Adaptive low-depth quantum algorithms for robust multiple-phase estimation

Haoya Li [1,*,†] Hongkang Ni,[2,*,‡] and Lexing Ying[1,2,§]

[1]*Department of Mathematics, Stanford University, Stanford, California 94305, USA*
[2]*Institute for Computational and Mathematical Engineering, Stanford University, Stanford, California 94305, USA*

This paper is an algorithmic study of quantum phase estimation with multiple eigenvalues. We present robust multiple-phase estimation (RMPE) algorithms with Heisenberg-limited scaling. The proposed algorithms improve significantly from the idea of single-phase estimation methods by combining carefully designed signal processing routines and an adaptive determination of runtime amplifying factors. They address both the *integer-power* model, where the unitary U is given as a black box with only integer runtime accessible, and the *real-power* model, where U is defined through a Hamiltonian H by $U = \exp(-2\pi i H)$ with any real runtime allowed. These algorithms are particularly suitable for early fault-tolerant quantum computers in the following senses: (1) a minimal number of ancilla qubits are used, (2) an imperfect initial state with a significant residual is allowed, (3) the prefactor in the maximum runtime can be arbitrarily small given that the residual is sufficiently small and a gap among the dominant eigenvalues is known in advance. Even if the eigenvalue gap does not exist, the proposed RMPE algorithms can achieve the Heisenberg limit while maintaining (1) and (2).

I. INTRODUCTION

Quantum phase estimation (QPE) is a fundamental problem in quantum computing. In this paper, we focus on the more complex scenario where there are multiple eigenvalues to be estimated. This problem is essentially different from the estimation of a single eigenvalue in many aspects, and most of the methods for single-mode QPE cannot be directly extended to the multiple-mode case. For example, if one directly applies Kitaev's algorithm or the robust phase estimation (RPE) algorithm [1–3], which are both commonly accepted benchmarks for single-mode QPE, then different eigenmodes lead to an aliasing effect such that the precision of the estimations cannot be improved in the iterative procedure. In this paper, we present an ensemble of adaptive methods using low-depth circuits that solve this problem successfully. The multiple-phase estimation problem also becomes particularly challenging when the gap between the dominant eigenvalues gets smaller. Our method also handles this issue effectively by incorporating a simple line spectrum estimation algorithm, which we developed in Ref. [4] that alleviates the constraint on spectral gaps.

A few key metrics are needed for evaluating the performance of multiple-phase algorithms. For example, one clearly prefers an algorithm with a small number of qubits and a low circuit depth while allowing for a high level of residual in the initial state. We elaborate on this afterward and show that the adaptive method we propose improves upon existing methods in terms of these metrics, making it well-suited for early fault-tolerant quantum computers.

A. Problem settings

Formally, this paper concerns the quantum phase estimation (QPE) problem with multiple eigenvalues as follows. For a unitary matrix U, let $\{(e^{-2\pi i\lambda_s}, |\psi_s\rangle)\}$ be the eigenpairs of U with $\lambda_s \in [0, 1]$. Suppose that $|\psi\rangle$ is an initial quantum state of the form

$$|\psi\rangle = \sum_{s=1}^{S} c_s |\psi_s\rangle + c_{\text{res}} |\psi_{\text{res}}\rangle, \qquad (1)$$

where S is the number of *dominant* eigenvalues and $c_{\text{res}} |\psi_{\text{res}}\rangle$ is the *residual*. Here, by dominant, we mean that the overlaps between these eigenstates $|\psi_s\rangle$ and $|\psi\rangle$ are bounded from below by a constant β, i.e., $\min_{1\leqslant s\leqslant S} |c_s|^2 \geqslant \beta > 0$. On the other hand, the energy of the residual $|c_{\text{res}}|^2$ is bounded from above by a constant ω less than β, i.e., $|c_{\text{res}}|^2 \leqslant \omega < \beta$. The goal is to estimate the set $\Lambda \equiv \{\lambda_s\}_{s=1}^{S}$ of dominant eigenvalues up to a prescribed accuracy ϵ.

B. Key metrics for performance evaluation

The difficulty level of multiple eigenvalue estimation depends on the gap between the dominant eigenvalues. If the gap is small, the problem becomes difficult. When the spectral gap is bounded from below by a constant independent of the desired precision ϵ, we refer to it as the *gapped* case. When no lower bound of the gap is assumed, we refer to it as the *gapless* case.

Several key complexity metrics to assess a multiple-phase estimation algorithm are listed as follows:

(1) The number of ancilla qubits required. The smaller, the better.

(2) The amount of residual ω allowed in the initial state $|\psi\rangle$, i.e., ω is the maximum $|c_{\text{res}}|^2$ allowed.

(3) The maximum runtime $T_{\max}$. It is defined as the maximum depth of the quantum circuits used by the algorithm.

---

*These authors contributed equally to this work.

†lihaoya@stanford.edu

‡hongkang@stanford.edu

§lexing@stanford.edu

(4) The total runtime T_{total}, i.e., the sum of the circuit depths over all executions. It has been shown in Refs. [5–7], for example, that T_{total} has a lower bound named the Heisenberg limit $T_{total} = \Omega(\epsilon^{-1})$.

(5) Finally, the minimum gap between the eigenvalues allowed.

Among these metrics, a small T_{max} is particularly important for early fault-tolerant quantum devices since these devices typically have a relatively short coherence time.

Based on these metrics, an ideal phase estimation algorithm should meet the following requirements:

(1) Using a small number of (even a single) ancilla qubits.

(2) Allowing the initial state $|\psi\rangle$ to be inexact and ideally ω to be proportional or even close to β.

(3) Satisfying $T_{max} = O(\epsilon^{-1})$ with the prefactor ideally proportional to ω/β.

(4) Achieving the Heisenberg-limited scaling $T_{total} = \tilde{O}(\epsilon^{-1})$, where $\tilde{O}$ means omitting the polylogarithmic term.

(5) Allowing the minimum gap to be arbitrarily small.

C. The integer-power model and the real-power model

The design of multiple eigenvalue estimation algorithms depends closely on the representation of U. In the most general model, U is represented by a quantum circuit or even a black box model, such as a quantum approximate optimization algorithm (QAOA) or variational quantum eigensolver (VQE). This model only allows access to integer powers U^j of U for $j \in \mathbb{Z}$, $j \geqslant 0$, and we refer to it as the *integer-power* model. A different model is $U = e^{-2\pi iH}$, where H is a quantum Hamiltonian with eigenvalues in [0,1]. Here, U is often implemented with Trotterization or the splitting method, where one of the main applications is to find the energy of the ground state and a few low-lying excited states of H. This model allows access to $U^t := e^{-2\pi iHt}$ for any $t \in \mathbb{R}_+$, and we refer to it as the *real-power* model.

D. Related work

As a fundamental primitive of quantum algorithms, quantum phase estimation has attracted a lot of research activity in the past few decades.

1. Single eigenvalue

Several existing methods can be applied when there is only one dominant eigenvalue, i.e., $S = 1$. The early algorithms require a perfect eigenstate, i.e., $c_{res} = 0$. One of the most fundamental algorithms is the Hadamard test, as shown in Fig. 1(a), which utilizes the I and S gates after the controlled-U gate. The Hadamard test provides estimations of $\text{Re}\langle\psi|U|\psi\rangle$ and $\text{Im}\langle\psi|U|\psi\rangle$, respectively, from the probability of getting $|0\rangle$ while measuring the ancilla qubit. For the Hadamard test, one needs $O(\epsilon^{-2})$ repetitions to reach precision ϵ, leading to a total gate complexity $T_{total} = O(\epsilon^{-2})$.

A quadratic improvement proposed by Kitaev [8,9] uses measurements of $\langle\psi|U^{2^j}|\psi\rangle$ for $j = 0, 1, 2, \ldots, J$ with $J = O(\log(\epsilon^{-1}))$, as illustrated by the circuit in Fig. 1(b). The total runtime $T_{total} = O(\epsilon^{-1})$ of Kitaev's method achieves the Heisenberg limit [5–7]. However, the original version of Kitaev's method only applies to perfect eigenstates, i.e., $c_{res} = 0$.

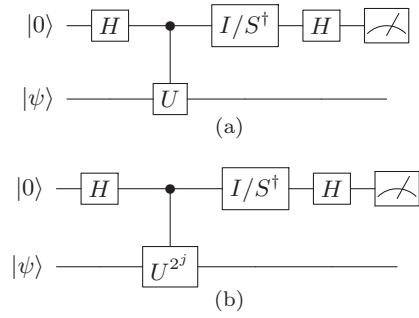
FIG. 1. (a) Illustration for the Hadamard test. Here H denotes the Hadamard gate. When estimating the real part of $\langle\psi|U|\psi\rangle$, we use I (the identity) for the gate after the controlled-U gate, while S^+ (the Hermitian conjugate of the phase gate S) is used for the estimation of the imaginary part. (b) Illustration for the Kitaev algorithm. For a sequence of powers j, the real and imaginary parts of $\langle\psi|U^{2^j}|\psi\rangle$ are estimated to reach a higher precision and improve the total complexity.

Another example that reaches the Heisenberg limit is the QPE algorithm with quantum Fourier transform (QFT) [10], which only involves a single execution but needs more ancilla qubits and a deeper circuit. Many alternatives have also been proposed in the recent literature [11–19]. For a more comprehensive overview about the QPE algorithms for a single eigenvalue, we refer to the detailed discussions in Refs. [18–20].

As we mentioned earlier, for early fault-tolerant quantum devices, besides using a small number of (or even only a single) qubits and reaching the Heisenberg limit $T_{total} = \tilde{O}(\epsilon^{-1})$ for the case $S = 1$ and $|c_{res}|^2 > 0$, it is also desired to allow the prefactor in $T_{max} = O(\epsilon^{-1})$ to depend on the magnitude of c_{res}. The authors of Ref. [19] proposed QCELS, an optimization subroutine that works with $|c_{res}|^2 \leqslant 0.29$ and allows the prefactor of $T_{max} = O(\epsilon^{-1})$ to scale as $\Theta(|c_{res}|)$. In a recent work [21], we showed that a modified version of the robust phase estimation (RPE) algorithm [1–3] can work with $|c_{res}|^2 \leqslant 2\sqrt{3} - 3 \approx 0.464$ and it gives the near-optimal prefactor $\Theta(|c_{res}|^2)$.

2. Multiple eigenvalues

The work [15] considers the problem of estimating multiple eigenvalues with a signal processing subroutine and adopts the matrix pencil method. The method can be sensitive to noise, and the Heisenberg-limited scaling is not achieved. The algorithm proposed in Ref. [22] also estimates multiple eigenvalues based on time series analysis. The total cost is $O(\epsilon^{-6})$, which is quite far from the Heisenberg limit. A more recent work [23] extends the idea of the RPE algorithm to multiple eigenvalues and achieves the Heisenberg limit. However, the residual term $|\psi_{res}\rangle$ is not allowed in Ref. [23], and the prefactor power of S is quite large.

The work [24] extended the QCELS algorithm to the multiple eigenvalue setting. This extension achieves the Heisenberg limit with a single ancilla qubit and allows for a residual. The maximum circuit depth can also be reduced when the residual amount ω is small compared with β. However, the performance of this optimization technique relies on the

existence of a spectral gap Δ between the multiple dominant eigenvalues, where the minimal runtime of the circuits is $\Omega(\Delta^{-1})$. Moreover, a grid search is generally needed to solve the optimization problem due to the highly complicated landscape, which leads to a classical cost exponential in S.

Quantum subspace diagonalization methods have also been used for multiple eigenvalue estimation [25–30], where the eigenvalues are obtained by solving certain projected eigenvalue problems or singular value problems. Compared with the numerical performance demonstrated, the theoretical analysis of the subspace diagonalization methods is still rather preliminary and pessimistic [31].

E. Contributions

This paper conducts a comprehensive study of the multiple eigenvalue quantum phase estimation. Our main contribution is a family of robust multiple-phase estimation (RMPE) algorithms for both the gapless and the gapped cases and for both the real-power and integer-power models. These RMPE algorithms build on the overall structure of the method proposed in Ref. [23].

For the gapless case, the proposed RMPE algorithm estimates the dominant $S \geqslant 1$ eigenvalues for any residual overlap $|c_{\mathrm{res}}|^2 < \beta$. It utilizes the measurement results from Hadamard tests to the unitary operators U^{M_ℓ} for an exponentially growing sequence $\{M_\ell\}$, one for each step ℓ. The estimated dominant eigenvalues of U^{M_ℓ} allow us to narrow down the intervals that contain the dominant eigenvalues of U, and the desired precision will be obtained after $O(\log(\epsilon^{-1}))$ steps. One key component is a simple but new signal processing routine that estimates eigenvalue locations from quantum measurements without any gap assumption. Another key component of the algorithm is the careful choice of $\{M_\ell\}$ to avoid potential collision of the intervals. To achieve this for the real-power model, we utilize noninteger $\{M_\ell\}$ that grows each time by a factor between two and four. For the integer-power model, we leverage results from prime number theory for the choice of $\{M_\ell\}$ to avoid collisions. Both the real-power and integer-power versions can tolerate a residual ω up to β and achieve the Heisenberg limit. In summary, this algorithm meets all five requirements, except that the prefactor of $T_{\max}$ does not scale like ω/β.

