

Risk-minimizing states for the quantum-phase-estimation algorithm

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 16 May 2025; accepted 23 October 2025; published 14 November 2025)

The quantum-phase-estimation algorithm (QPEA) is widely used to find estimates of unknown phases. The original algorithm relied on an input state in a uniform superposition of all possible bit strings. However, it is known that other input states can reduce certain Bayesian risks of the final estimate. Here, we derive a method to find the risk-minimizing input state for any risk. These states are represented by an eigenvector of a Toeplitz matrix with elements given by the Fourier coefficients of the loss function of interest. We show that, while the true optimal state does not have a closed form for a general loss function, it is well approximated by a state with a cosine form. When the cosine frequency is chosen appropriately, these states outperform the original QPEA and achieve the optimal theoretical quantum-advantage scaling for three common risks. Furthermore, we prove that the uniform input state is suboptimal for any reasonable loss function. Finally, we design methods to mitigate the impact of depolarizing noise on the performance of QPEA.

DOI: [10.1103/8n2r-mxlc](https://doi.org/10.1103/8n2r-mxlc)

I. INTRODUCTION

Quantum phase estimation is an important subroutine for quantum technologies. It is used in several contenders for nonclassical advantage. Examples of such contenders are quantum algorithms [1–5], quantum chemistry [6–8], and interferometry [9–11]. A vital component in these examples is the estimation of an unknown phase θ applied via a unitary evolution $U(\theta)$. Due to the random nature of quantum measurements, estimates of θ and $\hat{\theta}$ will have an intrinsic variance and hence an error. The goal of a phase-estimation algorithm is to minimize the error of $\hat{\theta}$ with respect to some available resource, typically how many times $U(\theta)$ is applied across the algorithm. The use of entanglement or coherence can achieve a theoretical quadratic improvement of estimator variance [12,13]. In particular, for N applications of the unknown unitary $U(\theta)$, a classically shot-noise-limited estimate's variance scales as $1/N$, while a nonclassically Heisenberg-limited estimate's variance scales as $1/N^2$. However, few algorithms achieve this quadratic advantage due to issues with a lack of point identification (the output quantum state or measurement data could arise from multiple distinct values of the unknown parameter θ) [14] or decoherence in the quantum system [15,16].

A point-identified estimate of θ can be found by measuring several different circuits and using classical postprocessing

on the observed outcomes, either relying on Bayesian inference or a classical discrete Fourier transform [17–21]. Thus, one can rule out all-but-one of the potential estimates of θ . Alternatively, one can achieve point identification via a measurement of a single quantum circuit that uses the inverse quantum Fourier transform, implemented either using two-qubit gates [22] or feedback-based methods [23,24]. The resulting, fully quantum algorithm is often known as the quantum-phase-estimation algorithm (QPEA) [25–27]. The QPEA is one of the most used subroutines in the field of quantum algorithms. The traditional QPEA circuit is initialized with a register of m qubits in an equal mix of the computational basis states. The error of the obtained estimate is often quoted as a Heisenberg-limited error because the QPEA gives a correct binary representation of $\theta/2\pi$ with probability greater than 40% [22]. However, this error is deceiving, as one cannot know if the algorithm has succeeded without first knowing the true value of θ . From a complexity-theoretic perspective, meaningful errors should consider the errors from all outcomes. An example of such an error measure is the Bayes risk. The Bayes risk is the average of the expected posterior loss over all possible measurement outcomes with respect to some loss function. Previous works have shown that the QPEA does not achieve the quadratic (Heisenberg-limited) improvement for risks calculated using the Holevo variance [18,28]. However, a quantum advantage can be reinstated by initializing the register of the same circuit in a different state than the initially suggested uniform-superposition state [18,29,30].

