Heisenberg-limited quantum phase estimation of multiple eigenvalues with few control qubits

A. Dutkiewicz[1], B.M. Terhal[2], and T. O'Brien[3,1]

[1]Instituut-Lorentz, Universiteit Leiden, 2300 RA Leiden, The Netherlands

[2]QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands and JARA Institute for Quantum Information, Forschungszentrum Juelich, D-52425 Juelich, Germany

[3]Google Quantum AI, 80636 Munich, Germany

arXiv:2107.04605v3 [quant-ph] 23 Sep 2022

Quantum phase estimation is a cornerstone in quantum algorithm design, allowing for the inference of eigenvalues of exponentially-large sparse matrices. The maximum rate at which these eigenvalues may be learned, –known as the Heisenberg limit–, is constrained by bounds on the circuit complexity required to simulate an arbitrary Hamiltonian. Single-control qubit variants of quantum phase estimation that do not require coherence between experiments have garnered interest in recent years due to lower circuit depth and minimal qubit overhead. In this work we show that these methods can achieve the Heisenberg limit, *also* when one is unable to prepare eigenstates of the system. Given a quantum subroutine which provides samples of a 'phase function' $g(k) = \sum_j A_j e^{i\phi_j k}$ with unknown eigenphases ϕ_j and overlaps A_j at quantum cost $O(k)$, we show how to estimate the phases $\{\phi_j\}$ with (root-mean-square) error δ for total quantum cost $T = O(\delta^{-1})$. Our scheme combines the idea of Heisenberg-limited multi-order quantum phase estimation for a single eigenvalue phase [1, 2] with subroutines with so-called dense quantum phase estimation which uses classical processing via time-series analysis for the QEEP problem [3] or the matrix pencil method. For our algorithm which adaptively fixes the choice for k in $g(k)$ we prove Heisenberg-limited scaling when we use the time-series/QEEP subroutine. We present numerical evidence that using the matrix pencil technique the algorithm can achieve Heisenberg-limited scaling as well.

Contents

1 Introduction **2**
1.1 Outline 3

2 The classical and quantum tasks of phase estimation **3**
2.1 Phase distance on the circle 5
2.2 Limits for single-eigenvalue phase estimation 6

3 Defining the task of multiple-eigenvalue phase estimation **8**
3.1 Methods of dense signal phase estimation 9

4 Multiple eigenvalues: multi-order estimation and the phase matching problem **10**
4.1 Heisenberg-limited algorithm for multiple phases 10
4.2 Bounding the error with and without failures 12
4.3 Algorithm 4.1 achieves the Heisenberg limit 18

5 Numerical implementation **20**

6 Conclusion **21**

Acknowledgements **23**

References **23**

A Phase extraction subroutine **25**
A.1 The Quantum Eigenvalue Estimation Problem 25
A.2 Conservative QEEP Eigenvalue Extraction 26
A.3 Proof of Lemma 3.2 27

B Proof of Lemma 4.2 **27**

C Range of shifted phase estimates **29**

1 Introduction

For quantum computers to overcome the 50^+ year head start in research and development enjoyed by their classical competition, quantum algorithms must eke out every inch of quantum speedup over their classical counterparts. Quantum phase estimation (QPE) of a unitary operator U, –a BQP-complete problem [4]–, plays a central or support role in many promising quantum applications [5–7].

However, not all flavours of quantum phase estimation are equally powerful. To perform QPE with accuracy (root-mean-square) error δ, initial implementations of quantum phase estimation [8, 9] used a $O(\log(\delta^{-1}))$-qubit control register, and required computation time scaling as $T = O(\delta^{-2})$, known as the Sampling Noise Limit, when contributions from outlying (unlikely) data are taken into account [1]. Much work has been undertaken over the succeeding years to improve QPE estimation rates to the theoretically optimal Heisenberg limit $T = O(\delta^{-1})$ [10–12] and reduce the control overhead.

The requirement to perform applications of U conditional on a large entangled control register is technically challenging and has strong coherence requirements. It has been known for a long time [13] that the control register in QPE can be replaced by a single qubit using classical feedback and re-preparation of the control qubit, also known as iterative QPE [14]. For estimating the phase of a single eigenstate, –assuming the preparation of this eigenstate–, the sampling noise limit can thus simply be achieved using a single control qubit. In [1] iterative-QPE was extended to achieve the Heisenberg limit $T = O(\delta^{-1})$. This Heisenberg limit can be shown, via Cramer-Rao bounds [1], to be a lower bound on the cost of phase estimation, assuming one cannot fast-forward the unitary U [15]. This type of estimation have additionally demonstrated a relative robustness to error [2, 16], which is of interest to NISQ applications. Other analyses of QPE use maximum-likelihood [17], or Bayesian [16, 18] inference.

The requirement to prepare eigenstates of the unitary U is not possible for most applications. It is well known that the 'textbook' QPE algorithm succeeds for any initial state, i.e. the output is always an accurate estimate of one of the eigenphases of U [8]. However, the performance of few-ancilla QPE on starting states that are not eigenstates has only been examined recently. In [19] it was demonstrated numerically that one may infer single eigenvalues from mixed or superposed initial eigenstates using standard classical signal processing techniques [20]. Under some additional constraints on the system, it was recently found that these techniques could be performed in the absence of any control qubits or the need to apply controlled unitary operations [21, 22], a further significant saving. Due to the need to 'densely sample' the phase function $g(k) = \sum_j A_j e^{i\phi_j k}$, and a lack of optimization of the classical post-processing techniques, Ref. [19] only achieved sampling-noise-limited scaling, but not Heisenberg-limited scaling. By dense sampling we mean that we draw samples from $g(k)$ with k a sequence of integers, $k = 0, 1, \ldots, K$ (as opposed to, say, choosing $k = 2^d$, i.e. exponentially increasing which is used in textbook QPE and iterative QPE). By a clever adjustment of the quantum phase estimation problem to target estimation of the spectral function, Eq. (3), of the input state, Ref. [3] was able to prove rigorous results, with bounds that were subsequently improved in Ref. [23]. Still, these results fall short of reaching "the Heisenberg limit for the problem of estimating multiple phases", however that should be defined.

In this work, we demonstrate single-control qubit quantum phase estimation at a so-called Heisenberg limit. We extend the methods used in Refs. [1, 2] that obtain Heisenberg-limited scaling for single eigenphases to the multiple-phase setting by the use of a multi-order scheme and phase matching subroutines between different orders. We show that to make this phase matching unambiguous requires the sampling scheme to be adaptive, i.e. the next choice for k of $g(k)$ depends on the current phase estimates. At each order the multi-order algorithm requires input from a dense phase estimation method: for a given order k we use samples from $g(k\mathsf{k})$ with $\mathsf{k} = 0, 1, \ldots, K$. As we require the freedom to choose k a real number, to be applicable to a completely general U we must invoke the quantum singular value transformation of Ref. [24], which requires $O(1)$ additional control bits. Us-

ing the time-series or QEEP analysis of Ref. [3] as classical processing subroutine, we are able to obtain a rigorous proof of Heisenberg-limited scaling of our multi-order scheme. Using the matrix pencil method analysed in Ref. [19], as such a dense subroutine, we are able to show numerical results consistent with the Heisenberg limit, with a performance improvement over the time-series analysis results.

In essence, our paper is concerned with what choices of k in $g(k)$ and what classical processing are needed to enable Heisenberg-limited scaling, i.e. scaling which minimizes the total number of applications T of (controlled) U (which we refer to as the quantum cost) given a targeted error δ with which to estimate multiple eigenvalue phases of U present in some input state $|\Psi\rangle$. It can thus be viewed as purely solving a problem of classical signal processing. This does not mean that such questions are trivial: for example, the question of how to estimate phases if one is allowed to only get single samples from $g(k)$ for a set of randomly chosen k relates to the dihedral hidden subgroup problem in quantum information theory [25]. We note that other work, based on a Monte Carlo extension of [3], achieving Heisenberg scaling (up to polylog factors) was recently presented in [26]. Another recent work also demonstrated numerical evidence for Heisenberg-limited phase estimation using Bayesian methods [27]. We also note that the information-theoretically optimal method in [17] which picks random k in $g(k)$ has a classical processing cost which is linear in the quantum cost T, and Theorem 1 in [17] can be converted to bound the mean-squared-error in estimating a single phase, see comments below Theorem 2.5 in Section 2.2. In principle this information-theoretic method can be extended to the case of n_ϕ eigenvalue phases, but the classical processing cost will be exponential in n_ϕ as one iterates over the possible values of the phases, while our Algorithm 4.1 has a polynomial (but superlinear) quantum and classical cost in terms of the number of phases n_ϕ.

1.1 Outline

We begin in Sec. 2 by defining a few mathematical objects. We separate the quantum part of a phase estimation problem, namely sampling of the phase (or signal) function $g(k)$ in Eq. (2) given an input unitary U and input state $|\Psi\rangle$ in Definition 2.2 through running some quantum circuits, and the classical processing of samples from $g(k)$ to extract the eigenvalue data of U. In Sec. 2.1 we state and prove some needed properties of the distance between phases. In Sec. 2.2 we prove several Cramer-Rao bounds on the scaling of the error versus the total quantum cost for the estimation of a single eigenvalue phase, Theorem 2.5. We state the previous result on getting Heisenberg-limited scaling for a single eigenvalue phase in Algorithm 2.6. Table 1 contains a glossary of the symbols used in this paper.

In Sec. 3 we properly define a multi-eigenvalue phase estimation problem (Def. 3.1). We discuss algorithms that can be used to extract multiple phases from densely-sampled signal $g(k)$. We state error bounds satisfied by the output of Alg. A.3 (Lemma 3.2) that is used as the fixed-order subroutine in our final Algorithm 4.1 which achieves Heisenberg scaling.

In Section 4 we present our Heisenberg-limited algorithm. We discuss a critical aliasing problem to be solved which occurs when estimating multiple eigenvalues. We show that an adaptive choice for k in $g(k)$ can solve this issue and we prove that such adaptive choice always exists in Lemma 4.2. In Lemma 4.3 and Lemma 4.4 we prove some properties about Algorithm 4.1 which will be helpful in proving the final Theorem 4.5.

Thus in Theorem 4.5 we prove Algorithm 4.1 achieves Heisenberg-limited scaling, which is the main result of this work. In Sec. 5 we numerically compare this rigorous implementation to an implementation using the matrix pencil method, used in Ref. [19], for which we are unable to find a rigorous proof of Heisenberg scaling. We finish the paper with a discussion, Section 6.

2 The classical and quantum tasks of phase estimation

One may separate quantum phase estimation into the extraction of a signal which consists of oscillations at eigenvalue frequencies ϕ_j at a chosen time k, and the processing of this signal to resolve the frequencies. Let us first define the following:

Definition 2.1 (Signal or Phase Function, and Spectral Function)**.** *Let $U \in \mathsf{U}(2^N)$ be an N-qubit unitary operator, and $|\Psi\rangle \in \mathbb{C}^{2^N}$ an N-qubit state. We label the eigenstates $|\phi_j\rangle$ of U by their*

Symbol	Term	Description
U	Unitary	The unitary whose eigenphases we wish to estimate.
T	Quantum cost	The total number of applications of controlled U over the course of the phase estimation algorithm.
ϕ_j	Phase	A number $\phi_j \in [0, 2\pi)$ such that the jth eigenvalue of U is $e^{i\phi_j}$.
A_j	Overlap	The overlap of the input state and the jth eigenstate of U; see Def. 2.1.
$g(k)$	Signal / Phase function	See Def. 2.1.
d	Order	Running index for the order of estimation. At each order we construct new phase estimates $\tilde{\phi}_j^{(d)}$ using new data and the previous estimates $\tilde{\phi}_j^{(d-1)}$.
$\tilde{\phi}_j^{(d)}$	Estimate	Estimate of ϕ_j obtained at the dth order; see step 4d of Alg. 4.1.
k_d	Exponent	At order d we perform the QEEP subroutine for $V = U^{k_d}$.
κ_d	Multiplier	Defines $k_d = \kappa_d k_{d-1}$.
$\theta_j^{(d)}$	Phase	Eigenphase of U^{k_d}.
$\tilde{\theta}_j^{(d)}$	Estimate	Estimate of $\theta_j^{(d)}$; see step 4b of Alg. 4.1.
ϵ	Single order error	Error parameter used for the QEEP subroutine at each order; see Alg. A.3.
δ	Final error	A bound for standard deviation of the final estimates; see Def. 3.1.
p_d	Confidence bound	The probability with which the QEEP subroutine at order d succeeds; see Eq. (28).
A	Overlap bound	We wish to estimate the phases for which the overlap is $A_j > A$; see Def. 3.1.

Table 1: Glossary of symbols used in the main text.

phase — $U|\phi_j\rangle = e^{i\phi_j}|\phi_j\rangle$. *We can decompose* $|\Psi\rangle$ *in terms of these eigenstates,*

$$|\Psi\rangle = \sum_j a_j |\phi_j\rangle, \tag{1}$$

and write the overlap $A_j := |a_j|^2, \sum_j A_j = 1$. *We define the phase function, – also called the signal–,* $g(k)$ *for* $k \in \mathbb{R}$ *of a state* $|\Psi\rangle$ *under* U *as*

$$g(k) = \sum_j A_j e^{ik\phi_j}. \tag{2}$$

The spectral function $A(\phi)$ *is defined as*

$$A(\phi) = \sum_j A_j \delta(\phi - \phi_j). \tag{3}$$

Note that $\int_0^{2\pi} d\phi A(\phi) = 1$, *and* $g(k) = \int_0^{2\pi} d\phi e^{ik\phi} A(\phi)$*; i.e. the phase function sets the Fourier coefficients of the spectral function.*

Note that one may change seamlessly between the description of a unitary U and its eigenvalues and a Hermitian operator H and its eigenvalues using the transform $U = e^{iHt}$ for an appropriate choice of t. Note that since $g(-k) = g^*(k)$ we can restrict ourselves to $k \geq 0$.

One may consider algorithms estimating $g(k)$ at integer $k \in \mathbb{Z}^+$, which require the quantum circuits using controlled-U^k in Fig. 1 with $k \in \mathbb{Z}^+$. In our final Alg. 4.1 we will however use $k \in \mathbb{R}^+$ (in practice $k \in \mathbb{Q}^+$). In order to implement U^k, we can write $k = \lfloor k \rfloor + \alpha$, and we can simulate U^k in time $O(k)$ if we can simulate U^α in time independent of k. The accuracy of this fractional query to U^α can be independent of the final error in our phase estimation (as long as it is sufficiently small). Simulating U^α is not a significant issue for Hamiltonian simulation methods such as Trotter decompositions [28], which allow simulation of e^{iHt} for arbitrary $t \in \mathbb{R}$. If we instead have access to a circuit implementation of a unitary U, or a block-encoding of a Hamiltonian H, we can implement a fractional query of U^α via the quantum singular value transform [24, Corollary 34]. The circuit to implement U^α to error ϵ requires $\mathcal{O}(1)$ additional ancilla qubits, and $\mathcal{O}(\Delta_{\max}^{-1} \log(1/\epsilon))$ implementations of controlled-U (where $\Delta_{\max}$ is the largest gap in the spectrum

of U) [1]. We will assume in this work that our states have support on at most n_ϕ phases, and so we can bound $\Delta \geq \pi/n_\phi$. The cost of implementing U^α will thus not be a significant part of the cost to implement U^k under the assumption $k >> A_0^{-1}\log(1/\epsilon)$. To simplify our remaining analysis, we assume herein that we can implement U^k at a total quantum cost k for all $k \in \mathbb{R}_+$.

The following task summarizes the quantum subroutine for phase estimation which is to be executed with the quantum circuits in Fig. 1:

Definition 2.2 (Phase Function Estimation, PFE)**.** *Let U be an N-qubit unitary operator and $|\Psi\rangle$ an N-qubit quantum state. Assume*

1. *A quantum circuit implementation of U (conditional on a control qubit) and,*
2. *A quantum circuit that prepares $|\Psi\rangle$.*

Given a $k \in \mathbb{Z}^+$, error $\epsilon > 0$, and confidence $0 < p \leq 1 \in \mathbb{R}$, PFE *outputs an estimate $\tilde{g}(k)$ of the phase function $g(k)$ of $|\Psi\rangle$ under U, with $\mathbb{P}(|\tilde{g}(k) - g(k)| \leq \epsilon) \geq 1 - p$ with quantum cost $T = M|k|$ where M is the number of repetitions of both experiments in Fig. 1 and $M = \Theta(|\ln(1-p)|\epsilon^{-2})$ via a Chernoff bound. Our assumption[2] that the cost of implementing U^k for $k \in \mathbb{R}^+$ scales as $O(k)$ implies that the above scaling for the cost of phase function estimation holds when $k \in \mathbb{R}^+$.*

Note that estimating the quantum cost of the subroutine in Def. 2.2 as linear in k is consistent with general no-fast forwarding statements [15] which state that for general Hamiltonians one cannot implement $U^t = \exp(itH)$ in time sublinear in t. It is expected that phase function estimation is hard to do efficiently on a classical computer as it allows one, via classical post-processing, to sample from the eigenvalue distribution from the input state which can be reformulated as a BQP-complete problem [4].

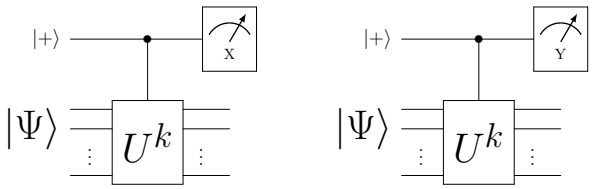

Figure 1: For an input state $|\Psi\rangle = \sum_j a_j |\phi_j\rangle$ and initial ancilla state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ the probability for the ancilla measurement outcome to be ± 1 is $\mathbb{P}(\pm 1) = \frac{1}{2}\sum_j A_j[1 \pm \cos(k\phi_j)]$ (left circuit measuring in the X-basis) and $\mathbb{P}(\pm 1) = \frac{1}{2}\sum_j A_j[1 \mp \sin(k\phi_j)]$ (right circuit measuring in the Y-basis).

The quantum subroutine for PFE proceeds by executing the circuits in Fig. 1 for the given k. The control qubit is prepared in the $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state, and is used to control k applications of the unitary U on the system register prepared in $|\Psi\rangle$. The reduced density matrix of the control qubit then takes the form

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & g(k) \\ g^*(k) & 1 \end{pmatrix}. \tag{4}$$

The phase function $g(k)$ is extracted by state tomography of the control qubit: one estimates the real and imaginary parts of $g(k) = g^{(r)}(k) + ig^{(i)}(k)$ by M repetitions of the circuit in Fig. 1 and measurements of the control qubit in the X- or Y-basis respectively. We will ignore any dependence of phase function estimation on N.