For the gapped case, we improve the algorithms with two key modifications. First, we adopt the ESPRIT algorithm [32], a different signal processing routine that allows for more accurate eigenvalue estimation when a gap is available. This allows us to build fairly accurate approximations to Λ even at the initial step. Second, with this better initial estimate, the number of iterations of the algorithm can be significantly reduced, resulting in $T_{\max}$ to scale like ω/β. Both the real-power and integer-power versions can tolerate residual $\omega = O(\beta)$ and reach the Heisenberg limit. In summary, this algorithm meets the first four requirements mentioned above.

Finally, for the case of a finite but small spectral gap, we propose hybrid algorithms for both the real-power and integer-power models. The scaling of maximal runtime is improved to $O(\Delta^{-1} + \epsilon^{-1})$, which is significantly better than the $O(\Delta^{-1}\epsilon^{-1})$ complexity of the algorithms for gapped case. Meanwhile, we still maintain the prefactor ω/β and reach

the Heisenberg limit. In a nutshell, the hybrid algorithm uses the gapless signal processing subroutine in the first several iterations as a burn-in period and then switches to the ESPRIT after the spectral gap is enlarged. To summarize, the hybrid algorithms meet the first four requirements mentioned above while improving the scaling of Δ compared with the algorithms for the gapped case.

The maximal runtime and total runtime of different algorithms in this paper are summarized in Table I.

Comparisons

Compared with [23], our RMPE algorithms have the following advantages: First, the proposed algorithms allow an imperfect initial state with a residual term. Second, by using a simpler signal processing routine in both the gapped and gapless cases, we obtain a smaller prefactor in terms of the sparsity level S: the prefactor of T_{total} is only quadratic in terms of the sparsity level S. In contrast, the dependency on S obtained in Ref. [23] is $O(S^{12})$. Third, the proposed RMPE algorithms work in the real-power and integer-power models, while Ref. [23] only works with the real-power model. Furthermore, when the dominant eigenvalues present a gap, $T_{\max}$ of our gapped algorithm has a prefactor of order ω/β, which is a property not shared by the algorithm of Ref. [23].

Compared with [24], the proposed RMPE algorithms have two advantages. First, we address the gapless case that is not visited in Ref. [24]. Second, the classical computational complexity is much lower: our RMPE algorithms scale polynomially in S while the optimization step of the method in Ref. [24] generally scales exponentially in S.

For clarity, a comparison of different multiphase estimation algorithms is presented in Table II.

F. Contents

The rest of the paper is organized as follows: Section II introduces the main structure of the proposed algorithms. Section III discusses the gapless case in detail. In Sec. IV, we present the easier gapped case while focusing on its difference from the gapless case. In Sec. V, we provide the hybrid algorithms for the gapped case. All these sections address both the real-power and integer-power models.

A few comments about the notations are in order here. For a measurable set T, we use $|T|$ and $|T|_{\mathbb{T}}$ to denote the Lebesgue measure of $T \subset \mathbb{R}$ and $T \subset \mathbb{T}$, respectively, where $\mathbb{T}$ is the 1-dimensional torus. The η neighborhood of a set $T \subset \mathbb{R}$, denoted by $B(T, \eta)$, is defined as

$$B(T, \eta) := \cup_{t \in T} [t - \eta, t + \eta]. \tag{2}$$

The η neighborhood of a set $T \subset \mathbb{T}$, denoted as $B_{\mathbb{T}}(T, \eta)$, is defined similarly on the torus $\mathbb{T}$. For a set $T \subset \mathbb{R}$, we denote the set $\{ct + d \,|\, t \in T\}$ for any $c, d \in \mathbb{R}$ by $cT + d$.

II. ALGORITHM OUTLINE

This section describes the main algorithmic structure. The following sections specialize this structure to the gapless and gapped cases, respectively.

Applying the Hadamard test to $U^1, \ldots, U^K$ for some integer K with input $|\psi\rangle$ provides a noisy measurement $\{y(k)\}$

TABLE I. The summary of different algorithms presented in this paper.

Algorithm	Setting	$T_{\max}$	T_{total}
Sec. III B	Gapless, noninteger	$O\left(\frac{1}{\epsilon}\log\frac{1}{\beta-\omega}\right)$	$\tilde{O}\left(\frac{S^2}{\epsilon(\beta-\omega)^2}\right)$
Sec. III C	Gapless, integer	$O\left(\frac{1}{\epsilon}\log\frac{1}{\beta-\omega}\right)$	$\tilde{O}\left(\frac{S^5}{\epsilon(\beta-\omega)^2}\right)$
Sec. IV B	Gapped, noninteger	$O\left(\frac{(\Delta^{-1}+S^2)S^2\omega}{\beta\epsilon}\right)$	$\tilde{O}\left(\frac{(\Delta^{-1}+S^2)S^2}{\omega\beta\epsilon}\right)$
Sec. IV C	Gapped, integer	$O\left(\frac{(\Delta^{-1}+S^4)S^2\omega}{\beta\epsilon}\right)$	$\tilde{O}\left(\frac{(\Delta^{-1}+S^4)S^2}{\omega\beta\epsilon}\right)$
Sec. V A	Gapped, noninteger	$O_{S,\beta}(\max\{\omega^{-1}\Delta^{-1},\omega\epsilon^{-1}\})$	$\tilde{O}_{S,\beta}(\omega^{-2}\Delta^{-1}+\omega^{-1}\epsilon^{-1})$
Sec. V B	Gapped, integer	$O_{S,\beta}(\max\{\omega^{-1}\Delta^{-1},\omega\epsilon^{-1}\})$	$\tilde{O}_{S,\beta}(\omega^{-2}\Delta^{-1}+\omega^{-1}\epsilon^{-1})$

of

$$\langle\psi|U^k|\psi\rangle = \sum_{s=1}^{S}|c_s|^2 e^{-2\pi i\lambda_s k} + |c_{\text{res}}|^2\langle\psi_{\text{res}}|U^k|\psi_{\text{res}}\rangle,$$

for $k = -K, \ldots, K$. This is the Fourier transform of a probability measure $f(x)$ defined as

$$f(x) \equiv \sum_{s=1}^{S}|c_s|^2\delta_{\lambda_s}(x) + \text{Residual},$$

where Residual is a positive measure with mass $|c_{\text{res}}|^2$. $\Lambda = \{\lambda_s\}_{s=1}^{S}$ can then be viewed as the dominant support of the measure $f(x)$.

At the first step, from the noisy approximation $y(k)$ to the Fourier coefficient $\hat{f}(k)$, which is obtained from averaging the output of N_{HR} repetitions of Hadamard test, one can first extract a rough estimation E_0 of Λ using an appropriate signal processing routine. Roughly speaking, at this first step, we want to estimate Λ to η precision. The signal processing routine and the determination of parameter η will be detailed in the following sections.

The algorithm then chooses a sequence of amplifying factors $(m_1, m_2, \ldots, m_\ell)$. Define $M_0 = 1$ and $M_\ell = m_1 m_2 \cdots m_\ell$. At the ℓth step, we start with $E_{\ell-1}$, the current approximation of Λ. Applying the Hadamard test to $U^{M_\ell}, \ldots, U^{M_\ell K}$ with input $|\psi\rangle$ gives rise to noisy measurements $\{y_\ell(k)\}$ of

$$\langle\psi|U^{M_\ell k}|\psi\rangle = \sum_{s=1}^{S}|c_s|^2 e^{-2\pi iM_\ell\lambda_s k} + |c_{\text{res}}|^2\langle\psi_{\text{res}}|U^{M_\ell k}|\psi_{\text{res}}\rangle. \tag{3}$$

This can be viewed as the Fourier transform of a probability measure $f_\ell(x)$:

$$f_\ell(x) \equiv \sum_{s=1}^{S}|c_s|^2\delta_{[\lambda_sM_\ell]}(x) + \text{Residual},$$

where Residual is again a positive measure with mass $|c_{\text{res}}|^2$ and $[x] \equiv (x \mod 1)$. The data $y_\ell(k)$ is a noisy approximation to the Fourier coefficient $\hat{f}_\ell(k)$. Again, with a signal processing subroutine, one can obtain an estimation Y_ℓ of the dominant support of $f_\ell(x)$, which is denoted by $\Lambda_\ell = \{[\lambda_sM_\ell]\}_{s=1}^{S}$. This estimation Y_ℓ provides information with an increased resolution and helps improve the current estimation $E_{\ell-1}$ to an update E_ℓ.

More specifically, with the spectrum estimation Y_ℓ, one divides it by M_ℓ to obtain the set $\frac{1}{M_\ell}Y_\ell$. From the property of the signal processing methods used, $\frac{1}{M_\ell}Y_\ell$ will be a union of at most S disjoint intervals $\frac{1}{M_\ell}Y_\ell = \cup_{i=1}^{S_\ell}I'_{\ell,i}$ with $\Lambda \subset \cup_{i=1}^{S_\ell}(I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell})$ for properly chosen integers $q_{\ell,i}$. Here $S_\ell \leqslant S$ is the number of intervals in $\frac{1}{M_\ell}Y_\ell$, and $I'_{\ell,i}$ is the ith disjoint interval in $\frac{1}{M_\ell}Y_\ell$. By carefully determining the amplifying factors $(m_1, m_2, \ldots, m_\ell)$, we show that at most one integer $q_{\ell,i}$ satisfies $(I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell}) \cap E_{\ell-1} \neq \varnothing$. Then we define $(I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell})$ as $I_{\ell,i}$ and form the updated estimation $E_\ell = \cup_{i=1}^{S_\ell}I_{\ell,i}$. The determination of m_ℓ will be discussed in detail in the following sections. Given that m_ℓ is chosen properly, the algorithm can be summarized as in Algorithm I.

We show in the next few sections that the determination of K, N_{HR}, m_ℓ, and η ensures that $E_\ell = \cup_{i=1}^{S_\ell}I_{\ell,i}$ enjoys the following properties:

(1) $\{I_{\ell,i}\}$ are disjoint, and $\sum_{i=1}^{S_\ell}|I_{\ell,i}| \leqslant S\eta/M_\ell$.

TABLE II. The comparison of different multiphase estimate algorithms.

Algorithm	Allow residual	Heisenberg limit	Gapless case	Integer power	Short depth	Remark
Ref. [15]	✗	✗	✗	✓	?	
Ref. [22]	✓	✗	✓	✓	✓	
Ref. [23]	✗	✓	✓	✗	✗	Large prefactor in power of S
Ref. [24]	✓	✓	✗	✗	✓	Large classical computational cost
Refs. [25–30]	✓	?	?	?	?	
Sec. III (this work)	✓	✓	✓	✓	✗	
Sec. V (this work)	✓	✓	✗	✓	✓	

Algorithm 1 Structure of robust multiple-phase estimation (RMPE) algorithm.

Input: ϵ: desired precision, β: the lower bound of dominant eigenvalues, S: the number of dominant spikes, ω: upper bound for the residual in the initial state $|\psi\rangle$

Set the initial estimation E_{-1}.

Set $M_0 = 1, \ell = 0$.

Calculate η, K, N_{HR}, and α according to ϵ, β, ω, and S.

while $\eta/M_\ell > \epsilon$ **do**

    Choose an $m_\ell \geqslant 2$ according to Lemma III.4 or Lemma III.8 and set $M_\ell = M_{\ell-1}m_\ell$ if $\ell > 0$.

    Run the circuit in Fig. 1(a) with U replaced by $U^{M_\ell k}$ for $N_{\text{HR}}/2$ times each for the real and imaginary parts and $0 \leqslant k \leqslant K$ and obtain the signal y_ℓ.

    Obtain Y_ℓ from a signal processing routine, and then several intervals $I'_{\ell,1}, \ldots, I'_{\ell,S_\ell}$ from $\frac{1}{M_\ell}Y_\ell$. The choice of m_ℓ ensures that for each $I'_{\ell,i}$, only one integer $q_{\ell,i}$ gives $(I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell}) \cap E_{\ell-1} \neq \varnothing$.

    Set $I_{\ell,i} = I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell}$ and $E_\ell = \bigcup_{i=1}^{S_\ell} I_{\ell,i}$

    $\ell \leftarrow \ell + 1$.

end while

Output: The final estimation E_L as an approximated support of Λ.

---

(2) For each $I_{\ell,i}$, the intersection $I_{\ell,i} \cap \Lambda \neq \varnothing$.

(3) $\Lambda \subset E_\ell \subset B(\Lambda, \frac{\eta}{M_\ell})$ (real-power model) or $\Lambda \subset E_\ell \subset B_{\mathbb{T}}(\Lambda, \frac{\eta}{M_\ell})$ (integer-power model).