In this article, we expand on previous works by constructing a methodology to find the state that the QPEA register should be prepared in to minimize a general risk. We find that the optimal state can be represented by the eigenvector of a Toeplitz matrix with elements determined by the Fourier coefficients of the loss function used to calculate the risk. We use this fact to prove that the traditional QPEA state

*Contact author: jgs46@cam.ac.uk

†Contact author: chwb101@cam.ac.uk

‡Contact author: drma2@cam.ac.uk

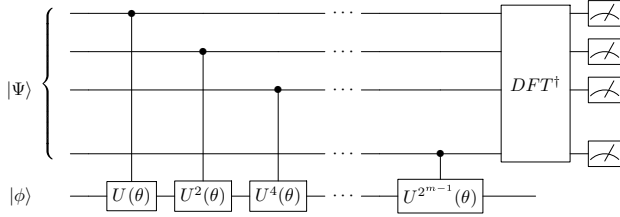


FIG. 1. The QPEA circuit.

only minimizes risks calculated using a trivial constant loss functions, a situation where every measurement process, even a random guess of θ , has the same risk. For general loss functions, we show that, while the optimal state does not have a closed form, such states are well approximated by a state whose amplitudes have a cosine form over the computational basis. By tuning the frequency of the cosine envelope, one can decrease the risk to values below that of the traditional QPEA. For risks calculated using the Holevo variance, our approximate states give results akin to previously derived optima [29]. We further demonstrate that the cosine states can obtain a quadratic improvement of three other common loss functions when compared to the shot-noise limits of those risks. The traditional QPEA cannot. Finally, we show that, even in the presence of depolarizing noise, risks better than the shot-noise limit can be obtained by repetitive measurements of circuits initialized with cosine states.

II. OPTIMAL QPEA REGISTER STATES

In quantum phase estimation one estimates an unknown phase θ of a unitary $U(\theta)$. To do so, one uses an eigenstate $|\phi\rangle$ such that $U(\theta)|\phi\rangle = e^{i\theta}|\phi\rangle$. For this state, the phases $\theta + 2\pi l$ for integer l are equivalent, so unique unitaries will have $\theta \in \Theta = [\theta_0 - \pi, \theta_0 + \pi]$. Here, θ_0 is the center of the interval Θ , commonly set to either 0 or π . The QPEA circuit is shown in Fig. 1. The input state is spread over two registers. The first register consists of m qubits initially prepared in the state

$$|\Psi\rangle = \sum_{j=0}^N c_j |j\rangle, \quad (1)$$

where $N = 2^m - 1$ and $c_j \in \mathbb{R}$. The second register is prepared in the state $|\phi\rangle$. The i th qubit of the first register is used as the control of 2^{i-1} controlled- $U(\theta)$ gates with the second register as the target. In total, N controlled- $U(\theta)$ gates are applied in the circuit. Throughout this work, we consider the total number of $U(\theta)$ gates used as the computational resource. A final measurement of the first register yields a bitwise estimate of $\theta/2\pi$.

In this work, we also consider the effect of applying some noise to the quantum state throughout the circuit. We adopt a simple and analytically tractable noise model in which the first register is subject to a depolarizing channel of strength λ after each $U(\theta)$ gate. The final state of the first register at the time of measurement with this noise is

$$\rho = (1 - \lambda)^N |\Phi\rangle\langle\Phi| + \frac{1 - (1 - \lambda)^N}{2^m} I, \quad (2)$$

where

$$|\Phi\rangle = \frac{1}{\sqrt{2^m}} \sum_{j=0}^N \sum_{k=0}^N c_k e^{ik(\theta - \frac{2\pi j}{2^m})} |j\rangle. \quad (3)$$

The probability that the first register is measured in the state $|y\rangle$, i.e., the likelihood function, is

$$p(y|\theta) = \frac{1}{2^m} + \frac{(1 - \lambda)^N}{2^m} \sum_{\substack{k=-N \\ k \neq 0}}^N \sum_{j=0}^N c_j c_{j-k} e^{ik(\theta - \frac{2\pi y}{2^m})}. \quad (4)$$