2.1 Phase distance on the circle

Quantum phase estimation describes a series of protocols to estimate the eigenphases ϕ_j. As these eigenvalues are defined on the circle $[0, 2\pi)$, we need a notion of distance which respects this periodicity:

Definition 2.3. *For $x \in \mathbb{R}$ we define the distance $|\cdot|_T \in [0, \pi]$ as*

$$|x|_T := \min_{m\in\mathbb{Z}}(|\Delta|), \text{ with } x = \Delta + 2\pi m, \text{ with } \Delta \in [-\pi, \pi). \tag{5}$$

[1] The $\Delta_{\max}$ dependence in our circuit comes from the requirement in [24, Corollary 34] that our unitary have spectrum on $[-\pi + \Delta_{\max}, \pi - \Delta_{\max}]$. Though this will not immediately be the case, we are free to rotate the spectrum of our unitary to satisfy this requirement, as long as the spectral gap exists. We also only require to consider those eigenstates with support on our initial state in this algorithm; eigenstates with zero weight can be adjusted as part of implementing the quantum singular value transformation without affecting the phase function $g(k)$.

[2] In practice, the cost of phase function estimation for $k \in \mathbb{R}^+$ will scale as $M(\lfloor k \rfloor + O(n_\phi \log(1/\epsilon))$ when using quantum signal processing techniques. One can confirm that the effect this has on our result is to effectively increase the cost of calling the QEEP subroutine, Alg. A.3, from $O(\epsilon^{-6})$ to $O(\epsilon^{-6}\log(1/\epsilon))$. This will only change the prefactor of our Heisenberg-limited phase estimation algorithm (Alg. 4.1), and does not prevent it achieving the Heisenberg limit.

Clearly, the distance obeys the triangle inequality: for $x_1, x_2 \in \mathbb{R}$

$$|x_1 + x_2|_T \leq |x_1|_T + |x_2|_T. \tag{6}$$

The following Lemma addresses a technical issue in the proof of performance of Algorithm 4.1. For integer $k \in \mathbb{Z}^+$, we have $|kx|_T = \min_{m\in\mathbb{Z}} |k\Delta + 2\pi m| = k \min_{m\in\mathbb{Z}} |\Delta + \frac{2\pi m}{k}|_T$, implying Eq. (8) below directly. However, for $k \in \mathbb{R}^+$ (or rational numbers $k \in \mathbb{Q}^+$) we need to specify a range of x for which such a statement holds, that is:

Lemma 2.4. *Suppose $k > 1$ and $\theta, \phi \in [0, 2\pi)$. If*

$$\frac{\pi}{k} \leq \phi \leq \frac{\pi(2\lfloor k \rfloor - 1)}{k}, \tag{7}$$

we have for any θ

$$\min_{n\in\{0,\ldots,\lfloor k\rfloor -1\}} \left|\phi - \frac{\theta}{k} - \frac{2\pi n}{k}\right|_T = \frac{1}{k}\left|k\phi - \theta\right|_T. \tag{8}$$

Proof. Let $x = \phi - \frac{\theta}{k}$, then Eq. (7) and $\theta \in [0, 2\pi)$ imply that

$$-\pi \leq kx \leq 2\pi\lfloor k \rfloor - \pi. \tag{9}$$

and thus $kx = \Delta + 2\pi m$ with $m \in \{0, \ldots \lfloor k \rfloor - 1\}$ and $\Delta \in [-\pi, \pi)$ and $|k\phi - \theta|_T = |\Delta|$. Hence $x = \frac{\Delta}{k} + \frac{2\pi m}{k}$ with $m \in \{0, \ldots \lfloor k \rfloor - 1\}$, implying Eq. (8) where the minimum can be achieved by $m = n$. □

2.2 Limits for single-eigenvalue phase estimation

For the special case of estimating a single eigenvalue phase, the Cramér-Rao theorem can be used to lower bound the quantum cost T to learn the phase with accuracy δ, known as the Heisenberg limit. The problem of estimating multiple phases ϕ_j, in the presence of unknown overlaps A_j, is not easily amenable to such Fisher information analysis as it is a multi-parameter estimation problem. However, it can be expected that the cost T of this task is at least as high as that of single phase estimation, hence it is of interest to review these bounds here. The following theorem also proves a dense signal limit which sits in between Heisenberg and sampling noise scaling:

Theorem 2.5 (The Heisenberg, Dense Signal and Sampling Limits)**.** *The (root-mean-square) error δ of an estimator $\tilde{\phi}$ of the eigenvalue phase ϕ employing the circuits in Fig. 1 on a eigenstate of U is always lower bounded as*

$$\text{Heisenberg Limit}: \delta \geq cT^{-1}, \tag{10}$$

where T is the quantum cost of implementing the circuits. If we choose to use only quantum circuits with $k = 1$, the sampling noise limit holds:

$$\text{Sampling Noise Limit}: \delta \geq T^{-1/2}. \tag{11}$$

If we choose circuits with $k = 0, 1, \ldots, K$ with a fixed number of repetitions M for each circuit we are bound by a so-called 'dense signal' limit:

$$\text{Dense Signal Limit}: \delta \geq cT^{-3/4}. \tag{12}$$

In these statements c is some constant.

Proof. Let $\tilde{\phi}$ be an estimator of ϕ which is inferred from the data $\mathbf{x}$. Here the data $\mathbf{x}$ is the string of outcomes of the ancilla qubit measurements for all the experiments using the left and right circuits in Fig. 1. We have

$$\delta^2 = \sum_{\mathbf{x}} \mathbb{P}(\mathbf{x}|\phi)(\phi - \tilde{\phi}(\mathbf{x}))^2 \geq I^{-1}(\phi), \tag{13}$$

by the Cramér-Rao theorem [29, 30] where the Fisher information is defined as

$$I(\phi) = \sum_{\mathbf{x}} \mathbb{P}(\mathbf{x}|\phi) \left\{\frac{\partial}{\partial\phi} \ln\left[\mathbb{P}(\mathbf{x}|\phi)\right]\right\}^2. \tag{14}$$

Thus $I(\phi)$ limits the information we may learn about ϕ given a dataset $\mathbf{x}$ drawn from $\mathbb{P}(\mathbf{x}|\phi)$ and we can calculate $I(\phi)$. Let M_k^r be the number of experiments, using the circuit with the X measurement, and M_k^i be the number of experiments using the circuit with the Y measurement with a certain chosen k. The Fisher information for all independent experiments together is additive, i.e. $I(\phi) = \sum_k [M_k^r I(\phi|k, r) + M_k^i I(\phi|k, i)]$ with $I(\phi|k, r)$ and $I(\phi|k, i)$ the Fisher information of a single experiment and $\sum_k$ is the sum over the chosen set of ks. For a single experiment we can calculate, using Eq. (14) and the probability for the output bit given in Fig. 1, that $I(\phi|k, r) = I(\phi|k, i) = k^2$ and thus

$$I(\phi) = \sum_k k^2 (M_k^r + M_k^i). \tag{15}$$

At the same time the total quantum cost of all experiments is

$$T = \sum_k k(M_k^r + M_k^i). \tag{16}$$

The key insight here is that the relative dependence on k is different between T and I and this implies that the trend of the number of experimental runs $M_k^{r/i}$ as a function of k will affect the maximum rate of estimation. If we choose only $k = 1$ we see that $\delta^2 \geq \frac{1}{T}$ which is the sampling noise limit.

The biggest value for $I(\phi)$ for a given T is obtained when we choose a single largest possible $k = K$ so that $T = KM_K$ and $I(\phi) = K^2 M_K = T^2/M_K$ with $M_K = M_K^r + M_K^i$. This implies a Heisenberg limit, i.e. $\delta \geq cT^{-1}$ where c is some constant depending on M_K. If we however make the 'dense signal' choice, that is, $M_k^r = M_k^i = M$ for $k = 1, 2 \ldots K$, then

$$I(\phi) = \frac{M}{3} K(K+1)(1+2K), \tag{17}$$

while the total quantum cost is $T = MK(K+1)$. Increasing K with M fixed, we have to leading order in K that $\delta \geq I(\phi)^{-\frac{1}{2}} = \sqrt{\frac{3}{2}} M^{-\frac{1}{2}} K^{-\frac{3}{2}} = \sqrt{\frac{3}{2}} M^{\frac{1}{4}} T^{-\frac{3}{4}}$, which is the dense signal limit. □

Remark: We note that a randomized version of the dense signal choice can potentially scale in near-Heisenberg-limited fashion. In this method, one would draw k at random from $1, \ldots, K$ and repeat this S times to generate random variables $k_1, \ldots, k_S$ and repeat experiments with fixed M for each such k_i. With the right choice of $S \times M = \text{polylog}(K)$, one can argue, using the Cramer-Rao lower bound analysis above, that the expected error $\mathbb{E}(\delta) \geq \frac{\text{polylog}(\mathbb{E}(T))}{\mathbb{E}(T)}$ where $\mathbb{E}(T)$ is the expected quantum cost. The algorithm in [17] uses such strategy with $M = 1$. Clearly, the sampling noise limit can be achieved by choosing $k = 1$ in the circuits of Fig. 1. However, one can ask whether the dense signal limit or the Heisenberg limit can also be achieved, in particular when we demand that the classical post-processing is computationally efficient, meaning that this processing is polynomial in the quantum cost T. For the dense signal limit one needs a classical method to process the estimates of $g(k)$ at $k = 1, \ldots, K$ to estimate $\tilde{\phi}$. Using perturbation theory in the noise, the matrix pencil method has been claimed to achieve this for a single eigenvalue [31].

Achieving the Heisenberg limit is non-trivial due to phase aliasing: the phase function $g(k)$ obtained by the experiments using U^k remains invariant if the phase ϕ is shifted by $\frac{2\pi}{k}$. This implies that a strategy of estimating ϕ from a single point $g(k)$ at large k will fail unless ϕ is already known to sit within a window of width $\frac{2\pi}{k}$. This issue is circumvented by sampling $g(k)$ at multiple orders $k = 2^d$ to 'gradually zoom in' on ϕ. To get Heisenberg scaling, one lets the number of samples M and thus the confidence, to depend on the order, so that the most significant bits of ϕ are determined with the highest confidence. Methods for doing this were first introduced in Ref. [1], and improved in Ref. [2] for the purpose of gate calibration. Here we state the result:

Algorithm 2.6 (Heisenberg Algorithm For Single Eigenvalue Phase [1, 2])**.** *Given an targeted error $\delta > 0$, and numbers $\alpha, \gamma \in \mathbb{Z}^+$. The Heisenberg algorithm which outputs an estimate $\tilde{\phi}$ for ϕ proceeds as follows:*

1. *Fix $d_f = \lceil \log_2(1/\delta) \rceil$.*
2. *For $d = 0, 1 \ldots, d_f$:*
   (a) *Use the* PFE *subroutine, Def. 2.2, circuits in Fig. 1, to obtain an estimate $\tilde{g}(k)$ of $g(k)$ for $k = 2^d$ using $M_d = \alpha + \gamma(d_f + 1 - d)$ repetitions of both experiments.*
   (b) *Compute $\tilde{\theta}^{(d)} = \text{Arg}[\tilde{g}(2^d)] \in [0, 2\pi)$*
   (c) *If $d = 0$, set $\tilde{\phi}^{(0)} = \tilde{\theta}^{(0)}$.*
   (d) *Else, set $\tilde{\phi}^{(d)}$ to be the unique value in the interval $[\tilde{\phi}^{(d-1)} - \frac{\pi}{2^d}, \tilde{\phi}^{(d-1)} + \frac{\pi}{2^d})$ (with periodic boundaries) such that*

$$2^d \tilde{\phi}^{(d)} = \tilde{\theta}^{(d)} \mod 2\pi. \tag{18}$$

3. *Return $\tilde{\phi} = \tilde{\phi}^{(d_f)}$ as an estimate for ϕ.*

It was proven in [2] that for some choices of α and γ the root-mean-square error δ on the final estimate $\phi^{(d_f)}$ is at most cT^{-1} for a constant c and total cost $T = 2\sum_{d=0}^{d_f} 2^d M_d$, thus reaching the Heisenberg limit.

One might consider the effect of experimental noise on these limits. The algorithm given in [2] was proven to be robust against noise that affected the estimation of any $g(k)$ by no more than

$\frac{1}{\sqrt{8}}$. However, realistic noise tends to scale with the circuit depth, eventually breaking this bound. In the presence of a uniform depolarizing channel, it is possible to extend the above calculation of Fisher information to optimize the recovery of a single phase ϕ, however in the limit that $\delta \to 0$ only the sampling-noise limit can be obtained:

Lemma 2.7. *The (root-mean-square) error δ of an estimator $\tilde{\phi}$ of the eigenvalue phase ϕ employing the circuits in Fig. 1 on an eigenstate of U in the presence of a pure depolarizing channel with a fixed failure probability $p = e^{-1/\tau}$ per iteration of U is bounded as*

$$\delta \geq \frac{\tau^{-\frac{1}{2}} T^{-\frac{1}{2}}}{2e}. \tag{19}$$

Proof. A pure depolarizing channel sends the off-diagonal element of the reduced density matrix in Eq. (4) to $p^k g(k)$. This adjustment can be propagated directly through to the Fisher information (Eq. (15)), which becomes

$$I(\phi) = \sum_k e^{-2k/\tau} k^2 (M_k^r + M_k^i), \tag{20}$$

while the total quantum cost (Eq. 16) remains the same. Let us consider the relative contribution to I versus the contribution to T of a single choice of k; if we write $I(\phi) = \sum_k I_k$ and $T = \sum_k T_k$, we have $I_k/T_k = ke^{-2k/\tau}$. Differentiating w.r.t. k and setting equal to zero yields

$$\frac{d}{dk}\frac{I_k}{T_k} = e^{-2k/\tau} - \frac{2k}{\tau} e^{-2k/\tau} = 0 \tag{21}$$

$$\to k = \frac{\tau}{2}. \tag{22}$$

Optimizing $I(\phi)$ with respect to T thus requires setting $k = \frac{\tau}{2}$ and increasing $M_k = M_k^r + M_k^i$, which yields

$$I(\phi) = \frac{\tau^2}{4e} M_k, T = \frac{\tau}{2} M_k. \tag{23}$$

Substituting this into the Cramér-Rao bound $\delta \geq I(\phi)^{-\frac{1}{2}}$ yields the desired result. □

Remark: Note that as we fix $T \sim \tau$ in the above, we in effect have Heisenberg-limited scaling in τ. However, if we treat τ as a constant this is only the sampling noise limit as defined above. This result holds only for the simplest-possible noise case; more complicated noise is difficult to analyse, but numerical results show it may prevent estimation beyond some minimum value using standard techniques [19]. We assume herein that all circuits are noiseless (and will not use Lemma 2.7 in the rest of this work).

3 Defining the task of multiple-eigenvalue phase estimation

In this section we define the goal of estimating multiple eigenvalue phases of some unitary U. When the input state $|\Psi\rangle$ is supported on multiple eigenstates, choosing a single k does not suffice, simply since knowing Eq. (2) at a single point k does not give a unique solution $\{A_j, \phi_j\}$ [19]. A simple way to circumvent this problem is thus to estimate 'densely': estimating $g(k)$ for all integers $0 \leq k \leq K$ would allow us to fit up to $O(K)$ (ϕ_j, A_j) pairs. However, this does not saturate the Heisenberg limit as shown in Theorem 2.5, hence we need to come up with a different method.

Separate from this, the full eigenspectrum of an arbitrary N-qubit unitary U has up to 2^N unique values, making it impossible to describe in polynomial time in N. In addition, the spectral content of the input state $|\Psi\rangle$ could be very dense, with many eigenvalues clustered together instead of separated by gaps, and the overlap for these eigenvalues, A_j, could be sharply concentrated or uniformly spread. To deal with general input states, Ref. [3] thus formulated the quantum eigenvalue estimation problem (QEEP): instead of estimating individual phases, the focus is on estimating the spectral function $A(\phi)$ in Eq. (3) with some resolution. We recall the precise definition of this problem in App. A, Def. A.1. In our case, we focus on the case where the initial state only has a non-zero overlap with a small number n_ϕ of eigenvectors of U, and we want to estimate eigenphases corresponding to each of them. Here is our precise definition of the problem to be solved:

Definition 3.1 (Multiple eigenvalue estimation problem)**.** *Fix an error bound $\delta > 0$, an overlap bound $A > 0$. For a unitary U and state $|\Psi\rangle$, we assume that $A_j > A$ for exactly n_ϕ phases ϕ_j and $A_j = 0$ for all other phases so that $n_\phi \leq A^{-1}$. Let $g(k) = \sum_j A_j e^{ik\phi_j}$ be the phase function in Def. 2.1 and assume access to the PFE quantum*

subroutine in Def. 2.2 for any $k \in \mathbb{R}^+$, generating data $\mathbf{x}$. The task is to output a set $\{\tilde{\phi}_l\}$ of n_ϕ or fewer estimates of the phases $\{\phi_j\}$ such that, if we take the closest estimate $\tilde{\phi}_j^{(\text{closest})}(\mathbf{x})$ of each phase ϕ_j given the data $\mathbf{x}$,

$$\tilde{\phi}_j^{(\text{closest})}(\mathbf{x}) = \arg\min_{\tilde{\phi}_l}(|\tilde{\phi}_l(\mathbf{x}) - \phi_j|_T), \quad (24)$$

the accuracy error

$$\delta_j = \sqrt{\sum_{\mathbf{x}} \mathbb{P}\left(\mathbf{x}|\{\phi_l, A_l\}\right) \left|\tilde{\phi}_j^{(\text{closest})}(\mathbf{x}) - \phi_j\right|_T^2}, \quad (25)$$

is bounded by $\delta_j \leq \delta$ for all $j = 1, \ldots, n_\phi$.

Remark: Def. 3.1 allows us the freedom to assign a single estimate to multiple phases when calculating the final mean-squared-error.

3.1 Methods of dense signal phase estimation

Achieving the Heisenberg limit for multiple eigenvalues requires solving the problem considered in Def. 3.1 with a total quantum cost $T = O(\delta^{-1})$. We intend to accomplish that goal with a multi-order estimation scheme. At each order d, we will use a data processing method to estimate multiple eigenphases $\theta_j^{(d)}$ of U^{k_d} to within some error ϵ, from data generated by PFE (analogous to step 2b of Alg. 2.6). (We will stitch the estimates of the phases $\theta_j^{(d)}$ together to give Heisenberg-limited estimates of the corresponding ϕ_j in a manner similar to step 2d of Alg. 2.6.) We can offload the estimation of $\theta_j^{(d)}$ to a subroutine; we will show later that we can afford a subroutine with superlinear scaling in ϵ as the final error δ in our multi-order scheme can be made arbitrarily small even at fixed ϵ. In this section we discuss the two subroutines that we will consider in this work: the matrix pencil method (Alg. 5.1) first studied for QPE in Ref. [19], and the 'time series analysis' proposed in Ref. [3] to solve the QEEP problem mentioned above.

