Iterative quantum-phase-estimation protocol for shallow circuits

Joseph G. Smith,[1,2,*] Crispin H. W. Barnes,[1,†] and David R. M. Arvidsson-Shukur[2,‡]
[1]*Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, United Kingdom*
[2]*Hitachi Cambridge Laboratory, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*

(Received 4 July 2022; accepted 2 December 2022; published 19 December 2022)

Given N_{tot} applications of a unitary operation, parametrized by an unknown phase, a phase-estimation protocol on a large-scale fault-tolerant quantum system can reduce the standard deviation of an estimate of the phase from scaling as $O[1/\sqrt{N_{\text{tot}}}]$ to scaling as $O[1/N_{\text{tot}}]$. Owing to the limited resources available to near-term quantum devices, protocols that do not entangle probes have been developed. Their mean absolute error scales as $O[\log(N_{\text{tot}})/N_{\text{tot}}]$. Here, we propose a two-step protocol for near-term phase estimation, with an improved error scaling. Our protocol's first step produces several low-standard-deviation estimates of θ, within θ's parameter range. The second step iteratively homes in on one of these estimates. Our protocol achieves a mean-absolute-error scaling of $O[\sqrt{\log(\log N_{\text{tot}})}/N_{\text{tot}}]$ and a root-mean-square-error scaling of $O[\sqrt{\log N_{\text{tot}}}/N_{\text{tot}}]$. Furthermore, we demonstrate a reduction in the constant scaling factor and the required circuit depths. This allows our protocol to outperform the asymptotically optimal quantum-phase-estimation algorithm for realistic values of N_{tot}.

DOI: 10.1103/PhysRevA.106.062615

I. INTRODUCTION

The task of finding an unknown parameter θ of a unitary operation $\hat{U}(\theta)$ requires *phase estimation*. Phase estimation is one of the most prominent tasks in quantum-information processing. Various forms of phase estimation occur in, for example, the subroutines of quantum algorithms [1–5], protocols to find ground-state energies [6], gravitational-wave detection [7], fixed-reference-frame sharing [8], synchronization of clocks [9], and, famously, the measurement of time [10]. To measure an unknown quantity of interest, θ, a quantum probe ψ_0 is subjected to the unitary operation $\hat{U}(\theta)$, such that the output probe ψ_θ carries useful information [11]. This information is then accessed via measurements. As quantum measurements are probabilistic in nature, statistics lead to a bound on the error of any estimate of θ, $\tilde{\theta}$. (Throughout this paper, estimates of the quantity X are distinguished using $\tilde{X}$.) By utilizing quantum phenomena, these bounds can be improved. In particular, if $\hat{U}(\theta)$ is queried N_{tot} times, each time using a separate probe, the error, $\Delta\tilde{\theta}$, scales asymptotically with the shot-noise limit $\Delta\tilde{\theta} \propto 1/\sqrt{N_{\text{tot}}}$. Using quantum coherence or entanglement, the scaling can be improved to the Heisenberg limit: $\Delta\tilde{\theta} \propto 1/N_{\text{tot}}$ [12–14]. The ability to decrease the error in this way constitutes one of the most tractable technological applications for quantum advantage.

An example of an algorithm that achieves the Heisenberg limit is the quantum-phase-estimation algorithm (QPEA). This algorithm uses the inverse Fourier transform on a set of entangled probes to provide an estimate of a phase [15,16]. However, the circuit depths, coherence times, and gate fidelities needed for practical use of this algorithm are far beyond the realistic regime of noisy intermediate-scale quantum devices [17]. Instead, one can use maximum likelihood estimators (MLEs) to analyze the measurement outcomes of a single, shallower circuit [18]. These quantum-classical strategies involve quantum-probe preparation followed by "classical" measurements, which sample individual probes separately [19]. A significant, but often overlooked, drawback of MLE strategies that sample only one circuit is that their error minimization leads to an estimate which is not *point identified* (see below). That is, the MLE cannot distinguish between several possible values of θ [20–22]. To combat this, MLE-based protocols have been introduced that iteratively measure multiple circuits to allow point identification [23].

In this paper, we consider quantum phase estimation with shallow circuits. We construct a two-step protocol that splits the phase-estimation problem into a quantum-classical strategy and a point-identification strategy. Compared with previous phase-estimation protocols, our protocol achieves better mean-error bounds with shallower circuits. When the point identification is conducted iteratively, our protocol achieves a mean-absolute-error scaling of $O(\sqrt{\log(\log N_{\text{tot}})}/N_{\text{tot}})$ and a root-mean-square-error scaling of $O(\sqrt{\log N_{\text{tot}}}/N_{\text{tot}})$. These scalings are better than those of previous iterative protocols [8,9]. Additionally, we show that our protocol, which requires no entanglement between probes or programmable phase shifts [23], achieves estimates with lower mean errors than those acquired by the QPEA, for experimentally realistic circuit depths and values of N_{tot}.

II. BACKGROUND

Throughout this paper, we focus on Stone's encoded unitaries with a fixed θ [24]: $\hat{U}(\theta) = e^{i\theta\hat{A}}$, where $\hat{A}$ is a known Hermitian generator independent of θ [25]. After ignoring a global phase factor and conducting a suitable parameter

