for the algorithm to succeed, only knowledge of the upper / lower bounds of the parameters is required (see Section 2). For optimal complexity, the parameters can be chosen according to $\Delta, \Delta_{\text{dom}}, p_{\text{tail}}, p_{\min}$ (see Section 3 for details).

The rest of this paper is organized as follows. In Section 2, we introduce the main idea of our method, present the algorithm, and provide a brief summary of related works. The complexity of our algorithm is detailed in Section 3. The proofs are included in Appendix A. Section 4 proposes several numerical examples and compares our algorithm with previous ones to justify its efficiency. A summary of this work and a discussion of possible extensions are provided in Section 5.
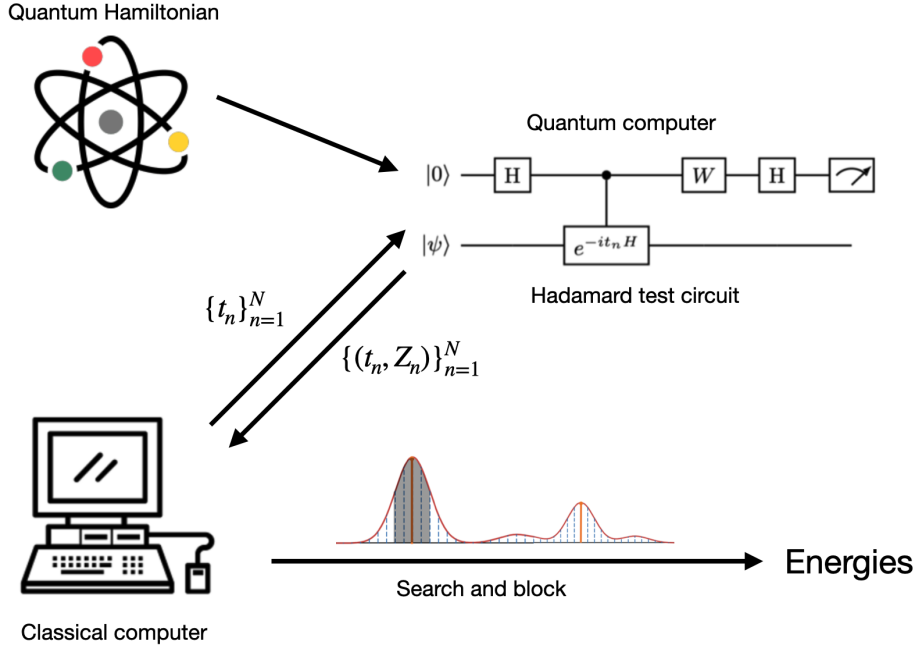
Figure 1: Flowchart of the main algorithm. The procedure involves three steps: Firstly, a sequence of t_n is generated from a truncated Gaussian using a classical computer. In the second step, the Hadamard test circuit is implemented on a quantum computer to produce the dataset (t_n, Z_n) as defined in (8). Within the Hadamard test circuit, we select $W = I$ or $W = S^\dagger$ (where S is the phase gate) to estimate the real or imaginary part of $\langle \psi | \exp(-itH) | \psi \rangle$. In the final step, postprocessing is performed on the quantum data (t_n, Z_n) for eigenvalue estimation.

2 Main algorithm and previous work

We first introduce our main idea informally. Given a Hamiltonian H and an initial state $|\psi\rangle$, our approach relies only on quantum access to the Hadamard test. Specifically, for any $t \in \mathbb{R}$, we can repeat the Hadamard test with $|\psi\rangle$ and $\exp(-iHt)$ several times to obtain an unbiased estimation $Z(t)$ of the following expression:

$$\mathcal{Z}(t) = \langle \psi | \exp(-iHt) | \psi \rangle = \sum_{m=1}^{M} p_m \exp(-i\lambda_m t) \,, \tag{1}$$

meaning $\mathbb{E}(Z(t)) = \mathcal{Z}(t)$.

The central subroutine of our algorithm involves a filtering-searching process. Given an even probability density $a(t)$, we independently generate N samples $\{t_n\}_{n=1}^{N}$ from this distribution. Subsequently, we obtain approximations $\{Z(t_n)\}_{n=1}^{N}$ and calculate the filtering function:

$$G(\theta) = \left| \frac{1}{N} \sum_{n=1}^{N} Z(t_n) \exp(i\theta t_n) \right| \,, \tag{2}$$

where $Z(t_n)$ is a random variable that can take values in $\{\pm 1 \pm i\}$; a detailed discussion can be found later.

Define the Fourier transform of the probability density function $a(t)$ as follows:

$$F(x) = \int_{-\infty}^{\infty} a(t) \exp(ixt) \mathrm{d}t \,. \tag{3}$$

Because $\{t_n\}$ are independently sampled from $a(t)$, it is straightforward to see

$$G(\theta) \approx |\mathbb{E}_{t \sim a}\left(Z(t)\exp(i\theta t)\right)| = \sum_{m=1}^{M} p_m F(\theta - \lambda_m) =: \mathcal{G}(\theta), \quad \text{when} \quad N \gg 1. \quad (4)$$

We observe that $\mathcal{G}(\theta)$ is an even function of real value, given that $a(t) = a(-t)$. In an ideal scenario, assuming that $F(x)$ reaches its maximum at $x = 0$ and decays rapidly as $|x|$ increases so that $\max_{1 \le m \le M-1} |F(\lambda_{m+1} - \lambda_m)| \ll 1$, we can derive the following approximations:

$$G(\theta) \approx \mathcal{G}(\theta) \approx \begin{cases} p_m F(\theta - \lambda_m), & \theta \text{ close to } \lambda_m, \\ 0, & \text{otherwise}. \end{cases} \quad (5)$$

Subsequently, we can implement the following search procedure to identify all dominant eigenvalues:

- First, we locate the maximum point $\boldsymbol{\theta}_1$ of $G(\theta)$. Following (5), we find $\boldsymbol{\theta}_1 \approx \lambda_{m_1}$, where $p_{m_1} = \max_{1 \le m \le M} p_m$.

- To approximate the next dominant eigenvalue, we establish a block interval $\mathcal{I}_{B,1} = [\boldsymbol{\theta}_1 - d, \boldsymbol{\theta}_1 + d]$ around $\boldsymbol{\theta}_1$ with a suitable d. Subsequently, we identify the second maximal point $\boldsymbol{\theta}_2$ of $\mathcal{G}(\theta)$ outside the block interval, denoted $\boldsymbol{\theta}_2 = \arg\max_{\theta \in \mathcal{I}_{B,1}^c} G(\theta)$.

  Given that F concentrates around 0, we have $p_{m_1} \exp(\theta - \lambda_{m_1}) \ll 1$ when $\theta \in \mathcal{I}_{B,1}^c$. Consequently, the impact of the first dominant eigenvalue is mitigated by blocking the interval in the second search step, allowing us to show $\boldsymbol{\theta}_2 \approx \lambda_{m_2}$, where $p_{m_2} = \max_{m \ne m_1} p_m$.

- After acquiring $\boldsymbol{\theta}_2$, we update the block interval by defining $\mathcal{I}_{B,2} = [\boldsymbol{\theta}_2 - d, \boldsymbol{\theta}_2 + d] \cup \mathcal{I}_{B,1}$ and identify the third maximal point of $G(\theta)$ outside $\mathcal{I}_{B,2}$.

  This searching and updating process is iteratively repeated until a set of $|\mathcal{D}|$ "maximal" points is discovered. Ultimately, we obtain an approximate set $\{\boldsymbol{\theta}_m\}_{m=1}^{|\mathcal{D}|}$ corresponding to the set of dominant eigenvalues $\{\lambda_m\}_{m \in \mathcal{D}}$.

To ensure the success of the process, it is crucial to choose a suitable probability density function $a(t)$ such that the function $F(x)$ concentrates around $x = 0$. In this paper, we adopt the *truncated Gaussian density function* for $a(t)$:

$$\begin{aligned} a(t) = &\left(1 - \int_{-\sigma T}^{\sigma T} \frac{1}{\sqrt{2\pi}T} \exp\left(-\frac{s^2}{2T^2}\right) \mathbf{1}_{[-\sigma T, \sigma T]}(s)\mathrm{d}s\right) \delta_0(t) \\ &+ \frac{1}{\sqrt{2\pi}T} \exp\left(-\frac{t^2}{2T^2}\right) \mathbf{1}_{[-\sigma T, \sigma T]}(t). \end{aligned} \quad (6)$$

Here, $\delta_0(t)$ is Dirac delta function at point 0. The choice of $a(t)$ is inspired by the fact that the Fourier transform of a Gaussian function remains a Gaussian function and a recent result of spike localization [23]. The parameter σ represents the level of truncation. Specifically, when $\sigma = \infty$, we have $F(x) = \exp\left(-\frac{T^2 x^2}{2}\right)$, reaching its maximum at $x = 0$ and exponentially decaying to zero with respect to $T|x|$. Furthermore, the use of the truncated Gaussian ensures that the maximum runtime $T_{\max} = \max_n |t_n|$ never exceeds σT. In the following part of the paper, we may also employ the notation $a_T(t) = a(t)$ and $F_T(x) = F(x)$ to emphasize the dependence on T.

Now, we are ready to introduce our main algorithm. With the motivation explained above, we propose the algorithm in two steps:

Accepted in 〈 〉uantum 2024-09-19, click title to verify. Published under CC-BY 4.0.

5

Step 1: Data generation.

We implement the Hadamard test quantum circuit as shown in Fig. 1 to obtain our data set. Specifically, we can set $W = I$ (or $W = S^\dagger$ with S being the phase gate), measure the ancilla qubit, and define a random variable X (or Y) such that $X = 1$ (or $Y = 1$) if the outcome is 0 and $X = -1$ (or $Y = -1$) if the outcome is 1. Then

$$\mathbb{E}(X + iY) = \langle\psi|\exp(-itH)|\psi\rangle . \tag{7}$$

Given a set of time points $\{t_n\}_{n=1}^N$ drawn from the probability density $a(t)$, we apply Hadamard tests to generate the following data set:

$$\mathcal{D}_H = \{(t_n, Z_n)\}_{n=1}^N := \{(t_n, X_n + iY_n)\}_{n=1}^N . \tag{8}$$

Here, X_n and Y_n are random variables that each take values of -1 or 1. Each evaluation of X_n (or Y_n) only requires running the Hadamard test circuit with $W = I$ (or $W = S^\dagger$) once at $t = t_n$. Referring to (7), we obtain:

$$\mathbb{E}(Z_n) = \langle\psi|\exp(-it_nH)|\psi\rangle, \quad Z_n \in \{\pm 1 \pm i\}, \quad |Z_n| = \sqrt{2}. \tag{9}$$

Hence, Z_n serves as an unbiased and bounded estimate of $\langle\psi|\exp(-it_nH)|\psi\rangle$. Furthermore, it should be noted that if we employ the aforementioned method to construct the data set, the maximum simulation time is $T_{\max} = \max_{1\leq n\leq N}|t_n|$, and the total simulation time is $T_{\text{total}} = \sum_{n=1}^N |t_n|$.

We summarize the data generation process in Algorithm 1.

Algorithm 1 Data generator

1: **Preparation:** Number of data pairs: N; Truncated Gaussian density: $a(t)$;
2: **Running:**
3: $n \leftarrow 1$;
4: **while** $n \leq N$ **do**
5:     Generate a random variable t_n with the probability density $a(t)$.
6:     **if** $t_n > 0$ **then**
7:         Run the quantum circuit (Figure 1) with $t = t_n$ and $W = I$ to obtain X_n.
8:         Run the quantum circuit (Figure 1) with $t = t_n$ and $W = S^\dagger$ to obtain Y_n.
9:         $Z_n \leftarrow X_n + iY_n$.
10:     **end if**
11:     **if** $t_n = 0$ **then**
12:         $Z_n \leftarrow 1$.
13:     **end if**
14:     $n \leftarrow n + 1$
15: **end while**
16: **Output:** $\{(t_n, Z_n)\}_{n=1}^N$

Step 2: Filtering and searching. After generating the dataset, we define the Gaussian filtering function as in (2). By selecting a suitable q, we define the set of candidates:

$$\theta_j = -\pi + \frac{jq}{T}, \quad 0 \leq j \leq J := \left\lfloor \frac{2\pi T}{q} \right\rfloor .$$

Subsequently, we iterate the previously described searching procedure K times using $G(\theta)$ and the set $\{\theta_j\}_{j=0}^J$ to obtain the approximation $\{\boldsymbol{\theta}_k\}_{k=1}^K$.

The main algorithm of this paper is described in Algorithm 2. It should be noted that our algorithm is flexible in terms of parameter selection. Specifically, a meticulous adjustment of the parameters σ, T, N, α, q based on prior knowledge of (p_m, λ_m) is unnecessary for the algorithm to work. When no prior knowledge is available, we recommend choosing larger values for σ, T, N. In practice, to avoid detection of repeated dominant eigenvalues during the search, we recommend choosing $q < \alpha \ll T$. In Section 3, we provide a lower bound for α, σ, T, N (no required upper bound) and an upper bound for q (no required lower bound) to ensure the algorithm's success.

Algorithm 2 Quantum Multiple Eigenvalue Gaussian filtered Search (QMEGS)

1: **Preparation:** Number of data pairs: N; Depth parameter: T; Block parameter: α; Searching parameter: q; Truncated Gaussian density: $a_T(t)$; Number of dominant eigenvalues (guess): K;
2: **Running:**
3: Generate a data set of size N ▷ Step 1: Generate data
$$\{(t_n, Z_n)\}_{n=1}^N$$
using Algorithm 1 with truncated Gaussian density $a_T(t)$.
4: $J \leftarrow \left\lfloor \frac{2\pi T}{q} \right\rfloor$.
5: Generate discrete candidates: $\theta_j \leftarrow -\pi + \frac{jq}{T}$.
6: Calculate   ▷ Step 2: Compute the filtered density function
$$G_j \leftarrow \left| \frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta_j t_n) \right|, \quad 0 \le j \le J.$$
7: Block set: $\mathcal{I}_{B,1} \leftarrow \emptyset$.   ▷ Step 3: Find peaks
8: $k \leftarrow 1$.
9: **while** $k \le K$ **do**
10:     $j_k = \operatorname{argmax}_{\theta_j \notin \mathcal{I}_{B,k}} G_j$.
11:     $\boldsymbol{\theta}_k \leftarrow \theta_{j_k}$.
12:     $\mathcal{I}_{B,k+1} \leftarrow \mathcal{I}_{B,k} \cup \left(\boldsymbol{\theta}_k - \frac{\alpha}{T}, \boldsymbol{\theta}_k + \frac{\alpha}{T}\right)$. ▷ Block interval to avoid finding the same peak
13:     $k \leftarrow k + 1$
14: **end while**
15: **Output:** $\{\boldsymbol{\theta}_k\}_{k=1}^K$

Finally, we emphasize that while the informal analysis provided above appears to be conceptually simple, there are still some issues that need to be addressed.