We detail our implementation of the matrix pencil method in Alg. 5.1; this is a well-known algorithm in signal processing, that is known to achieve the dense-sampling limit for a single eigenvalue [30, 31]. However, bounding the error of the matrix pencil method in estimating many phases typically requires a minimal gap Δ between these phases, $\Delta = \min_{i \neq j} |\phi_i - \phi_j|_T$, and that we query the PFE to obtain estimates of $g(k)$ at $k > \frac{1}{\Delta}$. This is a proven necessary condition to estimate multiple ϕ_j to error $\epsilon \leq c\Delta$ [32] for some constant c. In our case, we need to allow for the case where we are estimating two phases $\theta_0^{(d)}$, $\theta_1^{(d)}$ to an error $\epsilon \geq |\theta_0^{(d)} - \theta_1^{(d)}|$. In principle this is not forbidden by the result of [32] (and numerical simulation confirms that this indeed works), but we do not know of a formal statement about the scaling of the matrix pencil method in this situation. Instead, we opt for a different method for the purposes of forming a rigorous proof of the Heisenberg limit, and test the matrix pencil method in numerics only.

Ref. [3] proved rigorously the QEEP can be solved from the densely sampled signal $g(k)$ generated with the PFE in Def. 2.2 using a 'time-series analysis' algorithm. We review the results of Ref. [3] in detail in App. A. However, the solution to the QEEP is an estimation $\tilde{A}(\phi)$ of a discretization of the spectral function $A(\phi)$ (Eq. (3)) rather than a set of estimates $\{\tilde{\theta}_j^{(d)}\}$. To use the time-series analysis algorithm as a subroutine in our multi-order phase estimation algorithm then requires converting from one form to the other. This is achieved by Alg. A.3, the Conservative QEEP Eigenvalue Extraction algorithm. This algorithm is designed so that its output fulfills the following guarantees whenever the time-series analysis algorithm succeeds (defined as $\|\tilde{A}(\phi) - A(\phi)\|_1 \leq \epsilon$)

Lemma 3.2. *Fix a confidence bound $0 < p < 1$, an overlap bound A, a number of phases $n_\theta < \frac{1}{A}$, and an error bound $0 < \epsilon < \frac{A}{3}$. Let $g(k) = \sum_j A_j e^{ik\theta_j}$ be the phase function for a unitary V, with $A_j > A$ for exactly n_θ phases θ_j, and $A_j = 0$ for all other phases. Let $\{\tilde{\theta}_l\}$ be a set of estimates of $\{\theta_j\}$ generated by Alg. A.3 with error ϵ and confidence bound p. With probability at least p, the following statements are true:*

1. *For each phase θ_j with $A_j > 0$, there exists at least one estimate $\tilde{\theta}_l$ such that*
$$|\theta_j - \tilde{\theta}_l|_T \leq 2\epsilon.$$
2. *For each estimate $\tilde{\theta}_l$ there exists at least one phase θ_j with $A_j > 0$ such that*
$$|\theta_j - \tilde{\theta}_l|_T \leq 2\epsilon.$$
3. *The number of estimates $|\{\tilde{\theta}_l\}| \leq n_\theta$.*

See App A for a proof. The total quantum cost of Alg. A.3 is inherited directly from the cost of the time series analysis algorithm (as it involves no additional quantum circuitry), which is $O(\epsilon^{-6}|\log(1-p)|)$.

4 Multiple eigenvalues: multi-order estimation and the phase matching problem

To achieve Heisenberg-limited scaling for multiple phases, we combine the dense signal algorithms which can resolve multiple phases in the previous section with the single-phase Heisenberg limited algorithm, Algorithm 2.6, which achieves the correct scaling.

A natural way to achieve such combination is to estimate phases $\{\theta_j^{(d)}\}$ of $V = U^{2^d}$ for multiple orders $d = 0, 1 \ldots, d_f$ via a dense signal method (e.g. Alg. A.3), and then combine them in the same manner as in Algorithm 2.6.

If we would manage to get an estimate of $\theta_j^{(d)} = \phi_j 2^d$ at each order d with error ϵ *and* be able to combine these estimates in an unequivocal manner, then reaching the Heisenberg limit for multiple phases may be feasible. Note that the error in the final d_f^{th} estimate in this case would be $\delta \sim \epsilon/2^{d_f}$ with ϵ in Algorithm A.3. One could thus achieve arbitrarily small δ for fixed ϵ by making d_f arbitrarily large. This allows us to use (possibly non-optimal) routines such as the QEEP algorithm since the scaling with ϵ does not propagate to a scaling in δ for a sufficiently small ϵ.

However, the combination of phase information at different orders in the case of multiple phases may not be feasible when we use $V = U^{2^d}$ for increasing d as in Algorithm 2.6. The reason is that if we have multiple phases, the previous order estimates provide sets of 'ballpark' intervals and it may not clear or unambiguous which interval to choose in order to convert a new estimate $\tilde{\theta}_j^{(d)}$ to an updated $\tilde{\phi}_j^{(d)}$ (as in step 2d of Algorithm 2.6). We would like the choice of the next order to be such that this 'matching with a previous estimate' can be done unambiguously.

For this, we will estimate the multiple eigenphases of $V = U^{k_d}$ for $k_d = \prod_{d'=1}^{d} \kappa_{d'}$ with $\kappa_d \geq 2$ a, possibly non-integer, multiplier. For this algorithm to have a means of associating each dth order estimate $\tilde{\theta}_j^{(d)}$ with a single previous-order estimates $\tilde{\phi}_j^{(d-1)}$, we use an adaptive strategy for choosing the next multiplier κ_d in Alg. 4.1. That is, the algorithm will determine a κ_d based on the estimates $\tilde{\phi}_j^{(d-1)}$ from the previous round. Although this scheme requires some classical processing of the experimental data before the experiment is finished, it is not an adaptive scheme in the same sense as iterative QPE [14], as we do not require feedback within the coherent lifetime of a single experiment.

The generalization from using U^{2^d} to U^{k_d} for $k_d \in \mathbb{R}^+$ presents one small additional complication. In order to prove bounds on the estimation at each order we will require invoking Lemma 2.4. However, this requires that our phases ϕ_j satisfy Eq. (7) (unless $k_d \in \mathbb{Z}^+$). If a phase ϕ_j does not satisfy Eq. (7), one can construct a situation where two corresponding estimates $\tilde{\phi}_j^{(d)}$ are found on either side of the branch cut at 2π, and where we cannot guarantee that our algorithm would choose the 'correct' one (without knowledge of the hidden ϕ_j). To solve this issue, we note that one may shift the phases of U by a constant χ by performing phase estimation on $Ue^{-i\chi}$ instead of U. This need not even be done on the quantum device, as one simply multiplies estimates of $g(k)$ by $e^{-ik\chi}$. As we assume the existence of only n_ϕ phases, we can always find some $Ue^{i\chi}$ with phases in some window $[\phi_{\min}, \phi_{\max})$ with $\phi_{\min} \geq \frac{\pi}{k}$, $\phi_{\max} \leq \frac{\pi(2\lfloor k \rfloor - 1)}{k}$ when $k \geq 3n_\phi$. This will allow us to invoke Lem. 2.4 to match estimates of eigenphases of $Ue^{i\chi}$ and estimates of eigenphases of $(Ue^{i\chi})^k$ as we require. We also note that the above issue can be circumvented when $U = e^{iHt}$ by a suitable choice of t.

4.1 Heisenberg-limited algorithm for multiple phases

We now describe our Heisenberg-limited phase estimation algorithm. This algorithm targets a final error $\delta = O(\delta_c)$, where δ_c is a fixed input to the algorithm itself (We will calculate the constant of proportionality in the proof of Theorem 4.5). The Heisenberg limit will be achieved by making this δ_c smaller while keeping the error ϵ of the phase extraction subroutine, Alg. A.3, constant.

Algorithm 4.1. *[Adaptive multi-order phase estimation algorithm] We assume access to the con-*

servative QEEP eigenvalue extraction algorithm, Alg. A.3 for a unitary $V = U^k$ (for arbitrary $k \in \mathbb{R}^+$), and an initial state $|\Psi\rangle$. Fix a final error δ_c, an overlap bound A, a number of phases $n_\phi \leq A^{-1}$, and error parameters ϵ_0 and ϵ bounded as

$$\epsilon_0 \leq \epsilon_{\text{crit},0} \equiv \frac{2\pi}{300 n_\phi^4} \tag{26}$$

and

$$\epsilon \leq \epsilon_{\text{crit}} \equiv \frac{2\pi}{300 n_\phi^2}. \tag{27}$$

Let the confidence parameter p_d be

$$p_d = 1 - e^{-\alpha}\left(\frac{k_d \delta_c}{\pi}\right)^\gamma, \tag{28}$$

given some real numbers $\alpha > 0$, $\gamma > 2$ and k_d to be chosen below. The algorithm proceeds as follows:

1. *Let $d = 0$ and $k_{d=0} = 1$. Use Alg. A.3 to find a set of first estimates $\{\tilde{\phi}_j^{(0)}\}$ of eigenvalues of U with error parameter ϵ_0 in Eq. (26), overlap bound A, and confidence $p_{d=0}$ in Eq. (28). If this set is empty or has more than n_ϕ elements, return $\{0\}$ (this is a failure mode).*

2. *Find the point $\zeta \in [0, 2\pi)$ defined by*

$$\zeta = \underset{\zeta' \in [0,2\pi]}{\arg\max} \min_j \left|\tilde{\phi}_j^{(0)} - \zeta'\right|_T, \tag{29}$$

i.e. ζ is the midway point in the largest gap between the phase estimates $\tilde{\phi}_j^{(0)}$. Let

$$d_\zeta = \min_j \left|\tilde{\phi}_j^{(0)} - \zeta\right|_T, \tag{30}$$

i.e. d_ζ is half the size of the largest gap. Shift the unitary $U \to U e^{-i(\zeta + \frac{1}{2}d_\zeta - 8\epsilon_0)}$, $\tilde{\phi}_j^{(0)} \to \tilde{\phi}_j^{(0)} - \zeta - d_\zeta/2 + 8\epsilon_0 \mod 2\pi$.

3. *Choose $\kappa_1 = k_1$ with $k_1 \in [3n_\phi, 3n_\phi + 1]$ such that for all $\tilde{\phi}_j^{(0)} \neq \tilde{\phi}_l^{(0)}$, either*

$$|\tilde{\phi}_j^{(0)} k_1 - \tilde{\phi}_l^{(0)} k_1|_T > 4\epsilon_0(1 + k_1). \tag{31}$$

or

$$|\tilde{\phi}_j^{(0)} - \tilde{\phi}_l^{(0)}|_T < \frac{\pi}{k_1}. \tag{32}$$

4. *While $k_d < \frac{2\epsilon}{\delta_c}$:*

   (a) *Set $d \to d + 1$.*

   (b) *Use Alg. A.3 to find a set of estimates $\{\tilde{\theta}_l^{(d)}\}$ of eigenvalues of $V = U^{k_d}$ with error parameter ϵ in Eq. (27), overlap bound A, and confidence p_d in Eq. (28).*

   (c) *If there exists some $\tilde{\phi}_j^{(d-1)}$ such that*

$$\min_l |k_d \tilde{\phi}_j^{(d-1)} - \tilde{\theta}_l^{(d)}|_T > 2\epsilon(1 + \kappa_d), \tag{33}$$

   or there exists some $\tilde{\theta}_l^{(d)}$ such that

$$\min_j |k_d \tilde{\phi}_j^{(d-1)} - \tilde{\theta}_l^{(d)}|_T > 2\epsilon(1 + \kappa_d), \tag{34}$$

   or the number of estimates $|\{\tilde{\theta}_l^{(d)}\}| > n_\phi$, return $\{\tilde{\phi}_j^{(d-1)} + \zeta + d_\zeta/2 - 8\epsilon_0 \mod 2\pi\}$. This is a failure mode.

   (d) *If not, for each $\tilde{\theta}_l^{(d)}$, find the estimate $\tilde{\phi}_j^{(d-1)}$ and an integer $n \in [0, k_d)$ which minimizes*

$$|\tilde{\phi}_j^{(d-1)} - (\tilde{\theta}_l^{(d)} + 2\pi n)/k_d|_T, \tag{35}$$

   and set $\{\tilde{\phi}_l^{(d)}\} = \{(\tilde{\theta}_l^{(d)} + 2\pi n)/k_d\}$.

   (e) *If any $\tilde{\phi}_j^{(d)} \in [0, \frac{\pi}{k_d}) \cup (\frac{\pi(2\lfloor k_d \rfloor - 1)}{k_d}, 2\pi]$, return $\{\tilde{\phi}_j^{(d-1)} + \zeta + d_\zeta/2 - 8\epsilon_0 \mod 2\pi\}$. This is a failure mode.*

   (f) *Choose the multiplier $\kappa_{d+1} \in [2, 3]$ such that for all $\tilde{\phi}_j^{(d)} \neq \tilde{\phi}_l^{(d)}$, either*

$$|\tilde{\phi}_j^{(d)} k_d \kappa_{d+1} - \tilde{\phi}_l^{(d)} k_d \kappa_{d+1}|_T > 4\epsilon(1 + \kappa_{d+1}). \tag{36}$$

   or

$$|\tilde{\phi}_j^{(d)} - \tilde{\phi}_l^{(d)}|_T < \frac{\pi - 2\epsilon(1 + \kappa_{d+1})}{k_d \kappa_{d+1}}, \tag{37}$$

   and set $k_{d+1} = k_d \kappa_{d+1}$.

5. *Return $\{\tilde{\phi}_j^{(d)} + \zeta + d_\zeta/2 - 8\epsilon_0 \mod 2\pi\}$.*

In principle the first few orders d could be skipped given accurate prior knowledge of our phases ϕ_j. However, as the largest circuits are executed at the latter d values, this will only change the constant factor of the algorithm, rather than the asymptotic scaling with δ.

In the rest of this section, we prove that Alg. 4.1 can achieve the Heisenberg limit. The first use of the QEEP subroutine requires a potentially smaller error parameter (ϵ_0, bounded by Eq. (26)) than subsequent uses (where ϵ needs to

be only bounded by Eq. (27)). (Invoking Alg. A.3 requires that the error parameters ϵ and ϵ_0 are at most $A/3 \leq 1/(3n_\phi)$, which is fulfilled by both bounds.) This relates to a technical issue: we require $k_1 \geq 3n_\phi$ in order for Lemma C.1 in Appendix C and thus Lemma 2.4 to apply. For later rounds $k_d \geq 3n_\phi$ automatically, and the multiplier κ_d is no longer constrained, which indirectly allows us to relax the region of valid choices for ϵ. The first step in proving the performance of Algorithm 4.1 is to show that the multipliers can be chosen in the first (step 3 in Alg. 4.1), and subsequent rounds (step 4f in Alg. 4.1), which obey the desired conditions. This is accomplished by the following Lemma which is proved in Appendix B.

Lemma 4.2. *Let $\{\tilde{\phi}_j^{(0)}\} \in [0, 2\pi)$ be a set of at most n_ϕ phases. Assuming Eq. (26), for a randomly chosen $k_1 \in [3n_\phi, 3n_\phi + 1]$ with probability at least $1/2$, either Eq. (31) or Eq. (32) holds for all $\tilde{\phi}_j^{(0)} \neq \tilde{\phi}_l^{(0)}$. Fix a k_d. Let $\{\tilde{\phi}_j^{(d)}\} \in [0, 2\pi)$ be a set of at most n_ϕ phases. Assuming Eq. (27), for a randomly chosen $\kappa_{d+1} \in [2, 3]$ with probability at least $3/4$, either Eq. (36) or Eq. (37) holds for all $\tilde{\phi}_l^{(d)} \neq \tilde{\phi}_j^{(d)}$.*

Remarks: The probability with which a multiplier can be found which obeys the desired property is rather arbitrary in this Lemma and can be increased by choosing a smaller ϵ. Note that it is easy to verify whether for a randomly chosen multiplier the desired conditions hold or not. The validity of this Lemma importantly does not depend on whether the phase estimates are actually accurate, it only depends on the number of phases n_ϕ. In practice, we do not generate a random multiplier κ_{d+1} through this Lemma, but simply exhaustively search for a valid κ_{d+1} starting at the maximal value, see Section 5.

The reason to adaptively choose the multiplier κ_{d+1} for $d = 0, 1, \ldots$ is that two (estimated) phases in principle need to lead to separate estimates at the next order: this is expressed in Eq. (36). An exception to this occurs when the (estimated) phases are still close enough, as in Eq. (37), so that their next-order refined estimates could merge at the next order, see Fig. 2. Phase estimates can thus split and merge over the multiple orders. They split when sufficient accuracy is available at the next order to distinguish them, they can stay or are allowed to merge when such accuracy is not yet needed at the given order.

In what follows below we will assume, just for simplicity of the proof, that the error parameter ϵ is bounded by $\epsilon_{\text{crit},0}$ in Eq. (26) for all rounds, and ϵ is the same for all rounds, including the first one.

4.2 Bounding the error with and without failures

In this section we state and prove the two key intermediate lemmas, Lemma 4.3 and Lemma 4.4 on our way towards proving that Alg. 4.1 reaches the Heisenberg-limit. Together, these lemmas allow us to bound the error in Alg. 4.1, –assuming that the phase extraction subroutine succeeds for the first d rounds–, to within $O(\epsilon/k_d)$.

These Lemmas deal with the issue of 'aliasing' or the correct matching of new estimates with older estimates which is solved by the specific choice of κ_{d+1} in step 4f of Alg. 4.1, see also Fig. 2. It is important to note that there is no 1-1 relation between these estimates and the actual phases since the number of estimates is at most the number of phases.

Let d_f be the last order executed in Alg. 4.1, i.e. the last order for which we go through step 4b, construct the estimates $\{\tilde{\theta}_l^{(d_f)}\}$ and pass the tests at step 4c and 4e and output $\{\tilde{\phi}_l^{(d_f)}\}$. When none of the failure modes is encountered, d_f is set by the first k_d such that $k_d \geq \frac{\epsilon}{\delta_c}$ (since the next $k_{d+1} \geq 2\epsilon/\delta_c$ as $\kappa_{d+1} \geq 2$). Since $\kappa_d \geq 2$, we observe that

$$d_f \leq \log_2\left(\frac{2\epsilon}{\delta_c}\right). \tag{38}$$

In Corollary 4.3.1 we argue that when the QEEP subroutines, Alg. A.3, succeed up to order d_f, we indeed never exit via these failure modes.