After the first register is measured in the state $|y\rangle$, a posterior distribution $p(\theta|y)$ can be generated using Bayes' rule:

$$p(\theta|y) = \pi(\theta)p(y|\theta)/p(y). \quad (5)$$

Here, $\pi(\theta)$ is the prior distribution and

$$p(y) = \int_{\Theta} \pi(\theta)p(y|\theta)d\theta. \quad (6)$$

The prior distribution of most interest is the uniform prior, $\pi(\theta) = 1/2\pi$, which represents a complete lack of initial knowledge of the value of θ before any measurement is performed. For a uniform prior, $p(y) = 1/2^m$ for any initial state $|\Psi\rangle$. $p(\theta|y)$ allows one to construct an expected value of a loss function for any estimate $\hat{\theta}'$:

$$\mathcal{L}(y, \hat{\theta}') = \int_{\Theta} p(\theta|y)L(\hat{\theta}', \theta)d\theta, \quad (7)$$

where $L(\hat{\theta}', \theta)$ is a loss function. The estimate that we make is given by the value of $\hat{\theta}'$ that minimizes the expected loss:

$$\hat{\theta} = \operatorname{argmin}_{\hat{\theta}'} \mathcal{L}(y, \hat{\theta}'). \quad (8)$$

The expected value of $\mathcal{L}(y, \hat{\theta})$ for a given register state $|\Psi\rangle$ over all outcomes y is known as the risk:

$$\mathcal{R}(\Psi) = \sum_y p(y)\mathcal{L}(y, \hat{\theta}) = \sum_y \int_{\Theta} \pi(\theta)p(y|\theta)L(\hat{\theta}, \theta)d\theta. \quad (9)$$

Our work aims to establish a methodology that, for a given loss function and amount of resources, allows one to find the optimal state $|\Psi'\rangle$ that minimizes $\mathcal{R}(\Psi)$ over all initial-register states $|\Psi\rangle$.

Loss functions of particular interest are of the form $L(\hat{\theta}, \theta) = L(\delta)$, where $\delta = \theta - \hat{\theta}$ is the distance between θ and $\hat{\theta}$. In this paper, we restrict our analysis to loss functions that are nondecreasing with increasing $|\delta|$ and satisfy $L(\delta) = L(-\delta)$. In this case, $\mathcal{L}(y, \hat{\theta})$ measures the average distance between $\hat{\theta}$ and possible true values of θ . The four loss functions we study are summarized in Table I. While three of the loss functions are commonly studied, the 1-0 loss function is not. The 1-0 loss is 1 if $|\delta| > \epsilon$ and 0 otherwise, where $\epsilon > 0$. In this case, $\mathcal{R}(\Psi) = \Pr[|\hat{\theta} - \theta| > \epsilon]$. Due to equivalences of the phases $\theta + 2\pi l$ for integer l , every loss function should be periodic with period 2π . A nonperiodic loss function $L_{\text{NP}}(\delta)$ of the above form can be converted into a periodic function $L(\delta)$ by calculating

$$L(\delta) = \min_{l \in \mathbb{Z}} \{L_{\text{NP}}(\delta + 2\pi l)\}. \quad (10)$$

TABLE I. The Fourier coefficients of different loss functions, where $\delta = \theta - \hat{\theta}$.

Loss	$L(\delta)$	Nonzero Fourier coefficients
Absolute	$L(\delta) = \delta $	$L_k = \begin{cases} \pi/2, & k = 0 \\ -2/\pi k^2, & k = \text{odd} \end{cases}$
Squared	$L(\delta) = \delta^2$	$L_k = \begin{cases} \pi^2/3, & k = 0 \\ 2(-1)^k/k^2, & k \neq 0 \end{cases}$
Holevo	$L(\delta) = 4 \sin^2(\delta/2)$	$L_k = \begin{cases} 2, & k = 0 \\ -1, & k = 1 \end{cases}$
1-0	$L(\delta) = \begin{cases} 1, & \delta > \epsilon \\ 0, & \delta \leq \epsilon \end{cases}$	$L_k = \begin{cases} 1 - \epsilon/\pi, & k = 0 \\ -\sin(k\epsilon)/\pi k, & k \neq 0 \end{cases}$

If $\theta_0 = \hat{\theta}$ in Θ , $\delta \in [-\pi, \pi]$. One finds that $L(\delta) = L_{\text{NP}}(\delta)$, so this conversion can be ignored. Periodicity allows the loss function to be expanded as the Fourier series $L(\delta) = \sum_k L_k e^{ik\delta}$, where

$$L_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} L(\delta) e^{-ik\delta} d\delta. \quad (11)$$

The Fourier coefficients satisfy both $L_{-k} = L_k^*$ for real loss functions and $L_k \in \mathbb{R}$ for even loss functions.