Since the chosen m_ℓ satisfies $m_\ell \geqslant 2$, it can be deduced from the last property above that the while loop in Algorithm I ends in $O(\log(\frac{\eta}{\epsilon}))$ iterations with high probability. In Sec. III B, we show that for the real-power model, one can choose a proper $m_\ell \in [2, 4]$ according to the previous estimations. On the other hand, for the integer-power model, one can choose an appropriate m_ℓ with the help of prime numbers. Detailed explanations and analyses are provided in Sec. III C. Compared with the original version of Kitaev's method, where $m_\ell = 2$ for all ℓ, the adaptive calculation of these factors enables the proposed algorithm to address QPE problems with multiple dominant eigenvalues and a nonzero residual.

Here, we discuss some specific aspects of the signal processing routine used to extract Y_ℓ. At the ℓth iteration, we implement N_{HR} measurements of $\hat{f}_\ell(k)$ for each $k \in \{0, 1, \ldots, K\}$ such that the averaged measurement result $y_\ell(k)$ satisfies

$$|y_\ell(k) - \hat{f}_\ell(k)| \leqslant \alpha$$

for all k with high probability. The determination of parameters N_{HR}, K, and α will be elaborated in the following sections and is omitted for now. The problem of recovering $\Lambda_\ell = \{[\lambda_s M_\ell]\}_{s=1}^S$ from $\{y_\ell(k)\}_{k=0}^K$ has been extensively studied under the name line spectrum estimation, and plenty of established results for line spectrum estimations can be used. As explained earlier, Algorithm I requires that Y_ℓ satisfies the following three requirements.

(1) Y_ℓ is a union of at most S disjoint intervals such that each interval contains at least one dominant eigenvalue.

(2) Λ_ℓ is a subset of Y_ℓ.

(3) $|Y_\ell| \leqslant S\eta$ for a parameter η to be determined by the algorithm.

We provide a detailed description in Corollary III.2 and show that these requirements can indeed be satisfied for both the gapless case (Sec. III A) and the gapped case (Sec. IV A).

III. THE GAPLESS CASE

This section specializes Algorithm I to the gapless case, i.e., no gap is assumed among the dominant eigenvalues. The signal processing routine proposed in Ref. [4] satisfies the requirements 1, 2, and 3 listed above for Y_ℓ in Sec. II. After summarizing its main results in Sec. III A for completeness, we discuss the real-power model in Sec. III B and the integer-power model in Sec. III C. Figure 2 gives a graphical illustration of the gapless case algorithm.

A. Signal processing routine

Recall that y_ℓ are the averaged measurements result such that $|y_\ell(k) - \hat{f}_\ell(k)| \leqslant \alpha$, where $\hat{f}_\ell$ is defined in (3). Here $y_\ell(-k)$ is defined as the complex conjugate of $y_\ell(k)$ for $0 \leqslant k \leqslant K$, since $\hat{f}_\ell(-k)$ is the conjugate of $\hat{f}_\ell(k)$. Following [4], the spikes Λ_ℓ can be estimated by the following set:

$$X_\ell = \left\{ x : \left| \sum_{|k| \leqslant K} y_\ell(k) \hat{\phi}_p(k) e^{2\pi i k x} \right| > \frac{6\beta + 5\omega}{11} \phi_s \right\}, \quad (4)$$

where $\sigma = (\frac{1}{\pi} \ln \frac{12}{\beta - \omega})^{1/2}$, $\hat{\phi}_p(k) = \exp[-\pi(k\sigma/K)^2]$, $k \in \mathbb{Z}$, and $\phi_s = \sum_{k \in \mathbb{Z}} \hat{\phi}_p(k)$. The following theorem holds for this choice of X_ℓ:

Theorem III.1. Suppose $\alpha < \frac{\beta - \omega}{3}$ and $K \geqslant 3\tau$, then $\Lambda_\ell \subset X_\ell$ and $\max_{x \in X_\ell} \text{dist}(x, \Lambda_\ell) \leqslant \tau/K$, where $\tau = \frac{1}{\pi} \ln \frac{12}{\beta - \omega}$.

Note that it is possible that X_ℓ is a disjoint union of more than S intervals, and some of them may not contain a true spike. In the following corollary, we form a set Y_ℓ with the help of the estimations X_ℓ obtained above to meet the requirements listed in Sec. II. The proof can be found in Appendix A 1.

Corollary III.2. Using the set X_ℓ we obtained from the above signal processing routine, we can construct a set Y_ℓ that satisfies the following properties when $\eta > 3\tau/K$:

(1) $Y_\ell = \bigcup_{i=1}^{S_\ell} Y_{\ell,i}$ is the disjoint union of intervals, and $|Y_\ell| \leqslant S\eta$.

(2) For each interval $Y_{\ell,i}$, the intersection $I_{\ell,i} \cap \Lambda_\ell \neq \varnothing$.

(3) $\Lambda_\ell \subset Y_\ell \subset B_{\mathbb{T}}(\Lambda_\ell, \eta)$.

We emphasize that K dictates the estimation accuracy obtained. Even if ω is zero, the estimation error is proportional to $1/K$. Without making K larger, it is not possible to make the approximation more accurate in this signal-processing routine.

B. The real-power model

In this section, we aim to recover the eigenvalues of some Hamiltonian H assuming we have access to $e^{-2\pi iHt}|\psi\rangle$ for $t \in \mathbb{R}_+$, and U^t is an abbreviation of $e^{-2\pi iHt}$.

Without loss of generality, we can assume $\Lambda \subset [0, 0.9]$. Otherwise, one can prescale H appropriately. Hence the initialization E_{-1} is defined as $[0, 0.9]$. In this way, the property 2 can be guaranteed at step $\ell = 0$. Otherwise, if the signal processing subroutine gives an interval $I_{0,i}$ that is $[\gamma_1, \gamma_2]$ mod 1 with $\gamma_1 < 0$ and $\gamma_2 > 0$, one cannot tell whether there
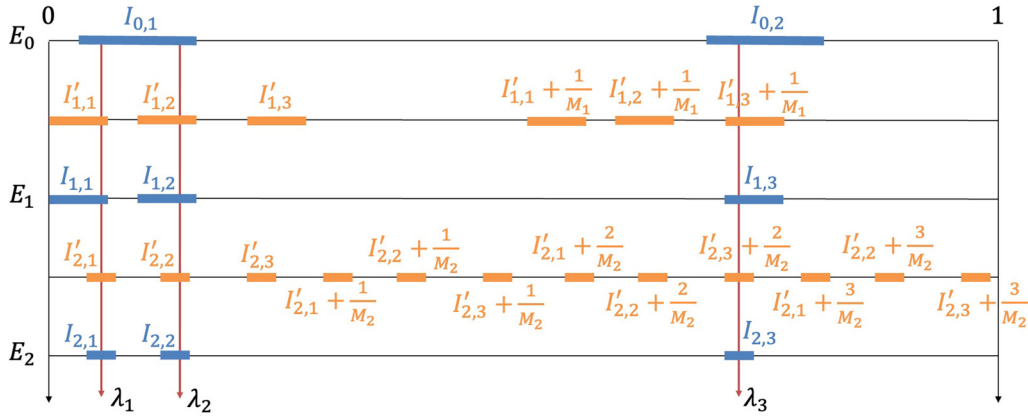
FIG. 2. Illustration of Algorithm I for the gapless case with $S = 3$. At each step, we maintain a union of at most S intervals as an estimation of Λ, shown by the horizontal blue sticks in the diagram. The yellow intervals represent $M_\ell^{-1} Y_\ell$ and its translates by integer multiples of M_ℓ^{-1}. The choice of M_ℓ ensures that only one integer $q_{\ell,i}$ satisfies $I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell} \cap E_{\ell-1} \neq \varnothing$. The union of these $I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell}$ then becomes the new estimation of Λ and is used for the next update. Note that the number of intervals in E_ℓ may increase. Due to the low resolution, λ_1 and λ_2 are covered by the same interval in E_0. As the resolution increases, they belong to two different intervals in E_1 and E_2.

is a true spike in $[0, \gamma_2]$ or $[1 + \gamma_1, 1]$. However, this can be avoided under this assumption because the eigenvalues are estimated to error level $\eta < 0.1$ at step $\ell = 0$, and the spikes near 0.9 and zero will not interfere with each other.

As previously mentioned, a vital step in Algorithm I is the determination of m_ℓ such that for each $I'_{\ell,i}$, only one integer $q_{\ell,i}$ gives $I'_{\ell,i} + \frac{q_{\ell,i}}{M_{\ell-1}m_\ell} \cap E_{\ell-1} \neq \varnothing$. Once such m_ℓ is chosen, then the properties 1, 2, and 3 are satisfied due to Corollary III.2.

Now assume that the properties 1, 2, and 3 are already satisfied at step $\ell - 1$. For any $\lambda_s \in \Lambda$, since $\lambda_s \in E_{\ell-1}$, there must be some $\lambda_s \in I'_{\ell,i} \mod \frac{1}{M_{\ell-1}m_\ell}$. Thus there exists at least one $q_{\ell,i}$ such that $I'_{\ell,i} + \frac{q_{\ell,i}}{M_{\ell-1}m_\ell} \cap E_{\ell-1} \neq \varnothing$. Choose an arbitrary such $q_{\ell,i}$ and let $I_{\ell,i} = I'_{\ell,i} + \frac{q_{\ell,i}}{M_{\ell-1}m_\ell}$, then $I_{\ell,i} \subset B(\Lambda, \frac{\eta}{2M_{\ell-1}})$ as long as $m_\ell \geqslant 2$. We also have $\Lambda \subset E_{\ell-1}$, which implies

$$I_{\ell,i} \subset B\left(\Lambda, \frac{\eta}{2M_{\ell-1}}\right) \subset B\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right).$$

Therefore, if the choice of m_ℓ satisfies

$$\left[B\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right) + \frac{q}{M_{\ell-1}m_\ell}\right] \cap E_{\ell-1} = \varnothing \ \forall q \in \mathbb{Z}\backslash\{0\}, \tag{5}$$

then we can deduce that $(I_{\ell,i} + \frac{q}{M_{\ell-1}}) \cap E_{\ell-1} = \varnothing$ for any $q \in \mathbb{Z}\backslash\{0\}$ and thus the choice of $q_{\ell,i}$ is unique. Moreover, we have

$$E_\ell = \bigcup_{i=1}^{S_\ell} I_{\ell,i} \subset B\left(\Lambda, \frac{\eta}{2M_{\ell-1}}\right) \subset B\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right).$$

Summarizing the discussion above, we obtain the following lemma:

Lemma III.3. Given the previous amplifying factor $M_{\ell-1}$ and estimation $E_{\ell-1}$ in Algorithm I, if the choice of m_ℓ satisfies (5), then the choice of $q_{\ell,i}$ in Algorithm I is unique, and the corresponding E_ℓ satisfies

$$E_\ell \subset B\left(\Lambda, \frac{\eta}{2M_{\ell-1}}\right) \subset B\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right). \tag{6}$$

Before establishing (5), we first state the following lemma, whose proof can be found in Appendix A 2.

Lemma III.4. Suppose M is a positive number and a set

$$G = \bigcup_{i=1}^{t}\left[\theta_i - \frac{\zeta_i}{M}, \theta_i + \frac{\zeta_i}{M}\right]$$

is the union of t disjoint intervals with $t \leqslant S$ and $\sum_{i=1}^{t} \zeta_i \leqslant S\zeta$. If $\zeta \leqslant \frac{1}{8S(2S-1)}$, then there must be some $m \in [2, 4]$ such that

$$\left[\theta_i - \frac{\zeta_i}{M} - \frac{q}{Mm}, \theta_i + \frac{\zeta_i}{M} - \frac{q}{Mm}\right] \cap \left[\theta_j - \frac{\zeta_j}{M}, \theta_j + \frac{\zeta_j}{M}\right]$$
$$= \varnothing \tag{7}$$

holds for all $1 \leqslant i, j \leqslant t$, and $q \in \mathbb{Z}\backslash\{0\}$. In other words, it holds that $B(G, \frac{q}{Mm}) \cap G = \varnothing$ for any $q \in \mathbb{Z}\backslash\{0\}$.