Lemma 4.3. *If each invocation of the QEEP subroutine, Alg. A.3, succeeds in Alg. 4.1 up to order d_f, then in this last round d_f in step 4d it holds that*

- *(Property 1a) For every phase ϕ_j there exists an estimate $\tilde{\phi}_l^{(d_f)}$ such that*

$$|\phi_j - \tilde{\phi}_l^{(d_f)}|_T \leq 2\epsilon/k_{d_f}.$$

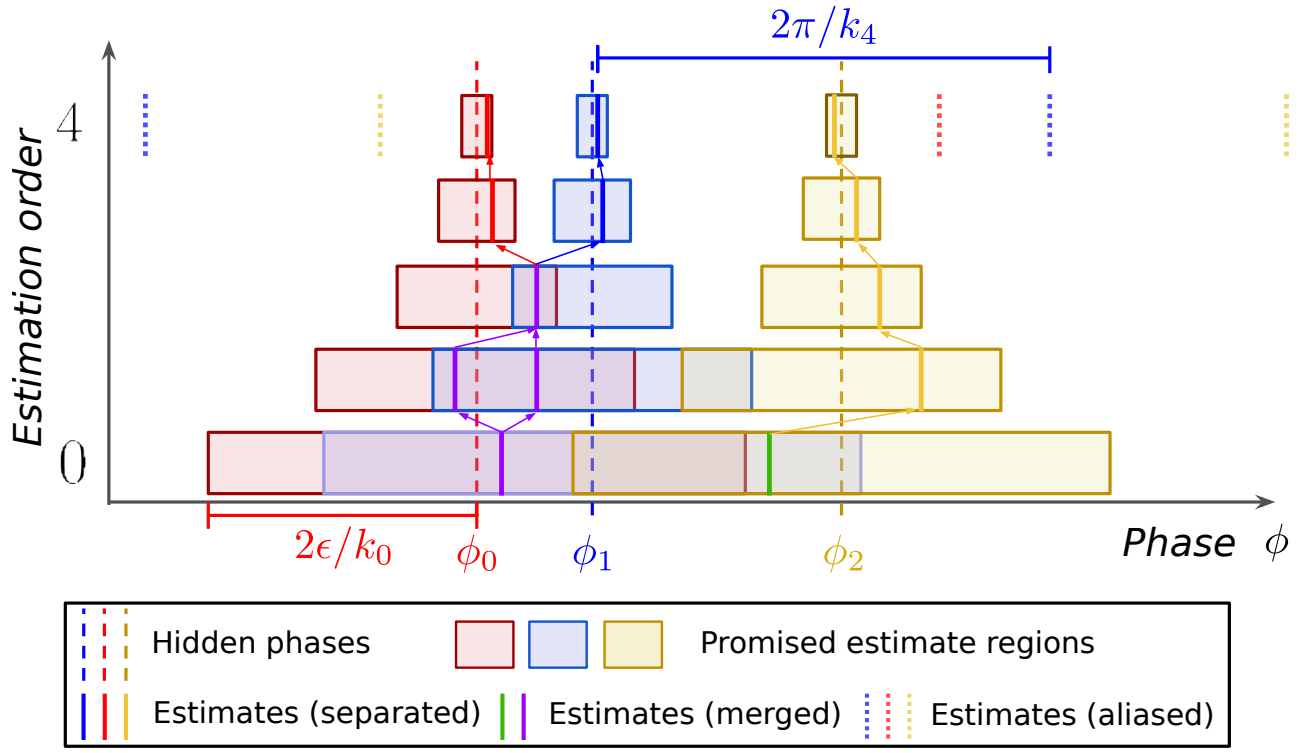

Figure 2: Schematic of the execution of Alg. 4.1 to estimate three hidden phases, ϕ_0, ϕ_1 and ϕ_2 (dashed lines). The algorithm progresses from bottom to top as the estimation order d increases. At each order d, the phase extraction subroutine (Alg. A.3) promises to return an estimate $\tilde{\theta}^{(d)}_{\to j}$ of each $\theta^{(d)}_j = k_d\phi_j \mod 2\pi$ that corresponds to an estimate $\tilde{\phi}^{(d)}_{\to j}$ (solid lines) lying within the promised estimate region about ϕ_j (coloured boxes). By matching phases at subsequent orders (arrows), the algorithm is able to converge to an ever-more accurate estimate of each ϕ_j. The phase extraction subroutine only promises that each region will contain at least one phase (and that the total number of estimates at each order is bounded by n_ϕ) - when two regions overlap, the subroutine may merge the phases to give a single estimate (green and purple lines). Estimates at subsequent orders may continue to separate and even re-merge until the regions separate, at which point the algorithm promises with high confidence a precise estimate for each hidden phase. The estimate $\tilde{\phi}^{(d)}_{\to j}$ at each order is only known $\mod (2\pi)/k_d$, leading to a set of potential aliases (dotted lines at $d = 4$) for each phase. We do not know a priori which alias is correct, and must rely on the fact that the true estimate $\tilde{\phi}^{(d)}_{\to j}$ needs to be close to a previous estimate $\tilde{\phi}^{(d-1)}_{\to j}$. By carefully choosing each k_d, we can guarantee that no alias will satisfy this condition (so long as Alg. A.3 succeeds), and our phase matching will be unambiguous.

- *(Property 1b) For every estimate $\tilde{\phi}^{(d_f)}_l$ there exists a phase ϕ_j such that*

$$|\phi_j - \tilde{\phi}^{(d_f)}_l|_T \leq 2\epsilon/k_{d_f}.$$

Proof. We prove this Lemma by induction. Consider the first round $d = 0$ ($k_{d=0} = 1$), i.e. step 1 of Alg. 4.1. If the QEEP subroutine, Alg. A.3, succeeds (with probability p_0) then Lemma 3.2 holds, namely for each ϕ_j there exists an estimate $\tilde{\phi}^{(0)}_l$ such that

$$|\phi_j - \tilde{\phi}^{(0)}_l|_T \leq 2\epsilon. \tag{39}$$

and for each estimate $\tilde{\phi}^{(0)}_l$ there exists at least one ϕ_j such that

$$|\phi_j - \tilde{\phi}^{(0)}_l|_T \leq 2\epsilon. \tag{40}$$

Hence the statement to be proven holds at $d = 0$. Now consider step 2 of Alg. 4.1 and invoke Lemma C.1 for which the assumptions are fulfilled by Eqs. (39),(40). This implies that with the choice of $k_1 \geq 3n_\phi$ in step 3 in Alg. 4.1 the shifted phases and their 0th-order estimates obey the technical condition in Lemma 2.4 and we can use Eq. (8). In the next steps we work with these shifted phases but for simplicity we don't use any new notation and refer to them as ϕ_j and estimates $\tilde{\phi}^{(d)}_j$ etc.

Now assume the statement to be proven holds

at order d i.e. let $\{\tilde{\phi}_l^{(d)}\}$ be a set of at most n_ϕ estimates of the phases $\{\phi_j\}$ with

- (Assumption 1a) For every phase ϕ_j there exists an estimate $\tilde{\phi}_l^{(d)}$ such that

$$|\phi_j - \tilde{\phi}_l^{(d)}|_T \leq 2\epsilon/k_d.$$

- (Assumption 1b) For every estimate $\tilde{\phi}_l^{(d)}$ there exists a phase ϕ_j such that

$$|\phi_j - \tilde{\phi}_l^{(d)}|_T \leq 2\epsilon/k_d.$$

Note that these assumptions certainly imply that one can apply Lemma 2.4 to the estimates $\tilde{\phi}_l^{(d)}$. That is, given that the real phases ϕ_j are $2\epsilon/k_d$ close to these estimates and that the (shifted) ϕ_j obey Eq. (110), it implies that Eq. (8) can be used with $k \geq 3n_\phi$ (which is the case for all rounds $d \geq 1$).

We consider the QEEP subroutine, Alg. A.3, with a given choice of κ_d obeying the conditions in step 3 (for $d = 1$) and step 4f (for higher d), executed in step 4b with confidence p_d. In the math below we refer to conditions on $\kappa_{d>1}$, namely Eq. (36) and Eq. (37), but the conditions on κ_1 in Eq. (31) and Eq. (32) are of identical form (so we don't make a separate argument for the $d = 0 \to d = 1$ induction step).

Let thus $\{\tilde{\theta}_l^{(d+1)}\}$ be a set of estimates of the eigenphases $\{\theta_j^{(d+1)}\}$ of $U^{k_{d+1}}$ corresponding to the set $\{\phi_j\}$, that is,

$$\theta_j^{(d+1)} = k_{d+1}\phi_j \mod 2\pi. \tag{41}$$

and $k_{d+1} = k_d\kappa_{d+1}$. By assuming that Alg. A.3 succeeds we can invoke Lemma 3.2, namely

- (Assumption 2a) For every phase $\theta_j^{(d+1)}$ there exists an estimate $\tilde{\theta}_l^{(d+1)}$ such that

$$|\theta_j^{(d+1)} - \tilde{\theta}_l^{(d+1)}|_T \leq 2\epsilon.$$

- (Assumption 2b) For every estimate $\tilde{\theta}_l^{(d+1)}$ there exists a phase $\theta_j^{(d+1)}$ such that

$$|\theta_j^{(d+1)} - \tilde{\theta}_l^{(d+1)}|_T \leq 2\epsilon.$$

To prove the induction step, we thus need to show that the set $\tilde{\phi}_j^{(d+1)}$ generated by step 4d of Alg. 4.1 satisfies the following two properties

- (Property 1a) For every phase ϕ_j there exists an estimate $\tilde{\phi}_l^{(d+1)}$ such that

$$|\phi_j - \tilde{\phi}_l^{(d+1)}|_T \leq 2\epsilon/k_{d+1}.$$

- (Property 1b) For every estimate $\tilde{\phi}_l^{(d+1)}$ there exists a phase ϕ_j such that

$$|\phi_j - \tilde{\phi}_l^{(d+1)}|_T \leq 2\epsilon/k_{d+1}.$$

First consider Assumption 2a. Assumption 2a implies that for every phase ϕ_j there exists a $\tilde{\theta}_l^{(d+1)}$ such that

$$|k_{d+1}\phi_j - \tilde{\theta}_l^{(d+1)}|_T \leq 2\epsilon. \tag{42}$$

In this proof we will use the label $l =\!\!\to j$ for this $\tilde{\theta}_l^{(d+1)}$ associated with ϕ_j. Thus, also using Eq. (8), Eq. (42) is equivalent to

$$\min_{n\in\{0,\ldots,\lfloor k_{d+1}\rfloor-1\}} |\phi_j - (\tilde{\theta}_{\to j}^{(d+1)} + 2\pi n)/k_{d+1}|_T \leq \frac{2\epsilon}{k_{d+1}}, \tag{43}$$

with $n_{j,\to j}^{\text{ideal}}$ the integer which achieves the minimum, i.e.

$$n_{j,\to j}^{\text{ideal}} = \operatorname*{arg\,min}_{n\in\{0,\ldots,\lfloor k_{d+1}\rfloor-1\}} \Big\{ |\phi_j - (\tilde{\theta}_{\to j}^{(d+1)} + 2\pi n)/k_{d+1}|_T \Big\}. \tag{44}$$

Similarly, by Assumption 1a, there is some $\tilde{\phi}_{l=\to j}^{(d)}$ which is $2\epsilon/k_d$-close to ϕ_j, again using a label which shows this association.

Consider the optimization in Eq. (35) at step 4d in Alg. 4.1. We define

$$\xi_l = \min_j \xi_{j,l}, \tag{45}$$

$$\xi_{j,l} \equiv \left|\tilde{\phi}_j^{(d)} - (\tilde{\theta}_l^{(d+1)} + 2\pi n_{j,l})/k_{d+1}\right|_T \tag{46}$$

with

$$n_{j,l} = \operatorname*{arg\,min}_{n\in\{0,\ldots,\lfloor k_{d+1}\rfloor-1\}} \Big\{ \left|\tilde{\phi}_j^{(d)} - (\tilde{\theta}_l^{(d+1)} + 2\pi n)/k_{d+1}\right|_T \Big\}, \tag{47}$$

$$n_l = \operatorname*{arg\,min}_{n_{j,l}} \xi_{j,l}. \tag{48}$$

The goal is thus to prove that for each ϕ_j, using the corresponding $\tilde{\theta}_{\to j}^{(d+1)}$, we have $n_{\to j} = n_{j,\to j}^{\text{ideal}}$ which directly implies Property 1a.

We can bound using Eq. (6) and then Eq. (8), Assumptions 1a and 2a and the optimality of $n_{\to j,\to j}$,

$$\begin{aligned}
&\xi_{\to j,\to j} \\
&= \left|\tilde{\phi}^{(d)}_{\to j} - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}\right|_T \\
&\leq \left|\tilde{\phi}^{(d)}_{\to j} - \phi_j\right|_T \\
&\quad + \left|\phi_j - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n^{\text{ideal}}_{j,\to j})/k_{d+1}\right|_T \\
&\leq \frac{2\epsilon}{k_d} + \frac{1}{k_{d+1}}\left|k_{d+1}\phi_j - \tilde{\theta}^{(d+1)}_{\to j}\right|_T \\
&= \frac{2\epsilon}{k_d} + \frac{1}{k_{d+1}}\left|\theta^{(d+1)}_j - \tilde{\theta}^{(d+1)}_{\to j}\right|_T \\
&\leq \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}}. \qquad (49)
\end{aligned}$$

Now if Eq. (36) holds for some other $m \not\to j$, we claim on the other hand that

$$\xi_{m,\to j} > \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}}, \qquad (50)$$

hence matching $\theta^{(d+1)}_{\to j}$ with such $\tilde{\phi}^{(d)}_m$, with $m \not\to j$ is non-optimal and will not be chosen in the Algorithm. To indeed see that Eq. (36) implies Eq. (50), we can calculate

$$\begin{aligned}
\frac{4\epsilon(1+\kappa_{d+1})}{k_{d+1}} &< \frac{1}{k_{d+1}}\left|k_{d+1}\tilde{\phi}^{(d)}_{\to j} - k_{d+1}\tilde{\phi}^{(d)}_m\right|_T \\
&\leq \frac{1}{k_{d+1}}\left|k_{d+1}\tilde{\phi}^{(d)}_{\to j} - \tilde{\theta}^{(d+1)}_{\to j}\right|_T \\
&\quad + \frac{1}{k_{d+1}}\left|\tilde{\theta}^{(d+1)}_{\to j} - k_{d+1}\tilde{\phi}^{(d)}_m\right|_T \\
&\leq \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}} + \min_{n\in\{0,\ldots,\lfloor k_{d+1}\rfloor-1\}}\Big\{ \\
&\qquad \left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n)/k_{d+1}\right|_T\Big\} \qquad (51) \\
\Rightarrow \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}} &< \\
\min_{n\in\{0,\ldots,\lfloor k_{d+1}\rfloor-1\}} &\left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n)/k_{d+1}\right|_T \\
&= \xi_{m,\to j}. \qquad (52)
\end{aligned}$$

Here we have used that Eq. (8) holds for the estimate $\tilde{\phi}^{(d)}_m$.

Alternatively, for those $m \not\to j$ for which Eq. (37) holds, we claim that

$$n_{m,\to j} = n_{\to j,\to j}, \qquad (53)$$

hence for those $\tilde{\phi}^{(d)}_m \neq \tilde{\phi}^{(d)}_{\to j}$ the algorithm produces a single new estimate equal to $(\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}$.

To see that Eq. (37) implies Eq. (53) indeed, note that it is sufficient to prove that

$$\left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}\right|_T < \frac{\pi}{k_{d+1}}, \qquad (54)$$

as one can then show that for $n' \neq n_{\to j,\to j} \in \{0,\ldots,\lfloor k_{d+1}\rfloor - 1\}$ that

$$\begin{aligned}
\frac{2\pi}{k_{d+1}} &\leq \left|\frac{2\pi}{k_{d+1}}(n_{\to j,\to j} - n')\right|_T \\
&= |(\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1} \\
&\qquad - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n')/k_{d+1}|_T \\
&\leq \left|(\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1} - \tilde{\phi}^{(d)}_m\right|_T \\
&\quad + \left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n')/k_{d+1}\right|_T \\
&< \frac{\pi}{k_{d+1}} + \left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n')/k_{d+1}\right|_T \qquad (55) \\
\Rightarrow \frac{\pi}{k_{d+1}} &< \left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n')/k_{d+1}\right|_T, \qquad (56)
\end{aligned}$$

so $n_{\to j,\to j}$ is optimal. Using Eq. (6), Eq. (37) and Eq. (60), we can prove Eq. (54) since

$$\begin{aligned}
&\left|\tilde{\phi}^{(d)}_m - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}\right|_T \\
&\leq \left|\tilde{\phi}^{(d)}_m - \tilde{\phi}^{(d)}_{\to j}\right| \\
&\quad + \left|\tilde{\phi}^{(d)}_{\to j} - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}\right|_T \\
&< \frac{\pi - 2\epsilon(1+\kappa_{d+1})}{k_{d+1}} + \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}} \\
&= \frac{\pi}{k_{d+1}}. \qquad (57)
\end{aligned}$$

We have thus shown that for each ϕ_j, there is a $\tilde{\theta}^{(d+1)}_{\to j}$, such that step 4d will output $(\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}$ with $n_{\to j,\to j}$ defined in Eq. (48), related to the previous order estimate $\phi^{(d)}_{\to j}$ which was already close to $\phi^{(d)}_j$. The last step is to show that $n_{\to j,\to j} = n^{\text{ideal}}_{j,\to j}$ using Property 1a. It holds that

$$\begin{aligned}
&|\phi_j - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}|_T \\
&\leq |\phi_j - \tilde{\phi}^{(d)}_{\to j}|_T \\
&\quad + |\tilde{\phi}^{(d)}_{\to j} - (\tilde{\theta}^{(d+1)}_{\to j} + 2\pi n_{\to j,\to j})/k_{d+1}|_T \\
&\leq \frac{2\epsilon}{k_d} + \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}} < \frac{\pi}{k_{d+1}}, \qquad (58)
\end{aligned}$$

where we used that $4\epsilon(\kappa_{d+1}+1) < \pi$. Indeed for $d > 1$, $\epsilon < \frac{\pi}{16}$ ($\kappa_{d+1} \leq 3$) and for $d = 0$, $\epsilon < \frac{\pi}{4(3n_\phi+2)}$ ($\kappa_1 \leq 3n_\phi + 1$) given the upper bounds on ϵ in Eqs. (27) and (26). This implies through the same argument as in Eq. (56) that

$n^{\text{ideal}}_{j,\to j}$ achieving the minimum in Eq. (43) equals $n_{\to j,\to j}$ and hence we obtain Property 1a.