*jgs46@cam.ac.uk
†chwb101@cam.ac.uk
‡drma2@cam.ac.uk

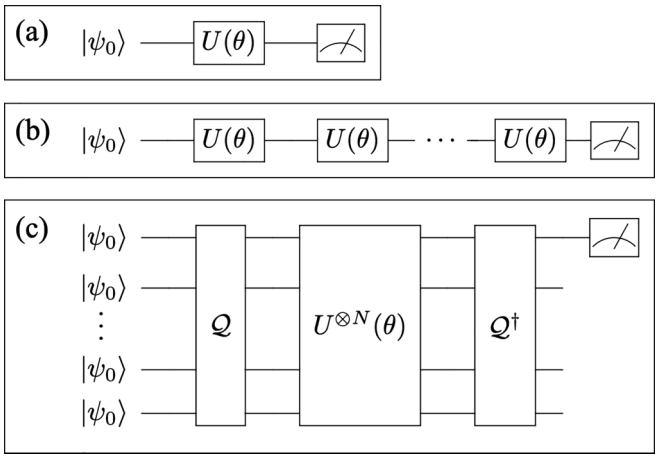

FIG. 1. Quantum circuits used to estimate θ with (a) one application of $\hat{U}(\theta)$ and (b) N coherent applications of $\hat{U}(\theta)$ in series. (c) Phase estimation via entanglement of N probes. The gate $\mathcal{Q}$ is used to entangle the probes into the GHZ state from an initial state $|\psi_0\rangle$.

rescaling, the action of $\hat{U}(\theta)$ on the eigenstates corresponding to the minimum and maximum eigenvalues of $\hat{A}$ is

$$\hat{U}(\theta)\,|a_{\min}\rangle = |a_{\min}\rangle\,,$$
$$\hat{U}(\theta)\,|a_{\max}\rangle = e^{i\theta}\,|a_{\max}\rangle\,. \tag{1}$$

We focus on optimal phase estimation, by setting the input-probe states to $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|a_{\min}\rangle + |a_{\max}\rangle)$, such that $\hat{U}(\theta)\,|\psi_0\rangle = \frac{1}{\sqrt{2}}(|a_{\min}\rangle + e^{i\theta}\,|a_{\max}\rangle)$. This state maximizes the acquired phase difference from $\hat{U}(\theta)$ [19]. For example, consider the measurement of the strength of a magnetic field, $\boldsymbol{B}$, aligned in the z direction. Then, $\hat{A}$ is the Pauli z matrix, $\theta \propto |\boldsymbol{B}|$, and $|\psi_0\rangle$ is prepared by polarizing, e.g., electron spins in the x direction.

We note that applying $\hat{U}(\theta)$ sequentially N times to $|\psi_0\rangle$ is equivalent to applying $\hat{U}(N\theta)$ to $|\psi_0\rangle$ once. The probability that the probe remains in the state $|\psi_0\rangle$ after N applications of $\hat{U}(\theta)$ is

$$p_0(N,\theta) = |\langle\psi_0|\hat{U}^N(\theta)|\psi_0\rangle|^2 = \tfrac{1}{2}[1+\cos(N\theta)]. \tag{2}$$

In this scenario, there is no entanglement generated between different probes. However, the probes themselves can be in internally entangled states. Alternatively, one could prepare N probes in a Greenberger-Horne-Zeilinger (GHZ) state and apply $\hat{U}(\theta)$ once to each probe in parallel [11] [see Figs. 1(b) vs 1(c)].

It is possible to estimate θ through an estimate of $p_0(N,\theta)$:

$$\theta = \pm\frac{1}{N}\arccos\left[2p_0(N,\theta)-1\right] + \frac{2\pi l}{N}, \tag{3}$$

for integer l. The estimate of $p_0(N,\theta)$ can be achieved by first preparing ν probes in state $|\psi_0\rangle$, then applying a $\hat{U}(\theta)$ operation N times to each probe, and finally measuring the probes in the $\{|\psi_0\rangle\,, |\psi_0^\perp\rangle\}$ basis. If x of these ν measurements correspond to the $|\psi_0\rangle$ outcome, MLEs [26] can be used to estimate $p_0(N,\theta)$: $\widetilde{p}_0(N,\theta) = \frac{x}{\nu}$. The associated standard deviation is $\sigma_{\widetilde{p}_0(N,\theta)} \geqslant \sqrt{\frac{p_0(N,\theta)(1-p_0(N,\theta))}{\nu}}$ [27]. [We distinguish the mean absolute error (MAE) of an estimate of X, $\Delta\widetilde{X}_{\text{MAE}}$, from the root-mean-square error (RMSE), $\Delta\widetilde{X}_{\text{RMS}}$, and from the standard deviation $\sigma_{\widetilde{X}}$.] From Eq. (3), we see that an estimate of θ has a lower bound on the standard deviation: $\sigma_{\widetilde{\theta}} \geqslant \frac{1}{N\sqrt{\nu}}$. This inequality saturates for large ν. The reduction in standard deviation by a factor N arises directly from quantum coherence [in Fig. 1(b)] or entanglement between probes [in Fig. 1(c)] [28]. Methods that do not use quantum phenomena [Fig. 1(a)] have $N = 1$ and achieve a standard deviation bounded by the standard quantum limit: $\sigma_{\widetilde{\theta}} \geqslant \frac{1}{\sqrt{\nu}}$.

An obvious problem with the aforementioned quantum methods is that for any given $p_0(N,\theta)$, $2N$ different values of $\theta \in [0, 2\pi)$ satisfy Eq. (3). Point identification [20–22] is needed to determine the correct l and yield an unambiguous estimate of θ. Even the classical method, where $N = 1$, cannot distinguish between a true underlying parameter of θ or $2\pi - \theta$. In this case, one can achieve point identification by carrying out also a second circuit in which $\hat{U}(\theta)$ is followed by $\hat{U}(\pi/2)$, a known, fixed unitary. For example, this may be a magnetic pulse of a known and fixed duration and strength. In the second circuit, $p_0(1,\theta)$ becomes $p_0(1,\theta+\pi/2) = \frac{1}{2}[1-\sin(\theta)]$. If $p_0(1,\theta+\pi/2) < 1/2$, $\theta \in [0,\pi)$, else $\theta \in [\pi, 2\pi)$ [29]. Thus the second circuit allows us to point-identify in which subspace of the parameter range the unknown parameter lies. In the general case, $N > 1$, point identification is not achieved by applying $\hat{U}(\pi/2)$ alone. One must iteratively increase N and conduct corresponding quantum-classical point-identification techniques until the target N is reached [8]. The point-identification procedures require measurements that do not necessarily decrease the error of the final estimate. Consequently, point identification leads to difficulties in reaching the Heisenberg limit.