- In the informal analysis, to ensure (5), we require

$$\max_{1 \le m \le M-1} |F(\lambda_{m+1} - \lambda_m)| \ll 1. \tag{10}$$

Given our choice of $a(t)$ as specified in (6), where $F(x) \approx \exp\left(-\frac{T^2 x^2}{2}\right)$, ensuring that (10) asks for

$$T_{\max} > T > \frac{1}{\min_{1 \le m \le M-1} |\lambda_{m+1} - \lambda_m|} = \Omega(M), \tag{11}$$

which poses an undesirable requirement in the context of quantum phase estimation[ii]. Here, the last equality is based on the assumption that all the eigenvalues of H belong to the interval $[-\pi, \pi]$.

[ii]For example, in ground state energy estimation, it is well known that $T_{\max}$ can be independent of the spectral gap $\lambda_2 - \lambda_1$.

- In the informal derivation, we ignored the finite sampling and measurement errors and replaced $G(\theta)$ directly with $\mathcal{G}(\theta)$. In practice, however, ensuring small finite sampling and measurement errors requires a large number of samples and is not ideal in quantum computing.

In Section 3, we present several results that demonstrate that our algorithm can autonomously address the two concerns mentioned above. We will show that under condition $p_{\min} > p_{\text{tail}}$, (5) and (11) are not necessary. Specifically, even if T does not satisfy (11), and (5) does not hold, the set $\{\boldsymbol{\theta}_k\}_{k=1}^{|\mathcal{D}|}$ still serves as an approximation to the set of dominant eigenvalues in a meaningful sense. Moreover, even when $G(\theta)$ is not $\mathcal{O}(\epsilon)$ close to $\mathcal{G}(\theta)$, the search process remains stable enough to ensure the accuracy of the approximation.

We also remark here that the choice of distribution $a(t)$ is not confined to the form specified in (6), and it can be tailored based on the setting of the problem. For instance, in scenarios where only a unitary operator U is provided as a black box, and the goal is to retrieve its eigenvalues $e^{i\lambda_m}$, the power t is constrained to integers and thus $a(t)$ needs to be a distribution on integers. A detailed discussion is given in Appendix C.

2.1 Comparison with previous work

In this section, we review previous multiple eigenvalue estimation algorithms. A summary of the comparison between different algorithms is listed in Table 1.

The first work using the Hadamard test circuit to achieve Heisenberg-limited scaling is [5]. Several works have been developed to achieve Heisenberg-limited scaling for multiple eigenvalue estimation and without relying on any gap assumptions (Properties (2) and (3)[iii]). For example, [17] introduced a method for estimating multiple eigenvalue phases that extends the idea of robust phase estimation (RPE) [11, 24, 25] to multiple eigenvalues and achieves Heisenberg-limited scaling without relying on any assumptions about the spectral gap. However, the theoretical analysis in [17] is based on the assumption that all non-dominant modes vanish, as defined in [17, Definition 3.1 and Theorem 4.5], specifically $p_{\text{tail}} = 0$. This drawback has been resolved by a recent work [19]. The robust multiple phase estimation (RMPE) method in [19] extends the Kitaev-type RPE method [11, 25] to the multiple eigenvalue estimation problem by using adaptive simulation time amplifying factors and suitable signal processing algorithms. In contrast to [17], the algorithm proposed in [19, Section III] can achieve Heisenberg-limited scaling without the requirement of a gap between dominant eigenvalues, even when dealing with an imperfect initial state ($p_{\text{tail}} > 0$) (Properties (1)-(3)). The design of the algorithm is based on a dedicated line spectrum estimation algorithm [23], which employs the same Gaussian filtering function as employed in this work. Moreover, in scenarios where a gap exists between dominant eigenvalues, the algorithm presented in [19, Section V] combines the line spectrum estimation algorithm with ESPRIT to estimate the dominant eigenvalues with a short circuit depth (Property (4)). Specifically, to achieve ϵ-accuracy, the algorithm requires $T_{\max} = \mathcal{O}_{p_{\min},|\mathcal{D}|}(\max\{p_{\text{tail}}^{-1}\Delta_{\text{dom}}^{-1}, p_{\text{tail}}\epsilon^{-1}\})$ and $T_{\text{total}} = \mathcal{O}_{p_{\min},|\mathcal{D}|}(p_{\text{tail}}^{-2}\Delta_{\text{dom}}^{-1} + p_{\text{tail}}^{-1}\epsilon^{-1})$, resulting in a short circuit depth as long as $\epsilon = \mathcal{O}(p_{\text{tail}}^2\Delta_{\text{dom}})$. Compared to the algorithms outlined in [19, Sections III, V], the quantum cost of QMEGS exhibits a logarithmic dependence on $|\mathcal{D}|$, and QMEGS can further achieve the "constant" depth property in the

---

[iii]Rigorously speaking, Heisenberg-limited scaling without gap requirement surpasses the strength of Properties (2) and (3) stated in Section 1. It implies that the algorithm can achieve Heisenberg-limited scaling without any gap assumption. Specifically, for any $\epsilon > 0$, the algorithm achieves ϵ-accuracy with $T_{\text{total}} = \tilde{\mathcal{O}}_{p_{\text{tail}}, p_{\min}}(1/\epsilon)$, where the constant is independent of $\Delta, \Delta_{\text{dom}}$. As demonstrated in Theorem 3.1, our algorithm achieves Heisenberg-limited scaling without relying on any gap assumptions.

presence of a spectral gap between dominant and other eigenvalues. In addition, QMEGS does not require meticulous parameter adjustment (e.g., σ, T, N, α, q) based on prior knowledge of $\{(p_m, \lambda_m)\}$, and can be more flexible than the algorithms proposed in [19].

An optimization-based signal processing method called the multi-modal, multi-level quantum complex exponential least squares (MM-QCELS) method has recently been proposed [18], which generalizes the quantum complex exponential least squares (QCELS) method [10] to the setting of multiple eigenvalues. When $\Delta_{\text{dom}} > 0$, MM-QCELS can approximate these dominant eigenvalues with Heisenberg limited scaling, "short", or "constant" circuit depth (Properties (1), (2), (4), (5)). Furthermore, the quantum cost of MM-QCELS only depends logarithmically on $|\mathcal{D}|$ (Property (6)). To achieve Heisenberg-limited scaling, MM-QCELS needs to generate a sequence of datasets on quantum computers. The classical optimization procedure of MM-QCELS to find K dominant eigenvalues solves an optimization problem in a K-dimensional space (see Appendix B for a brief overview). In the worst-case scenario, the classical post-processing cost can grow exponentially in ϵ^{-K}. Compared to MM-QCELS, QMEGS has a much simpler data generation and searching process. The algorithm only requires a single dataset generated by a single T. Additionally, leveraging block intervals in the searching process, QMEGS can achieve ϵ-accuracy with Heisenberg-limited scaling even when $\epsilon = \Omega(\Delta_{\text{dom}})$. Regarding the classical processing cost, QMEGS only requires evaluating the filter function at a finite number of discrete points in $[-\pi, \pi]$ and the number of evaluations is $\mathcal{O}(1/\epsilon)$ and is independent of K.

Quantum subspace diagonalization (QSD), quantum Krylov, matrix pencil, and ESPRIT methods, as highlighted in studies such as [15, 16, 20, 26–33], offer an alternative way to solve the eigenvalue estimation problem. These methods estimate eigenvalues by addressing specific projected eigenvalue problems or singular value problems and have proven valuable for estimating ground-state and excited-state energies across various scenarios. Despite classical perturbation theories suggesting potential challenges, such as ill-conditioned projected problems and sensitivity to noise, empirical observations indicate that these quantum methods often outperform pessimistic theoretical predictions. Recently, [14] proposed and investigated the complexity of a quantum phase estimation algorithm based on ESPRIT [34, 35]. It recovers the frequencies by computing the SVD of the Hankel matrix generated from samples of the signal. Assuming a sufficiently large dominant spectral gap Δ_{dom} and $p_{\text{tail}} = 0$, [14] establishes that ESPRIT can achieve ϵ precision with $T_{\max} = \widetilde{\mathcal{O}}(\text{poly}(|\mathcal{D}|)\epsilon^{-o(1)})$ and $T_{\text{total}} = \widetilde{\mathcal{O}}(\text{poly}(|\mathcal{D}|)(p_{\min}\epsilon)^{-2})$. Notably, the relationship $T_{\text{total}} = \Omega(T_{\max}^2)$, which arises from the aliasing issue, poses a challenge to the ESPRIT method (as well as other ESPRIT-like methods, such as the observable dynamical mode decomposition (ODMD) method [20]) in achieving the Heisenberg-limited scaling (see Appendix B for detail). More recently, [19, Section IV] presented a multi-level ESPRIT technique to overcome the restriction of $p_{\text{tail}} = 0$ and reach the Heisenberg limited scaling with short circuit depth (Properties (1), (2), (4)). Specifically, when $\epsilon \ll \Delta_{\text{dom}}$, the authors illustrate that multilevel ESPRIT can achieve ϵ-accuracy with $T_{\max} = \widetilde{\mathcal{O}}(\text{poly}(|\mathcal{D}|)p_{\text{tail}}(p_{\min}\epsilon)^{-1})$ and $T_{\text{total}} = \widetilde{\mathcal{O}}(\text{poly}(|\mathcal{D}|)(p_{\text{tail}}p_{\min})^{-1})$. In addition, despite the apparent importance of the gap assumption in the previous analysis of ESPRIT [14, 19], our numerical experiments in Section 4 suggest that this assumption might be relaxed in real-world applications.

It should be noted that filtering techniques have also been explored in previous studies. For example, in [9], the authors introduced the Gaussian derivative filter to estimate the ground-state energy (smallest eigenvalues) and achieve a "Constant" depth property (Property (5)). However, it remains unclear how to extend their algorithms to address the MEE problem while satisfying the first four properties (Properties (1)-(4)). More recently,

[12] proposed an algorithm that employs a filter function based on the derivative Heaviside function. In this approach, the eigenvalues of H can be approximated by identifying the local maxima of the approximate derivative of the cumulative distribution function (CDF). As $T_{\max}$ tends to infinity, their approach aligns with QMEGS in that both methods aim to identify the local maximal point in the sum of delta functions. However, when aiming to identify dominant eigenvalues with finite $T_{\max}$, the performance of their algorithm and its ability to achieve Heisenberg-limited scaling are unclear.

Algorithms	Properties				Comments
	Allow $p_{\text{tail}} > 0$	Heisenberg limit	No gap requirement	"Short" depth	
QEEA [13]	✓	✗	✓	✗	
ESPRIT [14]	?	✗	?	✗	
[17]	?	✓	✓	✗	poly($|\mathcal{D}|$) quantum cost
[19, Theorem III.5]	✓	✓	✓	✗	poly($|\mathcal{D}|$) quantum cost
[19, Theorem V.1]	✓	✓	✗	✓	
MM-QCELS [18]	✓	✓	✗	✓	"Constant" depth, $\log|\mathcal{D}|$ quantum cost large classical cost
QMEGS (this work)	✓	✓	✓	✓	"Constant" depth, $\log|\mathcal{D}|$ quantum cost

Table 1: Comparison of the existing theoretical analysis of different methods for multiple eigenvalue estimation. In the table, a question mark indicates that the method potentially satisfies this property, although there is no existing result that rigorously proves it. For simplicity, certain properties in this table represent a slightly relaxed version of those introduced in Section 1. Specifically, Allow $p_{\text{tail}} > 0$ means that the algorithm can deal with an imperfect initial state. The Heisenberg limit implies that the algorithm can achieve ϵ-accuracy with $T_{\text{total}} = \widetilde{\mathcal{O}}(1/\epsilon)$, assuming that all other parameters remain constant. No gap requirement means that the algorithm can achieve ϵ-accuracy with $T_{\text{total}} = \text{poly}(1/\epsilon)$, where the polynomial and constants are independent of $\Delta, \Delta_{\text{dom}}$. "Short" depth means that, when $\epsilon = \widetilde{\mathcal{O}}_{p_{\text{tail}}, p_{\min}, |\mathcal{D}|}(\Delta_{\text{dom}})$, the algorithm can achieve ϵ-accuracy with $T_{\max} = \widetilde{\mathcal{O}}_{p_{\min}, |\mathcal{D}|}(p_{\text{tail}}/\epsilon)$ and $T_{\text{total}} = \widetilde{\mathcal{O}}_{p_{\text{tail}}, p_{\min}, |\mathcal{D}|}(1/\epsilon)$. A more detailed comparison can be found in Section 2.1.

3 Statement of the main results

This section introduces the complexity result of Algorithm 2. The goal is to demonstrate that, across all three regimes categorized by our knowledge of the spectral gap, the algorithm consistently identifies a highly accurate approximation of the dominant eigenvalues through appropriate parameter selection. We provide the proof of the results in Appendix A.

The three cases are divided as follows depending on the values of $T, \Delta, \Delta_{\text{dom}}$:

- General case: In this case, the theoretical result does not require assumptions about T, Δ_{dom}, and Δ. Consequently, we refer to it as the *No gap requirement* regime. The result is summarized in the following theorem.

  Theorem 3.1 ($\forall T > 0$). *Assume $p_{\min} > p_{\text{tail}}$ and $K \geq |\mathcal{D}|$. Given the probability of failure $\eta > 0$, we choose the following parameters:*

  – *Block constant: $\alpha = \Omega\left(\log^{1/2}\left(\frac{1}{p_{\min} - p_{\text{tail}}}\right)\right)$,*

- *Searching parameter:* $q = \mathcal{O}\left(\log^{1/2}\left(\frac{p_{\min}}{p_{\text{tail}}+(p_{\min}-p_{\text{tail}})/2}\right)\right)$, $q < \alpha/3$, and $\alpha/q \in \mathbb{N}$,

- *Truncation parameter:* $\sigma = \Omega\left(\log^{1/2}\left(\frac{1}{p_{\min}-p_{\text{tail}}}\right)\right)$,

- *Number of samples:* $N = \Omega\left(\frac{1}{(p_{\min}-p_{\text{tail}})^2}\log\left(\left(\frac{T}{q}+|\mathcal{D}|\right)\frac{1}{\eta}\right)\right)$.

Then, with probability at least $1-\eta$, we have that for each $i \in \mathcal{D}$, there exists $1 \le k_i \le |\mathcal{D}|$ such that

$$|\lambda_i - \boldsymbol{\theta}_{k_i}| \le \frac{\alpha}{T}. \tag{12}$$

In particular, for any $\epsilon > 0$, to achieve

$$\{\lambda_m\}_{m\in\mathcal{D}} \subset \cup_k[\boldsymbol{\theta}_k - \epsilon, \boldsymbol{\theta}_k + \epsilon],$$

it suffices to choose

$$T_{\max} = \widetilde{\Theta}\left(\frac{1}{\epsilon}\right), \quad T_{\text{total}} = \widetilde{\Theta}\left(\frac{1}{(p_{\min}-p_{\text{tail}})^2\epsilon}\log\left(\frac{|\mathcal{D}|}{\eta}\right)\right),$$

The above theorem guarantees that Properties (1)-(3) are attained as outlined in Section 1. This result shares similarities with [19, Theorem III.5]. Since a spectral gap is not assumed, T may not be large enough to distinguish between close dominant eigenvalues, and it is possible for two dominant eigenvalues to fall within the same interval $[\boldsymbol{\theta}_k - \alpha/T, \boldsymbol{\theta}_k + \alpha/T]$. Also, some intervals may not include any dominant eigenvalues since $K \ge |\mathcal{D}|$. But (12) ensures that the dominant eigenvalues are covered by the union set $\cup_k[\boldsymbol{\theta}_k - \alpha/T, \boldsymbol{\theta}_k + \alpha/T]$.