Now let's prove Property 1b. Given a $\tilde{\theta}^{(d+1)}_l$, let $\phi_{j=\to l}$ be the real phase for which, by Assumption 2b, it holds that

$$\begin{aligned}|k_{d+1}\phi_{\to l} - \tilde{\theta}^{(d+1)}_l|_T \\ = k_{d+1}|\phi_{\to l} - (\tilde{\theta}^{(d+1)}_l + 2\pi n^{\text{ideal}}_{\to l,l})/k_{d+1}|_T \\ \leq 2\epsilon. \quad (59)\end{aligned}$$

To prove Property 1b, we need to show that $n_l = n^{\text{ideal}}_{\to l,l}$ with n_l defined in Eq. (48). Let also $\tilde{\phi}^{(d)}_{\to j}$ be the previous order $2\epsilon/k_d$-close estimate to $\phi_{j=\to l}$ by Assumption 1a. The idea is that $\tilde{\phi}^{(d)}_{\to j} = \tilde{\phi}^{(d)}_{\to(\to l)}$ is matched to $\tilde{\theta}^{(d+1)}_l$ in the optimization step of the algorithm, so that this leads to a better estimate for the phase $\phi_{j=\to l}$.

Given a $\tilde{\theta}^{(d+1)}_l$ we can deduce, as before, that

$$\begin{aligned}\xi_{\to j,l} &= \left|\tilde{\phi}^{(d)}_{\to j} - (\tilde{\theta}^{(d+1)}_l + 2\pi n_{\to j,l})/k_{d+1}\right|_T \\ &\leq \left|\tilde{\phi}^{(d)}_{\to j} - \phi_{\to l}\right|_T \\ &\quad + \left|\phi_{\to l} - (\tilde{\theta}^{(d+1)}_l + 2\pi n^{\text{ideal}}_{\to l,l})/k_{d+1}\right|_T \\ &\leq \frac{2\epsilon}{k_d} + \frac{1}{k_{d+1}}\left|k_{d+1}\phi_{\to l} - \tilde{\theta}^{(d+1)}_l\right|_T \\ &= \frac{2\epsilon}{k_d} + \frac{1}{k_{d+1}}\left|\theta^{(d+1)}_{\to l} - \tilde{\theta}^{(d+1)}_l\right|_T \\ &\leq \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}}. \quad (60)\end{aligned}$$

Using previous arguments, all other $\xi_{m,l}$ are either larger or give the same integer $n_{\to j,l}$ and thus $n_l = n_{\to j,l}$. In addition, we can bound, using this equality and Assumption 1a

$$\begin{aligned}&|\phi_{\to l} - (\tilde{\theta}^{(d+1)}_l + 2\pi n_{\to j,l})/k_{d+1}|_T \\ &\leq |\phi_{\to l} - \tilde{\phi}^{(d)}_{\to j}|_T + |\tilde{\phi}^{(d)}_{\to j} - (\tilde{\theta}^{(d+1)}_l + 2\pi n_{\to j,l})/k_{d+1}|_T \\ &\leq \frac{2\epsilon}{k_d} + \frac{2\epsilon(1+\kappa_{d+1})}{k_{d+1}} \\ &< \frac{\pi}{k_{d+1}}, \quad (61)\end{aligned}$$

implying that $n_l = n_{\to j,l} = n^{\text{ideal}}_{\to l,l}$ as desired. □

Algorithm 4.1 has a few failure modes, namely steps 1, 4c and 4e where we exit and return an estimate of lower order. Arguments in Lemma 4.3 show that these failure modes are only encountered when the QEEP subroutine, Alg. A.3, fails at some order. 'Failure' here is not complete failure; regardless of whether the algorithm fails, it will return a set of estimates $\tilde{\phi}_j$, and the error in these estimates will contribute to Eq. 25. To achieve the Heisenberg limit we must make sure that both the probability of failure is small, and that the estimates $\tilde{\phi}_j$ from a failed instance of the algorithm still lie close to the true values ϕ_j to minimize their contribution to Eq. (25). The probability p_d with which the QEEP subroutine succeeds at the dth order is bounded by the parameters α, γ, given as an input to Alg. 4.1 (Eq. 28). The probability that this achieves for up to and including the dth order is bounded by $\prod_{d'=0}^{d} p_{d'}$. It is crucial for the success of our algorithm that we do not encounter any exit modes other than those mentioned above, which we can now prove given the machinery developed in Lem. 4.3.

Corollary 4.3.1. *If each invocation of the QEEP subroutine, Alg. A.3, succeeds in Alg. 4.1, we never exit at step 1, 4c or 4e.*

Proof. Consider step 1 of Alg. 4.1 applying Alg. A.3 which obeys Lemma 3.2, showing that success of Alg. A.3 implies that the number of phases is at most n_ϕ. By assumption there is at least one phase with $A_j > 0$, and hence success means that Alg. A.3 cannot return the empty set due to statement 1. of Lemma 3.2. Hence if Alg. A.3 succeeds we do not exit at step 1 of Alg. 4.1. Now consider step 4c of Alg. 4.1: again success of Alg. A.3 implies that the number of estimates does not exceed n_ϕ. Consider Eq. (33) and Eq. (34); we wish to show that these will not hold if Alg. A.3 succeeds up to order $d-1$. If Alg. A.3 succeeds up to order $d-1$, the phase estimates $\tilde{\phi}^{(d-1)}_j$ obey Eq. (8) for k_d, hence the condition in Eq. (34) equals, for each $\tilde{\theta}^{(d)}_l$

$$\begin{aligned}\min_j \min_n |\tilde{\phi}^{(d-1)}_j - (\tilde{\theta}^{(d)}_l + 2\pi n)/k_d|_T \\ > \frac{2\epsilon(1+\kappa_d)}{k_d}, \quad (62)\end{aligned}$$

and we argued previously, via induction, that this does not happen when Alg. A.3 succeeds up to order d, as $\min_j \xi_{j,l}$ is upper-bounded as in Eq. (60) for all $d' \leq d$. Similarly, Eq. (33) implies the existence of a $\tilde{\phi}^{(d-1)}_j$ with

$$\begin{aligned}\min_l \min_n |\tilde{\phi}^{(d-1)}_j - (\tilde{\theta}^{(d)}_l + 2\pi n)/k_d|_T \\ > \frac{2\epsilon(1+\kappa_d)}{k_d}. \quad (63)\end{aligned}$$

which can not happen due to the success of Alg. A.3 which implies the bound in Eq. (49). Consider lastly step 4e which exits if the current dth order estimates do not lie in the region for which Eq. (7) holds with given k_d. We have argued in Lemma 4.3 that, assuming success of the subroutines implementing Alg. A.3, that Eq. (7) holds for the phase estimates at all orders. □

Now let us consider failures of the QEEP subroutine, Alg. A.3, which do not lead to exiting. Let's imagine that the first failure occurs at some order d_0. Now we want to make sure that continuing with higher orders after such failure still leads to an error of order $\sim \epsilon/k_{d_0-1}$, even though the failure (or any subsequent failure) is not detected.

To show this, we check that if Alg. 4.1 exits at some later round, namely during $d = d_f + 1$ and outputs estimates $\tilde{\phi}_j^{(d_f)}$ that these will be sufficiently close to the estimates right before failure, that is, the set of phases $\tilde{\phi}_j^{(d_0-1)}$.

Then, by Lem. 4.3, these estimates will also be sufficiently close to the true phases ϕ_j.

Lemma 4.4. *Let Alg. 4.1 exit at order $d = d_f+1$ and let the QEEP subroutine, Alg. A.3, of step 4b first fail at $d = d_0 \leq d_f + 1$. For each ϕ_j, there will be an estimate $\tilde{\phi}_l^{(d_f)}$, produced at step 4d in Alg. 4.1 which satisfies*

$$\left|\phi_j - \tilde{\phi}_l^{(d_f)}\right|_T \leq \frac{14\epsilon}{k_{d_0-1}}. \tag{64}$$

Vice-versa, for each estimate $\tilde{\phi}_l^{(d_f)}$ there exists a phase ϕ_j such that

$$\left|\phi_j - \tilde{\phi}_l^{(d_f)}\right|_T \leq \frac{14\epsilon}{k_{d_0-1}}. \tag{65}$$

Proof. Since each QEEP subroutine, Alg. A.3, in Alg. 4.1 succeeds up to order $d_0 - 1$, Lemma 4.3 guarantees that

- (Property 1a) For every phase ϕ_j there exists an estimate $\tilde{\phi}_l^{(d_0-1)}$ such that

$$|\phi_j - \tilde{\phi}_l^{(d_0-1)}|_T \leq \frac{2\epsilon}{k_{d_0-1}}.$$

- (Property 1b) For every estimate $\tilde{\phi}_l^{(d_0-1)}$ there exists a phase ϕ_j such that

$$|\phi_j - \tilde{\phi}_l^{(d_0-1)}|_T \leq \frac{2\epsilon}{k_{d_0-1}}.$$

Then, since the algorithm does not exit at step 4c through Eqs. (33) or step 4e for any order $d = d_0, \dots d_f$, for each estimate $\tilde{\phi}_l^{(d-1)}$ we can associate some $\tilde{\theta}_{m_l}^{(d)}$ that satisfies

$$\frac{1}{k_d}\left|k_d\tilde{\phi}_l^{(d-1)} - \tilde{\theta}_{m_l}^{(d)}\right|_T = \min_{n\in\{0,\dots,\lfloor k_{d+1}\rfloor-1\}}\left|\tilde{\phi}_l^{(d-1)} - (\tilde{\theta}_{m_l}^{(d)} + 2\pi n)/k_d\right|_T \leq \frac{2\epsilon(1+\kappa_d)}{k_d}, \tag{66}$$

where the second equality follows from being allowed to apply Eq. (8) (which is validated by passing the test at step 4e). This implies that in step 4d of Alg. 4.1 at round d, for a given $\tilde{\theta}_{m_l}^{(d)}$, the optimization of ξ_{n,m_l} over n will pick the integer n_{l,m_l}, i.e. the integer associated with matching $\tilde{\theta}_{m_l}^{(d)}$ with $\tilde{\phi}_l^{(d-1)}$. Next, similar as in the proof of Lemma 4.3, we can consider the possibility of matching to other estimates $\tilde{\phi}_k^{(d-1)} \neq \tilde{\phi}_l^{(d-1)}$. Since κ_d is chosen in step 4f of Alg. 4.1, we claim that either Eq. (36) holds, in which case

$$\min_{n\in\{0,\dots,\lfloor k_d\rfloor-1\}}\left|\tilde{\phi}_k^{(d-1)} - \frac{2\pi n + \tilde{\theta}_{m_l}^{(d)}}{k_d}\right|_T > \frac{2\epsilon}{k_d}(1+\kappa_d), \tag{67}$$

hence this ξ_{k,m_l} is not optimal, or that Eq. (37) holds, in which case

$$n_{l,m_l} = \operatorname*{arg\,min}_{n\in\{0,\dots,\lfloor k_d\rfloor-1\}}\left|\tilde{\phi}_k^{(d-1)} - \frac{2\pi n + \tilde{\theta}_{m_l}^{(d)}}{k_d}\right|_T = n_{k,m_l}. \tag{68}$$

The proofs of these claims are exactly the same as in the proof of Lemma 4.3, i.e. using Eqs. (54), (56),(57).

Now let us prove Eq. (64). Given a phase ϕ_j, we can use Property (1a) to find an associated estimate $\tilde{\phi}_{\to j}^{(d_0-1)}$ within $2\epsilon/k_{d_0-1}$. Then for this estimate let $\theta_{m_{\to j}}^{(d_0)}$ be the matched estimate in the next round for which Eq. (66) holds, so that the round produces a new estimate $\tilde{\phi}_{\to j}^{(d_0)} = \frac{\tilde{\theta}_{m_{\to j}}^{(d_0)} + 2\pi n_{\to j,m_{\to j}}}{k_{d_0}}$ (which we label with $\to j$ again) for which

$$|\tilde{\phi}_{\to j}^{(d_0-1)} - \tilde{\phi}_{\to j}^{(d_0)}|_T \leq \frac{2\epsilon}{k_{d_0}}(1+\kappa_{d_0}). \tag{69}$$

Then again for $\tilde{\phi}_{\to j}^{(d_0)}$ there exists some matching $\theta_{m\to j}^{(d_0+1)}$ etc. and this generates a series of estimates $\tilde{\phi}_{\to j}^{(d)}$ up to order d_f. For a given ϕ_j we can then bound, using this series of estimates and $\kappa_d \geq 2$ for all d,

$$\begin{aligned}\left|\tilde{\phi}_j^{(d_f)} - \phi_j\right|_T &\leq \left|\tilde{\phi}_{\to j}^{(d_0-1)} - \phi_j\right|_T \\ &\quad + \sum_{d=d_0}^{d_f} \left|\tilde{\phi}_{\to j}^{(d-1)} - \tilde{\phi}_{\to j}^{(d)}\right|_T \\ &\leq \frac{2\epsilon}{k_{d_0-1}} + \sum_{d=d_0}^{d_f} \frac{2\epsilon(1+\kappa_d)}{k_d} \\ &= \frac{2\epsilon}{k_{d_0-1}}\left(1 + \sum_{d=d_0}^{d_f} \frac{1+\kappa_d}{\kappa_{d_0}\kappa_{d_0+1}\ldots\kappa_d}\right) \\ &\leq \frac{2\epsilon}{k_{d_0-1}}\left(1 + \sum_{n=0}^{\infty} \frac{3}{2^n}\right) \\ &= \frac{14\epsilon}{k_{d_0-1}}. \end{aligned} \tag{70}$$

Now let's prove Eq. (65) and start with an estimate $\tilde{\phi}_l^{(d_f)}$ which was obtained from some $\tilde{\theta}_l^{(d_f)}$ matched with a previous estimate $\tilde{\phi}_l^{(d_f-1)}$ (just for convenience we again use the same label) such that $|\tilde{\phi}_l^{(d_f)} - \tilde{\phi}_l^{(d_f-1)}|_T \leq \frac{2\epsilon}{k_{d_f}}(1+\kappa_{d_f})$, using that we do not exit through Eq. (34). Then again for this previous estimate $\tilde{\phi}_l^{(d_f-1)}$ we can repeat the argument and create a sequence of estimates up to $\tilde{\phi}_l^{(d_0-1)}$. For the last estimate, we invoke Property (1b), namely that there is a nearby ϕ_j. Then we can upperbound for this ϕ_j: $|\tilde{\phi}_l^{(d_f)} - \phi_j|_T \leq |\tilde{\phi}_l^{(d_0-1)} - \phi_j|_T + \sum_{d=d_0}^{d_f} |\tilde{\phi}_l^{(d)} - \tilde{\phi}_l^{(d-1)}|_T$ etc., exactly as in Eq. (70), leading to Eq. (65). □

4.3 Algorithm 4.1 achieves the Heisenberg limit

We have seen that the success of the QEEP subroutines in Alg. 4.1 leads to an error scaling as $\epsilon/k_{d_f} \sim \delta_c$. Now we must choose the success probability p_d of these subroutines in Eq. (28), depending on α, γ so that the total mean-squared-error is bounded by some $\delta^2 = O(\delta_c^2)$ while the quantum cost $T = O(\delta^{-1})$. We note that the next theorem contains no logarithmic factors in δ^{-1}, as in [26], but achieves pure Heisenberg scaling.

Theorem 4.5. *Algorithm 4.1 solves the multiple eigenvalue estimation problem in Def. 3.1 with accuracy error δ and total quantum cost $T = O(\delta^{-1})$, given A, n_ϕ and a fixed ϵ_0 and ϵ obeying Eqs. (27) and (26), and some choice for the constants $\alpha > 0$ and $\gamma > 2$.*

Remarks: Note that the dependence on the number of phases n_ϕ is not made explicit in the statement of this Theorem, but this dependence will be polynomial in n_ϕ, not necessarily a very low-order polynomial. This dependence comes through the choice for ϵ_0 (and ϵ) via Eq. (26) (resp. Eq. (27)) which sets the error and thus the running time of the QEEP Algorithm A.2.

Proof. Our proof is motivated by the analysis in [2] for a single phase ϕ leading to Theorem 2.6. The idea is to bound the mean-squared-error in the final estimation of ϕ by summing over error contributions at each order d at which the phase extraction subroutine may fail (with probability $1 - p_d$).

In our case the multipliers κ_d (and k_d) at each order are not fixed (as in Theorem 2.6) but depend on phase estimates at previous orders and thus measurement data at previous orders. Our confidence parameter p_d in Eq. (28), which determines the number of repeats of experiments, and hence the cost, in Alg. A.3, depends on k_d and is thus a random variable depending on previous measurement data. All measurement data are denoted by $\mathbf{x}$ and thus we have random variables $k_d(\{\kappa_{d'}(\mathbf{x})\}_{d'=1}^d)$ and $p_d(\{\kappa_{d'}(\mathbf{x})\}_{d'=1}^d)$.

Consider the mean-squared-error δ_j^2 for the jth phase ϕ_j in Eq. (25) in Definition 3.1. We have three error contributions to consider given a choice for the random variable k_d.

1. With probability $1 - p_0$ the subroutine Alg. A.3 in Alg. 4.1 fails at step 1 ($d = 0$). In this case, as we always return *some* estimate, δ_j is bounded for all j by π.

2. With probability at most $(1 - p_{d_0})\prod_{d=0}^{d_0-1} p_d \leq 1 - p_{d_0} = e^{-\alpha}\left(\frac{k_{d_0}\delta_c}{\pi}\right)^\gamma$, the subroutine Alg. A.3 in Alg. 4.1 fails for the first time at some order $1 \leq d_0 \leq d_f$, and the algorithm proceeds in any way afterwards (by possibly exiting or not). In this case, Lemma 4.4 bounds δ_j for all j by $\frac{14\epsilon}{k_{d_0-1}}$ or Lemma 4.3 bounds δ_j for all j by $\frac{2\epsilon}{k_{d_0-1}} \leq \frac{14\epsilon}{k_{d_0-1}}$.

3. With probability less than $\prod_{d=0}^{d_f} p_d < 1$ the subroutine Alg. A.3 in Alg. 4.1 succeeds up to the final round d_f, and Lemma 4.3 implies that $\delta_j \leq 2\epsilon/k_{d_f} \leq 2\delta_c$ for all j as $k_{d_f} \geq \epsilon/\delta_c$.