Throughout this paper, we take the total number of applications of $\hat{U}(\theta)$, N_{tot}, as the resource of phase-estimation protocols. That is, we compare the mean errors of protocols with N_{tot} applications of the unknown unitary. To investigate the viability of protocols on noisy intermediate-scale quantum hardware, we also consider the protocols' maximum circuit depth $N_{\max}$.

III. TWO-STEP PROTOCOLS

We now introduce our two-step protocols, which split the phase estimation into two steps: First, a fine-tuning step that executes a circuit with N applications of $\hat{U}(\theta)$ to achieve several low-standard-deviation estimates of θ. Second, a point-identification step that disambiguates the estimate through either an iterative method or an application of the QPEA (see below). Given a point-identification method and a value of N_{tot}, N is chosen to minimize $\Delta\widetilde{\theta}$.

Consider a measurement of the circuit in Fig. 1(b) with $N = 2^m$, where $m \in \mathbb{N}$. This corresponds to the fine-tuning step of our protocol. By defining $\theta \equiv 2\pi T$, $T \in [0,1)$, and binary-expanding $T = \sum_{j=1}^{\infty} t_j 2^{-j}$, where t_j is the jth binary bit of T, the probability of measuring a $|\psi_0\rangle$ state, Eq. (2), becomes

$$p_0(2^m,\theta) = \tfrac{1}{2}[1 \pm \cos(\theta_{\text{FT}})]. \tag{4}$$

Here, $\theta_{\rm FT} \equiv 2\pi \sum_{j=m+2}^{\infty} t_j 2^{m-j} \in [0, \pi]$, and addition (subtraction) occurs if $t_{m+1} = 0$ ($t_{m+1} = 1$). We note that only the bits t_j with $j > m+1$ affect $p_0(2^m, \theta)$ in this fine-tuning step. The circuit with $N = 2^m$ is executed $\nu_{\rm FT}$ times and, upon counting $x_{\rm FT}$ probes in the state $|\psi_0\rangle$, we estimate $\widetilde{p}_0(2^m, \theta) = \frac{x_{\rm FT}}{\nu_{\rm FT}}$. We then invert Eq. (4) to estimate $\theta_{\rm FT}$. Fine-tuning involves $\hat{U}(\theta)$ being applied $\nu_{\rm FT} 2^m$ times and returns an estimate with $\sigma_{\widetilde{\theta}_{\rm FT}} = \frac{1}{\sqrt{\nu_{\rm FT}}}$ for large $\nu_{\rm FT}$.

The next step is point identification, which involves finding the bits t_j with $j \leqslant m+1$. These bits define the quantity $\theta_{\rm PI} \equiv 2\pi \sum_{j=1}^{m+1} t_j 2^{-j}$. An estimate of $\theta_{\rm PI}$ can be found by a number of methods. We give two examples below. In general, this step applies $\hat{U}(\theta)$ a total of $N_{\rm PI}$ times. A final estimate of θ is then given by

$$\widetilde{\theta} = \widetilde{\theta}_{\rm PI} + 2^{-m} \widetilde{\theta}_{\rm FT}, \tag{5}$$

with standard deviation $\sigma_{\widetilde{\theta}} = 2^{-m} \sigma_{\widetilde{\theta}_{\rm FT}}$ if the point identification was successful.

If $\hat{U}(\theta)$ is applied $N_{\rm tot}$ times over the two steps, $\nu_{\rm FT}$ can take a maximum value of $\lfloor 2^{-m}(N_{\rm tot} - N_{\rm PI}) \rfloor$, giving a standard deviation of

$$\sigma_{\widetilde{\theta}} = \frac{1}{2^m \sqrt{\lfloor 2^{-m}(N_{\rm tot} - N_{\rm PI}) \rfloor}}. \tag{6}$$

$\sigma_{\widetilde{\theta}}$ can then be combined with the other errors to form the MAE or RMSE of the estimate: If the protocol produces multiple errors, $\Delta\widetilde{\theta}_i$, with probabilities ϵ_i, the MAE and RMSE are given by

$$\Delta\widetilde{\theta}_{\rm MAE} \leqslant \sum_i \epsilon_i \Delta\widetilde{\theta}_i,$$
$$\Delta\widetilde{\theta}_{\rm RMS} \leqslant \sqrt{\sum_i \epsilon_i (\Delta\widetilde{\theta}_i)^2}. \tag{7}$$

A. Iterative method

Here, we outline how to estimate $\theta_{\rm PI}$ through iteration of many circuits. These circuits have varying depth, $N = 2^i$, for integers $i \in [0, 1, \ldots, m-1]$, such that the deepest circuit has depth $N_{\rm max} = 2^{m-1}$. The circuits are executed to estimate whether $p_0(2^i, \theta) > 1/2$, using the MLE method defined above. We set $m \to i$ in Eq. (4) and note that if $t_{i+2} = 1$, then $\cos(2\pi \sum_{j=i+2}^{\infty} t_j 2^{i-j}) \leqslant 0$, and $p_0(2^i, \theta) \leqslant 1/2$ or $p_0(2^i, \theta) \geqslant 1/2$ if $t_{i+1} = 0$ or $t_{i+1} = 1$, respectively. If instead $t_{i+2} = 0$, the relationship between $p_0(2^i, \theta)$ and t_{i+1} is the opposite. Therefore knowing the value of the bit t_{i+1} and estimating whether $p_0(2^i, \theta) < 1/2$ allows us to estimate the bit t_{i+2}. We then iterate by increasing i from 0 up to $m-1$ to estimate all of the first $m+1$ bits of T, bar the first bit, t_1. t_1 is estimated differently, by using an evolution of $\hat{U}(\theta + \pi/2)$ as described above. The whole iteration process is summarized in Fig. 2. The advantage of estimating θ from most- to least-significant bit, as opposed to from least- to most-significant bit [23], is twofold: First, there is no reliance on the programmability of a phase shift, only a fixed one of $\pi/2$. Second, the number of bits needed to describe θ need not be capped. In addition, because our phase-estimation problem is split into many circuits, it is suitable for parallel execution.