Since $E_{\ell-1} = \bigcup_{i=1}^{S_{\ell-1}} I_{\ell-1,i}$ is the union of at most S disjoint intervals and $|E_{\ell-1}| \leqslant \frac{S\eta}{M_{\ell-1}}$, its neighborhood $B(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}})$ must be t disjoint intervals with $t \leqslant S$ and $|B(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}})| \leqslant \frac{2S\eta}{M_{\ell-1}}$. This is exactly the case in Lemma III.4 when taking $G = B(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}})$, $M = M_{\ell-1}$, and $\zeta = \eta$. Therefore, the conclusion of Lemma III.4 guarantees that we can constructively find an m_ℓ such that for any $q \in \mathbb{Z}\backslash\{0\}$,

$$\left[B\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right) + \frac{q}{M_{\ell-1}m_\ell}\right] \cap B\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right) = \varnothing,$$

which is stronger than (5). Thus the requirement on m_ℓ in Algorithm I is satisfied, i.e., for each $I'_{\ell,i}$, only one integer $q_{\ell,i}$ gives $(I'_{\ell,i} + \frac{q_{\ell,i}}{M_\ell}) \cap E_{\ell-1} \neq \varnothing$ due to the choice of m_ℓ. According to Lemma III.4, one can choose any $\eta \leqslant \frac{1}{8S(2S-1)}$. We thus obtain the following theorem:

Theorem III.5. Define $\tau = \frac{1}{\pi} \log_2 \frac{12}{\beta-\omega}$ and let α be a constant that satisfies $\alpha < \frac{\beta-\omega}{3}$. For any $\eta \leqslant \frac{1}{8S(2S-1)}$,

if $K \geqslant 3\eta^{-1}\tau$ and

$$N_{\mathrm{HR}} = 2\left\lceil \frac{4}{\alpha^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K+1)\right]\right\rceil,$$

$$(8)$$

and the signal processing algorithm in Sec. III A is used for spectrum estimation in Algorithm I, then with probability at least $1 - \rho$, the output E_L satisfies

$$\Lambda \subset E_L \subset B(\Lambda, \epsilon). \tag{9}$$

The maximum runtime and total runtime are

$$T_{\max} = O(\eta K \epsilon^{-1}), \quad T_{\mathrm{total}} = O(\eta K^2 N_{\mathrm{HR}} \epsilon^{-1}). \tag{10}$$

Proof. For the N_{HR} defined above, one knows from Hoeffding's inequality that for any ℓ and k,

$$\mathbb{P}(|y_\ell(k) - \hat{f}_\ell(k)| < \alpha) > 1 - \frac{\rho}{\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right)(K+1)}. \tag{11}$$

Thus $|y_\ell(k) - \hat{f}_\ell(k)| < \alpha$ is true for all ℓ and k with probability at least $1 - \rho$ by the union bound, and the rest follows from Lemmas III.3 and III.4. ∎

Corollary III.6. By choosing $\eta = \frac{1}{8S(2S-1)}$ and $K = 3\eta^{-1}\tau$, the maximum runtime is $T_{\max} = O(\log(\frac{1}{\beta-\omega})\epsilon^{-1})$ and the total complexity is $T_{\mathrm{total}} = \tilde{O}((\beta-\omega)^{-2}S^2\epsilon^{-1})$, which achieves the Heisenberg limit.

C. The integer-power model

This section shows that if U is given as a black box and only (positive) integer powers of U can be accessed, Algorithm I can be applied with the help of prime numbers. As for the initialization, we take $E_{-1} = [0, 1]$. Similar with Sec. III C, we can verify the uniqueness of $q_{\ell,i}$ in Algorithm I with help of the following lemma:

Lemma III.7. Given integer $M_{\ell-1}$ and the set $E_{\ell-1}$, if the choice of integer m_ℓ satisfies

$$\left[B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right) + \frac{q}{M_{\ell-1}m_\ell}\right] \cap E_{\ell-1} = \varnothing \mod 1$$

for all $q \in \mathbb{Z}$ that $(m_\ell M_{\ell-1}) \nmid q$, $\tag{12}$

then the choice of $q_{\ell,i}$ in Algorithm I is unique, and the corresponding E_ℓ satisfies

$$E_\ell \subset B_{\mathbb{T}}\left(\Lambda, \frac{\eta}{2M_{\ell-1}}\right) \subset B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{2M_{\ell-1}}\right). \tag{13}$$

This lemma is the modulo-1 version of Lemma III.3, which can be directly obtained from Lemma III.3 since $\{m_\ell\}$ are all integers and all the sets we consider here are in $\mathbb{T}$. In what follows, we denote the ith prime number by p_i. Here we define $p_0 = 1$ to unify the notations. The factor m_ℓ in the algorithm above can be chosen with the help of the following lemma. The proof is provided in Appendix A 3.

Lemma III.8. For any $(\theta_1, \ldots, \theta_t) \in \mathbb{R}^t$ and $\frac{t(t-1)}{2} + 1$ prime numbers $p_1 < p_2 < \cdots < p_{\frac{t(t-1)}{2}+1}$, there is at least one

$p_l \in \{p_1, p_2, \ldots, p_{\frac{t(t-1)}{2}+1}\}$ such that

$$\min_{\substack{1 \leqslant i \leqslant j \leqslant t \\ p_l \nmid k}} \left| |\theta_i - \theta_j| - \frac{k}{p_l}\right| \geqslant \frac{1}{2p_{\frac{t(t-1)}{2}}p_{\frac{t(t-1)}{2}+1}}. \tag{14}$$

Remark III.9. The result also applies to mutually prime integers that are not necessarily prime themselves. Moreover, this procedure only involves classical computing with polynomial complexity in terms of t, so it can be implemented efficiently.

Based on Lemma III.8, one can show that there is at least one $m_\ell \in \{p_1, p_2, \ldots, p_{\frac{S(S-1)}{2}+1}\}$ that satisfies the requirements for Algorithm I if

$$\eta < \left(3Sp_{\frac{S(S-1)}{2}}p_{\frac{S(S-1)}{2}+1}\right)^{-1}.$$

Since $p_n = O(n\log n)$ [33,34], the requirement for η is $\eta = O(S^{-5}\log^{-2}(S))$. More precisely, in Ref. [35] it was proved that $p_n < n(\ln n + 2\ln\ln n)$ for $n \geqslant 4$, thus by direct calculation one knows $p_n \leqslant 2n(\ln n + 1)$ for any $n \geqslant 1$. Hence, it suffices to have

$$\eta \leqslant \min\left\{\frac{1}{6}, \frac{1}{3S^5(0.31 + 2\ln S)^2}\right\}.$$

To establish (13), we also need the following lemma, which is proved in Appendix A 4.

Lemma III.10 Suppose

$$\eta < \left(3Sp_{\frac{S(S-1)}{2}}p_{\frac{S(S-1)}{2}+1}\right)^{-1},$$

then for step ℓ ($\ell \geqslant 1$) in Algorithm I, one can choose a prime number m_ℓ such that

$$\left[B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}\right) + \frac{q}{m_\ell M_{\ell-1}}\right]$$

$$\cap B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}\right) = \varnothing \mod 1, \text{ if } (m_\ell M_{\ell-1}) \nmid q,$$

$$(15)$$

which guarantees the construction of E_ℓ in the algorithm.

Similar to Theorem III.5, one directly obtains the following theorem by applying Hoeffding's inequality and the union bound to Lemmas III.7 and III.10.

Theorem III.11 Define $\tau = \frac{1}{\pi}\ln\frac{12}{\beta-\omega}$ and let α be a constant that satisfies $\alpha < \frac{\beta-\omega}{3}$. For any

$$\eta \leqslant \min\left\{\frac{1}{6}, \frac{1}{3S^5(0.31 + 2\ln S)^2}\right\},$$

if $K \geqslant 3\eta^{-1}\tau$ and

$$N_{\mathrm{HR}} = 2\left\lceil \frac{4}{\alpha^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K+1)\right]\right\rceil,$$

$$(16)$$

and the signal processing algorithm in Sec. III A is used for spectrum estimation in Algorithm I, then with probability at least $1 - \rho$, the output E_L satisfies

$$\Lambda \subset E_L \subset B_{\mathbb{T}}(\Lambda, \epsilon). \tag{17}$$

The maximum runtime and total runtime are

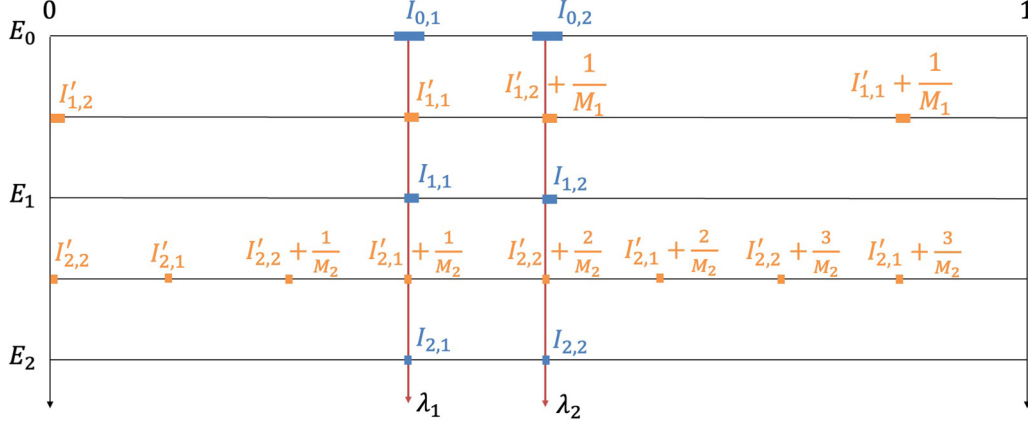$$T_{\max} = O(\eta K \epsilon^{-1}), \quad T_{\mathrm{total}} = O(\eta K^2 N_{\mathrm{HR}} \epsilon^{-1}). \tag{18}$$

FIG. 3. Illustration of Algorithm I in the gapped case with $S = 2$. Based on prior knowledge of the separation between λ_1 and λ_2, one can obtain much tighter initial-stage estimations than the gapless case (see Fig. 2) if the residual ω is small. As a result, much fewer steps are needed to reach the desired precision ϵ, and the maximum runtime is reduced.

Corollary III.12 By choosing

$$\eta = \min\left\{\frac{1}{6}, \frac{1}{3S^5(0.31 + 2\ln S)^2}\right\}$$

and $K = 3\eta^{-1}\tau$, the maximum runtime is $T_{\max} = \tilde{O}(\log(\frac{1}{\beta - \omega})\epsilon^{-1})$ and the total complexity is $T_{\text{total}} = \tilde{O}((\beta - \omega)^{-2}S^5\epsilon^{-1})$, which achieves the Heisenberg limit.

IV. THE GAPPED CASE

This section specializes Algorithm I to the simpler gapped case, i.e., there is a minimum separation among the dominant eigenvalues defined as

$$\Delta = \min_{1 \leqslant s \leqslant s' \leqslant S, \quad n \in \mathbb{Z}} |\lambda_s - \lambda_{s'} - n|. \quad (19)$$

We first review the famous signal processing routine ESPRIT in Sec. IV A. In particular, the number of frequencies K needed only depends on S and Δ. With the help of this specific version of ESPRIT, we show that Algorithm I meets the requirements 1, 2, 3, and 4. In particular, the maximum runtime $T_{\max}$ can scale like ω/β. The real-power and integer-power models are investigated in Secs. IV B and IV C, respectively. A graphical illustration of the algorithm for the gapless case is given in Fig. 3.

A. Signal processing routine

Since a finite-size gap between the dominant eigenvalues is available, multiple signal processing algorithms can be used in this setting, such as those in Refs. [36–38]. Here, we adopt a particular version of the ESPRIT algorithm discussed in Ref. [38].

Suppose that $K \geqslant 4S$ is even. Recall that, at the ℓth step, the data $y_\ell(k)$ collected from the Hadamard test results satisfy $|y_\ell(k) - \hat{f}_\ell(k)| \leqslant \alpha$, where $\hat{f}_\ell$ is defined in (3). The first step

in ESPRIT is to construct the following Hankel matrix:

$$H_\ell = \begin{bmatrix} y_\ell(0) & y_\ell(1) & \cdots & y_\ell\left(\frac{K}{2}\right) \\ y_\ell(1) & y_\ell(2) & \cdots & y_\ell\left(\frac{K}{2} + 1\right) \\ \vdots & \vdots & \ddots & \vdots \\ y_\ell\left(\frac{K}{2}\right) & y_\ell\left(\frac{K}{2} + 1\right) & \cdots & y_\ell(K) \end{bmatrix}.$$

Then one applies the singular value decomposition (SVD) to H_ℓ and obtains

$$H_\ell = [U_\ell, U_\ell^\perp]\Sigma[V_\ell, V_\ell^\perp]^*,$$

where U_ℓ has S columns. Let the first $K/2$ rows of U_ℓ be $U_\ell^{(0)}$ and let its last $K/2$ rows be $U_\ell^{(1)}$. The last step is to compute the eigenvalues $\{\mu_1, \ldots, \mu_S\}$ of the matrix $(U_\ell^{(0)})^\dagger U_\ell^{(1)}$, where $(U_\ell^{(0)})^\dagger$ is the Moore-Penrose pseudo-inverse of $U_\ell^{(0)}$. The output is the set

$$\tilde{\Lambda}_\ell := \left\{-\frac{1}{2\pi}\arg(\mu_1), \ldots, -\frac{1}{2\pi}\arg(\mu_S)\right\}.$$

The following theorem is proved in Ref. [38] for this particular version of ESPRIT:

Theorem 4.1. Suppose $K \geqslant 4S$ is even and $K \geqslant \frac{4}{\Delta}$, where $\tilde{\Delta} \geqslant \Delta_\ell := \min_{1 \leqslant i < j \leqslant S} |M_\ell\lambda_i - M_\ell\lambda_j|_s$. For any constant $C \in (2, \frac{K\tilde{\Delta}}{2})$, if

$$\omega + \alpha \leqslant \frac{C\beta}{8(C-1)\sqrt{2S}}\sqrt{1 - \frac{2(C-1)S}{CK}}, \quad (20)$$

then

$$\text{md}(\tilde{\Lambda}_\ell, \Lambda_\ell)$$

$$\leqslant \frac{40S^2}{\beta}\sqrt{\frac{C^3(2+K)}{(C-1)^3 K}}\left(1 - \frac{2CS}{(C-1)K}\right)^{-1}(\omega + \alpha), \quad (21)$$

where $\text{md}(\cdot, \cdot)$ denotes the matching distance between two finite sets with the same cardinality.