In terms of these Fourier coefficients, Eq. (7) becomes

$$\mathcal{L}(y, \hat{\theta}') = L_0 + (1 - \lambda)^N \sum_{k=1}^N \sum_{j=k}^N c_j c_{j-k} L_k \cos k\vartheta, \quad (12)$$

where $\vartheta = \hat{\theta}' - 2\pi y/2^m$. This loss is minimized by the estimate $\hat{\theta} = 2\pi y/2^m$:

$$\mathcal{L}(y, \hat{\theta}) = L_0 + (1 - \lambda)^N \sum_{k=1}^N \sum_{j=k}^N c_j c_{j-k} L_k. \quad (13)$$

$\mathcal{L}(y, \hat{\theta})$ is independent of the observed value of y . Equation (9) shows that the risk of the QPEA with a register prepared in the state $|\Psi\rangle$ is

$$\mathcal{R}(\Psi) = L_0 + (1 - \lambda)^N \sum_{k=1}^N \sum_{j=k}^N c_j c_{j-k} L_k. \quad (14)$$

If depolarizing noise is present in each $U(\theta)$ gate ($\lambda > 0$), $\mathcal{R}(\Psi) \rightarrow L_0$ for large N for any initial-register state. Therefore, the risk of estimates made using the QPEA with a register prepared in any state tends to a constant value: $\mathcal{R}(\Psi) = O(1)$.

The risk-minimizing state can be found by rewriting Eq. (14) into the matrix equation

$$\mathcal{R}(\Psi) = L_0 + (1 - \lambda)^N \mathbf{c}^\dagger \mathbf{R} \mathbf{c}. \quad (15)$$

Here, $\mathbf{c}$ is a unit vector of the values c_j and

$$\mathbf{R} = \begin{pmatrix} 0 & L_{-1} & L_{-2} & \cdots & L_{-N} \\ L_1 & 0 & L_{-1} & \cdots & L_{1-N} \\ L_2 & L_1 & 0 & \cdots & L_{2-N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_N & L_{N-1} & L_{N-2} & \cdots & 0 \end{pmatrix}. \quad (16)$$

For even loss functions, $\mathbf{R}$ is a symmetric banded Toeplitz matrix. The optimal register state $|\Psi'\rangle$ with a minimal risk of $\mathcal{R}(\Psi')$ is represented by the column vector $\mathbf{c}$ that is an eigenvector of $\mathbf{R}$ with the smallest eigenvalue. $|\Psi'\rangle$ is independent of λ .

When risk is determined through the Holevo variance $L(\delta) = 4 \sin^2(\delta/2)$, $\mathbf{R}$ is a tridiagonal matrix with the minimum eigenvector described by

$$c_i(\omega) = \sqrt{\frac{2}{N+2}} \cos \left[\left(\frac{N}{2} - i \right) \frac{\pi}{N+2} \right]. \quad (17)$$

This result is in agreement with the state proposed in Ref. [29]. The minimal Holevo risk with this state is

$$\mathcal{R}(\Psi') = 2 - 2(1 - \lambda)^N \cos \left(\frac{\pi}{N+2} \right). \quad (18)$$

Unfortunately, no closed-form expression exists for the minimum-eigenvalue eigenvector of general banded Toeplitz matrices [31,32]. Therefore, numerical methods are needed to find the optimal state for other risks, such as the absolute, squared, and 1-0 risks.

One can use the above analysis to solve an alternative task: Find the initial state $|\Psi\rangle$ required to prepare the m -qubit output register in the state $|[2^m \theta / 2\pi]\rangle$ with the highest probability. This output state is an m -bit binary representation of $\theta/2\pi$. The preparation of $|[2^m \theta / 2\pi]\rangle$ is integral to many quantum algorithms, such as Shor's algorithm [4] and the HHL algorithm [33], and is explored in detail in Ref. [34]. The task is equivalent to finding the state that minimizes the probability that $|\hat{\theta} - \theta| \geq \pi/2^m$ averaged over all θ , or, using the terminology above, the risk calculated using the 1-0 loss function with $\epsilon = \pi/2^m$. Therefore, the optimal state for the desired preparation is the eigenstate with the minimal eigenvalue of the matrix in Eq. (16) with $L_k = \sin(k\pi/2^m)/\pi k$. This agrees with the results in Ref. [34].