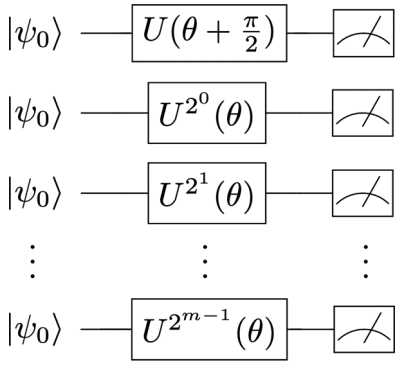

FIG. 2. Circuits that are executed to estimate the first $m+1$ bits of T, $\theta_{\rm PI}$. The ith bit is estimated by a circuit with $N = 2^{i-2}$ applications of $\hat{U}(\theta)$ for $i > 1$, and $N = 1$, with an additional phase shift, for $i = 1$.

To cap the probability, ϵ, that the entire point-identification step fails, we need to limit the probability, ϵ_i, that the ith bit of T is incorrectly assigned. Thus the circuit used to estimate the ith bit must be executed a minimum number of times, ν_i. A suitable ν_i can be calculated using the binomial distribution's Chernoff bound [8,16]:

$$\Pr[|\widetilde{p}_0(N, \theta) - p_0(N, \theta)| \geqslant \delta] \equiv \epsilon_i \leqslant 2e^{-\nu_i \delta^2/2}, \tag{8}$$

where δ is the maximum allowed absolute difference between the estimated $\widetilde{p}_0(N, \theta)$ and true $p_0(N, \theta)$. Failure occurs if $\widetilde{p}_0(N, \theta) > 1/2$ when $p_0(N, \theta) < 1/2$ (and vice versa): $|\widetilde{p}_0(N, \theta) - p_0(N, \theta)| \geqslant |\frac{1}{2} - p_0(N, \theta)|$. Hence we choose $\delta = |\frac{1}{2} - p_0(N, \theta)|$ when solving Eq. (8):

$$\nu_i \geqslant \frac{2\ln(2/\epsilon_i)}{(1/2 - p_0(N, \theta))^2} = \frac{8\ln(2/\epsilon_i)}{\cos^2(N\theta)}. \tag{9}$$

Problematically, one needs knowledge of θ to find ν_i. Furthermore, when $\cos N\theta \approx 0$, we require $\nu_i \to \infty$: An impossibly large number of samples is required to correctly identify the ith bit. Previous works [30,31] on iterative phase-estimation schemes have not considered this issue, which we now address.

If $N = 2^{k-2}$, then $\cos(N\theta) = 0$ when $N\theta = \frac{\pi l}{2}$ for integer l. Also, $\cos(N\theta) = 0$ when T is represented exactly in binary with k bits. In this case, $p_0(N, \theta) = \frac{1}{2}$, and there is an equal probability of estimating $p_0(N, \theta) < \frac{1}{2}$ as there is of estimating $p_0(N, \theta) > \frac{1}{2}$. If T is estimated up to the $(m+1)$th bit with $m+1 > k$, the bits after the kth bit also have a 50% probability of being incorrect [due to the $\pm$ sign in Eq. (4) being flipped], and $\theta_{\rm PI} - \widetilde{\theta}_{\rm PI} = \frac{\pi}{2^m}$. However, if $\pm \to \mp$ in Eq. (4), then also the fine-tuning step is affected, resulting in $\widetilde{\theta}_{\rm FT} = \pi - \theta_{\rm FT} = \pi$ instead of $\widetilde{\theta}_{\rm FT} = 0$. These two issues cancel out and lead to a correct estimate of θ. For values of θ where $\cos(N\theta) \approx 0$, $\theta_{\rm FT}$ is small, and the final estimate $\widetilde{\theta} = \theta - \frac{2\theta_{\rm FT}}{2^m}$. Furthermore, this mislabeling of the bit will occur with probability less than 50% provided the denominator, $\alpha \equiv 8\sec^2 N\theta$, is large enough. In simulations, $\alpha = 32$ was seen as sufficient for $N_{\rm tot} \leqslant 10^6$ (see below).

In summary, the circuit that estimates the ith bit is executed $\nu_i = \alpha \ln(2/\epsilon_i)$ times. Over the whole point-identification

step, we thus apply $\hat{U}(\theta)$ a total number

$$N_{\mathrm{PI}} = \nu_1 + \sum_{i=2}^{m+1} 2^{i-2}\nu_i = \alpha \ln(2/\epsilon_1) + \sum_{i=2}^{m+1} \alpha 2^{i-2} \ln(2/\epsilon_i), \tag{10}$$

times to estimate θ_{PI}.