We emphasize that the approximation error is controlled by the noise in the gapped case. The number K of frequencies only needs to be proportional to S, which is considered a fixed number. This means that the maximum depth of the circuit can

be kept fixed. Only more repetitions are needed to bring down the noise level.

When ω is sufficiently small compared with β, (20) can always be satisfied by properly choosing α. The set Y_ℓ needed in the main algorithm can then be defined as

$$Y_\ell = B_{\mathbb{T}}\left(\tilde{\Lambda}_\ell, \frac{\eta}{2}\right),$$

with η satisfying

$$\frac{80S^2}{\beta}\sqrt{\frac{C^3(2+K)}{(C-1)^3 K}}\left(1 - \frac{2CS}{(C-1)K}\right)^{-1}(\omega + \alpha) < \eta.$$

This Y_ℓ is guaranteed to satisfy the three properties stated in Corollary III.2.

In the following sections, we provide modified versions of the algorithms described in Sec. III and show that, when ω is close to zero, one can further reduce the maximal runtime by a factor $\Omega(\omega/\beta)$.

B. The real-power model

The following theorem states the $T_{\max}$ and T_{total} bounds for our RMPE algorithm for the real-power model under the gapped case. The proof can be found in Appendix A 5.

Theorem IV.2. Let $\tilde{\Delta} < \min\{\frac{1}{8S(2S-1)}, \Delta\}, K > 4/\tilde{\Delta}$, and η be any number such that

$$\frac{80S^2}{\beta}\sqrt{\frac{C^3(2+K)}{(C-1)^3 K}}\left(1 - \frac{2CS}{(C-1)K}\right)^{-1}\omega$$
$$< \eta < \frac{1}{8S(2S-1)}, \tag{22}$$

where C is a constant such that $C \in (2, \frac{K\tilde{\Delta}}{2})$. Suppose that α satisfy

$$\alpha < \min\left\{\frac{C\beta}{8(C-1)\sqrt{2S}}\sqrt{1 - \frac{2(C-1)S}{CK}},\right.$$
$$\left.\frac{\beta}{80S^2}\sqrt{\frac{(C-1)^3 K}{C^3(2+K)}}\left(1 - \frac{2CS}{(C-1)K}\right)\eta\right\} - \omega, \tag{23}$$

and N_{HR} satisfy

$$N_{\text{HR}} = 2\left\lceil \frac{4}{\alpha^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K+1)\right]\right\rceil, \tag{24}$$

and the ESPRIT routine described in Sec. IV A is used for spectrum estimation in Algorithm I, then with probability at least $1 - \rho$, the output E_L satisfies

$$\Lambda \subset E_L \subset B(\Lambda, \epsilon). \tag{25}$$

The maximum runtime and total runtime are

$$T_{\max} = O(\eta K\epsilon^{-1}), \quad T_{\text{total}} = O(\eta K^2 N_{\text{HR}}\epsilon^{-1}). \tag{26}$$

When the left inequality of (22) is tight, i.e., $\eta = \Theta(\frac{\omega}{\beta}S^2)$, we have the following corollary that gives a smaller prefactor of $T_{\max}$ compared with Corollary III.6:

Corollary IV.3. By setting $\eta = \Theta(\frac{\omega}{\beta}S^2)$, the complexity bounds given in the theorem are

$$T_{\max} = O((\Delta^{-1} + S^2)S^2\omega\beta^{-1}\epsilon^{-1}),$$
$$T_{\text{total}} = \tilde{O}((\Delta^{-1} + S^2)S^2\omega^{-1}\beta^{-1}\epsilon^{-1}). \tag{27}$$

The prefactor of $T_{\max}$ is proportional to $\omega\beta^{-1}$, the ratio between the energy of the residual in $|\psi\rangle$ and the energy lower bound of the dominant eigenmodes.

Therefore, the maximum circuit depth prefactor can be made small when this ratio is sufficiently small. However, T_{total} will increase for small ω, so there is a trade-off between maximal runtime and total runtime, and one should choose η and ω appropriately in practice.

It is worth noticing that ω is not fixed when the initial state $|\psi\rangle$ is given. Instead, we can adjust ω as long as $\omega \geqslant |c_{\text{res}}|^2$ and (22) is feasible.

C. The integer-power model

Similar to Sec. III C, one can implement the algorithm with only integer powers of U. The following theorem states the $T_{\max}$ and T_{total} bounds for our RMPE algorithm for the integer-power model under the gapped case. We give the proof of it in Appendix A 6.

Theorem IV.4. Let

$$\tilde{\Delta} < \min\left\{\frac{1}{2p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}}, \Delta\right\},$$

$K > 4/\tilde{\Delta}$, and η be any number such that

$$\frac{80S^2}{\beta}\sqrt{\frac{C^3(2+K)}{(C-1)^3 K}}\left(1 - \frac{2CS}{(C-1)K}\right)^{-1}\omega$$
$$< \eta < \frac{1}{4Sp_{\frac{S(S-1)}{2}}p_{\frac{S(S-1)}{2}+1}},$$

where C is a constant such that $C \in (2, \frac{K\Delta}{2})$. Let α satisfy

$$\alpha < \min\left\{\frac{C\beta}{8(C-1)\sqrt{2S}}\sqrt{1 - \frac{2(C-1)S}{CK}},\right.$$
$$\left.\frac{\beta}{80S^2}\sqrt{\frac{(C-1)^3 K}{C^3(2+K)}}\left(1 - \frac{2CS}{(C-1)K}\right)\eta\right\} - \omega,$$

and

$$N_{\text{HR}} = 2\left\lceil \frac{4}{\alpha^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K+1)\right]\right\rceil. \tag{28}$$

Suppose ESPRIT is used for spectrum estimation in Algorithm I and m_ℓ is chosen according to Lemma III.10, then with probability at least $1 - \rho$, the output E_L satisfies

$$\Lambda \subset E_L \subset B(\Lambda, \epsilon). \tag{29}$$

The maximum runtime and total runtime are

$$T_{\max} = O(\eta K\epsilon^{-1}), \quad T_{\text{total}} = O(\eta K^2 N_{\text{HR}}\epsilon^{-1}). \tag{30}$$

Corollary IV.5. When ω is sufficiently close to zero, one can take $\eta = \Theta(\frac{\omega}{\beta}S^2)$ and $K = \Theta(\frac{4}{\tilde{\Delta}})$, then we have

$$T_{\max} = \tilde{O}((\Delta^{-1} + S^4)S^2\omega\beta^{-1}\epsilon^{-1}),$$

$$T_{\text{total}} = \tilde{O}((\Delta^{-1} + S^4)S^2\omega^{-1}\beta^{-1}\epsilon^{-1}), \quad (31)$$

Here, the maximum circuit depth prefactor again scales like $O(\omega\beta^{-1})$.

V. HYBRID ALGORITHM WITH IMPROVED PREFACTOR

Assuming a finite spectral gap Δ, the results in Sec. IV give a prefactor $\omega\beta^{-1}$ in $T_{\max}$ (see Corollaries IV.3 and IV.5). However, when the gap Δ is a small, the prefactor Δ^{-1} in $T_{\max}$ is undesirable. On the other hand, the results in Sec. III work for an arbitrarily small spectral gap but cannot provide a prefactor as in Corollaries IV.3 and IV.5 due to the signal processing technique used.

In this section, we propose to combine the methods in Secs. III and IV under the general framework of Algorithm I. Intuitively, if the gap Δ is small, one first applies the signal processing technique in Sec. III A for certain iterations as a burn-in period and then switches to the ESPRIT routine in Sec. IV A. We prove that this hybrid method provides an improved prefactor in $T_{\max}$.

A. The real-power model

The following theorem gives the detailed description and complexity bound of the hybrid algorithm when we have access to the real powers of the given unitary U, which is proved in Sec. A 7.

Theorem V.1. Let $\tilde{\Delta}$ be any number in $(0, \frac{1}{8S(2S-1)})$, $\tilde{M} := \tilde{\Delta}/\Delta$, integer $K_2 > 4/\tilde{\Delta}$, and η be any number such that

$$\frac{80S^2}{\beta}\sqrt{\frac{C^3(2+K_2)}{(C-1)^3K_2}}\left(1 - \frac{2CS}{(C-1)K_2}\right)^{-1}\omega$$

$$< \eta < \frac{1}{8S(2S-1)}, \quad (32)$$

where C is a constant such that $C \in (2, \frac{K_2\tilde{\Delta}}{2})$.

In Algorithm I, when $M_\ell \leqslant \tilde{M}$, we use the signal processing subroutine in Sec. III with parameters $\tau = \frac{1}{\pi}\ln\frac{12}{\beta-\omega}$, $\alpha_1 = \frac{\beta-\omega}{3}$, $K_1 = \lceil 3\eta^{-1}\tau\rceil$, and

$$N_{\text{HR},1} = 2\left\lceil\frac{4}{\alpha_1^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K_1+1)\right]\right\rceil.$$

$$(33)$$

When $M_\ell > \tilde{M}$, we use ESPRIT as the signal processing subroutine with parameters $\tilde{\Delta}$, α_2, K_2, $N_{\text{HR},2}$, where

$$\alpha_2 < \min\left\{\frac{C\beta}{8(C-1)\sqrt{2S}}\sqrt{1 - \frac{2(C-1)S}{CK_2}},\right.$$

$$\left.\frac{\beta}{80S^2}\sqrt{\frac{(C-1)^3K_2}{C^3(2+K_2)}}\left(1 - \frac{2CS}{(C-1)K_2}\right)\eta\right\} - \omega, \quad (34)$$

$$N_{\text{HR},2} = 2\left\lceil\frac{4}{\alpha_2^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K_2+1)\right]\right\rceil.$$

$$(35)$$

Under this setting, the maximal runtime of the algorithm is

$$T_{\max} = O\left(\max\left\{\log\left(\frac{1}{\beta}\right)\eta^{-1}\tilde{\Delta}\Delta^{-1}, \eta\tilde{\Delta}^{-1}\epsilon^{-1}\right\}\right), \quad (36)$$

and the total runtime is

$$T_{\text{total}} = \tilde{O}(\eta^{-2}\beta^{-2}\tilde{\Delta}\Delta^{-1} + \eta\omega^{-2}\tilde{\Delta}^{-2}\epsilon^{-1}). \quad (37)$$

Corollary V.2 If we choose $\tilde{\Delta} = \frac{1}{8S(2S-1)}$ and $\eta = \Theta(\omega\beta^{-1}S^2)$ in Theorem V.1, then the maximal and total runtime of the algorithm are

$$T_{\max} = O_{S,\beta}(\max\{\omega^{-1}\Delta^{-1}, \omega\epsilon^{-1}\}),$$

$$T_{\text{total}} = \tilde{O}_{S,\beta}(\omega^{-2}\Delta^{-1} + \omega^{-1}\epsilon^{-1}). \quad (38)$$

In the case that S and $\frac{1}{\beta}$ are moderate values that can be treated as constants, this scaling O is better than the one in Corollary IV.3. This improves the complexity from $\Delta^{-1}\epsilon^{-1}$ to $\Delta^{-1} + \epsilon^{-1}$ while retaining the prefactor ω in front of ϵ^{-1} in $T_{\max}$.

B. The integer-power model

Now, we give the analysis of the hybrid algorithm for the integer-power model. The proof for the following result can be found in Appendix A 8.

Theorem V.3. Let $\tilde{M} \geqslant 2$ such that

$$\tilde{\Delta} := \tilde{M}\Delta < \frac{1}{4Sp_{\frac{S(S-1)}{2}}p_{\frac{S(S-1)}{2}+1}}.$$

Let $\eta \leqslant \frac{1}{6}$ such that

$$\frac{80S^2}{\beta}\sqrt{\frac{C^3(2+K_2)}{(C-1)^3K_2}}\left(1 - \frac{2CS}{(C-1)K_2}\right)^{-1}\omega$$

$$< \eta < \frac{1}{4Sp_{\frac{S(S-1)}{2}}p_{\frac{S(S-1)}{2}+1}}, \quad (39)$$

where K_2 is an integer that satisfies $K_2 > 4/\tilde{\Delta}$ and C is a constant such that $C \in (2, \frac{K_2\tilde{\Delta}}{2})$.