III. COSINE STATES

Instead of calculating the optimal state numerically and finding a state that may be difficult to prepare in practice, we find an approximation to these optimal states. The symmetry of the rows of $\mathbf{R}$ impose two conditions on the minimum eigenvector: $c_i = c_{N-i}$ and $c_0 < c_1 < \cdots < c_{N/2}$. These conditions are satisfied by the approximation

$$c_i(\omega) = \sqrt{\frac{2 \sin \omega}{2^m \sin \omega + \sin(2^m \omega)}} \cos \left[\left(\frac{N}{2} - i \right) \omega \right]. \quad (19)$$

We refer to the initial state with these coefficients as the cosine state with frequency ω due to the similarity to cosine window states [30]. Notice that the state in Eq. (17) is a cosine state with $\omega = \pi/(N+2)$. When using a cosine state, the risk of $\hat{\theta} = 2\pi y/2^m$ is

$$\mathcal{R}(\omega) = L_0 + 2(1 - \lambda)^N \sum_{k=1}^N L_k f_k, \quad (20)$$

where

$$f_k = \frac{(2^m - k) \cos k\omega \sin \omega + \sin[(2^m - k)\omega]}{2^m \sin \omega + \sin(2^m \omega)}. \quad (21)$$

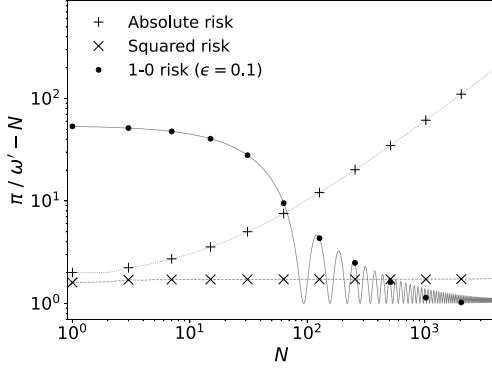


FIG. 2. Numerically calculated $\pi/\omega' - N$ for different values of N . The y axis function is chosen to aid visualization. The markers plot data for integer values of m with the lines plotting calculated data for noninteger values of m . For large N , linear regression shows that ω' scales as $\frac{\pi}{1.04N+22.27}$, $\frac{\pi}{N+1.72}$, and $\frac{\pi}{N+1.00}$ for the absolute, squared, and 1-0 loss functions, respectively.

The optimal cosine state has a frequency of $\omega' = \text{argmin}_{\omega} \mathcal{R}(\omega)$ and a risk of $\mathcal{R}(\omega')$. We show numerically calculated values of ω' for the absolute, squared, and 1-0 loss functions in Fig. 2.

Table II summarizes the noise-free ($\lambda = 0$) risk scaling of the absolute, squared, Holevo, and 1-0 losses using the optimal cosine states in the large- N limit. These risks are plotted for finite N in Fig. 3, alongside $\mathcal{R}(\Psi')$. $\mathcal{R}(\omega') \approx \mathcal{R}(\Psi')$ for the absolute, squared, and Holevo risks. $\mathcal{R}(\omega') \approx \mathcal{R}(\Psi')$ is also seen for the 1-0 risk when $m < 7$, although $\mathcal{R}(\omega') > \mathcal{R}(\Psi')$ when $m \geq 7$. This divergence is due to the cosine approximation of $|\Psi'\rangle$ breaking down because $|\Psi'\rangle$ takes roughly a normal-distribution-like form over the computational basis states. The corresponding risks of systems with differing values of λ are plotted in Fig. 4. One observes that when N increases for systems with $\lambda > 0$, both $\mathcal{R}(\omega')$ and $\mathcal{R}(\Psi')$ are approximately equal to the noiseless curves when N is small, decrease until they reach a single minimum before increasing, and tend to L_0 when $N \rightarrow \infty$. The value of this minimum risk decreases, and the value of N that obtains this minimum risk increases as λ decreases.