We now make the assertion that the whole point-identification protocol is incorrect with maximum probability ϵ, such that $1-\epsilon = \prod_{i=1}^{m}(1-\epsilon_i)$. Equivalently, $\epsilon \leqslant \sum_{i=1}^{m+1}\epsilon_i$. We use Lagrange multipliers to minimize Eq. (10) with this constraint. We find that $\epsilon_1 = 2^{-m}\epsilon$, $\epsilon_i = 2^{i-m-2}\epsilon$ for $i > 1$, and

$$N_{\mathrm{PI}} = \alpha 2^m \ln\left(\frac{8}{\epsilon}\right) - 2\alpha \ln 2. \tag{11}$$

To decrease ϵ, each circuit is sampled a larger number of times, proportional to $\ln(\frac{1}{\epsilon})$. The iteration described above succeeds with probability at least $1-\epsilon$. In these scenarios, the estimate has an error of $\Delta\widetilde{\theta}_{\mathrm{success}}$. However, if the ith bit of T is incorrectly identified, the subsequent bits are also incorrectly labeled. Therefore the final estimate of $\widetilde{\theta}_{\mathrm{PI}}$ differs from the true θ by up to twice the value of the ith bit: $\Delta\widetilde{\theta}_{\mathrm{fail},i} = \frac{\pi}{2^{i-2}}$. Each error $\Delta\widetilde{\theta}_{\mathrm{fail},i}$ will occur with probability ϵ_i. Therefore

$$\sum_{i=1}^{m+1} \epsilon_i \widetilde{\theta}_{\mathrm{fail},i} = (m+2)\frac{\pi\epsilon}{2^m},$$
$$\sum_{i=1}^{m+1} \epsilon_i (\widetilde{\theta}_{\mathrm{fail},i})^2 = \frac{2\pi^2\epsilon}{2^m}\left[3 - \frac{1}{2^m}\right]. \tag{12}$$

It is possible to take $\widetilde{\theta}_{\mathrm{PI}}$ as the final estimate of θ, i.e., have no fine-tuning step. We refer to this method as iteration alone. The error of a successful run of the iteration-alone protocol is the value of all the bits truncated: $\Delta\widetilde{\theta}_{\mathrm{success}} \leqslant \sum_{i=m+1}^{\infty}\frac{\pi}{2^m}$. The mean errors follow the bounds

$$\Delta\widetilde{\theta}_{\mathrm{MAE}} \leqslant (1+(m+1)\epsilon)\frac{\pi}{2^m},$$
$$\Delta\widetilde{\theta}_{\mathrm{RMS}} \leqslant \sqrt{\frac{2\pi^2}{2^m}\left[\frac{2}{2^m} - \frac{3\epsilon}{2^m} + 3\epsilon\right]}. \tag{13}$$

In the asymptotic limit, where m is large, a constant ϵ leads to $N_{\mathrm{tot}} = O(2^m)$, $\Delta\widetilde{\theta}_{\mathrm{MAE}} = O(m2^{-m}) = O(\log N_{\mathrm{PI}}/N_{\mathrm{PI}})$, and $\Delta\widetilde{\theta}_{\mathrm{RMS}} = O(2^{-m}) = O(1/\sqrt{N_{\mathrm{PI}}})$. However, allowing ϵ to be a function of m improves the scaling: The choice of $\epsilon = O(\frac{1}{m})$ leads to optimization of the MAE scaling of $\Delta\widetilde{\theta}_{\mathrm{MAE}} = O(\frac{\log(\log N_{\mathrm{PI}})}{N_{\mathrm{PI}}})$. Choosing $\epsilon = O(2^{-m})$ optimizes the RMSE with a scaling of $\Delta\widetilde{\theta}_{\mathrm{RMS}} = O(\frac{\log N_{\mathrm{PI}}}{N_{\mathrm{PI}}})$.

B. Iterative two-part protocol

In our protocol, we combine an iterative point-identification step with a fine-tuning step, where an additional circuit of depth $N_{\max} = 2^m$ is executed. This is outlined in AlgorithmI: Lines 1–14 are the point-identification step, and lines 15–17 are the fine-tuning step. If the iterative point identification succeeds, the error comes from the fine-tuning step: $\Delta\widetilde{\theta}_{\mathrm{success}} = \sigma_{\widetilde{\theta}}$ [Eq. (6)]. Combining the errors of the two steps of our protocol, the mean errors are bounded by

Algorithm 1. Iterative two-part protocol.

Input: Values of $N_{\mathrm{tot}}, m, \epsilon, \alpha$
Output: Estimate of θ

1: Execute $\hat{U}\left(\theta + \frac{\pi}{2}\right)$ circuit $\nu_1 = \alpha \ln\left(\frac{2^{m+1}}{\epsilon}\right)$ times
2: $x_1 \leftarrow$ number of $|\psi_0\rangle$ measured
3: $t_1 \leftarrow \begin{cases} 0, & 2x_1 < \nu_1 \\ 1, & 2x_1 \geqslant \nu_1 \end{cases}$
4: $\widetilde{\theta}_{\mathrm{PI}} \leftarrow \pi \times t_1$
5: $N_{\mathrm{left}} \leftarrow N_{\mathrm{tot}} - \nu_1$
6: $i \leftarrow 2$ ▷ i records the bit to be measured
7: **while** $i \leqslant m+1$
8: Execute depth $N = 2^{i-2}$ circuit $\nu_i = \alpha \ln\left(\frac{2^{m+3-i}}{\epsilon}\right)$ times
9: $x_i \leftarrow$ number of $|\psi_0\rangle$ measured
10: $t_i \leftarrow \begin{cases} \mathrm{NOT}\ t_{i-1}, & 2x_i < \nu_i \\ t_{i-1}, & 2x_i \geqslant \nu_i \end{cases}$
11: $\widetilde{\theta}_{\mathrm{PI}} \leftarrow \widetilde{\theta}_{\mathrm{PI}} + \frac{\pi}{2^{i-1}} \times t_i$
12: $N_{\mathrm{left}} \leftarrow N_{\mathrm{tot}} - 2^{i-2}\nu_i$
13: $i \leftarrow i+1$
14: **end while**
15: Execute depth $N = 2^m$ circuit $\nu_{\mathrm{FT}} = \left\lfloor \frac{N_{\mathrm{left}}}{2^m} \right\rfloor$ times
16: $x_{\mathrm{FT}} \leftarrow$ number of $|\psi_0\rangle$ measured
17: $\widetilde{\theta}_{\mathrm{FT}} \leftarrow (-1)^{t_{m+1}} \arccos\left(\frac{x_{\mathrm{FT}}}{\nu_{\mathrm{FT}}} - 1\right)$
18: **Return:** $\widetilde{\theta} = \widetilde{\theta}_{\mathrm{PI}} + \frac{\widetilde{\theta}_{\mathrm{FT}}}{2^m}$