Under the framework of Algorithm I, the hybrid algorithm uses the signal processing subroutine in Sec. III A with parameters $\tau = \frac{1}{\pi}\ln\frac{12}{\beta-\omega}$, $\alpha_1 = \frac{\beta-\omega}{3}$, $K_1 = \lceil 3\eta^{-1}\tau\rceil$, and

$$N_{\text{HR},1} = 2\left\lceil\frac{4}{\alpha_1^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil\log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K_1+1)\right]\right\rceil,$$

$$(40)$$

when $M_\ell \leqslant \tilde{M}$.

When $M_\ell > \tilde{M}$, the hybrid algorithm uses ESPRIT (see Sec. IV A) as the signal processing subroutine with parameters $\tilde{\Delta}$, α_2, K_2, and $N_{HR,2}$, where

$$\alpha_2 < \min\left\{ \frac{C\beta}{8(C-1)\sqrt{2S}}\sqrt{1 - \frac{2(C-1)S}{CK_2}}, \right.$$

$$\left. \frac{\beta}{80S^2}\sqrt{\frac{(C-1)^3 K_2}{C^3(2+K_2)}}\left(1 - \frac{2CS}{(C-1)K_2}\right)\eta \right\} - \omega, \quad (41)$$

and

$$N_{HR,2} = 2\left\lceil \frac{4}{\alpha_2^2}\left[\ln\frac{4}{\rho} + \ln\left(\left\lceil \log_2\frac{\eta}{\epsilon}\right\rceil + 1\right) + \ln(K_2 + 1)\right]\right\rceil. \quad (42)$$

Then with probability at least $1 - \rho$, the output E_L satisfies

$$\Lambda \subset E_L \subset B(\Lambda, \epsilon). \quad (43)$$

The maximal runtime of the algorithm is

$$T_{\max} = \tilde{O}_S(\max\{\eta^{-1}\tilde{\Delta}\Delta^{-1}, \eta\tilde{\Delta}^{-1}\epsilon^{-1}\}) \quad (44)$$

and the total runtime is

$$T_{\text{total}} = \tilde{O}_S(\eta^{-2}\tilde{\Delta}\Delta^{-1}N_{HR,1} + \eta\tilde{\Delta}^{-2}\epsilon^{-1}N_{HR,2}). \quad (45)$$

Corollary V.4. If we choose

$$\tilde{\Delta} = \frac{1}{4S p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}}$$

and $\eta = \Theta(\omega\beta^{-1}S^2)$ in Theorem V.3, then the maximal and total runtime of the algorithm are

$$T_{\max} = O_{S,\beta}(\max\{\omega^{-1}\Delta^{-1}, \omega\epsilon^{-1}\}),$$
$$T_{\text{total}} = \tilde{O}_{S,\beta}(\omega^{-2}\Delta^{-1} + \omega^{-1}\epsilon^{-1}). \quad (46)$$

The asymptotic scaling of Corollary V.4 is better than the one obtained in Corollary IV.5: the maximum runtime is improved from $\Delta^{-1}\epsilon^{-1}$ to $\Delta^{-1} + \epsilon^{-1}$, while the ϵ^{-1} term still enjoys the prefactor ω.

VI. DISCUSSIONS

This paper proposes robust multiple-phase estimation (RMPE) algorithms for estimating multiple eigenvalues. The proposed algorithms address both the gapless case in Sec. III and the gapped case in Secs. IV and V and for both the integer-power and real-power models.

When the problem is gapless, according to Sec. III, the number of frequencies K needs to be increased if one needs an estimation with a smaller error level η from the signal processing routine, which prevents us from reducing the maximum runtime even when the magnitude of the residual ω is close to zero. When a finite-size spectral gap is assumed, the number of frequencies K is a constant independent of η. An immediate direction is to explore signal processing algorithms that can improve K's dependence on η and the spectral gap. This enables us to reduce the prefactor in the maximum runtime when ω is small.

Another relevant problem comes from the bound in Lemma III.8:

$$\max_{\substack{1 \leqslant l \leqslant \frac{t(t-1)}{2}+1}} \min_{\substack{1 \leqslant i \leqslant j \leqslant t \\ p_l \nmid k}} \left| |\theta_i - \theta_j| - \frac{k}{p_l} \right| \geqslant \frac{1}{2 p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1}},$$

for any $\{\omega\}_{i=1}^t$. However, one can also show that

$$\inf_{(\theta_i)_{i=1}^t} \max_{1 \leqslant l \leqslant t'} \min_{\substack{1 \leqslant i \leqslant j \leqslant t \\ p_l \nmid k}} \left| |\theta_i - \theta_j| - \frac{k}{p_l} \right| = \gamma_t(t') > 0, \quad (47)$$

for any $t' > t$. The conclusion of Lemma III.8 is thus

$$\gamma_t\left(\frac{1}{2}t(t-1) + 1\right) \geqslant \frac{1}{2 p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1}}.$$

It would be interesting to find a sharper estimation for $\gamma_t(t')$ and, thereby, the optimal choice of t' in terms of the overall complexity.

APPENDIX: PROOFS

1. Proof of Corollary III.2

Proof. All the sets in the proof are in a modulo-1 sense because X_ℓ can be viewed as a subset of $\mathbb{T}$. Since X_ℓ is the level set of a continuous function, it can be written as the disjoint union of finitely many intervals:

$$X_\ell = \bigcup_{i=1}^q (a_i, b_i),$$

where $0 \leqslant a_1 \leqslant b_1 \leqslant a_2 \leqslant \cdots \leqslant a_q \leqslant b_q \leqslant 1$. Let $J = \{1 \leqslant j \leqslant q : a_j - b_{j-1} < \tau/K\}$, where $a_1 - b_0 < \tau/K$ means $a_1 + 1 - b_q < \tau/K$ since the intervals are in modulo-1 sense. Then we may define

$$Y_\ell = \left(\bigcup_{i=1}^q [a_i, b_i]\right) \cup \left(\bigcup_{j\in J} [b_{j-1}, a_j]\right).$$

In this way, we have $\Lambda_\ell \subset X_\ell \subset Y_\ell$, and $Y_\ell \subset B_\mathbb{T}(X_\ell, \frac{\tau}{2L}) \subset B_\mathbb{T}(\Lambda_\ell, \frac{3\tau}{2L}) \subset B_\mathbb{T}(\Lambda_\ell, \frac{\eta}{2})$, which means $|Y_\ell| \leqslant 3S\tau/K \leqslant S\eta$.

By definition, Y_ℓ is the disjoint union of some intervals of the form $[a_{i_n}, b_{j_n}]$, where $a_{i_n} - b_{j_{n-1}} \geqslant a_{i_n} - b_{i_{n-1}} \geqslant \tau/K$ and similarly $a_{i_{n+1}} - b_{j_n} \geqslant \tau/K$. If it does not contain any spike, then $\frac{1}{2}(a_{i_n} + b_{i_n})$ will be at least τ/K away from any spike. However, we know $\frac{1}{2}(a_{i_n} + b_{i_n}) \in X_\ell$, which contradicts with Theorem III.1. ∎

2. Proof of Lemma III.4

Proof. For fixed i and j, we define the following set:

$$R_{ij} := [2,4] \cap \left\{ m : \exists\, q \in \mathbb{Z} \setminus \{0\} \text{ such that } \left| \theta_i - \theta_j - \frac{q}{Mm} \right| \right.$$
$$\left. \leqslant \frac{1}{M}(\zeta_i + \zeta_j) \right\}, \tag{A1}$$

which is the set of all m that violate (7).

First, consider the case $i = j$. From the bound of ζ, we can deduce $\frac{1}{Mm} \geqslant \frac{1}{4M} \geqslant \frac{2S\zeta}{M} > \frac{1}{M}(2\zeta_i)$, and thus $R_{ii} = \varnothing$ for every i.

For the case $i \neq j$, we can assume $\theta_i > \theta_j$ without loss of generality because clearly $R_{ij} = R_{ji}$. Notice that the existence of such i and j implies $t \geqslant 2$. In this case, the bound of ζ again gives $\theta_i - \theta_j + \frac{1}{Mm} > \frac{1}{Mm} \geqslant \frac{1}{4M} \geqslant \frac{S\zeta}{M} > \frac{1}{M}(\zeta_i + \zeta_j)$, which means that only positive q can violate (7). Therefore, we can rewrite (A1) as

$$R_{ij} = [2,4] \cap \left(\bigcup_{q=1}^{\infty} \left[\frac{q}{M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)}, \right. \right.$$
$$\left. \left. \frac{q}{M(\theta_i - \theta_j) - (\zeta_i + \zeta_j)} \right] \right). \tag{A2}$$

We consider two cases to estimate $|R_{ij}|$. In the case that $M(\theta_i - \theta_j) \leqslant \frac{1}{4} - (\zeta_i + \zeta_j)$, we have

$$\frac{q}{M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)} \geqslant 4$$

for every q, so $|R_{ij}| = 0$. Another case is that $M(\theta_i - \theta_j) > \frac{1}{4} - (\zeta_i + \zeta_j)$, then the maximal q that

$$\frac{q}{M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)} \leqslant 4$$

is $q_{\max} \leqslant 4[M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)]$. Therefore,

$$|R_{ij}| \leqslant \sum_{q=1}^{q_{\max}} \left| \frac{q}{M(\theta_i - \theta_j) - (\zeta_i + \zeta_j)} - \frac{q}{M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)} \right|$$
$$= \frac{2(\zeta_i + \zeta_j) \sum_{q=1}^{q_{\max}} q}{[M(\theta_i - \theta_j) - (\zeta_i + \zeta_j)][M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)]}$$
$$\leqslant \frac{2(\zeta_i + \zeta_j)\{4[M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)]\}^2}{[M(\theta_i - \theta_j) - (\zeta_i + \zeta_j)][M(\theta_i - \theta_j) + (\zeta_i + \zeta_j)]}$$
$$= 32(\zeta_i + \zeta_j) \left(1 + \frac{2(\zeta_i + \zeta_j)}{M(\theta_i - \theta_j) - (\zeta_i + \zeta_j)} \right)$$
$$< 32(\zeta_i + \zeta_j) \left(1 + \frac{2(\zeta_i + \zeta_j)}{\frac{1}{4} - 2(\zeta_i + \zeta_j)} \right)$$
$$= \frac{32(\zeta_i + \zeta_j)}{1 - 8(\zeta_i + \zeta_j)}. \tag{A3}$$

Taking all pairs (i, j) into account, we have

$$\left| \bigcup_{1 \leqslant i,j \leqslant t} R_{ij} \right| = \left| \bigcup_{1 \leqslant i < j \leqslant t} R_{ij} \right| \leqslant \sum_{1 \leqslant i < j \leqslant t} \frac{32(\zeta_i + \zeta_j)}{1 - 8(\zeta_i + \zeta_j)}$$
$$< \sum_{1 \leqslant i < j \leqslant t} \frac{32(\zeta_i + \zeta_j)}{1 - 8S\zeta} = \frac{32(t-1)}{1 - 8S\zeta} \sum_{i=1}^{t} \zeta_i$$
$$\leqslant \frac{32(S-1)S\zeta}{1 - 8S\zeta} \leqslant 2 = 4 - 2, \tag{A4}$$

which means that

$$[2,4] \setminus \left(\bigcup_{1 \leqslant i,j \leqslant t} R_{ij} \right) \neq \varnothing, \tag{A5}$$

and any element m of this set satisfies (7). $\blacksquare$

3. Proof of Lemma III.8

Proof. When $i = j$, it is straightforward to see

$$\left| |\theta_i - \theta_j| - \frac{k}{p_l} \right| \geqslant \frac{1}{2 p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1}}$$

when $p_l \nmid k$, and the case $t = 1$ is thus verified. In the following, we assume $t \geqslant 2$ and prove

$$\min_{\substack{1 \leqslant i < j \leqslant t \\ p_l \nmid k}} \left| |\theta_i - \theta_j| - \frac{k}{p_l} \right| \geqslant \frac{1}{2 p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1}}$$

by contradiction. Suppose there is some $(\theta_1, \ldots, \theta_t) \in \mathbb{R}^t$ such that (14) does not hold. Then for any $p_l \in \{p_1, p_2, \ldots, p_{\frac{t(t-1)}{2}+1}\}$, there are some (i_l, j_l) such that

$$\left| |\theta_i - \theta_j| - \frac{k_l}{p_l} \right| < 1 \big/ \left(2 p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1} \right).$$

for some $p_l \nmid k_l$. Since there are at most $t(t-1)/2$ different values that $|\theta_i - \theta_j|$ can take, but there are $t(t-1)/2 + 1$ different prime numbers p_l, there must be some l and l' with $l \neq l'$ such that

$$|k_l/p_l - k_{l'}/p_{l'}| < 1 \big/ \left(p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1} \right).$$