We briefly describe the strategy of proving the theorem using proof by contradiction as follows: First, assume that there exists $\lambda_{m^\star}$ such that $\lambda_{m^\star} \notin \overline{\mathcal{I}}_{B,|\mathcal{D}|}$ (here $\overline{\mathcal{I}}$ denotes the closure of $\mathcal{I}$). Then one can show that there exists a grid point $\boldsymbol{\theta}^\star \notin \mathcal{I}_{B,|\mathcal{D}|}$ such that $|\boldsymbol{\theta}^\star - \lambda_{m^\star}| \le q/T$ using the condition of q, α. Thus the value of G at $\boldsymbol{\theta}^\star$ has the following lower bound due to (4):

$$G(\boldsymbol{\theta}^\star) \ge p_{\min}\exp\left(-\frac{q^2}{2}\right) + \text{small error}.$$

Then each $\left[\boldsymbol{\theta}_k - \frac{\alpha}{3T}, \boldsymbol{\theta}_k + \frac{\alpha}{3T}\right]$ must cover some dominant eigenvalue λ_m since otherwise

$$G(\boldsymbol{\theta}_k) \le p_{\text{tail}} + \exp\left(-\frac{\alpha^2}{18}\right) + \text{small error} < p_{\min}\exp\left(-\frac{q^2}{2}\right) + \text{small error} \le G(\boldsymbol{\theta}^\star),$$

which contradicts with the fact that $\boldsymbol{\theta}_k$ is the maximal point at k-th step. Finally, since $\lambda_{m^\star}$ is not covered by any block intervals, some dominant eigenvalue $\lambda_{m'}$ must be covered by at least two such intervals $\left[\boldsymbol{\theta}_k - \frac{\alpha}{3T}, \boldsymbol{\theta}_k + \frac{\alpha}{3T}\right]$ and $\left[\boldsymbol{\theta}_j - \frac{\alpha}{3T}, \boldsymbol{\theta}_j + \frac{\alpha}{3T}\right]$ ($j \ne k$). However, according to the definition of the block interval, we have $|\boldsymbol{\theta}_k - \boldsymbol{\theta}_j| \ge \frac{\alpha}{T}$, which implies $\left[\boldsymbol{\theta}_k - \frac{\alpha}{3T}, \boldsymbol{\theta}_k + \frac{\alpha}{3T}\right] \cap \left[\boldsymbol{\theta}_j - \frac{\alpha}{3T}, \boldsymbol{\theta}_j + \frac{\alpha}{3T}\right] = \emptyset$. This contradicts the existence of $\lambda_{m'}$ and concludes the proof.

In Appendix A, we account for both random and truncation errors and form a rigorous proof using this idea.

Accepted in 《Juantum 2024-09-19, click title to verify. Published under CC-BY 4.0.

11

- Gapped dominant eigenvalues case: In this case, we assume that T is large enough to allow the Gaussian filter to distinguish dominant eigenvalues. We show that Algorithm 2 attains Heisenberg-limited scaling with $T_{\max} = p_{\text{tail}}/(p_{\min}\epsilon)$. Similar results have been attained by previous algorithms, such as MM-QCELS [18] and [19, Theorem IV.2], and are referred to as *Short depth* regime (Property (4)). The results are summarized in the following theorem:

Theorem 3.2 $(T \gg \Delta_{\text{dom}}^{-1})$. *Assume that $p_{\min} > p_{\text{tail}}$ and define*

$$\zeta = \min\{(p_{\min} - p_{\text{tail}})/p_{\text{tail}}, 1\}.$$

Given failure probability $\eta > 0$, we choose $q < \alpha/3$, $\alpha/q \in \mathbb{N}$ such that

$$\sigma = \Omega\left(\log^{1/2}\left(\frac{1}{\zeta p_{\text{tail}}}\right)\right), \quad \alpha = \Omega(\sigma), \quad q = \mathcal{O}\left(\frac{\zeta p_{\text{tail}}}{\sigma}\right). \tag{13}$$

and

$$T = \Omega\left(\frac{\alpha}{\Delta_{dom}}\right), \quad N = \Omega\left(\frac{1}{(\zeta p_{\text{tail}})^2} \log\left(\left(\frac{T}{q} + |\mathcal{D}|\right)\frac{1}{\eta}\right)\right).$$

There exists

$$Q = \Theta\left(\exp\left(\Theta(\zeta^{-1})\right)(\sigma\zeta + 1)\frac{p_{\text{tail}}}{p_{\min}T}\right)$$

such that, with probability $1 - \eta$, for each $i \in \mathcal{D}$, there exists a unique $1 \leq k_i \leq |\mathcal{D}|$ such that

$$|\lambda_i - \boldsymbol{\theta}_{k_i}| \leq \frac{Q}{T}. \tag{14}$$

Furthermore, we have $k_i \neq k_j$ for $i \neq j$. In particular, when $p_{\min} > (1+o(1))p_{\text{tail}}$, for $\epsilon = \widetilde{\mathcal{O}}\left(p_{\text{tail}}\Delta_{dom}/p_{\min}\right)$, the following quantum cost is enough to achieve $\max_{m \in \mathcal{D}} |\lambda_m - \boldsymbol{\theta}_{k_m}| \leq \epsilon$:

$$T_{\max} = \widetilde{\Theta}\left(\frac{p_{\text{tail}}}{p_{\min}\epsilon}\right), \quad T_{\text{total}} = \widetilde{\Theta}\left(\frac{1}{p_{\text{tail}}p_{\min}\epsilon}\log\left(\frac{|\mathcal{D}|}{\eta}\right)\right).$$

In contrast to the general case, in this scenario,

$$\exp\left(-T^2(\lambda_{m_1} - \lambda_{m_2})^2/2\right) \ll 1, \quad m_1, m_2 \in \mathcal{D}.$$

This condition ensures that the Gaussian filter can effectively distinguish between different dominant eigenvalues, establishing a one-to-one correspondence between dominant eigenvalues λ_m and their corresponding maximal points $\boldsymbol{\theta}_{k_m}$.

To establish the short circuit depth result in (14), we begin by noting that, neglecting the influence of other eigenvalues and noise, we must have $|\lambda_m - \boldsymbol{\theta}_{k_m}| \leq q/T$ due to the decreasing property of the function $F(x)$. When accounting for the impact of other eigenvalues, the following observations hold: 1. The influence of other dominant eigenvalues is negligible when determining $\boldsymbol{\theta}_k$ since $T = \widetilde{\Omega}\left(\Delta_{\text{dom}}^{-1}\right)$; 2. The effect of the tail eigenvalues diminishes as p_{tail} approaches zero since $\left|\sum_{m \in \mathcal{D}^c} p_m \exp(-(\theta - \lambda_m)^2 T^2/2)\right| \leq p_{\text{tail}}$. By leveraging these observations and controlling the impact of random noise, we establish the validity of (14) in Appendix A.

- The case where a gap is known between dominant eigenvalues and the tail: In this case, we assume that $T \gg \Delta^{-1}$. The Gaussian filter thus becomes sufficiently sharp to distinguish between dominant eigenvalues and tail eigenvalues. In this scenario, we demonstrate that Algorithm 2 can achieve Heisenberg-limited scaling, with $T_{\max}\epsilon$ arbitrarily close to zero. Furthermore, Algorithm 2 is capable of achieving ϵ accuracy with $T_{\max} = \mathcal{O}(\log(1/\epsilon))$ and $T_{\text{total}} = \mathcal{O}(1/\epsilon^2)$. Since the maximum running time $T_{\max}$ depends only logarithmically on ϵ, we refer to this regime as the *"Constant" depth* regime (Property (5)). It should be noted that similar results have been achieved previously in the task of ground-state energy estimation algorithms, such as QCELS [10, 22], and [9]. Here, we extend this result to the case of multiple eigenvalue estimation.

Theorem 3.3 $(T \gg \Delta^{-1})$. *Assume that $p_{\min} > p_{\text{tail}}$. Choose all parameters such that they satisfy the condition of Theorem 3.1. Furthermore, given any $0 < \zeta < p_{\min}/4$ and failure probability $\eta > 0$, we choose*

$$\sigma = \Omega\left(\log^{1/2}\left(\frac{1}{\zeta}\right)\right), \quad \alpha = \Omega(\sigma), \quad q = \mathcal{O}\left(\frac{\zeta}{\sigma}\right),$$

and

$$T = \Omega\left(\frac{\alpha}{\Delta}\right), \quad N = \Omega\left(\frac{1}{\zeta^2}\log\left(\left(\frac{T}{q} + |\mathcal{D}|\right)\frac{1}{\eta}\right)\right).$$

Let $Q = \mathcal{O}\left(\frac{\zeta\sigma}{p_{\min}}\right)$. Then, with high probability, for each $i \in \mathcal{D}$, there exist $1 \leq k_i \leq |\mathcal{D}|$ such that

$$|\lambda_i - \boldsymbol{\theta}_{k_i}| \leq \frac{Q}{T}. \tag{15}$$

Furthermore, it holds that $k_i \neq k_j$ for $i \neq j$. In particular, for $\epsilon = \widetilde{\mathcal{O}}(\Delta/p_{\min})$ and $\delta = \Omega(p_{\min}\epsilon/\Delta)$, the following quantum cost is enough to achieve $\max_{m \in \mathcal{D}} |\lambda_m - \boldsymbol{\theta}_{k_m}| \leq \epsilon$:

$$T_{\max} = \Theta\left(\frac{\delta}{p_{\min}\epsilon}\log\left(\frac{1}{\delta}\right)\right), \quad T_{\text{total}} = \widetilde{\Theta}\left(\frac{1}{p_{\min}\delta\epsilon}\log\left(\frac{|\mathcal{D}|}{\eta}\right)\right).$$

According to the above theorem, we can increase the number of samples N to decrease $T_{\max}\epsilon$ to an arbitrarily small value. However, it is crucial to acknowledge the tradeoff, as an increase in the number of samples N also amplifies the total running time. In the extreme scenario where $N^{-1/2} = \widetilde{\Theta}(\epsilon)$, Algorithm 2 achieves ϵ accuracy with $T_{\max} = \Theta(\log(1/\epsilon))$.

The proof strategy employed in Theorem 3.3 closely mirrors that of Theorem 3.2. The key difference lies in the influence of tail eigenvalues. As $T = \Omega(\alpha/\Delta)$, we observe that

$$\left|\sum_{m \in \mathcal{D}^c} p_m \exp(-(\boldsymbol{\theta} - \lambda_m)^2 T^2/2)\right| \leq \exp(-\alpha^2/2) = \mathcal{O}(\zeta^2).$$

This implies that the impact of tail eigenvalues can be arbitrarily small. By combining this observation with the proof strategy applied in Theorem 3.2, we can establish the validity of (15).

The algorithm's ability to achieve a "Constant" depth comes from the rapid decay of the filter function $F(x)$ as $|x|$ increases. A similar concept is employed in QCELS [10,

22], where solving the optimization problem is nearly equivalent to a filtering and searching process using the Gaussian filter function. It is also important to note that the choice of the Gaussian filter function might not be exclusive to achieve the "Constant" depth property. For instance, in the algorithm presented in [9], the filtering and searching process is applied using the Gaussian derivative function to approximately locate the 'zero point' of the filtered signal, which corresponds to the *ground state* energy. Exploring the possibility of extending other filter functions to MEE and achieving similar results would be an interesting avenue for further investigation.

4 Numerical experiments

In this section, we demonstrate the efficiency of QMEGS numerically by comparing it with previously established quantum phase estimation algorithms: MM-QCELS [18], QPE (textbook version [1]), and ESPRIT [14]. A brief summary of these three methods is provided in Appendix B. We share the code on Github (`https://github.com/zhiyanding/phase_estimation_methods`).

We consider three models: 1. Toy Hamiltonian with almost zero dominant spectral gap between dominant eigenvalues (Section 4.1); 2. TFIM model (Section 4.2); 3. Hubbard model (Section 4.3). In all cases, we normalize the Hamiltonian spectrum so that the eigenvalues lie within the range of $[-\pi/4, \pi/4]$ for our numerical experiments. Specifically, we use the normalized Hamiltonian in the experiment.

$$\widetilde{H} = \frac{\pi H}{4\|H\|_2} \,. \tag{16}$$

In our test, we construct an initial state $|\psi\rangle$ with $p_1 = p_2 = 0.4$. Therefore, we let λ_1, λ_2 be the dominant eigenvalues and set $\mathcal{D} = \{1, 2\}$. In addition, the dominant spectral gap is $\Delta_{\text{dom}} = \lambda_2 - \lambda_1$. To test the four algorithms, we set $T = 100 \times 2^n$ with $n = 1, 2, \cdots, 7$. For QMEGS (Algorithm 2), we choose $N = 500$, $K = 2$, $\alpha = 5$, $\sigma = 1$, $q = 0.05$. For MM-QCELS [18, Algorithm 2], the parameters are set to $K = 2, T_0 = 100, N_0 = 10^3, N_{j \geq 1} = 500$ and $\sigma = 1$. For ESPRIT (see Appendix B), we set $N = \lfloor T \rfloor$ and $K = 2$. For QPE, we consider the specific version designed for estimating the ground state energy λ_1, which is a variation of the version proposed in [1, Chapter 5.2]. Specifically, we sample its distribution $N_{\text{QPE}} = \lceil 6/p_1 \rceil$ times, and take the minimal of the measures as the approximation to λ_1 (see Appendix B). The error QPE is defined as the error for estimating this single eigenvalue.

We then use QMEGS (Algorithm 2), MM-QCELS, and ESPRIT to estimate the two dominant eigenvalues and measure the max-min error:

$$\text{error} = \max_{m \in \mathcal{D}} \min_{\sigma \in \{1, 2, \cdots, K\}^{\otimes K}} |\boldsymbol{\theta}^*_{\sigma_i} - \lambda_m| \,. \tag{17}$$

In our case, $D = \{1, 2\}, K = 2$. This testing criterion gives QPE with an advantage. Nonetheless, we illustrate that the performance of QMEGS, MM-QCELS, and ESPRIT can all to surpass that of QPE.

4.1 Almost zero dominant spectral gap

In this test, we randomly generate a Hamiltonian H

$$H = \sum_{m=1}^{M} \lambda_m |\psi_m\rangle \langle\psi_m| \,,$$

with dimension $M = 20$, $\|H\| = 1$, and $\lambda_1 - \lambda_0 = 10^{-3}$. Thus, we have $\Delta_{\text{dom}} = \lambda_1 - \lambda_0 = 10^{-3}$. We then randomly generate an initial state $|\psi\rangle$ such that $p_1 = p_2 = 0.4$.

We apply QMEGS (Algorithm 2), MM-QCELS, ESPRIT, and QPE to estimate the dominant eigenvalues (λ_1, λ_2) of the normalized Hamiltonian $\widetilde{H}$ according to (16). A comparison of the results is shown in Fig. 2.