The above risks can be compared to the theoretically optimal scalings achieved by Fisher-information-maximizing measurements of noiseless circuits [12,13]. These

TABLE II. The large- N scalings of different risks in noiseless systems.

Loss	$\mathcal{R}_C$	$\mathcal{R}_Q$	$\mathcal{R}(u)$	$\mathcal{R}(\omega')$
Absolute loss	$O(\frac{1}{\sqrt{N}})$	$O(\frac{1}{N})$	$O(\frac{\ln N}{N})$	$O(\frac{1}{N})$
Squared loss	$O(\frac{1}{N})$	$O(\frac{1}{N^2})$	$O(\frac{1}{N})$	$O(\frac{1}{N^2})$
Holevo variance	$O(\frac{1}{N})$	$O(\frac{1}{N^2})$	$O(\frac{1}{N})$	$O(\frac{1}{N^2})$
1-0 loss	$O(e^{-N\epsilon/2})$	$O(e^{-N^2\epsilon/2})$	$O(\frac{1}{N\epsilon})$	$O(\frac{1}{N^3\epsilon})$

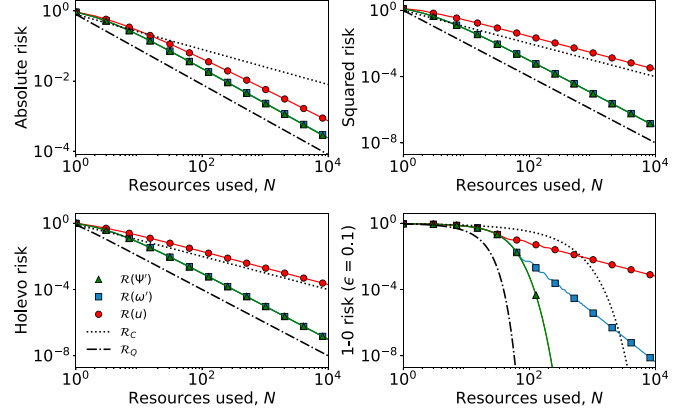


FIG. 3. The four common risks from measurements of a noiseless QPEA circuit with a register prepared in different states. $\mathcal{R}_C$ and $\mathcal{R}_Q$ are computed using $\sigma^2 = 1/N$ and $\sigma^2 = 1/N^2$, respectively.

measurements consist of a number ν measurements of a single circuit that apply n copies of $U(\theta)$ to a single qubit coherently. Estimates of θ achieved by these measurements tend to a normal distribution with mean θ and variance $\sigma^2 = 1/n^2\nu$ for large ν [17]. As a function of the total resource use $N = n \times \nu$, one finds that $\sigma^2 = O(1/N)$ if $\nu \propto N$ and n is constant. This is the shot-noise-limited variance seen in classical statistical inference [12]. If $n \propto N$ and ν is constant instead, $\sigma^2 = O(1/N^2)$. This is the Heisenberg limit, a quadratic quantum advantage [12]. Difficulties with point identification prevent the Heisenberg limit from being obtained exactly in practice [14], so this limit is often used as a theoretical best-case-scaling limit only. These variance limits can be used to calculate equivalent limits in risks by taking an expected value of a loss function over the corresponding normal distribution:

$$\mathcal{R} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\delta^2/2\sigma^2} L(\delta) d\delta. \quad (22)$$

The asymptotic risk scalings calculated using a shot-noise scaling σ^2 , $\mathcal{R}_C$, and a Heisenberg limited σ^2 , $\mathcal{R}_Q$, are presented in Table II. For risks calculated using the absolute,