Hence, $k_l/p_l = k_{l'}/p_{l'}$, otherwise

$$|k_l/p_l - k_{l'}/p_{l'}| \geqslant 1/p_l p_{l'} \geqslant 1 \big/ \left(p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1} \right),$$

which contradicts

$$|k_l/p_l - k_{l'}/p_{l'}| < 1 \big/ \left(p_{\frac{t(t-1)}{2}} p_{\frac{t(t-1)}{2}+1} \right).$$

But p_l and $p_{l'}$ are different prime numbers, so $k_l/p_l = k_{l'}/p_{l'}$ if and only if $k_l = m p_l$ and $k_{l'} = m p_{l'}$ for some integer m, which contradicts with $p_l \nmid k_l$ and $p_{l'} \nmid k_{l'}$. The contradiction indicates that (14) must hold. $\blacksquare$

4. Proof of Lemma III.10

Proof. Since $|E_{\ell-1}|_{\mathbb{T}} \leqslant \frac{S\eta}{M_{\ell-1}}$, one has $|M_{\ell-1} E_{\ell-1}|_{\mathbb{T}} \leqslant S\eta$ and

$$\left| M_{\ell-1} B_{\mathbb{T}} \left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}} \right) \right|_{\mathbb{T}} = |B_{\mathbb{T}}(M_{\ell-1} E_{\ell-1}, \eta)|_{\mathbb{T}} \leqslant 3S\eta.$$

Let $B_{\mathbb{T}}(M_{\ell-1}E_{\ell-1}, \eta) = \cup_{1 \leqslant i \leqslant t}[\theta_i - \eta_i, \theta_i + \eta_i]$, then $\sum_i \eta_i \leqslant \frac{3}{2}S\eta$. According to Lemma III.8, one can choose $m_\ell \in \{p_1, p_2, \ldots, p_{\frac{s_{\ell-1}(s_{\ell-1}-1)}{2}+1}\}$ such that

$$\left|\theta_i - \theta_j - \frac{q}{m_\ell}\right|_{\mathbb{T}} \geqslant \frac{1}{2p_{\frac{s_{\ell-1}(s_{\ell-1}-1)}{2}} p_{\frac{s_{\ell-1}(s_{\ell-1}-1)}{2}+1}} > \frac{3S\eta}{2} \geqslant \eta_i + \eta_j, \quad \text{if } m_\ell \nmid q.$$

This implies that

$$\left(B_{\mathbb{T}}(M_{\ell-1}E_{\ell-1}, \eta) + \frac{q}{m_\ell}\right) \cap B_{\mathbb{T}}(M_{\ell-1}E_{\ell-1}, \eta) = \varnothing \mod 1, \quad \text{if } m_\ell \nmid q,$$

which is equivalent to

$$\left[B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}\right) + \frac{q}{m_\ell M_{\ell-1}}\right] \cap B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}\right) = \varnothing \mod 1, \quad \text{if } m_\ell \nmid q. \tag{A6}$$

Now we prove (15) inductively. For $\ell = 1$, it is implied by (A6) since $M_0 = 1$. Suppose (15) already holds for $\ell - 1$, then from Lemma III.7 we know $E_{\ell-1} \subset B_{\mathbb{T}}(E_{\ell-2}, \frac{\eta}{2M_{\ell-2}})$, and thus $B_{\mathbb{T}}(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}) \subset B_{\mathbb{T}}(E_{\ell-2}, \frac{\eta}{M_{\ell-1}} + \frac{\eta}{2M_{\ell-2}}) \subset B_{\mathbb{T}}(E_{\ell-2}, \frac{\eta}{M_{\ell-2}})$. Therefore, from the induction hypothesis, one can deduce that

$$\left[B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}\right) + \frac{q'}{M_{\ell-1}}\right] \cap B_{\mathbb{T}}\left(E_{\ell-1}, \frac{\eta}{M_{\ell-1}}\right) = \varnothing \mod 1, \quad \text{if } M_{\ell-1} \nmid q', \tag{A7}$$

where we used $m_{\ell-1}M_{\ell-2} = M_{\ell-1}$. If we let $q = m_\ell q'$ in (A7), then $M_{\ell-1} \nmid q'$ is equivalent to $m_\ell \mid q$, $(m_\ell M_{\ell-1}) \nmid q$, so (15) is proved combining (A6) and (A7).

5. Proof of Theorem IV.2

Proof. We prove that Algorithm I will work at each step ℓ. The procedure is almost identical to Theorem III.5. We only need to check that after step $\ell - 1$ we can choose an $m_\ell \in [2, 4]$ that satisfies both (5) and the gap $\Delta_\ell := \min_{1 \leqslant i < j \leqslant S} |M_\ell \lambda_i - M_\ell \lambda_j|_s$ is greater than $\tilde{\Delta}$ so that ESPRIT can work within the error bound given in Theorem IV.1. In step $\ell - 1$ of Algorithm I, we can already bound λ_i in a interval with length less than $\eta/M_{\ell-1}$, which can be denoted by $\lambda_i \in [\theta_i - \frac{\eta}{2M_{\ell-1}}, \theta_i + \frac{\eta}{2M_{\ell-1}}]$. Therefore, $\Delta_\ell \geqslant \tilde{\Delta}$ is equivalent to make the S intervals

$$\left\{\left[\theta_i - \frac{\eta}{2M_{\ell-1}} - \frac{\tilde{\Delta}}{m_\ell M_{\ell-1}}, \theta_i + \frac{\eta}{2M_{\ell-1}} + \frac{\tilde{\Delta}}{m_\ell M_{\ell-1}}\right]\right\}_{i=1}^{S} \tag{A8}$$

disjoint in the modulo-$1/M_{\ell-1}m_\ell$ sense. Since $m_\ell \geqslant 2$, we can relax the term $\tilde{\Delta}/m_\ell M_{\ell-1}$ as $\tilde{\Delta}/2M_{\ell-1}$. Note that (5) is equivalent to saying that the S intervals

$$\left[\theta_i - \frac{3\eta}{4M_{\ell-1}}, \theta_i + \frac{3\eta}{4M_{\ell-1}}\right] \quad (1 \leqslant i \leqslant S) \tag{A9}$$

are disjoint in the modulo-$1/M_{\ell-1}m_\ell$ sense. Therefore, we conclude that we can find a proper m_ℓ by letting $M = M_{\ell-1}$ and $\zeta_i = \max\{\frac{\eta+\tilde{\Delta}}{2}, \frac{3\eta}{4}\}$ in Lemma III.4, since $\sum_{i=1}^{S} \zeta_i = S\max\{\frac{\eta+\tilde{\Delta}}{2}, \frac{3\eta}{4}\} \leqslant \frac{1}{8S(2S-1)}$. After proving that the ESPRIT algorithm described in Sec. IV A can be used in Algorithm I given the existence of the spectral gap, the rest of the proof of the complexity bounds are the same as Theorem III.5. ■

6. Proof of Theorem IV.4

Proof. From ESPRIT's assumptions and properties, E_1 already has S disjoint intervals, each containing an actual spike. All steps in the proof of Lemma III.10 go through except that one needs to check the assumptions for Theorem IV.1 for each iteration in Algorithm I. In other words, one must check that the minimum separation among Λ_ℓ is bounded from below by $\tilde{\Delta}$. We prove this by enhancing the induction hypothesis in the proof of Lemma III.10 with an additional condition:

$$\min_{1 \leqslant s \leqslant s' \leqslant S, \quad n \in \mathbb{Z}} |M_\ell(\lambda_s - \lambda_{s'}) - n| \geqslant \tilde{\Delta}. \tag{A10}$$

When $\ell = 0$, this follows from the definition of $\tilde{\Delta}$. Now assume that this holds for $\ell - 1$. By the choice of m_ℓ in the proof of Lemma III.10, one has

$$\min_{\substack{1 \leqslant s \leqslant s' \leqslant S \\ q \in \mathbb{Z}, m_\ell \nmid q}} \left|\theta_s - \theta_{s'} - \frac{q}{m_\ell}\right| \geqslant \frac{1}{2p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}},$$

where θ_s is the center of the sth interval in $E_{\ell-1}$, which is also the sth element in $\tilde{\Lambda}_{\ell-1}$ defined in Theorem IV.1. By the property of Y_ℓ, one has

$$\min_{\substack{1 \leqslant s \leqslant s' \leqslant S \\ q \in \mathbb{Z}, m_\ell \nmid q}} \left|M_{\ell-1}(\lambda_s - \lambda_{s'}) - \frac{q}{m_\ell}\right|$$

$$\geqslant \min_{\substack{1 \leqslant s \leqslant s' \leqslant S \\ q \in \mathbb{Z}, m_\ell \nmid q}} \left|\theta_s - \theta_{s'} - \frac{q}{m_\ell}\right| - S\eta \geqslant \frac{1}{4p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}}.$$

Thus

$$\min_{\substack{1 \leqslant s \leqslant s' \leqslant S \\ q \in \mathbb{Z}, m_\ell \nmid q}} |M_\ell(\lambda_s - \lambda_{s'}) - q| \geqslant \frac{m_\ell}{4p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}}$$

$$\geqslant \frac{1}{2p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}} \geqslant \tilde{\Delta}. \tag{A11}$$

By the induction hypothesis, one also has

$$\min_{1\leqslant s\leqslant s'\leqslant S, n\in\mathbb{Z}} |M_{\ell-1}(\lambda_s - \lambda_{s'}) - n| \geqslant \tilde{\Delta},$$

which means

$$\min_{1\leqslant s\leqslant s'\leqslant S, q'\in\mathbb{Z}} |M_\ell(\lambda_s - \lambda_{s'}) - m_\ell q'| \geqslant m_\ell \tilde{\Delta} > \tilde{\Delta},$$

and combining with (A11) one obtains

$$\min_{1\leqslant s\leqslant s'\leqslant S, n\in\mathbb{Z}} |M_\ell(\lambda_s - \lambda_{s'}) - n| \geqslant \tilde{\Delta}.$$

The other steps in the proof of Lemma III.10 can be used directly to show that the algorithm works, and the arguments for the complexity bounds are the same as those in Theorem III.11. ∎

7. Proof of Theorem V.1

Proof. Here, we only need to take care of the choice of the sequence of $\{m_\ell\}$, since the rest of the proof is the same as Theorems III.5 and IV.2. According to Lemma III.4, we need to choose m_ℓ such that (5) holds. In addition, when $M_\ell > \tilde{M}$, we also have to guarantee the spectral gap $\Delta_\ell = \min_{1\leqslant i<j\leqslant S} |M_\ell\lambda_i - M_\ell\lambda_j|_s$ is greater than $\tilde{\Delta}$ to meet the requirements of ESPRIT in Theorem IV.1.

At the stage of choosing m_ℓ, we know that $E_{\ell-1} = \bigcup_{i=1}^{S_{\ell-1}} I_{\ell-1,i}$ is the union of at most S disjoint intervals and $|E_{\ell-1}| \leqslant S\eta/M_{\ell-1}$. Therefore, its neighborhood $B(E_{\ell-1}, \frac{\max\{\tilde{\Delta},\eta\}}{2M_{\ell-1}})$ must be t disjoint intervals with $t \leqslant S$ and

$$\left| B\left(E_{\ell-1}, \frac{\max\{\tilde{\Delta}, \eta\}}{2M_{\ell-1}}\right) \right| \leqslant \frac{2S\max\{\tilde{\Delta}, \eta\}}{M_{\ell-1}}.$$

This is exactly the case in Lemma III.4 when taking $G = B(E_{\ell-1}, \frac{\max\{\tilde{\Delta},\eta\}}{2M_{\ell-1}})$, $M = M_{\ell-1}$, and $\zeta = \max\{\tilde{\Delta}, \eta\}$. Therefore, the conclusion of Lemma III.4 guarantees that we can constructively find an m_ℓ such that

$$\left[B\left(E_{\ell-1}, \frac{\max\{\tilde{\Delta}, \eta\}}{2M_{\ell-1}}\right) + \frac{q}{M_{\ell-1}m_\ell} \right]$$
$$\cap B\left(E_{\ell-1}, \frac{\max\{\tilde{\Delta}, \eta\}}{2M_{\ell-1}}\right) = \varnothing \ \forall q \in \mathbb{Z}\backslash\{0\}, \quad (A12)$$

which is stronger than (5).