We observe that the QPE error consistently decreases with increasing T_{max}. However, with the same T_{max}, the error for QPE is consistently higher than that of the other three methods. For QMEGS (Algorithm 2) and ESPRIT, errors do not change much with small values of T_{max} and gradually decrease when T_{max} is large enough. This behavior aligns with expectations, as for small T_{max}, QMEGS and ESPRIT struggle to distinguish between two dominant eigenvalues. When T_{max} becomes sufficiently large, these algorithms automatically differentiate between λ_1 and λ_2, leading to more accurate estimates. In contrast, the performance of MM-QCELS is different. Despite a large T_{max}, MM-QCELS struggles to differentiate between two dominant eigenvalues, resulting in the error persisting around $\Delta_m = 10^{-3}$. Additionally, when comparing the total evolution time (T_{total}), QMEGS (Algorithm 2) is shown to be more time efficient than both QPE and ESPRIT.
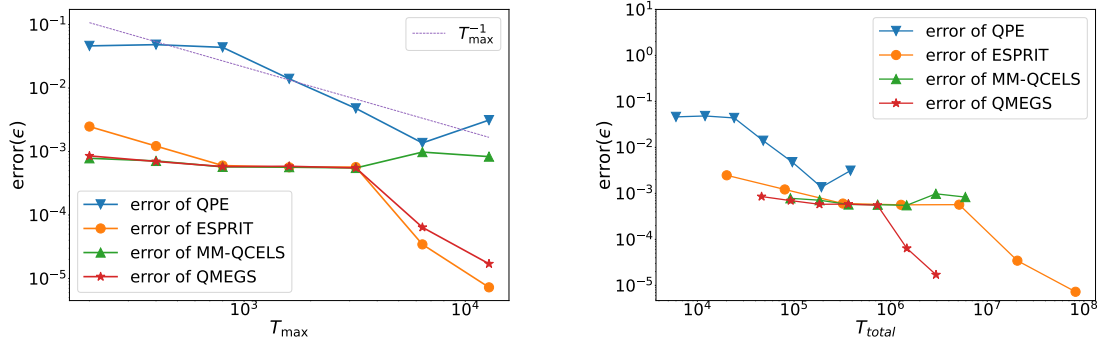

Figure 2: QMEGS (Algorithm 2) vs. MM-QCELS vs. ESPRIT vs. QPE in almost zero dominant spectral gap Hamiltonian with $p_1 = p_2 = 0.4$. We highlight that when the spectral gap is small ($\Delta_m = 10^{-3}$), MM-QCELS struggles to differentiate between two prominent eigenvalues even when T is large. In contrast, QMEGS and ESPRIT exhibit the ability to automatically distinguish between λ_1 and λ_2, achieving high accuracy as T increases.

4.2 Ising model

Consider the one-dimensional transverse field Ising model (TFIM) model defined on L sites with periodic boundary conditions:

$$H = -\left(\sum_{i=1}^{L-1} Z_i Z_{i+1} + Z_L Z_1 \right) - g \sum_{i=1}^{L} X_i \tag{18}$$

where g is the coupling coefficient, Z_i, X_i are Pauli operators for the i-th site and the dimension of H is 2^L. We choose $L = 8, g = 4$ and apply four algorithms to estimate the dominant eigenvalues (λ_1, λ_2) of the normalized Hamiltonian $\widetilde{H}$ according to (16). We note that in this case, $\Delta_{\text{dom}} = \lambda_1 - \lambda_0 \approx 0.2$. A comparison of the results is shown in Fig. 3.

We observe that the errors of QMEGS, MM-QCELS, and QPE exhibit a proportionality to the inverse of the maximal evolution time (T_{max}). In particular, the constant factor $\delta =$

$T\epsilon$ for QMEGS and MM-QCELS is considerably smaller than that of QPE. As illustrated in Fig. 3, QMEGS and MM-QCELS significantly reduce the maximal evolution time by two orders of magnitude. In contrast to the other three methods, the error of ESPRIT decreases with $T_{\max}^{-1.5\text{iv}}$, enabling it to achieve a smaller error than the other methods when $T_{\max}$ is sufficiently large. However, in terms of the total evolution time T_{total}, ESPRIT incurs a cost that is one order of magnitude higher than QMEGS.
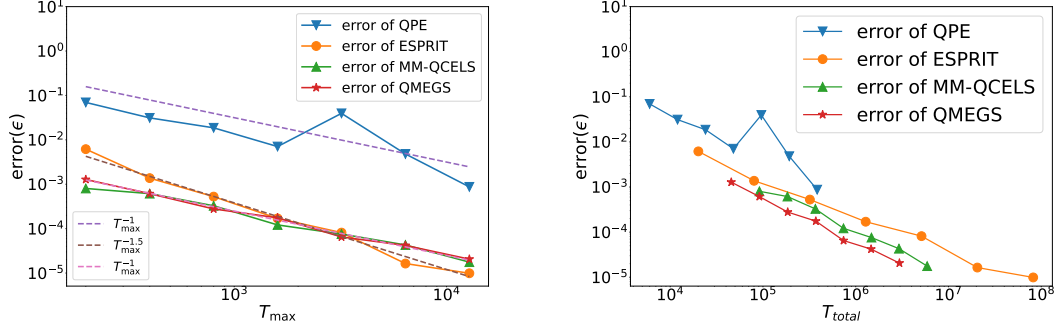

Figure 3: QMEGS (Algorithm 2) vs. MM-QCELS vs. ESPRIT vs. QPE in TFIM model with 8 sites with $p_1 = p_2 = 0.4$. Left: Depth ($T_{\max}$); Right: Cost (T_{total}). The errors of QPE, QMEGS, and MM-QCELS methods exhibit a linear scaling with $1/T_{\max}$, where the constant factor $\delta = T\epsilon$ for QMEGS and MM-QCELS is notably smaller than that of QPE. On the other hand, ESPRIT's error scales as $T_{\max}^{-1.5}$. In terms of the total evolution time T_{total}, QMEGS stands out with the smallest T_{total} compared to other methods and is shown to be one order of magnitude more cost-effective than ESPRIT.

4.3 Hubbard model

Consider the one-dimensional Hubbard model defined on L spinful sites with open boundary conditions:

$$H = -t \sum_{j=1}^{L-1} \sum_{\sigma \in \{\uparrow, \downarrow\}} c_{j,\sigma}^\dagger c_{j+1,\sigma} + U \sum_{j=1}^{L} \left(n_{j,\uparrow} - \frac{1}{2} \right) \left(n_{j,\downarrow} - \frac{1}{2} \right).$$

Here $c_{j,\sigma}(c_{j,\sigma}^\dagger)$ denotes the fermionic annihilation (creation) operator on the site j with spin σ, $n_{j,\sigma} = c_{j,\sigma}^\dagger c_{j,\sigma}$ denotes the number operator for $\sigma \in \{\uparrow, \downarrow\}$, and $t, U \in \mathbb{R}$ are parameters. We choose $L = 4, 8$, $t = 1$, $U = 10$ and test four algorithms to estimate λ_1, λ_0. We note that the dominant spectral gap $\Delta_{\text{dom}} = \lambda_1 - \lambda_0 \approx 0.02$.

The numerical results are presented in Fig. 4. Similar to the TFIM model, the maximal evolution time of QMEGS and MM-QCELS is nearly two orders of magnitude smaller than that of QPE, while the error of ESPRIT scales as $T_{\max}^{-1.5}$. Additionally, the overall computational cost of QMEGS is lower than that of other methods and is almost one order of magnitude lower than the cost of ESPRIT.

5 Discussions

In this paper, we introduce a novel algorithm, Quantum Multiple Eigenvalue Gaussian filtered Search (QMEGS), designed for addressing the multiple eigenvalues estimation prob-

---

iv A very recent work [36] rigorously proves this based on a highly nontrivial matrix perturbation argument. Specifically, when $T_{\max} \gg \Delta_m$, the authors show that the error of ESPRIT scales as $T_{\max}^{-1.5}$.
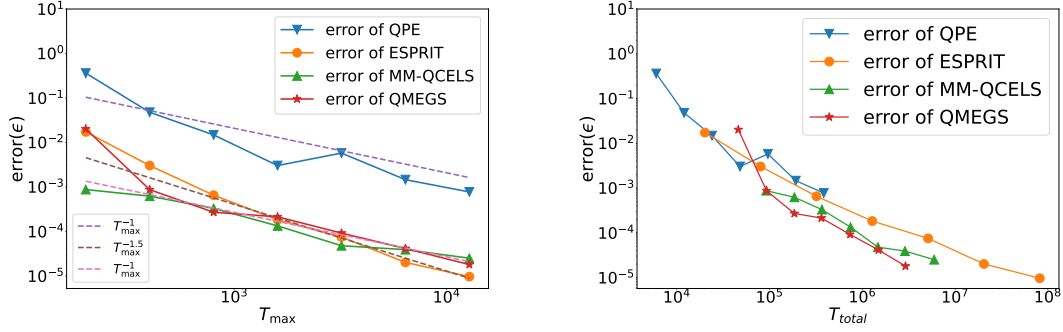
Figure 4: QMEGS (Algorithm 2) vs. MM-QCELS vs. ESPRIT vs. QPE in TFIM model with 8 sites with $p_1 = p_2 = 0.4$. Left: Depth ($T_{\max}$); Right: Cost (T_{total}). QMEGS and MM-QCELS methods exhibit significantly smaller $T_{\max}$ values compared to QPE. Moreover, QMEGS achieves the smallest total evolution time and is one order of magnitude more cost-effective than ESPRIT.

lem. In contrast to preceding algorithms for multiple eigenvalue estimation, QMEGS is the first algorithm capable not only of consistently achieving Heisenberg-limited scaling but also of attaining "short" and "constant" depth in the presence of gaps between eigenvalues. Additionally, QMEGS offers flexibility in parameter selection. It is not imperative to meticulously set the parameters based on prior knowledge of (p_m, λ_m) for the algorithm to operate effectively. Numerous numerical results substantiate the efficiency and flexibility of the algorithm.

There are several directions to extend this work.

- While QMEGS has small quantum computational complexity, the classical cost of the algorithm is linearly dependent on the inverse of the precision ϵ^{-1} and the size of the domain (assuming $\lambda_m \subset [-\pi, \pi]$). Exploring more efficient search algorithms is an intriguing direction to reduce this linear dependence to logarithmic potentially.

- QMEGS demonstrates flexibility in parameter selection, achieving accurate estimations when $\alpha, \sigma, T, N, q^{-1}$ are chosen to be sufficiently large. The challenge of adaptively selecting these parameters to achieve optimal complexity without prior knowledge of (p_m, λ_m) remains an open problem.

- In this study, we assume the exactness in the Hamiltonian simulation $\exp(-iHt)$ for simplicity. However, this assumption may not hold in the context of early fault-tolerant quantum computers. In cases where the Hamiltonian simulation is not exact, Z_n becomes a biased estimate for $\langle \psi | \exp(-it_n H) | \psi \rangle$. Investigating the impact of different noise models on QMEGS performance is an intriguing problem, and the results depend on the noise models chosen and the assumptions about Z_n. For example, when the bias size is independent of t and sufficiently small, such as in the case of small additive noise, a similar analysis to the proof of Theorem 3.1 (Appendix A) can be employed. This allows us to absorb the error into statistical error and non-dominant parts, demonstrating that the method still achieves Heisenberg-limited scaling.

- It is important to emphasize that while ESPRIT-type algorithms [14, 20] may not achieve Heisenberg-limited scaling and theoretically require gap assumptions, these algorithms still perform effectively in the absence of gaps between dominant eigenvalues and show rapid error decay with $T_{\max}$. In the presence of a gap, numerical

results indicate that the decay rate of the error appears to be $\mathcal{O}(T_{\max}^{-1.5})$. To our knowledge this phenomenon was not previously reported in the literature. The theoretical underpinning of these phenomena remains an open question, which makes it an intriguing direction for further exploration into the complexity of ESPRIT-type algorithms.

- As highlighted in Section 2.1, the Gaussian filtering function is not the only option for addressing the MEE problem. Some other filtering functions, such as the Gaussian derivative filter [9] and derivative Heaviside function [12], might also be used to solve the MEE problem. The exploration of alternative filtering functions that offer improved complexity performance and flexibility remains an open problem.

References

[1] Michael A Nielsen and Isaac Chuang. "Quantum computation and quantum information". Cambridge Univ. Pr. (2000).

[2] Daniel S Abrams and Seth Lloyd. "Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors". Phys. Rev. Lett. **83**, 5162 (1999).

[3] Yimin Ge, Jordi Tura, and J Ignacio Cirac. "Faster ground state preparation and high-precision ground energy estimation with fewer qubits". J. Math. Phys. **60**, 022202 (2019).

[4] Lin Lin and Yu Tong. "Near-optimal ground state preparation". Quantum **4**, 372 (2020).

[5] Lin Lin and Yu Tong. "Heisenberg-limited ground state energy estimation for early fault-tolerant quantum computers". PRX Quantum **3**, 010318 (2022).

[6] Yulong Dong, Lin Lin, and Yu Tong. "Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices". PRX Quantum **3**, 040305 (2022).

[7] Kianna Wan, Mario Berta, and Earl T Campbell. "Randomized quantum algorithm for statistical phase estimation". Phys. Rev. Lett. **129**, 030503 (2022).

[8] Ruizhe Zhang, Guoming Wang, and Peter Johnson. "Computing ground state properties with early fault-tolerant quantum computers". Quantum **6**, 761 (2022).

[9] Guoming Wang, Daniel Stilck França, Ruizhe Zhang, Shuchen Zhu, and Peter D Johnson. "Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision". Quantum **7**, 1167 (2023).

[10] Zhiyan Ding and Lin 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).

[11] Hongkang Ni, Haoya Li, and Lexing Ying. "On low-depth algorithms for quantum phase estimation". Quantum (2023).

[12] Nick S. Blunt, Laura Caune, Róbert Izsák, Earl T. Campbell, and Nicole Holzmann. "Statistical phase estimation and error mitigation on a superconducting quantum processor". PRX Quantum **4**, 040341 (2023).

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

[14] M E Stroeks, J Helsen, and B M Terhal. "Spectral estimation for hamiltonians: a comparison between classical imaginary-time evolution and quantum real-time evolution". New J. Phys. **24**, 103024 (2022).

[15] Cristian L. Cortes and Stephen K. Gray. "Quantum Krylov subspace algorithms for ground- and excited-state energy estimation". Phys. Rev. A **105**, 022417 (2022).

[16] Ethan N. Epperly, Lin Lin, and Yuji Nakatsukasa. "A theory of quantum subspace diagonalization". SIAM J. Matrix Anal. Appl. **43**, 1263–1290 (2022).

[17] Alicja Dutkiewicz, Barbara M. Terhal, and Thomas E. O'Brien. "Heisenberg-limited quantum phase estimation of multiple eigenvalues with few control qubits". Quantum **6**, 830 (2022).

[18] Zhiyan Ding and Lin Lin. "Simultaneous estimation of multiple eigenvalues with short-depth quantum circuit on early fault-tolerant quantum computers". Quantum **7**, 1136 (2023).

[19] Haoya Li, Hongkang Ni, and Lexing Ying. "Adaptive low-depth quantum algorithms for robust multiple-phase estimation". Phys. Rev. A **108**, 062408 (2023).