$$\Delta\widetilde{\theta}_{\mathrm{MAE}} \leqslant (1-\epsilon)\sigma_{\widetilde{\theta}} + \frac{(m+2)\pi\epsilon}{2^m},$$
$$\Delta\widetilde{\theta}_{\mathrm{RMS}} \leqslant \sqrt{(1-\epsilon)\sigma_{\widetilde{\theta}}^2 + \frac{2\pi^2\epsilon}{2^m}\left[3 - \frac{1}{2^m}\right]}. \tag{14}$$

Optimization of the MAE occurs for the choice $\epsilon = O(m^{-3/2})$, and optimization of the RMSE occurs for the choice $\epsilon = O(2^{-m})$. Using these values of ϵ, the bounds above scale as

$$\Delta\widetilde{\theta}_{\mathrm{MAE}} = O\left(\frac{\sqrt{\log(\log N_{\mathrm{tot}})}}{N_{\mathrm{tot}}}\right),$$
$$\Delta\widetilde{\theta}_{\mathrm{RMS}} = O\left(\frac{\sqrt{\log N_{\mathrm{tot}}}}{N_{\mathrm{tot}}}\right), \tag{15}$$

respectively. These mean-error scalings are better than what one would achieve by using the iteration-alone protocol. Furthermore, in the simulations below (see Fig. 3), we see that our protocol benefits from a significant reduction in the constant factor of the error scaling and in the required circuit depth.

C. Point identification using the QPEA

The QPEA employs inverse Fourier transforms instead of MLEs to estimate θ [32]. To gain a b-bits estimate of $T = \theta/2\pi$ with an expected failure probability of ϵ, $t = b + \lceil \log_2(2 + \frac{1}{2\epsilon}) \rceil$ probes are manipulated with $N_{\mathrm{tot}} = 2^t - 1$ applications of $\hat{U}(\theta)$. The ith probe is subject to 2^{i-1} coherent