Moreover, (A12) means that $|\lambda_i + \frac{q}{M_\ell} - \lambda_j| \geqslant \tilde{\Delta}/M_{\ell-1}$ for $\lambda_i, \lambda_j \in \Lambda$ and $q \in \mathbb{Z}\backslash\{0\}$, since $\Lambda \in E_{\ell-1}$. Therefore, it holds that $|M_\ell\lambda_i - M_\ell\lambda_j + q| \geqslant \tilde{\Delta}$ for $q \in \mathbb{Z}\backslash\{0\}$. When $M_\ell > \tilde{M}$, we also have $|M_\ell\lambda_i - M_\ell\lambda_j| \geqslant M_\ell\Delta \geqslant \tilde{\Delta}$, which means

$$\Delta_\ell = \min_{1\leqslant i<j\leqslant S} |M_\ell\lambda_i - M_\ell\lambda_j|_s \geqslant \tilde{M}\Delta = \tilde{\Delta},$$

and this meets the requirement of ESPRIT.

At the stage of $M_\ell \leqslant \tilde{M}$, we have reached the accuracy of $\epsilon_1 = \tilde{M}^{-1}$, so Theorem III.5 gives that the maximal runtime is

$$T_{\max,1} = O(\eta K_1 \epsilon_1^{-1}) = O\left(\log\left(\frac{1}{\beta}\right)\eta^{-1}\tilde{\Delta}\Delta^{-1}\right),$$

and the total runtime is

$$T_{\text{total},1} = O(\eta K_1^2 \epsilon_1^{-1} N_{HR,1}) = \tilde{O}(\eta^{-2}\beta^{-2}\tilde{\Delta}\Delta^{-1}).$$

At the stage of $M_\ell > \tilde{M}$, according to Theorem IV.2 the maximal runtime is

$$T_{\max,2} = O(\eta K_2 \epsilon^{-1}) = O(\eta\tilde{\Delta}^{-1}\epsilon^{-1}),$$

and the total runtime is

$$T_{\text{total},2} = O(\eta K_2^2 \epsilon^{-1} N_{HR,2}) = \tilde{O}(\eta\omega^{-2}\tilde{\Delta}^{-2}\epsilon^{-1}).$$

Then we can conclude by $T_{\max} = \max\{T_{\max,1}, T_{\max,2}\}$ and $T_{\text{total}} = T_{\text{total},1} + T_{\text{total},2}$. ∎

8. Proof of Theorem V.3

Proof. Let $\tilde{\ell}$ be the largest ℓ such that $M_\ell \leqslant \tilde{M}$. From the definition of $N_{HR,1}$ above, one knows from Hoeffding's inequality that for any $\ell \leqslant \tilde{\ell}$ and $|k| \leqslant K_1$,

$$\mathbb{P}(|y_\ell(k) - \hat{f}_\ell(k)| < \alpha) > 1 - \frac{\rho}{(\lceil \log_2 \frac{\eta}{\epsilon} \rceil + 1)(K_1 + 1)}, \quad (A13)$$

and for any $\ell > \tilde{\ell}$ and $|k| \leqslant K_2$,

$$\mathbb{P}(|y_\ell(k) - \hat{f}_\ell(k)| < \alpha) > 1 - \frac{\rho}{(\lceil \log_2 \frac{\eta}{\epsilon} \rceil + 1)(K_2 + 1)}, \quad (A14)$$

Thus $|y_\ell(k) - \hat{f}_\ell(k)| < \alpha$ is true for all ℓ and k with probability at least $1 - \rho$ by the union bound. We condition on this event in the rest of the proof.

We first show that the spectral gap can be sufficiently enlarged by applying a burn-in process using the method in Sec. III. The proof is similar to that of Theorem IV.4. We enhance the induction hypothesis in Lemma III.10 by an additional condition

$$\Delta_\ell \geqslant \min\left\{ M_\ell\Delta, \frac{1}{2p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}} \right\}. \quad (A15)$$

When $\ell = 0$, this follows from the definition of Δ_ℓ. Now assume that this holds for $\ell - 1$. By the choice of m_ℓ in the proof of Lemma III.10, one has

$$\min_{\substack{1\leqslant s\leqslant s'\leqslant t \\ q\in\mathbb{Z}, m_\ell\nmid q}} \left| \theta_s - \theta_{s'} - \frac{q}{m_\ell} \right| \geqslant \frac{1}{2p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}},$$

where θ_s is the center of the sth interval in $E_{\ell-1}$, and t is the number of intervals in $E_{\ell-1}$. For any λ and λ' in Λ, if λ and λ' belong to different intervals indexed by s and s', one has

$$\min_{q\in\mathbb{Z}, m_\ell\nmid q} \left| M_{\ell-1}(\lambda - \lambda') - \frac{q}{m_\ell} \right| \geqslant \min_{\substack{1\leqslant s\leqslant s'\leqslant t \\ q\in\mathbb{Z}, m_\ell\nmid q}} \left| \theta_s - \theta_{s'} - \frac{q}{m_\ell} \right|$$
$$-S\eta \geqslant \frac{1}{4p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}},$$

by the property of Y_ℓ. Thus,

$$\min_{q\in\mathbb{Z}, m_\ell\nmid q} |M_\ell(\lambda - \lambda') - q| \geqslant \frac{m_\ell}{4p_{\frac{1}{2}S(S-1)}p_{\frac{1}{2}S(S-1)+1}}$$
$$\geqslant \min\left\{ M_\ell\Delta, \frac{1}{2p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}} \right\}. \quad (A16)$$

On the other hand, if λ and λ' belong to the same interval, then $|M_{\ell-1}(\lambda - \lambda')| \leqslant S\eta$ and $|M_\ell(\lambda - \lambda')| \leqslant m_\ell S\eta \leqslant 1/4$, so

$$\min_{q \in \mathbb{Z}, m_\ell \nmid q} |M_\ell(\lambda - \lambda') - q|$$

$$= |M_\ell(\lambda - \lambda')| \geqslant M_\ell \Delta \geqslant \min\left\{M_\ell \Delta, \frac{1}{2 p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}}\right\}. \tag{A17}$$

By the induction hypothesis, one also has

$$\min_{n \in \mathbb{Z}} |M_{\ell-1}(\lambda - \lambda') - n| \geqslant \min\left\{M_{\ell-1}\Delta, \frac{1}{2 p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}}\right\},$$

which means

$$\min_{q' \in \mathbb{Z}} |M_\ell(\lambda - \lambda') - m_\ell q'| \geqslant \min\left\{M_\ell \Delta, \frac{1}{2 p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}}\right\},$$

and combining with (A16) and (A17) one obtains

$$\min_{n \in \mathbb{Z}} |M_\ell(\lambda - \lambda') - n| \geqslant \min\left\{M_\ell \Delta, \frac{1}{2 p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}}\right\}.$$

When $\ell = \tilde{\ell} + 1$, one has

$$\min_{n \in \mathbb{Z}} |M_\ell(\lambda - \lambda') - n| \geqslant \min\left\{M_\ell \Delta, \frac{1}{2 p_{\frac{S(S-1)}{2}} p_{\frac{S(S-1)}{2}+1}}\right\} \geqslant \tilde{\Delta}.$$

With the same proof as Theorem IV.4, one can show that, for any $\ell > \tilde{\ell}$, E_ℓ has S disjoint intervals, each containing an actual spike and that

$$\min_{1 \leqslant s \leqslant s' \leqslant S, n \in \mathbb{Z}} |M_\ell(\lambda_s - \lambda_{s'}) - n| \geqslant \tilde{\Delta},$$

which ensures that the ESPRIT subroutine can be applied for any $\ell > \tilde{\ell}$.

In the first stage, i.e., when $M_\ell \leqslant \tilde{M}$, the maximal runtime is

$$T_{\max,1} = O(M_{\tilde{\ell}}K_1) = O_S\left(\eta^{-1}\log\left(\frac{1}{\beta}\right)\tilde{\Delta}\Delta^{-1}\right),$$

and the total runtime is

$$T_{\text{total},1} = O\left(M_{\tilde{\ell}}K_1^2 N_{HR,1}\right) = O_S\left(\eta^{-2}\log^2\left(\frac{1}{\beta}\right)\tilde{\Delta}\Delta^{-1}N_{HR,1}\right).$$

For the second stage, i.e., when $M_\ell > \tilde{M}$, according to Theorem IV.4 the maximal runtime is

$$T_{\max,2} = O\left(\eta K_2\epsilon^{-1}\right) = O(\eta\tilde{\Delta}^{-1}\epsilon^{-1}),$$

and the total runtime is

$$T_{\text{total},2} = O\left(\eta K_2^2\epsilon^{-1}N_{HR,2}\right) = O(\eta\tilde{\Delta}^{-2}\epsilon^{-1}N_{HR,2}).$$

Thus, the overall maximum runtime is

$$T_{\max} = \tilde{O}_S(\max\{\eta^{-1}\tilde{\Delta}\Delta^{-1}, \eta\tilde{\Delta}^{-1}\epsilon^{-1}\}),$$

and the overall total runtime is

$$T_{\text{total}} = \tilde{O}_S(\eta^{-2}\tilde{\Delta}\Delta^{-1}N_{HR,1} + \eta\tilde{\Delta}^{-2}\epsilon^{-1}N_{HR,2}).$$

∎

---

[1] S. Kimmel, G. H. Low, and T. J. Yoder, Robust calibration of a universal single-qubit gate set via robust phase estimation, Phys. Rev. A **92**, 062315 (2015).

[2] F. Belliardo and V. Giovannetti, Achieving Heisenberg scaling with maximally entangled states: An analytic upper bound for the attainable root-mean-square error, Phys. Rev. A **102**, 042613 (2020).

[3] A. E. Russo, K. M. Rudinger, B. C. A. Morrison, and A. D. Baczewski, Evaluating energy differences on a quantum computer with robust phase estimation, Phys. Rev. Lett. **126**, 210501 (2021).

[4] H. Li, H. Ni, and L. Ying, A note on spike localization for line spectrum estimation, Appl. Comput. Harmonic Anal. **67**, 101577 (2023).

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

[6] M. Zwierz, C. A. Pérez-Delgado, and P. Kok, General optimality of the Heisenberg limit for quantum metrology, Phys. Rev. Lett. **105**, 180402 (2010).

[7] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, Nat. Commun. **9**, 78 (2018).

[8] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, arXiv:quant-ph/9511026.

[9] A. Y. Kitaev, A. Shen, M. N. Vyalyi, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Soc., 2002), No. 47.

[10] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, Proc. R. Soc. London, Ser. A **454**, 339 (1998).

[11] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Simulating Hamiltonian dynamics with a truncated Taylor series, Phys. Rev. Lett. **114**, 090502 (2015).

[12] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, Nature (London) **450**, 393 (2007).

[13] E. Knill, G. Ortiz, and R. D. Somma, Optimal quantum measurements of expectation values of observables, Phys. Rev. A **75**, 012328 (2007).

[14] D. Poulin and P. Wocjan, Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer, Phys. Rev. Lett. **103**, 220502 (2009).

[15] T. E. O'Brien, B. Tarasinski, and B. M. Terhal, Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments, New J. Phys. **21**, 023022 (2019).

[16] Y. Dong, L. Lin, and Y. 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).

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

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

[19] Z. Ding and L. Lin, Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation, PRX Quantum **4**, 020331 (2023).

[20] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[21] H. Ni, H. Li, and L. Ying, On low-depth algorithms for quantum phase estimation, Quantum **7**, 1165 (2023).

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

[23] A. Dutkiewicz, B. M. Terhal, and T. E. O'Brien, Heisenberg-limited quantum phase estimation of multiple eigenvalues with few control qubits, Quantum **6**, 830 (2022).

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

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

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

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

[28] K. Klymko, C. Mejuto-Zaera, S. J. Cotton, F. Wudarski, M. Urbanek, D. Hait, M. Head-Gordon, K. B. Whaley, J. Moussa, N. Wiebe, W. A. de Jong, and N. M. Tubman, Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware, PRX Quantum **3**, 020323 (2022).

[29] R. M. Parrish and P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation, arXiv:1909.08925.

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

[31] E. N. Epperly, L. Lin, and Y. Nakatsukasa, A theory of quantum subspace diagonalization, SIAM J. Matrix Anal. Appl. **43**, 1263 (2022).

[32] R. Roy and T. Kailath, Esprit-estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech, Signal Process. **37**, 984 (1989).

[33] J. Hadamard, Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques, Bull. Soc. Math. Fr. **24**, 199 (1896).

[34] C. J. d. L. V. Poussin, *Recherches Analytiques sur la Théorie des Nombres Premiers* (Hayez, Bruxelles, 1897), Vol. 1.

[35] J. B. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, Ill. J. Math. **6**, 64 (1962).

[36] V. I. Morgenshtern and E. J. Candes, Super-resolution of positive sources: The discrete setup, SIAM J. Imaging Sci. **9**, 412 (2016).

[37] Q. Denoyelle, V. Duval, and G. Peyré, Support recovery for sparse deconvolution of positive measures, arXiv:1506.08264.

[38] W. Li, W. Liao, and A. Fannjiang, Super-resolution limit of the esprit algorithm, IEEE Trans. Inf. Theory **66**, 4593 (2020).