[20] Yizhi Shen, Daan Camps, Aaron Szasz, Siva Darbha, Katherine Klymko, David B. Williams-Young, Norm M. Tubman, and Roel Van Beeumen. "Estimating eigenenergies from quantum dynamics: A unified noise-resilient measurement-driven approach". In 2023 QCE. Pages 302–303. (2023).

[21] Amara Katabarwa, Katerina Gratsea, Athena Caesura, and Peter D. Johnson. "Early fault-tolerant quantum computing". PRX Quantum **5**, 020101 (2024).

[22] Zhiyan Ding, Yulong Dong, Yu Tong, and Lin Lin. "Robust ground-state energy estimation under depolarizing noise" (2023).

[23] Haoya Li, Hongkang Ni, and Lexing Ying. "A note on spike localization for line spectrum estimation". Appl. Comput. Harmon. Anal. **67**, 101577 (2023).

[24] B L Higgins, D W Berry, S D Bartlett, M W Mitchell, H M Wiseman, and G J Pryde. "Demonstrating heisenberg-limited unambiguous phase estimation without adaptive measurements". New J. Phys. **11**, 073023 (2009).

[25] Shelby Kimmel, Guang Hao Low, and Theodore J. Yoder. "Robust calibration of a universal single-qubit gate set via robust phase estimation". Phys. Rev. A **92**, 062315 (2015).

[26] William Huggins, Joonho Lee, Unpil Baek, Bryan O'Gorman, and K Whaley. "A non-orthogonal variational quantum eigensolver". New J. Phys.**22** (2020).

[27] Katherine Klymko, Carlos Mejuto-Zaera, Stephen J. Cotton, Filip Wudarski, Miroslav Urbanek, Diptarka Hait, Martin Head-Gordon, K. Birgitta Whaley, Jonathan Moussa, Nathan Wiebe, Wibe A. de Jong, and Norm M. Tubman. "Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware". PRX Quantum **3**, 020323 (2022).

[28] Jarrod R. McClean, Mollie E. Kimchi-Schwartz, Jonathan Carter, and Wibe A. de Jong. "Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states". Phys. Rev. A **95**, 042308 (2017).

[29] Mario Motta, Chong Sun, Adrian Tan, Matthew O'Rourke, Erika Ye, Austin Minnich, Fernando Brandão, and Garnet Chan. "Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution". Nat. Phys **16**, 1–6 (2020).

[30] Robert M. Parrish and Peter Leonard McMahon. "Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation". preprint (2019).

[31] Kazuhiro Seki and Seiji Yunoki. "Quantum power method by a superposition of time-evolved states". PRX Quantum **2**, 010333 (2021).

[32] Nicholas H. Stair, Renke Huang, and Francesco A. Evangelista. "A multireference quantum Krylov algorithm for strongly correlated electrons". J. Chem. Theory Comput. **16**, 2236–2245 (2020).

[33] Thomas E O'Brien, Brian Tarasinski, and Barbara M Terhal. "Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments". New J. Phys. **21**, 023022 (2019).

[34] Richard Roy and Thomas Kailath. "Esprit-estimation of signal parameters via rotational invariance techniques". IEEE Trans. Acoust. **37**, 984–995 (1989).

[35] Weilin Li, Wenjing Liao, and Albert Fannjiang. "Super-resolution limit of the esprit algorithm". IEEE Trans. Inf. Theory **66**, 4593–4608 (2020).

[36] Zhiyan Ding, Ethan N. Epperly, Lin Lin, and Ruizhe Zhang. "The ESPRIT algorithm under high noise: Optimal error scaling and noisy super-resolution". preprint (2024).

[37] Dominic W. Berry, Yuan Su, Casper Gyurik, Robbie King, Joao Basso, Alexander Del Toro Barba, Abhishek Rajput, Nathan Wiebe, Vedran Dunjko, and Ryan Babbush. "Analyzing prospects for quantum advantage in topological data analysis". PRX Quantum **5**, 010319 (2024).

[38] Sam McArdle, András Gilyén, and Mario Berta. "Quantum state preparation without coherent arithmetic". preprint (2022).

[39] Samuel L. Braunstein, Carlton M. Caves, and Gerard J. Milburn. "Generalized uncertainty relations: Theory, examples, and lorentz invariance". Ann. Phys. **247**, 135–173 (1996).

[40] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. "Quantum metrology". Phys. Rev. Lett. **96**, 010401 (2006).

A   Proofs

In this section, we prove Theorems 3.1 to 3.3. We first introduce a lemma that helps us bound the error caused by time truncation and random noise. Define the error term

$$
E_j = |E(\theta_j)|, \quad E(\theta) =: \frac{1}{N} \sum_{n=1}^{N} Z_n \exp(i\theta t_n) - \sum_{m=1}^{M} p_m \exp\left(-\frac{T^2(\lambda_m - \theta)^2}{2}\right),
$$

where θ_j is defined in Algorithm 2. We note that E_j contains errors caused by time truncation and finite samples of t_n. We demonstrate the smallness of E_j in the following lemma:

Lemma A.1. *Given $\delta > 0$ and the pair of overlap-eigenvalues $\{(p_i, \lambda_i)\}_{m=1}^M$, if $\sigma = \Omega\left(\log^{1/2}(1/\delta)\right)$, we have*

$$\left| \mathbb{E}\left(\frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta t_n) \right) - \sum_{m=1}^M p_m \exp\left(-\frac{T^2(\lambda_m - \theta)^2}{2} \right) \right| \leq \delta^2. \tag{19}$$

Furthermore, define $E(\theta) = \frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta t_n) - \sum_{m=1}^M p_m \exp\left(-\frac{T^2(\lambda_m - \theta)^2}{2} \right)$. Given $\eta > 0$ and $\Theta = \{\theta_j\}_{j=1}^J \cup \{\lambda_m\}_{m \in \mathcal{D}}$, if $N = \Omega\left(\frac{1}{\delta^2} \log\left((J + |\mathcal{D}|) \frac{1}{\eta} \right) \right)$, we have

$$\mathbb{P}\left(\max_{\theta \in \Theta} |E(\theta)| \leq \delta \right) \geq 1 - \eta, \tag{20}$$

and

$$\mathbb{P}\left(\cap_{\theta, \theta' \in \Theta, \theta \neq \theta'} \left\{ |E(\theta) - E(\theta')| \leq \sigma T \delta |\theta - \theta'| + \delta^2 \right\} \right) \geq 1 - \eta. \tag{21}$$

Proof. According to Algorithm 1, when $\sigma = \Omega\left(\log^{1/2}(1/\delta)\right)$, we have

$$\left| \mathbb{E}\left(\frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta t_n) \right) - \sum_{m=1}^M p_m \exp\left(-\frac{T^2(\lambda_m - \theta)^2}{2} \right) \right|$$

$$\leq \int_{|t| > \sigma T} \frac{1}{\sqrt{2\pi}T} e^{-\frac{t^2}{2T^2}} \left| \sum_{m=1}^M p_m e^{-it\lambda_m} e^{i\theta t} \right| \mathrm{d}t = 2 \int_{\sigma T}^\infty \frac{1}{\sqrt{2\pi}T} e^{-\frac{t^2}{2T^2}} \mathrm{d}t \tag{22}$$

$$= \sqrt{\frac{2}{\pi}} \int_\sigma^\infty e^{-\frac{s^2}{2}} \mathrm{d}s < e^{-\sigma^2} < \delta^2,$$

where the inequality $\int_x^\infty e^{-\frac{1}{2}t^2} \mathrm{d}t \leq \sqrt{\frac{\pi}{2}} e^{-\frac{x^2}{2}}$ is used in the second last inequality. This proves (19). The proof of (20) and (21) is the same as the proof of [22, Appendix B.3 Lemma 4 eqn. (B8)], thus, we omit it here. $\qquad\square$

Recall that

$$G(\theta) = \left| \frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta t_n) \right|.$$

We are ready to prove Theorem 3.1.

Proof of Theorem 3.1. According to Lemma A.1 (20) with $\delta = \mathcal{O}(p_{\min} - p_{\text{tail}})$ and $J = \mathcal{O}(T/q)$, we obtain

$$\mathbb{P}\left(\max_{\theta \in \Theta} |E(\theta)| < \frac{p_{\min} - p_{\text{tail}}}{8} \right) \geq 1 - \eta. \tag{23}$$

Thus, to prove Theorem 3.1, it suffices to show (12) assuming $E_j < \frac{p_{\min} - p_{\text{tail}}}{8}$. Under this condition, we consider two classes of candidates:

- When $\theta_j \in \cup_{m \in \mathcal{D}} \left[\lambda_m - \frac{q}{T}, \lambda_m + \frac{q}{T} \right]$, we obtain

$$G_j = G(\theta_j) \geq \left| \sum_{m=1}^M p_m \exp\left(-\frac{T^2(\lambda_m - \theta_j)^2}{2} \right) \right| - E_j$$

$$> p_{\min} \exp\left(-\frac{q^2}{2} \right) - \frac{p_{\min} - p_{\text{tail}}}{8}$$

$$\geq p_{\text{tail}} + \frac{p_{\min} - p_{\text{tail}}}{2} - \frac{p_{\min} - p_{\text{tail}}}{8} \geq p_{\text{tail}} + \frac{3(p_{\min} - p_{\text{tail}})}{8}.$$

where we use $E_j < \frac{p_{\min} - p_{\text{tail}}}{8}$ in the second inequality, and the condition of q in the last inequality.

- When $\theta_j \notin \cup_{m \in \mathcal{D}} \left[\lambda_m - \frac{\alpha}{3T}, \lambda_m + \frac{\alpha}{3T}\right]$, we obtain

$$
\begin{aligned}
G_j &\leq \left| \sum_{m \in \mathcal{D}} p_m \exp\left(-\frac{T^2(\lambda_m - \theta_j)^2}{2}\right) \right| + \left| \sum_{m \in \mathcal{D}^c} p_m \exp\left(-\frac{T^2(\lambda_m - \theta_j)^2}{2}\right) \right| + E_j \\
&\leq \exp\left(-\frac{\alpha^2}{18}\right) + p_{\text{tail}} + \frac{p_{\min} - p_{\text{tail}}}{8} \leq p_{\text{tail}} + \frac{p_{\min} - p_{\text{tail}}}{4},
\end{aligned}
$$

where $\mathcal{D}^c := [M] \backslash \mathcal{D}$, and we use $E_j < \frac{p_{\min} - p_{\text{tail}}}{8}$ in the second inequality and the condition of α in the last inequality.

Because $q < \alpha/3$, the above two inequalities imply

$$
\max_{\theta_j \notin \cup_{m \in \mathcal{D}}\left[\lambda_m - \frac{\alpha}{3T}, \lambda_m + \frac{\alpha}{3T}\right]} G_j < \min_{\theta_j \in \cup_{m \in \mathcal{D}}\left[\lambda_m - \frac{q}{T}, \lambda_m + \frac{q}{T}\right]} G_j. \tag{24}
$$

Now, we prove (12) using the proof by contradiction argument. Assuming that there exists $m^\star \in \mathcal{D}$ such that

$$
\lambda_{m^\star} \notin \cup_{k=1}^K \left[\boldsymbol{\theta}_k - \frac{\alpha}{T}, \boldsymbol{\theta}_k + \frac{\alpha}{T}\right],
$$

then the grid point $\boldsymbol{\theta}^*$ closest to $\lambda_{m^\star}$ is not in the block set since $\alpha/q \in \mathbb{N}$ and the block set is a union of open intervals. According to (24), we must have $\boldsymbol{\theta}_k \in \cup_{m \in \mathcal{D}, m \neq m^\star}\left[\lambda_m - \frac{\alpha}{3T}, \lambda_m + \frac{\alpha}{3T}\right]$ for all $1 \leq k \leq |\mathcal{D}|$. By the pigeonhole principle, there exist $k_1 < k_2$ and $m' \in \mathcal{D}$ such that $\{\boldsymbol{\theta}_{k_1}, \boldsymbol{\theta}_{k_2}\} \subset \left[\lambda_{m'} - \frac{\alpha}{3T}, \lambda_{m'} + \frac{\alpha}{3T}\right]$. Thus, $|\boldsymbol{\theta}_{k_1} - \boldsymbol{\theta}_{k_2}| \leq \frac{2\alpha}{3T}$. However, this contradicts the blocking process that ensures

$$
\boldsymbol{\theta}_{k_2} \notin \left(\boldsymbol{\theta}_{k_1} - \frac{\alpha}{T}, \boldsymbol{\theta}_{k_1} + \frac{\alpha}{T}\right).
$$

This concludes the proof. $\qquad\square$

Next, we give the proof of Theorem 3.2. Define the tail term that contains the impact of non-dominant eigenvalues:

$$
G_{\text{tail}}(\theta) = \sum_{m \in \mathcal{D}^c} p_m \exp\left(-\frac{T^2(\lambda_m - \theta)^2}{2}\right). \tag{25}
$$

The proof of Theorem 3.2 is as follows.

Proof of Theorem 3.2. Because $p_{\min} > p_{\text{tail}}$, it is straightforward to see that the condition of Theorem 3.1 is satisfied. In addition, using the condition of N with Lemma A.1 by setting $\delta = \zeta p_{\text{tail}}/10$ and $J = \mathcal{O}(T/q)$, we obtain

$$
\mathbb{P}\left(\max_{\theta \in \Theta} |E(\theta)| \leq \zeta p_{\text{tail}}/10\right) \geq 1 - \eta/2, \tag{26}
$$

and

$$
\mathbb{P}\left(\cap_{\theta, \theta' \in \Theta, \theta \neq \theta'} \left\{|E(\theta) - E(\theta')| = \sigma T \zeta p_{\text{tail}} |\theta - \theta'|/10 + \zeta^2 p_{\text{tail}}^2/100\right\}\right) \geq 1 - \eta/2. \tag{27}
$$

where $\Theta = \{\theta_j\}_{j=1}^J \cup \{\lambda_m\}_{m \in \mathcal{D}}$.

To prove the theorem, it suffices to show that when (26) and (27) hold, (14) holds. Fixed $1 \leq k \leq |\mathcal{D}|$. Define $\mathcal{D}_{k_-} := \{m \in \mathcal{D} : \lambda_m \in \mathcal{I}_{B,k}\}$, where $\mathcal{I}_{B,k}$ is the block interval defined in the algorithm. We note that $\mathcal{D}_{k_-}$ is the set of indices that have already been covered by the blocked set. Because $T = \Omega\left(\frac{\alpha}{\Delta_{\text{dom}}}\right)$, we find that each interval $[\boldsymbol{\theta}_j - \frac{\alpha}{T}, \boldsymbol{\theta}_j + \frac{\alpha}{T}]$ contains exactly one dominant eigenvalue for all $1 \leq j \leq k-1$ and $\mathcal{D} \backslash \mathcal{D}_{k_-} \neq \emptyset$. Then, to prove (14), it suffices to show that, if $\left|\boldsymbol{\theta}_j - \lambda_{m_j}\right| \leq \frac{Q}{T}$ for $j < k^{\text{v}}$, then there exists $m_k \in \mathcal{D} \setminus \mathcal{D}_{k_-}$ such that

$$|\boldsymbol{\theta}_k - \lambda_{m_k}| \leq \frac{Q}{T}. \tag{28}$$

The proof of (28) consists of two main steps. In the first step, we directly control the error of the filter function to establish a loose bound $T|\lambda_{m_k} - \boldsymbol{\theta}_k| = \mathcal{O}(\zeta^{-1/2})$. When $p_{\min} \gg p_{\text{tail}}$, this results in $T|\lambda_{m_k} - \boldsymbol{\theta}_k| = \mathcal{O}((p_{\text{tail}}/p_{\min})^{1/2})$, a bound that is weaker than the one in (28). This weaker bound allows the confinement of $\boldsymbol{\theta}_{m_k}$ to a narrow region around λ_i. Then we employ a Taylor expansion of the filter function within this restricted region and refine the bound to $T|\lambda_i - \boldsymbol{\theta}_{k_i}| = \mathcal{O}(p_{\text{tail}}/p_{\min})$.