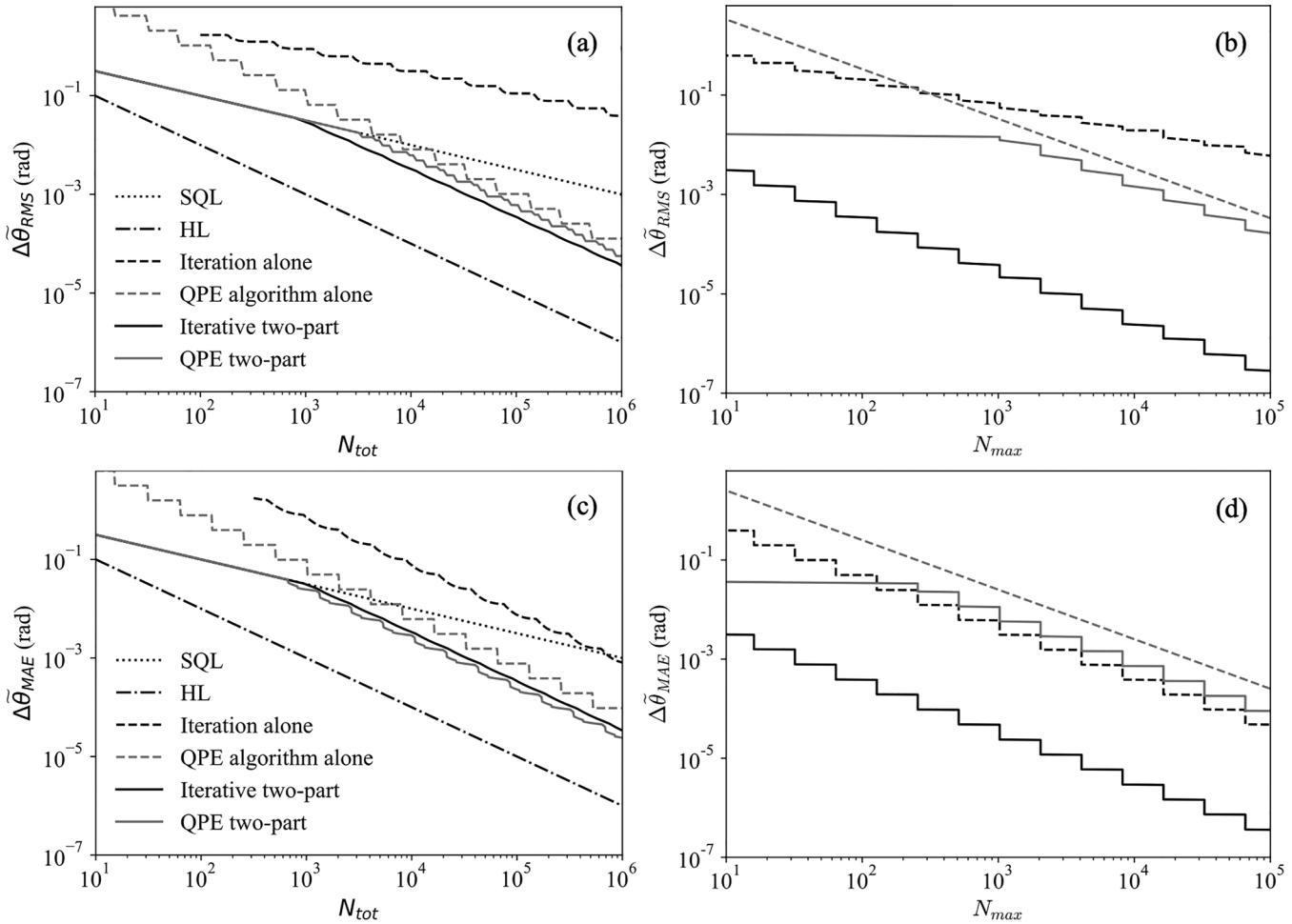

FIG. 3. Numerical simulations of the performance of each protocol discussed in the text, with $\alpha = 32$. The value of ϵ is chosen such that the minimum upper bound to either the RMSE [(a) and (b)] or the MAE [(c) and (d)] is achieved at different values of N_{tot}. (a) $\Delta\widetilde{\theta}_{\text{RMS}}$ vs N_{tot} for each protocol. (b) $\Delta\widetilde{\theta}_{\text{RMS}}$ vs N_{max} for each protocol. (This is the value of N_{max} that achieves the optimal bound on $\Delta\widetilde{\theta}_{\text{RMS}}$ for a given N_{tot}.) (c) $\Delta\widetilde{\theta}_{\text{MAE}}$ vs N_{tot} for each protocol. (d) $\Delta\widetilde{\theta}_{\text{MAE}}$ vs N_{max} for each protocol. QPE, quantum phase estimation.

applications of $\hat{U}(\theta)$ [16]. In the QPEA, $N_{\text{max}} = 2^{t-1} + O(t)$, where the linear term comes from applying the quantum Fourier transform. The algorithm succeeds with a probability $1-\epsilon$. In these cases, the error equals the value of the truncated bits: $\Delta\widetilde{\theta}_{\text{success}} \leqslant \frac{\pi}{2^{b-1}}$. The algorithm fails with probability ϵ. In these cases, the error scales as $\Delta\widetilde{\theta}_{\text{fail}} = O(\frac{1}{2^b})$. The probability of success is increased by increasing N_{tot} and N_{max}, such that $N_{\text{tot}} \propto N_{\text{max}} \propto \frac{1}{\epsilon}$. Combining the individual errors into means [Eq. (7)] gives the Heisenberg scaling: $\Delta\widetilde{\theta}_{\text{MAE/RMS}} = O(\frac{1}{N_{\text{tot}}})$ [33]. Despite the optimal scaling of the QPEA, the constant factor before the scaling is large. See Fig. 3. Thus the QPEA cannot be implemented, in general, with shallow circuits.

When using the QPEA in the point-identification step of our two-step protocol, we choose $b = m+1$. Consequently, $\hat{U}(\theta)$ is applied $N_{\text{PI}} = 2^{m+1+\lceil \log_2(2+1/2\epsilon)\rceil} - 1$ times in the point-identification step. We then use Eq. (6), Eq. (7), and $\Delta\widetilde{\theta}_{\text{fail}} = O(\frac{1}{2^b})$ to find that $N_{\text{tot}} = O(2^m)$. Consequently,

$$\Delta\widetilde{\theta}_{\text{MAE}} = O\left(\frac{1}{N_{\text{tot}}}\right), \quad \Delta\widetilde{\theta}_{\text{RMS}} = O\left(\frac{1}{N_{\text{tot}}}\right). \tag{16}$$

The mean-error scaling still follows the Heisenberg limit, but the constant before the scaling is smaller than for the QPEA alone. See Fig. 3. The largest circuit depth exceeds $N_{\text{max}} = 2^{m+\lceil \log_2(2+1/2\epsilon)\rceil} = O(N_{\text{tot}})$ in the asymptotic limit. Circuits are thus deeper than the iterative techniques described above and require many-probe entanglement.

IV. SIMULATIONS

In order to compare the performance of our two-step protocol with previous protocols, we provide numerical simulations. These are presented in Fig. 3. When producing the plots in Fig. 3, we optimized m and ϵ to find the supremum of $\Delta\widetilde{\theta}_{\text{RMS}}$ and $\Delta\widetilde{\theta}_{\text{MAE}}$ as a function of N_{tot}. Mean-error bounds are calculated from the worst-case value of θ for a given m and N_{tot}. We numerically study the following protocols: the iterative point identification alone, the QPEA alone, the iterative point identification with fine-tuning, and the QPEA with fine-tuning. The standard quantum limit (SQL) and Heisenberg limit (HL) have also been plotted. We also provide plots

of these suprema against their corresponding maximum circuit depth $N_{\max}$. We have set $\alpha = 32$ to facilitate comparisons with previous work [8].

Figures 3(a) and 3(b) plot the suprema of $\Delta\widetilde{\theta}_{\mathrm{RMS}}$ for $N_{\mathrm{tot}} \in [10^1, 10^6]$. For small N_{tot}, the two-part protocols overlap with the SQL because $m = 0$ is optimal, i.e., no point identification is required. The iterative two-part protocol's RMSE diverges from the SQL at $N_{\mathrm{tot}} \approx 900$, requiring $N_{\max} = 2$. The QPE-algorithm two-part protocol RMSE diverges from the SQL at $N_{\mathrm{tot}} \approx 2200$, when $N_{\max} = 1024$. Contrast this with the QPEA alone, which has a larger RMSE than the SQL for $N_{\mathrm{tot}} < 8191$ and $N_{\max} < 4096$. The performance of the two protocols that use the QPEA suffers from a higher RMSE due to the high resource requirements to reduce the failure probability ϵ.