We can now bound the mean-squared-error as a sum over the above three contributions weighted by their relevant overlaps:

$$\begin{aligned}
\delta_j^2 &\leq (1-p_0)\pi^2 + \sum_{\mathbf{x}} \mathbb{P}(\mathbf{x}) \left[\sum_{d_0=1}^{d_f-1} \mathbb{P}(\kappa_1, \ldots, \kappa_{d_0}|\mathbf{x})(1 - p_{d_0}(\{\kappa'_d\}_{d'=1}^{d_0})) \left[\frac{14\epsilon}{k_{d_0-1}(\{\kappa'_d\}_{d'=1}^{d_0-1})} \right]^2 \right] + 4\delta_c^2 \\
&= \pi^2 e^{-\alpha} \left(\frac{\delta_c}{\pi} \right)^\gamma + \sum_{\mathbf{x}} \mathbb{P}(\mathbf{x}) \left[\sum_{d_0=1}^{d_f-1} \mathbb{P}(\kappa_1, \ldots, \kappa_{d_0}|\mathbf{x}) e^{-\alpha} \left(\frac{k_{d_0}\delta_c}{\pi} \right)^\gamma \frac{196\epsilon^2}{k_{d_0-1}^2} \right] + 4\delta_c^2 \\
&\leq \pi e^{-\alpha} \left(\frac{\delta_c}{\pi} \right)^\gamma + 0.15 \times \delta_c^2 \sum_{\mathbf{x}} \mathbb{P}(\mathbf{x}) \mathbb{P}(\kappa_1, \ldots, \kappa_{d_0}|\mathbf{x}) \sum_{d_0=1}^{d_f-1} \left(\frac{k_{d_0}\delta_c}{\pi} \right)^{\gamma-2} + 4\delta_c^2. \qquad (71)
\end{aligned}$$

Here we have removed the dependency of k_{d_0} and k_{d_0-1} on the previous multipliers for notational simplicity. For $d = 1$ we have $e^{-\alpha}\kappa_{d_0}^2 196(\epsilon/\pi)^2 = e^{-\alpha}k_1^2 196(\epsilon_0/\pi)^2 \leq \frac{196\times16\times4}{(300)^2} \leq 0.15$ due to Eq. (26). For $d > 1$, $e^{-\alpha}\kappa_{d_0}^2 196(\epsilon/\pi)^2 \leq \frac{9\times196\times4}{(300)^2} = 0.08$, due to Eq. (27).

To evaluate the middle term, we write $k_{d_0} = k_{d_f}\frac{k_{d_0}}{k_{d_f}}$, and note that as $k_d = \prod_{d'=1}^{d} \kappa_{d'}$, we have $\frac{k_{d_0}}{k_{d_f}} \leq 2^{d_0-d_f}$ as the multiplier $\kappa_d \geq 2$. As $k_{d_f} \leq \frac{2\epsilon}{\delta_c} < \frac{\pi}{2\delta_c}$, we have

$$\begin{aligned}
\sum_{d_0=1}^{d_f-1} \left(\frac{k_{d_0}\delta_c}{\pi} \right)^{\gamma-2} &\leq \frac{1}{2^{\gamma-2}} \sum_{d_0=1}^{d_f-1} (2^{\gamma-2})^{(d_0-d_f)} \\
&\leq \frac{2^{4-\gamma}}{2^\gamma - 4}, \qquad (72)
\end{aligned}$$

where the last inequality holds since $\gamma > 2$. By letting the upper bound be independent of the κ_ds, we can remove the dependence on $\mathbf{x}$ in Eq. (71), using that $\sum_{\mathbf{x}} \mathbb{P}(\mathbf{x})\mathbb{P}(\kappa_1, \ldots, \kappa_{d_0}|\mathbf{x}) = 1$. This yields a final bound on δ_j of

$$\delta_j^2 \leq \delta_c^2 \left[\pi^{1-\gamma} e^{-\alpha} \delta_c^{\gamma-2} + \frac{0.15 \times 2^{4-\gamma}}{2^\gamma - 4} + 4 \right]. \qquad (73)$$

As $\gamma > 2$, this scales as δ_c^2 as $\delta_c \to 0$.

Let us now calculate the cost of executing Alg. 4.1 in terms of the number of unitary applications. Again this depends on the choice of multiplier κ_d at each step. Let us fix a sequences of k_ds, and let d_f be the final round of estimation in this algorithm, i.e. the final round for which we invoked the quantum subroutine in Alg. A.3. At each order d we use $V^{\mathsf{k}} = U^{k_d \mathsf{k}}$, where $\mathsf{k} = 0, 1, \ldots, K$, with $2M_d$ samples where K is a function of ϵ as in Theorem A.2. The cost of each experiment is $k_d\mathsf{k}$.

We can calculate

$$\begin{aligned}
T &= \sum_{d=0}^{d_f} \sum_{\mathsf{k}=1}^{K} (2M_d k_d \mathsf{k}) \\
&= \sum_{d=0}^{d_f} M_d k_d K(K+1). \qquad (74)
\end{aligned}$$

The QEEP algorithm in Theorem A.2 requires $M_d = \tilde{O}(|\ln(1-p_d)|\epsilon^{-4})$ with p_d in Eq. (28) and $K = \tilde{O}(\epsilon^{-1})$.

We may bound

$$\begin{aligned}
T &\leq \tilde{O}(\epsilon^{-6}) \sum_{d=0}^{d_f} k_d \left| -\alpha + \gamma \ln\left(\frac{k_d \delta_c}{\pi} \right) \right| \\
&= \tilde{O}(\epsilon^{-6}) \sum_{d=0}^{d_f} k_d \left[\alpha - \gamma \ln\left(\frac{k_d \delta_c}{\pi} \right) \right]. \qquad (75)
\end{aligned}$$

We again bound $k_d = \frac{k_d}{k_{d_f}} k_{d_f} \leq 2^{d-d_f} \frac{\pi}{2\delta_c}$, which yields

$$\begin{aligned}
T &\leq \delta_c^{-1} \tilde{O}(\epsilon^{-6}) \left[\frac{\pi}{2} \right. \\
&\quad \left. \times \sum_{d=0}^{d_f} 2^{d-d_f} \left[\alpha - \gamma(d - d_f - 1)\ln(2) \right] \right] \\
&\leq \delta_c^{-1} \tilde{O}(\epsilon^{-6})(\alpha + 2\gamma \ln(2)). \qquad (76)
\end{aligned}$$

Combining our bounds then yields

$$\delta \leq \delta_c \left[e^{-\alpha} \pi^{1-\gamma} + 4 + \frac{0.15 \times 2^{4-\gamma}}{2^\gamma - 4} \right]^{\frac{1}{2}} \tag{77}$$

$$\leq T^{-1} \left[e^{-\alpha} \pi^{1-\gamma} + 4 + \frac{0.15 \times 2^{4-\gamma}}{2^\gamma - 4} \right]^{\frac{1}{2}} \times \tilde{O}(\epsilon^{-6})[\alpha + 2\gamma \ln(2)]$$

$$= O(T^{-1}), \tag{78}$$

which is the Heisenberg limit. Note that a dependence on the number of phase n_ϕ enters the scaling of δ via ϵ which needs to be bounded as in Eqs. (26) and (27). □

5 Numerical implementation

Thm. 4.5 requires using the QEEP algorithm (Theorem A.2 and Algorithm A.3) in order to obtain provable bounds. Instead of analytic bounds, we now turn to a numerical demonstration, giving the opportunity to implement and test Algorithm 4.1 with a few modifications. We test the algorithm using two different sub-routines, one based on the matrix pencil method [31], and one based on the QEEP time-series analysis of Theorem A.2, as described in Algorithm A.3. Code to implement all simulations can be found at https://github.com/alicjadut/qpe.

To improve the practical performance of Alg. 4.1, we make the following two small changes. Firstly, instead of choosing κ_{d+1} in step 4f in the ranges declared in Lem. 4.2, we choose the largest κ_{d+1} consistent with Eq. (36) and Eq. (37) for all $\tilde{\phi}_j^{(d)}, \tilde{\phi}_l^{(d)}$. We note that the maximum such κ_{d+1} is bounded above by $\frac{\pi}{2\epsilon} - 1$, as the left-hand side of Eq. (36) is bounded above by 2π and the left-hand side of Eq. (37) is bounded below by 0. (In practice, tighter bounds can be found by checking the boundaries of the regions $R_{jl}^{(n)}$ defined in Eq. (92), and we find the largest possible κ_{d+1} by iterating backwards through these boundaries till a gap is found.) Secondly, as the bounds for ϵ and ϵ_0 in Lem. 4.2 are rather loose, and our performance scales rather badly in both, we choose the largest $\epsilon = \epsilon_0$ that allows all simulations to find a value of $\kappa_{d+1} > 2$ at each order.

When using the matrix pencil processing subroutine, we follow the implementation described in Ref. [19]:

Algorithm 5.1. *The matrix pencil method takes as input estimates of the phase function $g(\mathsf{k}) = \sum_j A_j e^{i\mathsf{k}\theta_j}$ for a unitary V at points $\mathsf{k} = 0, 1, \ldots K$ and an overlap bound A, and proceeds as follows:*

1. *Construct the $L_K \times (2K - L_K + 1)$ Hankel matrices $G^{(0)}$, $G^{(1)}$, where $G^{(a)}_{i,j} = g(i + j + a - K)$ for $i \in \{0, 1, \ldots L_K - 1\}, j \in \{0, 1, \ldots, 2K - L_K\}$, $a = 0, 1$, with $L_K = \lfloor (K+1)/2 \rfloor$, and using $g(-\mathsf{k}) = g^*(\mathsf{k})$.*

2. *Construct the $L_K \times L_K$ shift matrix T by least-squares minimization of the matrix 2-norm $\|TG^{(0)} - G^{(1)}\|$.*

3. *Calculate the eigenvalues of T, $\lambda_j = |\lambda_j| e^{i\tilde{\theta}_j}$ and from there the phase estimates $\tilde{\theta}_j$.*

4. *Calculate the overlap estimates $\tilde{A}_j$ by least-squares minimization of the vector 2-norm $\|BA - g\|$, where B is the $(K+1) \times L_K$ matrix*

$$B_{k,j} = \lambda_j^{\mathsf{k}}, \tag{79}$$

and $g = [g(0), \ldots g(K)]^T$.

5. *Return the phase estimates $\tilde{\theta}_j$ for which the corresponding overlap estimate $\tilde{A}_j \geq A$.*

To use this algorithm as a subroutine in Alg. 4.1 (in place of Alg. A.3), we implement it on the matrix $V = U^{k_d}$, which requires implementing $V^{\mathsf{k}} = U^{\mathsf{k}k_d}$ for a range of integer k on a quantum device.

To isolate the performance of the estimation routine from the generation of the signal itself, we do not test our protocols on data generated from simulating or approximating a particular unitary. Instead, we test the ability of the protocols to estimate $n_\phi = 2$ and $n_\phi = 4$ randomly-chosen phases $\phi_j \in [0, 2\pi]$ when sampling from the true phase function $g(k)$. We take all phases with equal weight — $A_j = 1/n_\phi$. We simulate the sampling from $g(k)$ in Algorithm 5.1 or Algorithm A.2 for some $V = U^{k_d}$ by simulating the readout of a control qubit with the reduced density matrix of Eq. (4). (In practice this would be generated by the quantum circuit in Fig. 1.) We first draw M_d i.i.d. samples from the two

Bernoulli distributions

$$\mathbb{P}_k^r(+1) = \frac{1}{2}\sum_{j=1}^{n_\phi} A_j(1+\cos(\theta_j k)), \tag{80}$$

$$\mathbb{P}_k^i(+1) = \frac{1}{2}\sum_{j=1}^{n_\phi} A_j(1-\sin(\theta_j k)), \tag{81}$$

where $\theta_j = k_d\phi_j \mod 2\pi$ are the eigenvalues of V. Then, we return the fraction of +1s drawn as estimates for the real and imaginary parts of $g(k)$ respectively. This is then repeated at all points $k = k_d\mathsf{k}$ for $\mathsf{k} = 0, 1, \ldots, K$. Following the discussion in Sec. 2.2 and using the notation from Eq. (74), we sum the total quantum cost for the algorithm over all requested $g(k)$ queries; $T = 2\sum_d \sum_{\mathsf{k}=1}^{K} \mathsf{k} k_d M_d$. (We ignore the subleading correction from the final term in Eq. (74) as this will not affect the scaling of our result.) For the signal length K and number of points M_d to sample each $g(k)$ at, we follow the bounds given in Ref. [3] (both when using the QEEP and matrix pencil subroutines):

- signal length: $K = \lceil 0.1 L \ln^2 L \rceil$, with $L = \lceil \frac{2\pi}{\epsilon} \rceil$ the number of bins used in the QEEP subroutine (Def. A.1).
- number of measurements of each circuit: $M_d = \lceil |\ln(1-p_d)|\, \epsilon^{-4} \rceil$.

Here, p_d is given in Eq. (28) for a given k_d. This equation requires fixing a choice of α and γ — across all experiments we take $\alpha = 2$ and $\gamma = 2.1$.

To demonstrate that our methods achieve the Heisenberg limit, in Fig. 3 we plot the error as a function of the total quantum cost for a set of simulations using the methods described above. Each simulation draws a different set of n_ϕ random phases, and a final error $\delta_c \in [10^{-5}, 10^{-2}]$ (for each choice of δ_c we use the same 50 sets of phases). We plot the error for each phase estimate separately in Fig. 3 (i.e. each simulation corresponds to n_ϕ points in the plot). As both the total quantum cost and the error is different between simulations, we bin all experiments within a range of T values, and calculate the root mean square error and root mean square total quantum cost. This gives a good approximation to the accuracy error defined in Def. 3.1 for the restricted data set used.

For the QEEP subroutine, we observe a clear fit of the data (blue points) to a $\delta \sim T^{-1}$ trend, as expected from Thm. 4.5, but with a rather large constant factor; we find $T \sim 10^{10}\delta^{-1}$ for estimating 2 phases and $T \sim 10^{15}\delta^{-1}$ for estimating 4. Further optimization of the QEEP algorithm for these purposes may yet improve this constant factor. However, as the methods of Ref. [3] were not designed for estimating individual phases, it may be expected that this method performs somewhat badly for this purpose, so we have not pursued this further.

Simulations using the matrix pencil subroutine outperform simulations using the QEEP subroutine by a factor of $10^4 - 10^6$, and clearly demonstrate Heisenberg-limited scaling $\delta \sim T^{-1}$ as well. We take this result instead of an analytic proof as strong numerical evidence for Heisenberg-limited scaling when a version of Alg. 4.1 is constructed using the matrix pencil method as a subroutine. We notice that the error in two phase estimates in the first bin of the matrix pencil method for $n_\phi = 2$ is significantly above the remainder of the population (by about a factor 100×), which blows up our error bars for this bin. Further investigation shows that the two phases in question are from the same simulation, and separated by only 1.5×10^{-4}. By contrast, for the simulation in question (at $d = 1$) our algorithm sets $k_1 K \sim 3 \times 10^3 < (1.5 \times 10^{-4})^{-1}$ (where $k_1 K$ is the largest value of the phase function $g(k)$ sampled during this simulation). This implies that our signal lies within the region where improving our estimation accuracy by increasing the number of shots M_d is exponentially hard [32]. In latter simulations with these two phases we see that our estimation error regresses to similar results as all other estimates.

6 Conclusion

In this work we studied Heisenberg-limited quantum phase estimation using a single control qubit. In this form of phase estimation, we rely on classical signal processing to extract eigenvalue data from the phase function $g(k)$ in Eq. (2).

It has been an open question whether these methods can achieve the Heisenberg limit in the case of multiple phases: Ref. [26] answered this question up to log factors with a Heisenberg-limited Monte Carlo algorithm, providing a sampling of the spectral function $A(\phi)$ in Eq. (3) from which to estimate the phases. In this work we

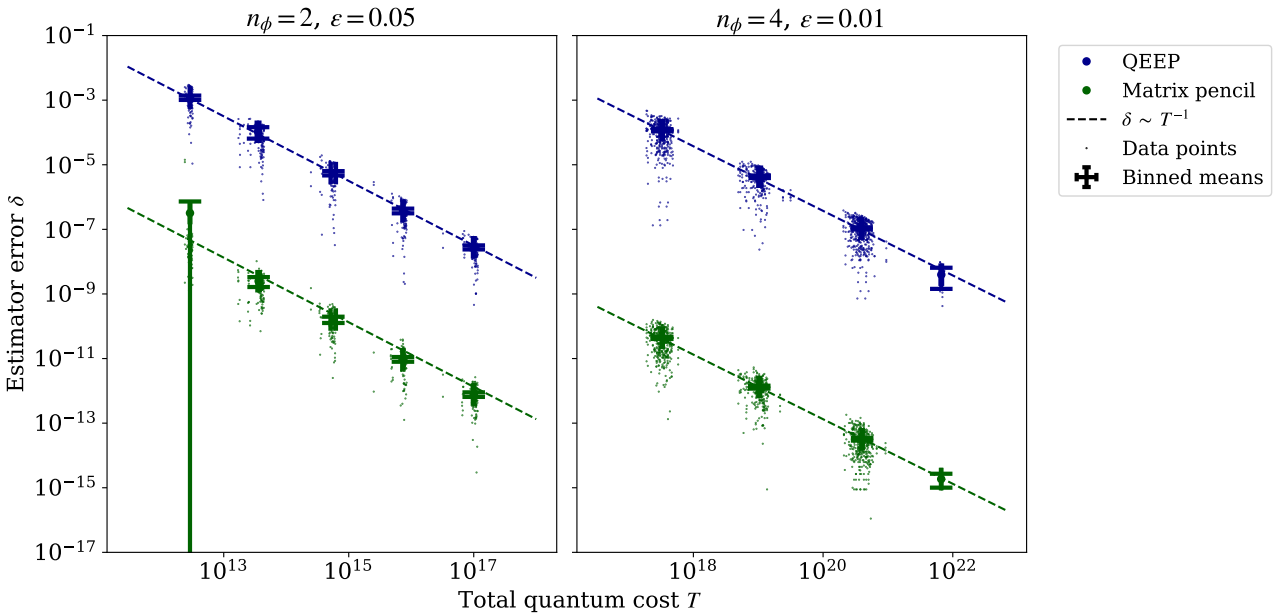

Figure 3: Convergence of the algorithm with total quantum cost T. Phase estimates were obtained with either QEEP (blue) or matrix pencil (green) subroutines with parameters described in the text. Individual points show the error $\delta = |\tilde{\phi}_j - \phi_j|_T$ on individual phases in each simulation. This data is binned in the x-axis, and for each bin a root-mean-square error and standard deviation (error bars) are plotted in the x- and y-direction. Dotted lines show a fit of these means to $\delta \sim T^{-1}$.

also answered this question in the affirmative exactly with a new adaptive multi-order phase estimation algorithm, for which we prove Heisenberg scaling if the algorithm uses a QEEP phase extraction subroutine. We numerically show the performance of this algorithm, also when instead of using a QEEP subroutine, one uses the matrix pencil method to extract phase estimates from the phase function $g(k)$. This result complements the previous work discussed in the introduction by closing the question of whether there exists a gap between classical post-processing of phase function data and fully quantum phase estimation.

In obtaining our results we encountered at least two details of quantum phase estimation that we have not seen discussed in the literature. The first is the dense signal limit, Eq. (12) in Thm. 2.5: sampling $g(k)$ at *all* integer point $k = 1, \ldots, K$ is sub-optimal regardless of what method is used to process the data. However, we also briefly argued that by picking points among $k = 1, \ldots, K$ at random one may go beyond this, and one could interpret this as allowing the results obtained in [26] in which such randomized choices for k are taken.