Define $m_k^\star = \operatorname{argmin}_{m \in \mathcal{D}} |\boldsymbol{\theta}_k - \lambda_m|$, $\lambda_k^\star = \lambda_{m_k^\star}$, $p_k^\star = p_{m_k^\star}$, and $\boldsymbol{\theta}_k^\star = \operatorname{argmin}_{\theta_j} \left|\theta_j - \lambda_{m_k^\star}\right|$. It is straightforward to see $\lambda_k^\star \notin \mathcal{I}_{B,k}$ because each interval $[\boldsymbol{\theta}_j - \frac{\alpha}{T}, \boldsymbol{\theta}_j + \frac{\alpha}{T}]$ contains exactly one dominant eigenvalue for every $j \in [k-1]$. Next, because the candidates θ_j are chosen with the step size q/T, we have

$$|\boldsymbol{\theta}_k^\star - \lambda_k^\star| \leq \frac{q}{T}.$$

Furthermore, because $\alpha/q \in \mathbb{N}$, and $\lambda_k^\star$ is not covered by the blocked set, we must have $\boldsymbol{\theta}_k^\star \notin \mathcal{I}_{B,k}$. In addition, we have $G(\boldsymbol{\theta}_k^\star)$ is close to $G(\lambda_k^\star)$, as given by

$$
\begin{aligned}
&|G(\boldsymbol{\theta}_k^\star) - G(\lambda_k^\star)| \\
&\leq \left| p_k^\star \left(1 - \exp\left(-\frac{T^2 (\lambda_k^\star - \boldsymbol{\theta}_k^\star)^2}{2} \right) \right) \right| \\
&\quad + \left| \sum_{m \in \mathcal{D} \setminus \{m_k^\star\}} p_m \exp\left(-\frac{T^2 (\lambda_m - \lambda_k^\star)^2}{2} \right) - p_m \exp\left(-\frac{T^2 (\lambda_m - \boldsymbol{\theta}_k^\star)^2}{2} \right) \right| \\
&\quad + |G_{\text{tail}}(\boldsymbol{\theta}_k^\star) - G_{\text{tail}}(\lambda_k^\star)| + |E(\boldsymbol{\theta}_k^\star) - E(\lambda_k^\star)|,
\end{aligned} \tag{29}
$$

where G_{tail} is defined in (25). For the first term in (29), since $|\boldsymbol{\theta}_k^\star - \lambda_k^\star| \leq \frac{q}{T}$, we have

$$\left| p_k^\star \left(1 - \exp\left(-\frac{T^2 (\lambda_k^\star - \boldsymbol{\theta}_k^\star)^2}{2} \right) \right) \right| = \mathcal{O}(q^2).$$

For the second term in (29), since $T = \Omega(\alpha/\Delta_{\text{dom}})$ and $|\boldsymbol{\theta}_k^\star - \lambda_m| \geq \frac{\Delta_{\text{dom}}}{2}$, we have

$$\left| \sum_{m \in \mathcal{D} \setminus \{m_k^\star\}} p_m \exp\left(-\frac{T^2 (\lambda_m - \lambda_k^\star)^2}{2} \right) - p_m \exp\left(-\frac{T^2 (\lambda_m - \boldsymbol{\theta}_k^\star)^2}{2} \right) \right| = \mathcal{O}(\exp(-\Theta(\alpha^2))).$$

For the third term in (29), by the $p_{\text{tail}}T$-Lipschitz property of $G_{\text{tail}}(\theta)$ and $|\boldsymbol{\theta}_k^\star - \lambda_k^\star| \leq \frac{q}{T}$, we have

$$|G_{\text{tail}}(\boldsymbol{\theta}_k^\star) - G_{\text{tail}}(\lambda_k^\star)| = \mathcal{O}(p_{\text{tail}}q).$$

$^{\text{v}}$When $k = 1$, we don't need this condition.

For the last term in (29), by (27) and $|\boldsymbol{\theta}_k^\star - \lambda_k^\star| \leq \frac{q}{T}$, we have

$$|E(\boldsymbol{\theta}_k^\star) - E(\lambda_k^\star)| \leq \zeta p_{\text{tail}}(\sigma q/10 + \zeta p_{\text{tail}}/100).$$

Combining them, we get that

$$|G(\boldsymbol{\theta}_k^\star) - G(\lambda_k^\star)| \leq \frac{\zeta p_{\text{tail}}^2}{10}, \tag{30}$$

which follows by employing employ (13) to control the parameters σ, α, q.

Now we show that $\boldsymbol{\theta}_k$ is close to $\lambda_k^\star$ using $\boldsymbol{\theta}_k^\star$ as a bridge. First, we notice

$$|\boldsymbol{\theta}_k - \lambda_j| \geq \frac{\Delta_{\text{dom}}}{2}, \quad \forall j \in \mathcal{D} \setminus \{m_k^\star\}.$$

Combining this with the condition that $T = \Omega\left(\log^{1/2}(1/(\zeta p_{\text{tail}}))\Delta_{\text{dom}}^{-1}\right)$, we have

$$\sum_{m \in \mathcal{D} \setminus \{m_k^\star\}} p_m \exp\left(-\frac{T^2 \left(\lambda_m - \lambda_k^\star\right)^2}{2}\right) \leq \frac{\zeta p_{\text{tail}}^2}{10}, \qquad \sum_{m \in \mathcal{D} \setminus \{m_k^\star\}} p_m \exp\left(-\frac{T^2 \left(\lambda_m - \boldsymbol{\theta}_k\right)^2}{2}\right) \leq \frac{\zeta p_{\text{tail}}^2}{10}. \tag{31}$$

Next, we use $|G_{\text{tail}}| \leq p_{\text{tail}}$, (26), and (31) to obtain

$$\begin{aligned} G\left(\boldsymbol{\theta}_k\right) &\leq p_k^\star \exp\left(-T^2 \left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2\right) + p_{\text{tail}} + \frac{\zeta p_{\text{tail}} + \zeta p_{\text{tail}}^2}{10} \\ &\leq p_k^\star \exp\left(-T^2 \left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2\right) + p_{\text{tail}} + \frac{\zeta p_{\text{tail}}}{5} \end{aligned} \tag{32}$$

and

$$G\left(\lambda_k^\star\right) \geq p_k^\star + \sum_{m \neq \{m_k^\star\}} p_m \exp\left(-\frac{T^2 \left(\lambda_m - \lambda_k^\star\right)^2}{2}\right) - |E(\lambda_k^\star)| \geq p_k^\star - \frac{\zeta p_{\text{tail}}}{10}. \tag{33}$$

Furthermore, (30) and the assumption that $\boldsymbol{\theta}_k$ is the maximal point imply that

$$G\left(\boldsymbol{\theta}_k\right) \geq G\left(\boldsymbol{\theta}_k^\star\right) \geq G\left(\lambda_k^\star\right) - \frac{\zeta p_{\text{tail}}}{10}. \tag{34}$$

Combining (32), (33), and (34), we obtain

$$p_k^\star \left(1 - \exp\left(-T^2 \left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2\right)\right) \leq (1 + 2\zeta/5) p_{\text{tail}},$$

which implies

$$|T(\lambda_k^\star - \boldsymbol{\theta}_k)| \leq \sqrt{-2\ln\left(1 - \frac{(1 + 2\zeta/5) p_{\text{tail}}}{p_{\min}}\right)} = \mathcal{O}\left(\zeta^{-1/2}\right). \tag{35}$$

We note that, in the case where $\zeta^{-1} = \mathcal{O}\left(p_{\text{tail}}/p_{\min}\right)$, the upper bound proved above scales with $(p_{\text{tail}}/p_{\min})^{1/2}$, while the goal is to improve the upper bound to Q, which linearly depends on $p_{\text{tail}}/p_{\min}$. Now we show how to improve (35). Similar to (32) and (33),

$$\begin{aligned} &\left|G^2\left(\boldsymbol{\theta}_k\right) - (p_k^\star)^2 \exp\left(-T^2 \left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2\right)\right. \\ &\left. -2p_k^\star \text{Re}\left(\exp\left(-T^2 \left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2\right)(G_{\text{tail}}(\boldsymbol{\theta}_k) + E(\boldsymbol{\theta}_k))\right)\right| = \mathcal{O}(p_{\text{tail}}^2) \end{aligned} \tag{36}$$

and

$$\left| G^2\left(\lambda_k^\star\right) - \left[(p_k^\star)^2 + 2p_k^\star \mathrm{Re}\left((G_{\mathrm{tail}}(\lambda_k^\star) + E(\lambda_k^\star)) \right) \right] \right| = \mathcal{O}(p_{\mathrm{tail}}^2). \tag{37}$$

Furthermore, (30) and the assumption that $\boldsymbol{\theta}_k$ is the maximal point imply that

$$G^2\left(\boldsymbol{\theta}_k\right) \geq G^2\left(\boldsymbol{\theta}_k^\star\right) \geq G^2\left(\lambda_k^\star\right) - \mathcal{O}(p_{\mathrm{tail}}^2). \tag{38}$$

Combining (36), (37), and (38), we obtain

$$p_k^\star \left(\exp\left(-T^2 \left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2 \right) - 1 \right)$$
$$\geq 2\mathrm{Re}\left(-\exp\left(-T^2\left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2 \right) (G_{\mathrm{tail}}(\boldsymbol{\theta}_k) + E(\boldsymbol{\theta}_k)) + (G_{\mathrm{tail}}(\lambda_k^\star) + E(\lambda_k^\star)) \right) - O(p_{\mathrm{tail}}^2/p_{\min})$$
$$\geq 2\mathrm{Re}\left(\left(1 - \exp\left(-T^2\left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2 \right) \right) (G_{\mathrm{tail}}(\boldsymbol{\theta}_k) + E(\boldsymbol{\theta}_k)) \right)$$
$$\quad + 2\mathrm{Re}\left(-(G_{\mathrm{tail}}(\boldsymbol{\theta}_k) + E(\boldsymbol{\theta}_k)) + (G_{\mathrm{tail}}(\lambda_k^\star) + E(\lambda_k^\star)) \right) - O(p_{\mathrm{tail}}^2/p_{\min})$$
$$\geq 2\mathrm{Re}\left(\left(1 - \exp\left(-T^2\left(\lambda_k^\star - \boldsymbol{\theta}_k\right)^2/2 \right) \right) E(\boldsymbol{\theta}_k) \right)$$
$$\quad + 2\mathrm{Re}\left(-(G_{\mathrm{tail}}(\boldsymbol{\theta}_k) + E(\boldsymbol{\theta}_k)) + (G_{\mathrm{tail}}(\lambda_k^\star) + E(\lambda_k^\star)) \right) - O(p_{\mathrm{tail}}^2/p_{\min}).$$

Define $F_k(\theta) = \exp\left(-\frac{T^2\left(\lambda_k^\star - \theta\right)^2}{2} \right)$[vi]. Then, the above inequality can be rewritten as

$$p_k^\star(1 - F_k(\boldsymbol{\theta}_k))(1 + F_k(\boldsymbol{\theta}_k)) + 2\mathrm{Re}\left((1 - F_k(\boldsymbol{\theta}_k))E(\boldsymbol{\theta}_k) \right)$$
$$\leq 2\left| G_{\mathrm{tail}}(\boldsymbol{\theta}_k) - G_{\mathrm{tail}}(\lambda_k^\star) \right| + 2\left| E(\boldsymbol{\theta}_k) - E(\lambda_k^\star) \right| + \mathcal{O}(p_{\mathrm{tail}}^2/p_{\min}).$$

Because $|E(\boldsymbol{\theta}_k)| \leq \zeta p_{\mathrm{tail}}/10$ according to (26), the above inequality further implies

$$\left(p_{\min} - \zeta p_{\mathrm{tail}}/5 \right) (1 - F_k(\boldsymbol{\theta}_k))$$
$$\leq 2\left| G_{\mathrm{tail}}(\boldsymbol{\theta}_k) - G_{\mathrm{tail}}(\lambda_k^\star) \right| + 2\left| E(\boldsymbol{\theta}_k) - E(\lambda_k^\star) \right| + \mathcal{O}(p_{\mathrm{tail}}^2/p_{\min}) \tag{39}$$
$$\leq 2(\sigma\zeta + 1)T p_{\mathrm{tail}} \left| \boldsymbol{\theta}_k - \lambda_k^\star \right| + \mathcal{O}\left(p_{\mathrm{tail}}^2/p_{\min} \right).$$

where we use the fact that $F_k(\boldsymbol{\theta}_k) \in [0, 1]$ in the first inequality, and use (27) and the fact that G_{tail} is $p_{\mathrm{tail}}T$-Lipschitz in the last inequality.

Using (35), we obtain

$$1 - F_k(\boldsymbol{\theta}_k) = 1 - \exp\left(-\frac{T^2\left(\boldsymbol{\theta}_k - \lambda_k^\star\right)^2}{2} \right) \geq \frac{\exp\left(-\Theta(\zeta^{-1}) \right) T^2\left(\boldsymbol{\theta}_k - \lambda_k^\star\right)^2}{2}.$$

Plugging this into (39), we obtain

$$\left(p_{\min} - \zeta p_{\mathrm{tail}}/5 \right) \frac{\exp\left(-\Theta(\zeta^{-1}) \right)}{2} T^2\left(\boldsymbol{\theta}_k - \lambda_k^\star\right)^2 - 2(\sigma\zeta + 1)p_{\mathrm{tail}}T \left| \boldsymbol{\theta}_k - \lambda_k^\star \right| = \mathcal{O}\left(p_{\mathrm{tail}}^2/p_{\min} \right).$$

Viewing the left-hand side as a quadratic function with respect to $T \left| \boldsymbol{\theta}_k - \lambda_k^\star \right|$, we obtain

$$\left| \boldsymbol{\theta}_k - \lambda_k^\star \right| = \mathcal{O}\left(\frac{\exp\left(\Theta(\zeta^{-1}) \right) (\sigma\zeta + 1)p_{\mathrm{tail}}}{\left(p_{\min} - \zeta p_{\mathrm{tail}}/5 \right) T} \right) \leq \frac{Q}{T}. \tag{40}$$

This concludes the proof. $\qquad\square$

[vi]We note that $F_k(\theta) \approx \int_{-\infty}^{\infty} a(t)\exp(i(\lambda_k^\star - \theta)t)dt = F(\lambda_k^\star - \theta)$, where F is defined as the Fourier transform of $a(t)$ (refer to (3)).