Figures 3(c) and 3(d) plot the suprema of $\Delta\widetilde{\theta}_{\mathrm{MAE}}$ for $N_{\mathrm{tot}} \in [10^1, 10^6]$. Again, for small N_{tot}, the two-part protocols overlap with the SQL. The QPEA two-part protocol's MAE now diverges from the SQL before the iterative two-part protocol's MAE. These divergences occur at the values $N_{\mathrm{tot}} \approx 680$ with $N_{\max} = 256$ for the QPEA two-part protocol, and $N_{\mathrm{tot}} \approx 900$ with $N_{\max} = 2$ for the iterative two-part protocol. The QPEA alone requires $N_{\mathrm{tot}} > 4095$ and $N_{\max} \geqslant 2048$ to obtain an MAE bound below the SQL.

Our simulations demonstrate that the iterative two-part protocol, despite having worse asymptotic RMSE and MAE scalings than the QPEA (or the two-part QPEA protocol introduced here) achieves a lower supremum for smaller values of N_{tot}. Furthermore, this iterative two-part protocol requires significantly shallower circuits ($N_{\max}$) than the QPEA.

V. CONCLUSION

We have proposed a two-step phase-estimation protocol. In the first step, our protocol produces several contending precise estimates of an unknown phase. It does so by sampling from a circuit with many applications of the unknown phase. Then, the protocol point-identifies which estimate is in the correct parameter regime by independently sampling multiple circuits, each of which doubles in depth. For a given total number of applications of the unitary operation, N_{tot}, our protocol achieves lower upper bounds on the mean absolute error (MAE) and the root-mean-square error (RMSE) than previous iterative protocols. Asymptotically, our protocol's MAE scales as $O(\sqrt{\log(\log N_{\mathrm{tot}})}/N_{\mathrm{tot}})$, which is to be compared with a previously published, best iterative scaling of $O(\log N_{\mathrm{tot}}/N_{\mathrm{tot}})$ [8,9]. The asymptotic scaling of the RMSE of our protocol is $O(\sqrt{\log N_{\mathrm{tot}}}/N_{\mathrm{tot}})$, which outperforms previous iterative methods. Furthermore, when compared with the QPEA, our protocol's circuits are shallower, independent of the failure probability, and they do not require interprobe entanglement. Our protocol also achieves lower MAE and RMSE bounds than the QPEA for currently realistic values of N_{tot}, despite having a worse asymptotic scaling. The achievement of low phase-estimation errors with shallow circuits suggests that our protocol is more practical to implement in hardware-limited situations, such as noisy intermediate-scale quantum hardware. The effect of noise present on the performance of our protocol is left for future work.

The authors thank A. Gottfries, C. K. Long, N. Mertig, and W. Salmon for enlightening discussions. The authors also acknowledge support from Hitachi, Lars Hierta's Memorial Foundation, and Girton College, Cambridge.

[1] P. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (IEEE, New York, 1994), pp. 124–134.
[2] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Contemp. Math. **305**, 53 (2002).
[3] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
[4] K. Plekhanov, M. Rosenkranz, M. Fiorentini, and M. Lubasch, Quantum **6**, 670 (2022).
[5] S. Lloyd, S. Bosch, G. De Palma, B. Kiani, Z.-W. Liu, M. Marvian, P. Rebentrost, and D. M. Arvidsson-Shukur, arXiv:2006.00841.
[6] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Mol. Phys. **109**, 735 (2011).
[7] B. P. Abbott, Phys. Rev. Lett. **116**, 061102 (2016).
[8] T. Rudolph and L. Grover, Phys. Rev. Lett. **91**, 217905 (2003).
[9] M. de Burgh and S. D. Bartlett, Phys. Rev. A **72**, 042301 (2005).
[10] D. Oblak, P. G. Petrov, C. L. G. Alzar, W. Tittel, A. K. Vershovski, J. K. Mikkelsen, J. L. Sorensen, and E. S. Polzik, Phys. Rev. A **71**, 043807 (2005).
[11] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).
[12] H. Lee, P. Kok, and J. P. Dowling, J. Mod. Opt. **49**, 2325 (2002).
[13] S. L. Braunstein, Phys. Rev. Lett. **69**, 3598 (1992).
[14] L. Pezzé and A. Smerzi, Phys. Rev. Lett. **102**, 100401 (2009).
[15] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. London A **454**, 339 (1998).
[16] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[17] E. G. Brown, O. Goktas, and W. Tham, arXiv:2006.14145.
[18] F. Chapeau-Blondeau and E. Belin, Ann. Telecommun. **75**, 641 (2020).
[19] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[20] A. Lewbel, J. Econ. Lit. **57**, 835 (2019).
[21] C. Bontemps and T. Magnac, Annu. Rev. Econ. **9**, 103 (2017).
[22] E. Tamer, Annu. Rev. Econ. **2**, 167 (2010).
[23] M. Dobšíček, G. Johansson, V. Shumeiko, and G. Wendin, Phys. Rev. A **76**, 030306(R) (2007).
[24] M. H. Stone, Ann. Math. **33**, 643 (1932).
[25] Unitaries where $\hat{A}$ has explicit θ dependence can often be recast as $e^{i\theta'\hat{A}'}$ such that $\hat{A}'$ has no θ dependence.
[26] J. Rice, *Mathematical Statistics and Data Analysis* (W. Ross MacDonald School Resource Services Library, Brantford, ON, Canada, 2015).

[27] A. Ly, M. Marsman, J. Verhagen, R. P. Grasman, and E.-J. Wagenmakers, J. Math. Psychol. **80**, 40 (2017).
[28] L. Maccone, Phys. Rev. A **88**, 042109 (2013).
[29] C. J. O'Loan, J. Phys. A: Math. Theor. **43**, 015301 (2010)..
[30] Z. Ji, G. Wang, R. Duan, Y. Feng, and M. Ying, IEEE Trans. Inf. Theory **54**, 5172 (2008).
[31] F. Belliardo and V. Giovannetti, Phys. Rev. A **102**, 042613 (2020).
[32] A. Y. Kitaev, arXiv:quant-ph/9511026.
[33] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde, New J. Phys. **11**, 073023 (2009).