The second is the need for adaptive choices of k_d to solve the phase matching problem. It is unclear to us how far this problem extends; although Lemma 4.2 provides a practical solution, others may still exist. Another open question with respect to Algorithm 4.1 is whether one can remove the need to choose real-valued multipliers κ_d and restrict to integer choices. Restricting $\kappa_d \in \mathbb{N}$ would significantly simplify some technical issues, i.e. the applicability of Lemma 2.4 and the need for shifting phases in step 2 of Alg. 4.1, but we don't know how to prove a version of Lemma 4.2 for $\kappa_d \in \mathbb{N}$. In fact, we don't know whether there is a fundamental difference in performance between only using data obtained with integer k in $g(k)$ versus data obtain with real-valued, –in practice rational–, k in $g(k)$.

We have assumed in our problem description, Def. 3.1, that the spectrum in the input state is discrete consisting of n_ϕ phases with amplitudes above some cut-off. In practice this condition may not be fulfilled and thus studying the performance of the algorithms on more typical spectra induced by many-body Hamiltonians and easy-to-prepare input states will be of interest. The scaling of our (provable) Heisenberg-limited algorithm in n_ϕ is also rather poor [$O(n_\phi^{24})$ as described], as we have not attempted to optimize

this aspect. This should in principle be immediately reducible to $O(n_\phi^3 \delta^{-1} + n_\phi^6)$ under the assumption that the matrix pencil method continues to achieves the dense sampling limit when estimating multiple eigenvalue, however we do not know a proof of this. In principle linear scaling with n_ϕ should be achievable (or even sub-linear if the methods of [33] could be applied in this setting); optimizing this is a clear target for future work.

A direction for future research is to make this algorithm efficient in practice (i.e. improve the parameter dependence and the practical run time) or devise yet-different Heisenberg-scaling algorithms and examine their performance in the presence of experimentally-noisy signals $g(k)$.

Acknowledgements

AD acknowledges financial support in the form of a scholarship from the Den Adel Fund.

References

[1] 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(7):073023, 2009. DOI: 10.1088/1367-2630/11/7/073023. URL https://arxiv.org/abs/0809.3308.

[2] 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. DOI: 10.1103/PhysRevA.92.062315. URL https://arxiv.org/abs/1502.02677.

[3] Rolando D. Somma. Quantum eigenvalue estimation via time series analysis. *New J. Phys.*, 21:123025, 2019. DOI: 10.1088/1367-2630/ab5c60. URL https://iopscience.iop.org/article/10.1088/1367-2630/ab5c60/pdf.

[4] Pawel Wocjan and Shengyu Zhang. Several natural BQP-complete problems. *ArXiv:quant-ph/0606179*, 2006. DOI: 10.48550/arXiv.quant-ph/0606179. URL https://arxiv.org/abs/quant-ph/0606179.

[5] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Sci. Stat. Comp.*, 26:1484, 1997. DOI: 10.1137/S0097539795293172. URL https://arxiv.org/abs/quant-ph/9508027.

[6] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for solving linear systems of equations. *Phys. Rev. Lett.*, 15(103):150502, 2009. DOI: 10.1103/PhysRevLett.103.150502. URL https://arxiv.org/abs/0811.3171.

[7] James D. Whitfield, Jacob Biamonte, and Alán Aspuru-Guzik. Simulation of electronic structure Hamiltonians using quantum computers. *Mol. Phys.*, 109:735–750, 2011. DOI: 10.1080/00268976.2011.552441. URL https://arxiv.org/abs/1001.3855.

[8] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000. ISBN 9780521635035. DOI: 10.1017/CBO9780511976667. URL https://books.google.de/books?id=65FqEKQOfP8C.

[9] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354, 1998. DOI: 10.1098/rspa.1998.0164. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1998.0164.

[10] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum metrology. *Physical review letters*, 96(1): 010401, 2006. DOI: 10.1103/PhysRevLett.96.010401. URL https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.96.010401.

[11] Wim van Dam, G. Mauro D'Ariano, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Optimal quantum circuits for general phase estimation. *Phys. Rev. Lett.*, 98:090501, Mar 2007. DOI: 10.1103/PhysRevLett.98.090501. URL https://link.aps.org/doi/10.1103/PhysRevLett.98.090501.

[12] Dominic W Berry, Brendon L Higgins, Stephen D Bartlett, Morgan W Mitchell,

Geoff J Pryde, and Howard M Wiseman. How to perform the most accurate possible phase measurements. *Physical Review A*, 80(5):052114, 2009. DOI: 10.1103/PhysRevA.80.052114.

[13] Robert B. Griffiths and Chi-Sheng Niu. Semiclassical Fourier transform for quantum computation. *Physical Review Letters*, 76 (17):3228–3231, Apr 1996. ISSN 1079-7114. DOI: 10.1103/physrevlett.76.3228. URL 10.1103/PhysRevLett.76.3228.

[14] A. Yu. Kitaev. Quantum measurements and the Abelian stabilizer problem. *ArXiv:quant-ph/9511026*, 1995. DOI: 10.48550/arXiv.quant-ph/9511026. URL https://arxiv.org/abs/quant-ph/9511026.

[15] Dominic W. Berry, Graeme Ahokas, Richard Cleve, and Barry C. Sanders. Efficient quantum algorithms for simulating sparse Hamiltonians. *Comm. Math. Phys.*, 270 (359), 2007. DOI: 10.1007/s00220-006-0150-x. URL https://arxiv.org/abs/quant-ph/0508139.

[16] Nathan Wiebe and Chris Granade. Efficient Bayesian phase estimation. *Phys. Rev. Lett.*, 117:010503, 2016. DOI: 10.1103/PhysRevLett.117.010503. URL https://arxiv.org/abs/1508.00869.

[17] Krysta M. Svore, Matthew B. Hastings, and Michael Freedman. Faster phase estimation. *Quant. Inf. Comp.*, 14(3-4):306–328, 2013. DOI: 10.48550/arXiv.1304.0741. URL https://arxiv.org/abs/1304.0741.

[18] Ewout van den Berg. Efficient Bayesian phase estimation using mixed priors. *ArXiv:2007.11629*, 2020. DOI: 10.22331/q-2021-06-07-469. URL https://arxiv.org/abs/2007.11629.

[19] 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. DOI: 10.1088/1367-2630/aafb8e. URL https://iopscience.iop.org/article/10.1088/1367-2630/aafb8e.

[20] David C. Rife and Robert R. Boorstyn. Single-tone parameter estimation from discrete-time observations. *IEEE Trans. Inf. Th.*, 20(5):591–598, 1974. DOI: 10.1109/TIT.1974.1055282. URL https://ieeexplore.ieee.org/document/1055282.

[21] Sirui Lu, Mari Carmen Bañuls, and J. Ignacio Cirac. Algorithms for quantum simulation at finite energies. *PRX Quantum*, 2:020321, 2020. DOI: 10.1103/PRXQuantum.2.020321. URL https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.2.020321.

[22] T.E. O'Brien, S. Polla, N.C. Rubin, W.J. Huggins, S. McArdle, S. Boixo, J.R. McClean, and R. Babbush. Error mitigation via verified phase estimation. *ArXiv:2010.02538*, 2020. DOI: 10.1103/PRXQuantum.2.020317. URL https://arxiv.org/abs/2010.02538.

[23] Alessandro Roggero. Spectral density estimation with the Gaussian integral transform. *ArXiv:2004.04889*, 2020. DOI: 10.1103/PhysRevA.102.022409. URL https://arxiv.org/abs/2004.04889.

[24] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 193–204, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367059. DOI: 10.1145/3313276.3316366. URL 10.1145/3313276.3316366.

[25] O. Regev. A subexponential time algorithm for the dihedral hidden subgroup problem with polynomial space. *ArXiv:quant-ph/0406151*, 2004. DOI: 10.48550/arXiv.quant-ph/0406151. URL https://arxiv.org/abs/quant-ph/0406151.

[26] Lin Lin and Yu Tong. Heisenberg-limited ground state energy estimation for early fault-tolerant quantum computers. *ArXiv:2102.11340*, 2021. DOI: 10.1103/PRXQuantum.3.010318. URL https://arxiv.org/abs/2102.11340.

[27] Valentin Gebhart, Augusto Smerzi, and Luca Pezzè. Heisenberg-limited bayesian multiphase estimation algorithm. *ArXiv:2010.09075*, 2020. DOI:

10.1103/PhysRevApplied.16.014035. URL https://arxiv.org/abs/2010.09075.

[28] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. Theory of trotter error with commutator scaling. *Phys. Rev. X*, 11:011020, Feb 2021. DOI: 10.1103/PhysRevX.11.011020. URL https://link.aps.org/doi/10.1103/PhysRevX.11.011020.

[29] Harald Cramér. *Mathematical Methods of Statistics*. Princeton University Press, 1946. ISBN 0691080046. DOI: 10.1515/9781400883868. URL https://archive.org/details/in.ernet.dli.2015.223699.

[30] Calyampudi Radakrishna Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, 37:81–89, 1945. DOI: 10.1007/978-1-4612-0919-5_16. URL https://link.springer.com/chapter/10.1007/978-1-4612-0919-5_16.

[31] Yingbo Hua and Tapan Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustic Speech and Signal Processing*, 38(5), 1990. DOI: 10.1109/29.56027. URL https://ieeexplore.ieee.org/document/56027.

[32] Ankur Moitra. Super-resolution, extremal functions and the condition number of Vandermonde matrices. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '15, page 821–830, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450335362. DOI: 10.1145/2746539.2746561. URL 10.1145/2746539.2746561.

[33] Lin Lin and Yu Tong. Near-optimal ground state preparation. *Quantum*, 4: 372, December 2020. ISSN 2521-327X. DOI: 10.22331/q-2020-12-14-372. URL 10.22331/q-2020-12-14-372.

A Phase extraction subroutine

For our Heisenberg-limited algorithm, at each order d we need to extract eigenphases of $V = U^{k_d}$ from the signal generated by PFE. We require that the phase estimates satisfy the promises of Lem. 3.2. This is achieved by Alg. A.3, which is based on the solution to QEEP of Ref. [3].

In this appendix, we first summarize the results of Ref. [3] by giving a precise definition of the QEEP (Def. A.1) and performance guarantees of the time-series analysis (Def. A.2). Then we describe the Conservative QEEP Eigenvalue Extraction algorithm (Alg. A.3) and give the proof of Lem. 3.2.

A.1 The Quantum Eigenvalue Estimation Problem

Definition A.1 (Quantum Eigenvalue Estimation Problem, QEEP)**.** *Let $A(\phi)$ be the spectral function defined in Eq. (3) for a unitary U and $|\Psi\rangle$. Given is an error parameter $\epsilon > 0$, a confidence parameter $1 \geq p > 0$, and a set of non-negative (approximate indicator) functions $f^l(\phi)$ for $\phi \in [0, 2\pi)$ for $l = 0, \ldots, L-1$, $L = \lceil \frac{2\pi}{\epsilon} \rceil$, where $f^l(.)$ has support on only the interval bin*

$$\mathcal{V}_l = [(l-1)\epsilon, (l+1)\epsilon]_T, \tag{82}$$

and $f^l(\phi) + f^{(l-1)}(\phi) = 1$ for all $\phi \in \mathcal{V}_l \cap \mathcal{V}_{l-1}$. Assuming access to the PFE *subroutine, Def. 2.2, the goal is to output an approximation $\tilde{b}_l$ for $l = 0, \ldots, L-1$ to the integral*

$$b_l = \int_0^{2\pi} d\phi A(\phi) f^l(\phi), \tag{83}$$

which satisfies

$$\sum_{l=0}^{L-1} |\tilde{b}_l - b_l| \leq \epsilon, \tag{84}$$

with probability at least p.

Note that the bins $\mathcal{V}_l$ have width 2ϵ and overlap on a region of width ϵ and $\sum_{l=0}^{L-1} b_l = 1$.

Theorem A.2 (QEEP Algorithm [3])**.** *One can solve the QEEP problem in Definition A.1 with $\{f^l(.)\}$ a set of 'bump' functions*

$$f^l(\phi) = \frac{2a}{\epsilon} \int_{l\epsilon - \frac{\epsilon}{2}}^{l\epsilon + \frac{\epsilon}{2}} d\phi' \times \exp\left\{ -\left[1 - \frac{4}{\epsilon^2}(\phi - \phi')^2\right]^{-1} \right\}, \tag{85}$$

with normalization constant $a \approx 2.252$, using PFE *in Def. 2.2 with $k = 0, 1, \ldots, K$ with $K =$*

$O\left(\epsilon^{-1}\ln^2(\epsilon^{-1})\right)$ *and* $M = \tilde{O}(|\ln(1-p)|\epsilon^{-4})$ *for each* $k = 1, 2\ldots, K$. *The total quantum cost for* U *is then bounded as* $T = O(MK^2) = \tilde{O}(|\ln(1-p)|\epsilon^{-6})$.

We note that the approximate indicator functions $f^l(.)$ in Eq. (85) are designed to have a quickly decaying Fourier series, which is required to achieve polynomial-time scaling. We also refer the reader to Ref. [23], which has extended the result by relaxing the requirement that $f^l(\phi) + f^{(l-1)}(\phi) = 1$ on the interval $\mathcal{V}_l \cap \mathcal{V}_{l-1}$. The idea of this QEEP algorithm is as follows. Since $b_l = \sum_j A_j f^l(\phi_j)$, using Eq. (3), periodically extending $f^l(\phi)$ beyond $[0, 2\pi)$ and Fourier decomposing gives $b_l = \sum_j A_j \sum_{k\in\mathbb{Z}} e^{ik\phi_j}\tilde{f}^l(k) = \sum_{k\in\mathbb{Z}} g(k)\tilde{f}^l(k)$. At the same time, the fact that $f^l(.)$ is an indicator function ensures that $b_l \approx \sum_{\phi_j\in\mathcal{V}_l} A_l$. Thus knowledge of $g(k)$ and the Fourier coefficients $\tilde{f}^l(k)$ for a range of k allows one to estimate the weights b_l. The requirement to estimate the spectral function to within a 1-norm ϵ, Eq. (84), is very stringent, hence the scaling of T with error ϵ is quite costly, $T = O(\epsilon^{-6})$. It is possible that one can improve the scaling by re-examining the analysis in [3].

A.2 Conservative QEEP Eigenvalue Extraction

Algorithm A.3 (Conservative QEEP Eigenvalue Extraction)**.** *Fix an overlap bound* A, *an error bound* $0 < \epsilon < \frac{A}{3}$, *and a confidence bound* $0 < p < 1$. *Assume access to a QEEP Algorithm A.2 for a unitary* V. *The algorithm proceeds as follows:*

1. *Use the QEEP subroutine with error* ϵ, *Alg. A.2, and confidence* p *to generate an estimate* $\tilde{b}_l$ *for* b_l *as defined in Eq.* (83).
2. *Construct the set*
$$S = \{l \in \{0, \ldots, L-1\} | \tilde{b}_l \geq A/3\}. \quad (86)$$
3. *Find the smallest* $l \in \{0, \ldots, L\}$ *with* $l \notin S$ *and call it* $l_{\min}$.
4. *For* $l' = l_{min}, \ldots, (l_{min} + L - 1) \bmod L$:
   (a) *if* $l' \in S$ *and* $(l'-1) \bmod L \in S$, *remove* l' *from* S.
5. *Return the set* $\{\tilde{\theta}_l = l\epsilon\}_{l\in S}$ *as a set of estimates of eigenphases of* V.

Let us now motivate Algorithm A.3. One may identify phases with sufficient probability in the output of the QEEP algorithm as the bins l with $b_l > b_{\text{cutoff}}$ with b_l in Eq. (83). Then to convert this into an estimate of a phase θ_j, for each such bin above cut-off we could output the estimate ϵl, i.e. in the middle of the corresponding bin $\mathcal{V}_l$.

We calculate appropriate values of b_{cutoff} and the QEEP error ϵ to guarantee that we output an estimate for each θ_j with $A_j > A$ with a provable confidence, and to guarantee no estimate in the absence of any θ_j. Def. A.1 states that when there exists such a $\theta_j \in \mathcal{V}_l \cap \mathcal{V}_{l-1}$, we are guaranteed with confidence p that $\tilde{b}_l + \tilde{b}_{l-1} + \epsilon > A$ (as $b_l + b_{l-1} > A$). To guarantee that at least one of $\tilde{b}_l$ or $\tilde{b}_{l-1}$ is larger than b_{cutoff} with the same confidence, we thus require $b_{\text{cutoff}} \leq (A-\epsilon)/2$.

Similarly, Def. A.1 states that when there exists no $\theta_j \in \mathcal{V}_l$ with $A_j > 0$, we are guaranteed with confidence p that $b_l < \epsilon$. To prevent a spurious estimate in this case, we require $b_{\text{cutoff}} \geq \epsilon$. Solving this to maximise ϵ (which minimizes the cost of the QEEP routine) yields $b_{\text{cutoff}} = \epsilon = A/3$ which is what we use in Alg. A.3.

A further small complication exists in solving the problem posed in Def. 3.1: we require that one outputs at most n_ϕ phases. This will be satisfied if we can guarantee at most one phase estimate per θ_j with $A_j > A$, as we know there are $n_\theta = n_\phi$ such estimates. As the bins $\mathcal{V}_l$ (Def. (82)) overlap, a phase θ_j may participate in up to two neighbouring bins — corresponding to amplitudes $b_l, b_{l+1} > A/3$. To ensure that this does not result in two estimates being generated, one could take all contiguous sets $b_{l_1}, b_{l_1+1}, \ldots, b_{l_2} > A/3$ and prune away every second index l. In doing this we need to respect the periodicity of the $\mathcal{V}_l$: $\mathcal{V}_{L-1} \cap \mathcal{V}_0 \neq \emptyset$, so these contiguous sets may wrap around the circle. Pruning every second index when starting from the middle of one of these contiguous sets may result in two neighbouring $l, l+1$ being removed, which is not what we desire. Instead, in the following pseudocode, after generating the set of all l with sufficient b_l, we find the first gap (in l) between these regions (corresponding to the first $b_l < A/3$). We then iterate (from this point $l_{\min}$ to $L-1$ and then from 0 to $l_{\min}$) over the b_l, and remove each l from our set if $b_{l-1} > A/3$ and

$l-1$ was not itself removed.

A.3 Proof of Lemma 3.2

Lemma 3.2 relates the performance of Alg. A.3 to the performance of the QEEP subroutine. We now prove this Lemma. Note that $\{\theta_j\}$ is an ordered list of n_ϕ phases, while $\{\tilde{\theta}_l\}$ is an ordered list (that we will show contains at most n_ϕ phases).