We prove Theorem 3.3 as follows.

Proof of Theorem 3.3. The proof is similar to the proof of Theorem 3.2. For simplicity, we omit some details in this proof.

By Lemma A.1 with $\delta = \zeta$ and $J = \mathcal{O}(T/q)$, it holds that

$$\mathbb{P}\left(\max_{\theta \in \Theta} |E(\theta)| = \mathcal{O}(\zeta)\right) \geq 1 - \eta, \tag{41}$$

and

$$\mathbb{P}\left(\cap_{\theta, \theta' \in \Theta, \theta \neq \theta'} \{|E(\theta) - E(\theta')| = \mathcal{O}\left(\sigma T \zeta |\theta - \theta'| + \zeta\right)\}\right) \geq 1 - \eta, \tag{42}$$

where $\Theta = \{\theta_j\}_{j=1}^J \cup \{\lambda_m\}_{m \in \mathcal{D}}$. Thus, it suffices to prove (15) assuming (41) and (42) hold.

Because $T = \Omega\left(\frac{\alpha}{\Delta}\right)$, for any $j \in [k-1]$, the interval $\left[\boldsymbol{\theta}_j - \frac{\alpha}{T}, \boldsymbol{\theta}_j + \frac{\alpha}{T}\right]$ contains exactly one dominant eigenvalue, and those intervals that contain $\{\lambda_j\}_{j \in \mathcal{D}}$ will not contain any $\lambda_{j'}$ for $j' \in \mathcal{D}^c$. Define $m_k^\star := \operatorname{argmin}_{m \in \mathcal{D}} |\boldsymbol{\theta}_k - \lambda_m|$, $\lambda_k^\star := \lambda_{m_k^\star}$, $p_k^\star := p_{m_k^\star}$, and $\boldsymbol{\theta}_k^\star := \operatorname{argmin}_{\theta_j} \left|\theta_j - \lambda_{m_k^\star}\right|$. Similar to the argument as the proof of Theorem 3.2, we know that $|\boldsymbol{\theta}_k^\star - \lambda_k^\star| \leq \frac{q}{T}$ and $\boldsymbol{\theta}_k^\star \notin \mathcal{I}_{B,k}$. In the following proof, we show that $|\boldsymbol{\theta}_k - \lambda_k^\star| \leq \frac{Q}{T}$. Define the filter function that contains the impact of other dominant eigenvalues:

$$G_{\mathrm{dom}}(\theta) := \sum_{m \in \mathcal{D} \setminus \{m_k^\star\}} p_m \exp\left(-\frac{T^2 (\lambda_m - \theta)^2}{2}\right).$$

We first prove that $G_{\mathrm{dom}}(\lambda_k^\star)$, $G_{\mathrm{dom}}(\boldsymbol{\theta}_k^\star)$, and $G_{\mathrm{dom}}(\boldsymbol{\theta}_k)$ are of order $O(\zeta^2)$ similar to the proof of Theorem 3.2. Because $\lambda_k^\star$ is the closest dominant eigenvalue for $\boldsymbol{\theta}_k^\star$ and $\boldsymbol{\theta}_k$ and $\alpha/T < \Delta/2$, we obtain

$$|\lambda_k^\star - \lambda_j| \geq \Delta, \quad |\boldsymbol{\theta}_k - \lambda_j| \geq \frac{\Delta}{2}, \quad |\boldsymbol{\theta}_k - \lambda_j| \geq \frac{\Delta}{2}, \quad \forall j \in \mathcal{D} \setminus \{m_k^\star\}.$$

Combining this with $T = \Omega(\log^{1/2}(1/\zeta)/\Delta)$, we obtain

$$\max\{G_{\mathrm{dom}}(\lambda_k^\star), G_{\mathrm{dom}}(\boldsymbol{\theta}_k^\star), G_{\mathrm{dom}}(\boldsymbol{\theta}_k)\} \leq \sum_{m \in \mathcal{D} \setminus \{m_k^\star\}} p_m \exp\left(-\Omega(T^2 \Delta^2)\right) = O(\zeta^2). \tag{43}$$

Then, we upper bound the tail errors $G_{\mathrm{tail}}(\lambda_k^\star)$ and $G_{\mathrm{tail}}(\boldsymbol{\theta}_k)$. For $G_{\mathrm{tail}}(\lambda_k^\star)$, since $|\lambda_k^\star - \lambda_m| \geq \Delta$ for any $m \in [M]$, using the same calculation as (43), it holds that

$$G_{\mathrm{tail}}(\lambda_k^\star) = O(\zeta^2 p_{\mathrm{tail}}). \tag{44}$$

where we also use $\sum_{m \in \mathcal{D}^c} p_m \leq p_{\mathrm{tail}}$. For $G_{\mathrm{tail}}(\boldsymbol{\theta}_k)$, according to Theorem 3.1, we have $|\boldsymbol{\theta}_k - \lambda_k^\star| \leq \alpha/T \leq \frac{\Delta}{2}$. This implies $|\boldsymbol{\theta}_k - \lambda_m| > \frac{\Delta}{2}$ for any $m \in \mathcal{D}^c$. Therefore, using a similar calculation as (43), we can show that

$$G_{\mathrm{tail}}(\boldsymbol{\theta}_k) = O(\zeta^2 p_{\mathrm{tail}}). \tag{45}$$

Now, we are ready to upper bound $|\boldsymbol{\theta}_k - \lambda_k^\star|$. Similar to (30), we can show $|G(\boldsymbol{\theta}_k^\star) - G(\lambda_k^\star)| = O(\zeta^2)$. This implies

$$G(\boldsymbol{\theta}_k)^2 \geq G(\boldsymbol{\theta}_k^\star)^2 \geq G(\lambda_k^\star)^2 - O(\zeta^2). \tag{46}$$

Define $F_k(\theta) = \exp\left(-\frac{T^2(\lambda_k^\star - \theta)^2}{2}\right)$. Since $G(\theta) = |p_k^\star F_k(\theta) + E(\theta) + G_{\mathrm{dom}}(\theta) + G_{\mathrm{tail}}(\theta)|$, we have

$$
\begin{aligned}
G(\boldsymbol{\theta}_k)^2 &= p_k^{\star 2} F_k(\boldsymbol{\theta}_k)^2 + |E(\boldsymbol{\theta}_k)|^2 + G_{\mathrm{dom}}(\boldsymbol{\theta}_k)^2 + G_{\mathrm{tail}}(\boldsymbol{\theta}_k)^2 \\
&\quad + 2p_k^\star F_k(\boldsymbol{\theta}_k)(\mathrm{Re}(E(\boldsymbol{\theta}_k)) + G_{\mathrm{dom}}(\boldsymbol{\theta}_k) + G_{\mathrm{tail}}(\boldsymbol{\theta}_k)) \\
&\quad + 2(\mathrm{Re}(E(\boldsymbol{\theta}))(G_{\mathrm{dom}}(\boldsymbol{\theta}_k) + G_{\mathrm{tail}}(\boldsymbol{\theta}_k)) + 2G_{\mathrm{dom}}(\boldsymbol{\theta}_k)G_{\mathrm{tail}}(\boldsymbol{\theta}_k) \\
&= p_k^{\star 2} F_k(\boldsymbol{\theta}_k)^2 + G_{\mathrm{tail}}(\boldsymbol{\theta}_k)^2 + 2p_k^\star F_k(\boldsymbol{\theta}_k)(\mathrm{Re}(E(\boldsymbol{\theta}_k)) + G_{\mathrm{tail}}(\boldsymbol{\theta}_k)) + 2\mathrm{Re}(E(\boldsymbol{\theta}_k))G_{\mathrm{tail}}(\boldsymbol{\theta}_k) + O(\zeta^2) \\
&= p_k^{\star 2} F_k(\boldsymbol{\theta}_k)^2 + 2p_k^\star F_k(\boldsymbol{\theta}_k)\mathrm{Re}(E(\boldsymbol{\theta}_k)) + O(\zeta^2),
\end{aligned}
$$

where the second step follows from (41) and (43), and the third step follows from (45). Similarly, for $G(\lambda_k^\star)^2$, by (41), (43), and (44), we have the following:

$$
G(\lambda_k^\star)^2 \geq p_k^{\star 2} + 2p_k^\star \mathrm{Re}(E(\lambda_k^\star)) - O(\zeta^2)
$$

Similar to the (39), the above two inequalities, (42), and (46) imply

$$
(p_{\min} - 2\zeta)(1 - F_k(\boldsymbol{\theta}_k)) \leq O(\zeta) \cdot \min\{\sigma T|\boldsymbol{\theta}_k - \lambda_k^\star|, 1\} + O(\zeta^2/p_{\min}). \tag{47}
$$

Similar to (39)-(40), using $\zeta < p_{\min}/4$, we further obtain

$$
|\boldsymbol{\theta}_k - \lambda_k^\star| = O\left(\frac{\zeta\sigma}{p_{\min} - 2\zeta}\frac{1}{T}\right) \leq \frac{Q}{T}.
$$

The theorem is then proved. $\qquad\square$

B Other quantum phase estimation algorithms

In this section, we give a brief summary of the previous quantum phase estimation algorithms that are tested in Section 4: MM-QCELS [18], QPE (textbook version [1]), and ESPRIT [14].

- (MM-QCELS [18]): The dataset used in MM-QCELS is similar to QMEGS (Algorithm 2) and is also generated by Algorithm 1. The main subroutine of MM-QCELS is called quantum complex exponential least squares (QCELS): Given a data set $\{(t_n, Z_n)\}_{n=1}^N$ generated from Algorithm 1, MM-QCELS obtains an estimate for the dominant eigenvalues by solving the following optimization problem:

$$
\left(\{r_k^*\}_{k=1}^K, \{\theta_k^*\}_{k=1}^K\right) = \underset{r_k \in \mathbb{C}, \theta_k \in \mathbb{R}}{\arg\min} L_K\left(\{r_k\}_{k=1}^K, \{\theta_k\}_{k=1}^K\right). \tag{48}
$$

  with loss function

$$
L_K\left(\{r_k\}_{k=1}^K, \{\theta_k\}_{k=1}^K\right) = \frac{1}{N}\sum_{n=1}^N \left| Z_n - \sum_{k=1}^K r_k \exp(-i\theta_k t_n) \right|^2. \tag{49}
$$

  Choosing $K = |\mathcal{D}|$ and a proper generated data set $\{(t_n, Z_n)\}_{n=1}^N$[vii], [18] shows that the solution $\{\theta_k^*\}_{k=1}^K$ of (48) is a good approximation to the set of dominant eigenvalues $\{\lambda_m\}_{m \in \mathcal{D}}$. MM-QCELS can reach Heisenberg limit scaling and small circuit depth when $p_{\mathrm{tail}} \ll 1$. However, the algorithm requires a spectral gap assumption, meaning $T = \Omega(1/\epsilon) > \Delta_{\mathrm{dom}}^{-1}$, to ensure that the optimization problem can differentiate between the dominant eigenvalues.

[vii] In [18, Algorithm 2], the authors need to generate a sequence of data set $\{(t_n, Z_n)\}_{n=1}^N$ with different T and N to ensure Heisenberg limit scaling of the algorithm theoretically.

- (QPE (textbook version [1])): We provide a brief review of QPE for ground state energy estimation. The quantum process involves a sequence of controlled time evolution operations e^{-iH} on a state $|0^d\rangle |\psi\rangle = \sum_{m=0}^{M-1} c_m |0^d\rangle |\psi_m\rangle$. Here, $|\psi_m\rangle$ represents the eigenstates associated with eigenvalues λ_m. The resulting quantum state from these operations, prior to applying the inverse Quantum Fourier Transform (QFT), is expressed as follows:

$$|\Psi\rangle = \frac{1}{\sqrt{N_t}} \sum_{j=-N_t/2}^{N_t/2-1} |j\rangle e^{-ijH} |\psi\rangle ,$$

where $N_t = 2^d$. After applying the inverse QFT and measuring the ancilla register, the probability of obtaining outcome k is given by:

$$P(k) = \sum_{m=0}^{M-1} |c_m|^2 K_{N_t} \left(\frac{2\pi k}{N_t} - \lambda_m \right), \tag{50}$$

where $-N_t/2 \leq k \leq N_t/2 - 1$, and K_{N_t} is the squared and normalized Dirichlet kernel defined as $K_{N_t}(\theta) = \frac{\sin^2(\theta N_t/2)}{N_t^2 \sin^2(\theta/2)}$. To simulate QPE classically, we sample this distribution N_{QPE} times to obtain a set of samples $\{k_i\}_{i=1}^{N_{QPE}}$. The ground state energy can then be approximated as $\widetilde{\lambda}_0 = \frac{2\pi \min_i k_i}{N_t}$.

We note that the original textbook version of QPE algorithm in [1, Chapter 5.2] is designed to estimate general eigenvalues, not necessarily the ground state energy. The version presented above is a variation specifically designed for ground state energy estimation. There are also other variations of QPE, such as the Gaussian/Kaiser window-based QPE [37, 38], where the authors implement different resource states to produce a more concentrated kernel and reduce T_{total}. In addition, it might be possible to extend QPE to estimate multiple eigenvalues simultaneously: For example, in the ideal case where $p_{\text{tail}} = 0$ and $\{2\pi\lambda_m\}_{m\in\mathcal{D}}$ are finite-digit numbers. In this ideal scenario, when N_t is sufficiently large, we have

$$P(k) = \sum_{m\in\mathcal{D}} |c_m|^2 \delta_0 \left(\frac{2\pi k}{N_t} - \lambda_m \right) .$$

Therefore, measuring the ancilla qubits $\mathcal{O}(\text{poly}(K))$ times is sufficient to obtain all the eigenvalues exactly. In the non-ideal case, we need to design an efficient method to post-process the output of QPE and obtain accurate estimations of multiple eigenvalues. To the best of our knowledge, we are unaware of any such procedure that has clear complexity analysis. Thus, for simplicity, we mainly consider the simple version of QPE in our paper.

- (ESPRIT [14]): The dataset used in ESPRIT is similar to Algorithm 2 and is also generated by Algorithm 1. The algorithm of ESPRIT relies on the construction and manipulation of the Hankel matrix: Given $T > 0$ and an odd integer $N > 0$, we first set $t_n = n\tau$ for $0 \leq n \leq N$, where $\tau = T/N$, and construct Hankel matrix $H \in \mathbb{C}^{\frac{N+1}{2} \times \frac{N+1}{2}}$ with $H_{i,j} = Z_{t_{i+j}}$. Here, Z_{t_n} is generated by Algorithm 1. We then find the singular value decomposition $H = U\Sigma V^\dagger$ and define

$$U_0 = U[:-1, :K], \quad U_1 = [1:, :K].$$

Here, U_0 contains first $\frac{N-1}{2}$ rows and first K columns of U and U_1 contains the last $\frac{N-1}{2}$ rows and first K columns of U. Finally, we find the eigenvalues $\{\mu_k\}_{k=1}^K$ of

$U_1^{-1}U_0$ (U_1^{-1} is the pseudoinverse of U_1) and define $\{\boldsymbol{\theta}_k = \text{angle}(\mu_k)/\tau\}_{k=1}^K$. According to the results of classical signal processing [35], we can demonstrate that, when $K = |\mathcal{D}|$ and T, N are chosen properly, $\{\boldsymbol{\theta}_k = \text{angle}(\mu_k)/\tau\}_{k=1}^K$ is a set that is close to the set of dominant eigenvalues $\{\lambda_m\}_{m \in \mathcal{D}}$.

We would like to highlight that the original ESPRIT method falls short of achieving the Heisenberg limit scaling in the context of quantum phase estimation. In the original ESPRIT framework, the choice of $\tau = 1$ is imperative to mitigate aliasing issues[viii]. Consequently, this requires $N = T$, $T_{\max} = T$, and results in $T_{\text{total}} = \Theta(T_{\max}^2)$. However, according to the generalized uncertainty relation [39], there exists a uniform complexity lower bound for phase estimation [40], asserting that the square of the error is at least $\Omega\left(T_{\text{total}}^{-1}T_{\max}^{-1}\right)$ in expectation. By combining $T_{\text{total}}^{-1}T_{\max}^{-1} = \mathcal{O}(\epsilon^2)$ and $T_{\text{total}} = \Theta(T_{\max}^2)$, we deduce $T_{\text{total}} = \Omega(\epsilon^{-4/3})$. The crux of the matter, as deduced from the previous analysis, is that the necessary selection of $\tau = 1$ to avoid aliasing poses a significant hurdle to ESPRIT in achieving the Heisenberg limit.

More recently, [19] proposes a multilevel ESPRIT approach to circumvent aliasing issues without enforcing $\tau = 1$. Notably, [19] generates a sequence of datasets with carefully chosen values for T and N, progressively refining the estimation of dominant eigenvalues. The successful application of multilevel techniques enables them to achieve Heisenberg-limited scaling and shorter circuit depth. In our numerical simulations presented in Section 4, for simplicity, we only consider the original ESPRIT and choose $N = T$ and $\tau = 1$ to mitigate aliasing issues associated with ESPRIT.

C Algorithm for integer-power setting

In certain phase estimation tasks, only a black box unitary U represented by a quantum circuit can be accessed. Under this setting, querying an arbitrary real power of U is not feasible. Instead, only integer powers of it can be acquired. To maintain consistency with the notation used in the case where $U = e^{-iH}$, we still assume $U|\psi_m\rangle = e^{-i\lambda_m}|\psi_m\rangle$. In this case, we aim to recover the phases $\lambda_m \mod 2\pi$. All other parameters, such as $p_{\min}$, p_{tail}, and $\mathcal{D}$, retain the same definitions as in the real-power setting unless explicitly stated otherwise. We define the $\mod 2\pi$ distance of two numbers u and v as

$$|u - v|_{2\pi} := \min\{|u - v \mod 2\pi|, |v - u \mod 2\pi|\}.$$

The idea is similar to the real power setting, which is to leverage the Gaussian-filtered spectral density. The difference is that we will use the periodic Gaussian

$$\phi_p(x) = W \sum_{j \in \mathbb{Z}} e^{-\frac{(x+2j\pi)^2 T^2}{2}}, \tag{51}$$

where $W < 1$ is a normalizing constant such that $\phi_p(0) = 1$. It is clear that ϕ_p is 2π-periodic, and its Fourier coefficients are

$$\widehat{\phi}_p(k) = \frac{1}{2\pi}\int_0^{2\pi}\phi_p(x)e^{-ikx}\mathrm{d}x = \frac{W}{\sqrt{2\pi}T}e^{-\frac{k^2}{2T^2}}. \tag{52}$$

[viii]It is not possible to differentiate between λ_k and $\lambda_k + 2\pi\tau$ in ESPRIT, as they produce the same data.

Therefore, we have $\sum_{k \in \mathbb{Z}} \widehat{\phi}_p(k) = \phi_p(0) = 1$. Let

$$a(k) = \begin{cases} \widehat{\phi}_p(0) + \sum_{|j| > \sigma T} \widehat{\phi}_p(j) & k = 0, \\ \widehat{\phi}_p(k) & 1 \leq |k| \leq \sigma T, \\ 0 & |k| > \sigma T. \end{cases}$$

be a distribution over $\mathbb{Z}$, and t be a random variable sampled from this distribution. Z_t denote the unbiased estimation of $\langle \psi | U^t | \psi \rangle$ obtained by the Hadamard test. Therefore, the maximal quantum runtime $T_{\max}$ is bounded by σT, as similar to the real-power setting. Except for the distribution of t, the rest part of the phase estimation algorithm goes the same as in Algorithm 2. We may also establish theorems that guarantee the performance of this slightly modified algorithm. First, we establish a lemma that gives several properties of the periodic Gaussian function $\phi_p(x)$.

Lemma C.1. *If $T \geq 1$, then $\phi_p(x)$ is increasing on $[-\pi, 0]$ and decreasing on $[0, \pi]$. For $x \in [-\frac{2\pi}{3}, \frac{2\pi}{3}]$,*

$$e^{-\frac{x^2 T^2}{2}} \leq \phi_p(x) \leq 1.01 e^{-\frac{x^2 T^2}{2}}, . \tag{53}$$

Proof. The monotonicity part is proved in [23, Lemma 2], where only some normalizing constants differ. Without loss of generality, we will prove the rest of the lemma assuming $x \geq 0$ since $\phi_p(x)$ is an even function. For the left part of (53), we need to notice that when $x = 0$, the equality holds. Moreover, when $x \in [0, \pi]$, we have

$$\frac{\mathrm{d}}{\mathrm{d}x} \phi_p(x) = WT^2 \left(-xe^{-\frac{T^2 x^2}{2}} + \sum_{j=1}^{+\infty} \left(-(2j\pi + x)e^{-\frac{T^2(2j\pi + x)^2}{2}} + (2j\pi - x)e^{-\frac{T^2(2j\pi - x)^2}{2}} \right) \right)$$

$$\geq WT^2 \left(-xe^{-\frac{T^2 x^2}{2}} \right) \geq -T^2 xe^{-\frac{T^2 x^2}{2}} = \frac{\mathrm{d}}{\mathrm{d}x} e^{-\frac{T^2 x^2}{2}},$$

where in the first inequality, we used the fact that each term of the summation is positive, and in the second inequality, we used $W \leq 1$. To see this, we may introduce the function $h(y) := ye^{-\frac{T^2 y^2}{2}}$, which is decreasing when $y \geq 1 \geq \frac{1}{T}$, and thus $h(2j\pi - x) \geq h(2j\pi + x)$ for all $j \geq 1$.

Next, we prove the right part of (53). This can be done by the following calculation.

$$\phi_p(x)e^{\frac{x^2 T^2}{2}} = W \sum_{j \in \mathbb{Z}} e^{-2T^2 \pi(\pi j^2 + xj)} \leq \sum_{j \in \mathbb{Z}} e^{-2T^2 \pi(\pi j^2 + xj)}$$

$$= 1 + e^{-2T^2 \pi(\pi - x)} + \sum_{j=1}^{+\infty} e^{-2T^2 \pi(\pi j^2 + xj)} + \sum_{j=2}^{+\infty} e^{-2T^2 \pi(\pi j^2 - xj)}$$

$$\leq 1 + e^{-2\pi(\pi/3)} + \sum_{j=1}^{+\infty} e^{-2\pi j} + \sum_{j=2}^{+\infty} e^{-2\pi j} < 1.01.$$

$\square$

Now, we may define the error function

$$E(\theta) = \frac{1}{N} \sum_{n=1}^{N} Z_n \exp(i\theta t_n) - \sum_{m=1}^{M} p_m \phi_p(\theta - \lambda_m)$$

and $E_j := E(\theta_j)$. Then we have the following lemma similar to Lemma A.1.

Lemma C.2. *Given $\delta > 0$ and the pair of overlap-eigenvalues $\{(p_i, \lambda_i)\}_{m=1}^M$, if $\sigma = \Omega\left(\log^{1/2}(1/\delta)\right)$, we have*

$$\left| \mathbb{E}\left(\frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta_j t_n) \right) - \sum_{m=1}^M p_m \phi_p(\theta - \lambda_m) \right| \leq \delta^2 \,. \tag{54}$$

Furthermore, given $\eta > 0$ and $\Theta = \{\theta_j\}_{j=1}^J \cup \{\lambda_m\}_{m \in \mathcal{D}}$, if $N = \Omega\left(\frac{1}{\delta^2} \log\left(\left(\frac{T}{q} + |\mathcal{D}|\right)\frac{1}{\eta}\right)\right)$, we have

$$\mathbb{P}\left(\max_{\theta \in \Theta} |E(\theta)| \leq \delta \right) \geq 1 - \eta \,. \tag{55}$$

Proof. According to Algorithm 1, when $\sigma = \Omega\left(\log^{1/2}(1/\delta)\right)$, we have

$$
\begin{aligned}
&\left| \mathbb{E}\left(\frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta_j t_n) \right) - \sum_{m=1}^M p_m \phi_p(\theta - \lambda_m) \right| \\
&= \left| \sum_{m=1}^M p_m \sum_{|k| > \sigma T} \widehat{\phi}_p(k) e^{ik(\theta - \lambda_m)} \right| \leq \sum_{|k| > \sigma T} \widehat{\phi}_p(k) e^{-\frac{t^2}{2T^2}} \left| \sum_{m=1}^M p_m e^{ik(\theta - \lambda_m)} \right| \leq \sum_{|k| > \sigma T} \widehat{\phi}_p(k) \\
&= \sum_{|k| > \sigma T} \frac{W}{\sqrt{2\pi} T} e^{-\frac{k^2}{2T^2}} \leq 2 \int_{\sigma T}^\infty \frac{1}{\sqrt{2\pi} T} e^{-\frac{t^2}{2T^2}} \, \mathrm{d}t = \sqrt{\frac{2}{\pi}} \int_\sigma^\infty e^{-\frac{s^2}{2}} \, \mathrm{d}s < e^{-\sigma^2} < \delta^2,
\end{aligned}
\tag{56}
$$

where we used the fact that $W < 1$ and bounded the summation using integration. This proves (54). The proof of (55) is the same as the proof of [22, Appendix B.3 Lemma 4 eqn. (B8)], thus, we omit it here. $\quad\square$

Define the magnitude function:

$$G(\theta) = \left| \frac{1}{N} \sum_{n=1}^N Z_n \exp(i\theta t_n) \right| \,.$$

and $G_j := G(\theta_j)$.

Finally, we have a similar theorem as Theorem 3.1. This shows this algorithm restricted to integer powers of U can also achieve the Heisenberg limit without any gap assumptions.

Theorem C.3 $(\forall T \geq 1)$. *Assume $p_{\min} > p_{\text{tail}}$ and $|\mathcal{D}| \leq K$. Given the probability of failure $\eta > 0$, we choose the following parameters:*

- *Block constant: $\alpha = \Omega\left(\log^{1/2}\left(\frac{1}{p_{\min} - p_{\text{tail}}}\right)\right)$,*

- *Searching parameter: $q = \mathcal{O}\left(\log^{1/2}\left(\frac{p_{\min}}{p_{\text{tail}} + (p_{\min} - p_{\text{tail}})/2}\right)\right)$, $q < \alpha$, and $\alpha/q \in \mathbb{N}$,*

- *Time truncation parameter: $\sigma = \Omega\left(\log^{1/2}\left(\frac{1}{p_{\min} - p_{\text{tail}}}\right)\right)$,*

- *Number of samples: $N = \Omega\left(\frac{1}{(p_{\min} - p_{\text{tail}})^2} \log\left(\left(\frac{T}{q} + |\mathcal{D}|\right)\frac{1}{\eta}\right)\right)$.*

Then, with probability at least $1 - \eta$, we have that for each $i \in \mathcal{D}$, there exists $1 \leq k_i \leq |\mathcal{D}|$ such that

$$|\lambda_i - \boldsymbol{\theta}_{k_i}|_{2\pi} \leq \frac{\alpha}{T} \,. \tag{57}$$

In particular, for any $\epsilon > 0$, to achieve

$$\{\lambda_m\}_{m \in \mathcal{D}} \subset \cup_k [\boldsymbol{\theta}_k - \epsilon, \boldsymbol{\theta}_k + \epsilon] \mod 2\pi,$$

it suffices to choose

$$T_{\max} = \widetilde{\Theta}\left(\frac{1}{\epsilon}\right), \quad T_{\text{total}} = \widetilde{\Theta}\left(\frac{1}{(p_{\min} - p_{\text{tail}})^2 \epsilon}\right),$$

where the logarithmic factor is omitted.

Proof. Similar to the proof of Theorem 3.1, we only need to prove

$$\max_{\theta_j \notin \cup_{m \in \mathcal{D}}\left[\lambda_m - \frac{\alpha}{3T}, \lambda_m + \frac{\alpha}{3T}\right]} G_j < \min_{\theta_j \in \cup_{m \in \mathcal{D}}\left[\lambda_m - \frac{q}{T}, \lambda_m + \frac{q}{T}\right]} G_j \tag{58}$$

under the assumption $E_j < \frac{p_{\min} - p_{\text{tail}}}{8}$. The rest of the proof is the same as in Theorem 3.1. We also consider two classes of candidates:

- When $\theta_j \in \cup_{m \in \mathcal{D}} \left[\lambda_m - \frac{q}{T}, \lambda_m + \frac{q}{T}\right] \mod 2\pi$, we obtain

$$\begin{aligned} G_j = G(\theta_j) &\geq \left|\sum_{m=1}^{M} p_m \phi_p (\lambda_m - \theta_j)\right| - E_j \\ &> p_{\min} \exp\left(-\frac{q^2}{2}\right) - \frac{p_{\min} - p_{\text{tail}}}{8} \\ &\geq p_{\text{tail}} + \frac{p_{\min} - p_{\text{tail}}}{2} - \frac{p_{\min} - p_{\text{tail}}}{8} \geq p_{\text{tail}} + \frac{3(p_{\min} - p_{\text{tail}})}{8}. \end{aligned}$$

  where we used (53) and $E_j < \frac{p_{\min} - p_{\text{tail}}}{8}$ in the second inequality, and the condition of q in the last inequality.

- When $\theta_j \notin \cup_{m \in \mathcal{D}} \left[\lambda_m - \frac{\alpha}{3T}, \lambda_m + \frac{\alpha}{3T}\right] \mod 2\pi$, we obtain

$$\begin{aligned} G_j &\leq \left|\sum_{m \in \mathcal{D}} p_m \phi_p (\lambda_m - \theta_j)\right| + \left|\sum_{m \in \mathcal{D}^c} p_m \phi_p (\lambda_m - \theta_j)\right| + E_j \\ &\leq 1.01 \exp\left(-\frac{\alpha^2}{18}\right) + p_{\text{tail}} + \frac{p_{\min} - p_{\text{tail}}}{8} \leq p_{\text{tail}} + \frac{p_{\min} - p_{\text{tail}}}{4}, \end{aligned}$$

  where we used Lemma C.1 and $E_j < \frac{p_{\min} - p_{\text{tail}}}{8}$ in the second inequality and the condition of α in the last inequality.

Therefore, (58) is proved, and we complete the proof of the theorem.

$\square$