Proof of Lemma 3.2 Our proof follows by showing that the output from Alg. A.3 satisfies these statements whenever Eq. (84) holds. This yields our confidence bound as Eq. (84) holds with probability p by Def. A.1.

To see that statement 2 holds when Eq. (84) is satisfied, note that if $\mathcal{V}_l$ contains no phases θ_j with $A_j > A$ and Eq. (84) is satisfied, $\tilde{b}_l < \epsilon < A/3$, and l will not be added to the set S in Alg. A.3. This implies that when Eq. (84) holds, if $l \in S$ there exists some $\theta_j \in \mathcal{V}_l$ with $A_j > A$, in which case $|\theta_j - \tilde{\theta}_l|_T = |\theta_j - \epsilon l|_T \leq 2\epsilon$.

To see that statement 3 holds when Eq. (84) is satisfied, note that $\theta_j \in \mathcal{V}_l \cap \mathcal{V}_{(l+1) \bmod L}$ for exactly one l (and $\theta_j \notin \mathcal{V}_m$ for $m \neq l, (l+1) \bmod L$). Then, when Eq. (84) holds, $\tilde{b}_l + \tilde{b}_{(l+1) \bmod L} > 2A/3$, so $\max(\tilde{b}_l, \tilde{b}_{(l+1) \bmod L}) > A/3$, and either l, $(l+1) \bmod L$ or l and $(l+1) \bmod L$ will be added to the set S during step 2 of Alg. A.3 for each phase θ_j. Then, step 4 of Alg. A.3 will remove $(l+1) \bmod L$ if l remains in S, so each phase θ_j can contribute to only one final estimate $\tilde{\theta}_l = l\epsilon$, and the number of estimates is bounded from above by the number of phases.

To see that statement 1 holds, we use the point in the previous paragraph that, when Eq. (84) is satisfied, each phase θ_j with $A_j > A$ adds either l, $(l+1) \bmod L$, or l and $(l+1) \bmod L$ to the set S during step 2. of Alg. A.3. Then in step 4. of Alg. A.3, l is removed from S only if $(l-1) \bmod L$ remains in S, and $(l+1) \bmod L$ is removed from S only if l remains in S. This implies that the distance from θ_j to an estimate is bounded by

$$\max_{\theta_j \in \mathcal{V}_l \cap \mathcal{V}_{(l+1)}} \max_{l'=l-1,l,l+1 \bmod L} |\theta_j - \epsilon l'|_T = \\ = \epsilon(l+1) \bmod L - \epsilon(l-1)) \bmod L = 2\epsilon, \quad (87)$$

as required. □

B Proof of Lemma 4.2

Let us first prove the existence of $\kappa_{d+1} \in [2, \kappa_{\max}]$ with $\kappa_{\max} = 3$ that satisfies our conditions (the proof for k_1 is similar), for small enough ϵ in Eq. (27). Note that Eq. (27) implies

$$\epsilon \leq \frac{\pi}{300} \approx 0.01. \quad (88)$$

Given some pair $\tilde{\phi}_j^{(d)} \neq \tilde{\phi}_l^{(d)}$, $j < l$, let $\Delta_{j,l} = |\tilde{\phi}_j^{(d)} - \tilde{\phi}_l^{(d)}|$ and let $R_{j,l}$ be a set of κ_{d+1} such that neither Eq. (37) nor Eq. (36) holds for the chosen phases, that is,

$$R_{j,l} = \Big\{ \kappa_{d+1} \in [2, \kappa_{\max}] : \\ |\Delta_{j,l}|_T \geq \frac{\pi - 2\epsilon(1+\kappa_{d+1})}{k_d \kappa_{d+1}} \\ \wedge |k_d \kappa_{d+1} \Delta_{j,l}|_T \leq 4\epsilon(1+\kappa_{d+1}) \Big\}. \quad (89)$$

We call $\cup_{j,l} R_{j,l}$ the forbidden region and want to show that we can choose a value for $\kappa_{d+1} \in [2,3]$ outside this forbidden region if ϵ is sufficiently small. We do this by bounding the size of the forbidden region above and showing that this is smaller than the region $[2,3]$, leaving room to choose κ_{d+1}.

Note that $R_{j,l}$ is nonempty only if

$$\begin{aligned} k_d \Delta_{j,l} &\geq k_d |\Delta_{j,l}|_T \\ &\geq \frac{\pi - 2\epsilon(1+\kappa_{d+1})}{\kappa_{d+1}} \\ &\geq \frac{\pi - 2\epsilon(1+\kappa_{\max})}{\kappa_{\max}} \\ &> \frac{73\pi}{225}, \end{aligned} \quad (90)$$

using Eq. (88).

We may write the set $R_{j,l}$ as

$$R_{j,l} = \left[\max\left(2, \frac{\pi - 2\epsilon}{k_d |\Delta_{j,l}|_T + 2\epsilon}\right), \kappa_{\max} \right] \\ \cap \bigcup_{n \in \mathbb{N}} R_{j,l}^{(n)}, \quad (91)$$

where $R_{j,l}^{(n)}$ is the set of κ_{d+1} for which

$$|k_d \kappa_{d+1} \Delta_{j,l}|_T = \left| k_d \kappa_{d+1} \Delta_{j,l} - 2\pi n \right| \\ \leq 4\epsilon(1+\kappa_{d+1}), \quad (92)$$

for some $n \in \mathbb{N}$.

Solving this equation for κ_{d+1} yields

$$R_{j,l}^{(n)} = \left[\frac{2\pi n - 4\epsilon}{k_d \Delta_{j,l} + 4\epsilon}, \frac{2\pi n + 4\epsilon}{k_d \Delta_{j,l} - 4\epsilon}\right]. \tag{93}$$

The size of each interval $R_{j,l}^{(n)}$ can then be calculated

$$\left|R_{j,l}^{(n)}\right| = \frac{8\epsilon(2n\pi + k_d\Delta_{j,l})}{k_d^2\Delta_{j,l}^2 - 16\epsilon^2}. \tag{94}$$

We can bound

$$\begin{aligned}|R_{j,l}| &\leq \left|\bigcup_{n=1}^{n_{\max}} R_{j,l}^{(n)}\right| \leq \sum_{n=1}^{n_{\max}} \left|R_{j,l}^{(n)}\right| \\ &= \frac{8\epsilon}{k_d^2\Delta_{j,l}^2 - 16\epsilon^2}\Big(\pi n_{\max}(n_{\max}+1) \\ &\qquad + k_d\Delta_{j,l} n_{\max}\Big).\end{aligned} \tag{95}$$

Here, $n_{\max} = n_{\max}(j,l)$ is the largest index of a set $R_{j,l}^{(n)}$ in Eq. (92) for which $\kappa_{d+1} \in [2, \kappa_{\max}]$. Since $\kappa_{d+1} \leq 3$, Eq. (92) implies that

$$n_{\max} \leq \frac{3(k_d\Delta_{j,l} + 4\epsilon) + 4\epsilon}{2\pi}, \tag{96}$$

Now using Eq. (90) and Eq. (88) gives

$$\begin{aligned}|R_{j,l}| &\leq \frac{1}{1 - (4\epsilon/k_d\Delta_{j,l})^2} \\ &\times \frac{30\epsilon(864000\epsilon^2 + 248160\epsilon\pi + 11899\pi^2)}{5329\pi^3} \\ &\leq \frac{3\epsilon(864000\epsilon^2 + 248160\epsilon\pi + 11899\pi^2)}{532\pi^3} \\ &\leq 0.24,\end{aligned} \tag{97}$$

where the last inequality used Eq. (88). As there are $n_\phi \geq 2$ phases the length of the total forbidden region $\cup_{j,l} R_{j,l}$ is bounded from above by $\frac{n_\phi^2}{2}|R_{j,l}|$. We want to this interval to be, say, at most 1/4, so that by choosing κ_{d+1} randomly we have a 75% change of not landing in the forbidden interval. For larger n_ϕ we thus should use

$$\epsilon \leq \epsilon_{\text{crit}} = \frac{2\pi}{300 n_\phi^2}, \tag{98}$$

leading to Eq. (27).

We now repeat the above approach for the special case of finding the multiplier in the first round $\kappa_1 = k_1$. Consider thus $k_1 \in [3n_\phi, \kappa_{\max}]$ with $\kappa_{\max} = 3n_\phi + 1$. The key difference here is that there is a stricter lower bound on this multiplier $\kappa_{\max}$ so that ϵ needs to be chosen smaller, depending on n_ϕ, namely we choose

$$\epsilon_0 \leq \epsilon_{\text{crit},0} = \frac{2\pi}{300 n_\phi^4}, \tag{99}$$

as expressed in Eq. (26).

Given some pair $\tilde{\phi}_j^{(0)} \neq \tilde{\phi}_l^{(0)}$, $j < l$, and again let $\Delta_{j,l} = |\tilde{\phi}_j^{(0)} - \tilde{\phi}_l^{(0)}|$. Then, let $R_{j,l}$ be the set of k_1 such that neither Eq. (37) nor Eq. (36) holds for the chosen phases. That is,

$$\begin{aligned}R_{j,l} = \Big\{&k_1 \in [3n_\phi, \kappa_{\max}] : \\ &|\Delta_{j,l}|_T \geq \frac{\pi - 2\epsilon_0(1+k_1)}{k_1} \\ &\wedge |k_1\Delta_{j,l}|_T \leq 4\epsilon_0(1+k_1)\Big\}.\end{aligned} \tag{100}$$

We again call $\cup_{j,l} R_{j,l}$ the forbidden region and want to show that we can choose a value for $k_1 \in [3n_\phi, 3n_\phi + 1]$ outside this forbidden region, assuming that ϵ_0 is small enough. Note that the logic of the first few inequalities in Eq. (90) still holds in this new calculation, leading to

$$\begin{aligned}\Delta_{j,l} &\geq \frac{\pi - 2\epsilon_0(1+\kappa_{\max})}{\kappa_{\max}} \\ &\geq \frac{\pi(1 - \frac{2+3n_\phi}{75n_\phi^4})}{1+3n_\phi},\end{aligned} \tag{101}$$

where the second inequality used Eq. (99) and the value for $\kappa_{\max}$. Note that for large n_ϕ this allows $\Delta_{j,l}$ to decrease like $\sim 1/n_\phi$, while previously $\Delta_{j,l}$ was lowerbounded by a constant, Eq. (90).

This time, we may write the set $R_{j,l}$ as

$$\begin{aligned}R_{j,l} = &\left[\max\left(3n_\phi, \frac{\pi - 2\epsilon_0}{|\Delta_{j,l}|_T + 2\epsilon_0}\right), \kappa_{\max}\right] \\ &\cap \bigcup_{n\in\mathbb{N}} R_{j,l}^{(n)},\end{aligned} \tag{102}$$

where $R_{j,l}^{(n)}$ is the set of k_1 satisfying

$$\left|k_1\Delta_{j,l} - 2\pi n\right| \leq 4\epsilon_0(1+k_1), \tag{103}$$

for some $n \in \mathbb{N}$. Solving this equation for k_1 yields

$$R_{j,l}^{(n)} = \left[\frac{2\pi n - 4\epsilon_0}{\Delta_{j,l} + 4\epsilon_0}, \frac{2\pi n + 4\epsilon_0}{\Delta_{j,l} - 4\epsilon_0}\right]. \tag{104}$$

with length

$$\left|R_{j,l}^{(n)}\right| = \frac{8\epsilon_0(2n\pi + \Delta_{j,l})}{\Delta_{j,l}^2 - 16\epsilon_0^2}. \tag{105}$$

We can then bound

$$\begin{aligned}|R_{j,l}| &\leq \left|\bigcup_{n=n_{\min}}^{n_{\max}} R_{j,l}^{(n)}\right| \leq \sum_{n=n_{\min}}^{n_{\max}} \left|R_{j,l}^{(n)}\right| \\ &= \frac{8\epsilon_0}{\Delta_{j,l}^2 - 16\epsilon_0^2}\Big[\pi(n_{\max}^2 - n_{\min}^2) \\ &\quad + (\Delta_{j,l} + \pi)(n_{\max} - n_{\min}) \\ &\quad + 2\pi n_{\min} + \Delta_{j,l}\Big]. \end{aligned} \tag{106}$$

Here, $n_{\max} = n_{\max}(j,l)$ and $n_{\min} = n_{\min}(j,l)$ are the largest and smallest indices of sets $R_{j,l}^{(n)}$ in Eq. (103) for which $k_1 \in [3n_\phi, 3n_\phi + 1]$. Finding the minimal and the maximal value for n in Eq. (103) given the bounds on k_1 gives

$$\begin{aligned}\frac{3n_\phi(\Delta_{j,l} - 4\epsilon_0) - 4\epsilon_0}{2\pi} &\leq n_{\min} \leq n_{\max} \\ &\leq \frac{(3n_\phi + 1)(\Delta_{j,l} + 4\epsilon_0) + 4\epsilon_0}{2\pi}. \end{aligned} \tag{107}$$

As there are n_ϕ phases, the length of the total forbidden region $\cup_{j,l} R_{j,l}$ is bounded from above by $\frac{n_\phi^2}{2}|R_{j,l}|$. By plugging the bound for ϵ_0 in Eq. (99) and the bound for $\Delta_{j,l}$ in Eq. (101) together into Eq. (106), one can verify that

$$|R_{j,l}|\frac{n_\phi^2}{2} \leq \frac{-16 - 168n_\phi - 621n_\phi^2 + 552n_\phi^4 + 14400n_\phi^5 + 44550n_\phi^6 + 40500n_\phi^7 + 73125n_\phi^8 + 270000n_\phi^9 + 202500n_\phi^{10}}{450n_\phi^3(-4 - 9n_\phi - 100n_\phi^3 - 150n_\phi^4 + 1875n_\phi^7)}, \tag{108}$$

which can be verified to be less than 0.5 for all n_ϕ. Hence a random choice for k_1 in the interval $[3n_\phi, 3n_\phi + 1]$ gives at least a 50% chance to not land in the forbidden region. □

C Range of shifted phase estimates

In this Appendix we prove that when the unitary U is shifted as in step 2 of Alg. 4.1, and as long as the output of the phase extraction subroutine (Alg. A.3) meets the promises given in Lem. 3.2, all phase estimates of U^{k_d} will lie in the region for which Lem. 2.4 holds. This allows us to invoke Lem. 2.4 as required during Lem. 4.3 and Lem. 4.4. Note that if we were to shift the spectrum such that the middle of the largest gap would sit at 0, we would do $U \to Ue^{-i\zeta}$. However, the 'stay-away-from-the-boundary' condition of Lem. 2.4 is not symmetric, hence we shift by a different amount which also depends on the error ϵ_0 in the phase estimates.

Lemma C.1. *Let $\{\phi_j\}$ be the list of eigenphases of unitary U, and let n_ϕ, $\{\tilde{\phi}_l^{(0)}\}$, ζ, d_ζ be as defined in steps 1 and 2 of Alg. 4.1. Assume:*

- *(Assumption 1a) For every phase ϕ_j, there exists an estimate $\tilde{\phi}_l^{(0)}$ such that $|\phi_j - \tilde{\phi}_l^{(0)}|_T \leq 2\epsilon_0$.*
- *(Assumption 1b) For every estimate $\tilde{\phi}_l^{(0)}$, there exists a phase ϕ_j such that $|\phi_j - \tilde{\phi}_l^{(0)}|_T \leq 2\epsilon_0$.*

Then for all $k \geq 3n_\phi$, for the estimated shifted eigenphases $\{\tilde{\varphi}_l^{(0)}\}$ it holds that

$$\frac{\pi}{k} + 16\epsilon_0 \leq \tilde{\varphi}_l^{(0)} \leq \frac{\pi(2\lfloor k \rfloor - 1)}{k} - 16\epsilon_0. \tag{109}$$

which implies Eq. (8) for $\tilde{\varphi}_l^{(0)} = \phi$. In addition, the eigenphases $\{\varphi_j\}$ of the shifted unitary $Ue^{-i(\zeta + d_\zeta/2 - 8\epsilon_0)}$ satisfy

$$\frac{\pi}{k} + 14\epsilon_0 \leq \varphi_j \leq \frac{\pi(2\lfloor k \rfloor - 1)}{k} - 14\epsilon_0. \tag{110}$$

which again implies Eq. (8) for $\varphi_j = \phi$.

Proof. Let

$$\tilde{\varphi}_l^{(0)} = \left(\tilde{\phi}_l^{(0)} - \zeta - \frac{d_\zeta}{2} + 8\epsilon_0\right) \bmod 2\pi. \tag{111}$$

Let us first show that

$$\tilde{\varphi}_l^{(0)} \in \left[\frac{\pi}{2n_\phi} + 8\epsilon_0, 2\pi - \frac{3\pi}{2n_\phi} + 8\epsilon_0\right]. \tag{112}$$

By definition of ζ (as the midway point in the largest gap) and d_ζ (as half the largest gap) we have

$$\left(\tilde{\phi}_l^{(0)} - \zeta - \frac{d_\zeta}{2}\right) \bmod 2\pi \in \left[\frac{d_\zeta}{2}, 2\pi - \frac{3d_\zeta}{2}\right]. \tag{113}$$

We have $d_\zeta \geq \pi/n_\phi$ (with equality corresponding to n_ϕ uniformly distributed estimates). By Eq. (26) it follows that

$$\epsilon_0 \leq \frac{\pi}{48 n_\phi}, \tag{114}$$

and thus $\epsilon_0 < \frac{d_\zeta}{16}$, leading to Eq. (112). By the assumptions the shifted phases φ_j lie within $2\epsilon_0$ from the estimates $\tilde{\varphi}_l^{(0)}$. Thus for each φ_j there exists a $\tilde{\varphi}_l^{(0)}$ such that

$$\begin{aligned} \varphi_j - 14\epsilon_0 &\geq \tilde{\varphi}_l^{(0)} - 16\epsilon_0 \\ &\geq \frac{\pi}{2n_\phi} - 8\epsilon_0 \\ &\geq \frac{\pi}{2n_\phi} - \frac{8\pi}{48 n_\phi} \\ &= \frac{\pi}{3n_\phi} \geq \frac{\pi}{k}. \end{aligned} \tag{115}$$

and

$$\begin{aligned} \varphi_j + 14\epsilon_0 &\leq \tilde{\varphi}_l^{(0)} + 16\epsilon_0 \\ &\leq 2\pi - 3\left(\frac{\pi}{2n_\phi} - 8\epsilon_0\right) \\ &\leq 2\pi - \frac{3\pi}{k} \\ &\leq \frac{\pi(2\lfloor k \rfloor - 1)}{k}. \end{aligned} \tag{116}$$

where we have used Eq. (112), Eq. (114), $k \geq 3n_\phi$ and $\lfloor k \rfloor > k - 1$. This implies Eq. (110) and also Eq. (109). □