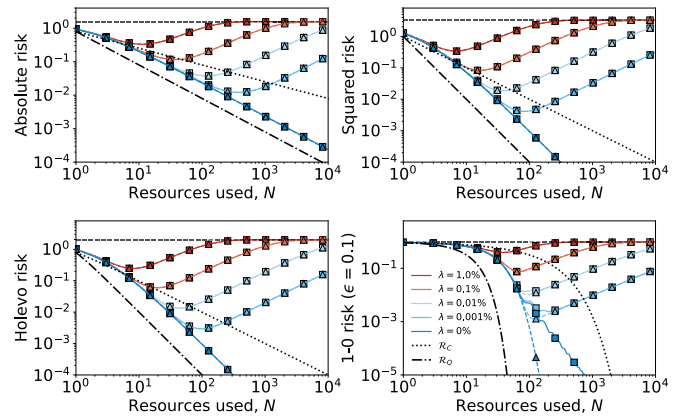


FIG. 4. The four common risks from measurements of the QPEA circuit with different values of λ . $\mathcal{R}_C$ and $\mathcal{R}_Q$ are computed using $\sigma^2 = 1/N$ and $\sigma^2 = 1/N^2$, respectively. Triangles joined by dashed lines show $\mathcal{R}(\Psi')$. Squares joined by solid lines show $\mathcal{R}(\omega')$.

squared, and Holevo loss functions, the scalings of $\mathcal{R}(\omega')$ and $\mathcal{R}_Q$ are equal when depolarization is absent, indicating quantum advantage. For risks calculated using the 1-0 loss function, $\mathcal{R}(\omega')$ scales worse than $\mathcal{R}_C$. The same is true for all the risks when depolarizing noise is present. A quantum advantage is possible for finite values of N in systems prone to depolarizing noise. For systems with λ less than some critical value, there exists a register size such that both $\mathcal{R}(\omega')$ and $\mathcal{R}(\Psi')$ are smaller than the corresponding shot-noise risk $\mathcal{R}_C$. For the values of $\mathcal{R}_C$ plotted in Fig. 4, numerical calculations give these critical values as $\lambda = 0.31\%$, $\lambda = 0.06\%$, $\lambda = 0.13\%$, and $\lambda = 3.26\%$ for risks calculated using the absolute, squared, Holevo, and 1-0 losses, respectively.

A second comparison can be made to measurements with the uniform-superposition state used by the traditional QPEA: $c_i = 2^{-m/2}$ [22,25]. The uniform state is equal to a cosine state with $\omega = 0$. The risk of estimates with this state is

$$\mathcal{R}(u) = L_0 + 2(1 - \lambda)^N \sum_{k=1}^N L_k \left(1 - \frac{k}{2^m}\right). \quad (23)$$

Because the minimization to find ω' is taken over all ω including $\omega = 0$, $\mathcal{R}(\omega') \leq \mathcal{R}(u)$. Therefore, estimates made with the optimal cosine state have an equal or smaller risk than estimates made with the uniform state. Evaluating Eq. (23) for noiseless circuits with the absolute, squared, Holevo, or 1-0 loss function gives

$$\begin{aligned} \mathcal{R}(u) &= \frac{\ln 4 + \psi^{(0)}\left(\frac{N+2}{2}\right) + \gamma_C}{\pi 2^{m-1}} + \frac{\psi^{(1)}\left(\frac{N+2}{2}\right)}{\pi}, \\ \mathcal{R}(u) &= \frac{\ln 2 + \Phi(2^m)}{2^{m-2}} + \psi^{(1)}\left(\frac{2^m + 1}{2}\right) - \psi^{(1)}(2^{m-1}), \\ \mathcal{R}(u) &= \frac{2}{2^m}, \\ \mathcal{R}(u) &= \frac{\sin\left(\frac{N\epsilon}{2}\right) \sin(2^{m-1}\epsilon)}{\pi 2^{m-1} \sin\left(\frac{\epsilon}{2}\right)} + 2 \sum_{k=0}^{\infty} \frac{\sin[(k + 2^m)\epsilon]}{\pi(k + 2^m)}, \end{aligned} \quad (24)$$

respectively. Here, $\psi^{(n)}(z)$ is the n th derivative of the digamma function, γ_C is Euler's constant, and $\Phi(a) = \int_{t=0}^{\infty} \frac{e^{-at}}{1+e^{-t}} dt$. For large N , these risks tend to the limits displayed in Table II. These four risks scale worse than $\mathcal{R}_Q$. Only the absolute risk scales better than $\mathcal{R}_C$. All of these risks tend to L_0 if $\lambda > 0$ for the reasons given above.

One may ask when the uniform state is optimal. This optimality occurs when the column vector of one is the minimum eigenvector of $\mathbf{R}$. When the register consists of one qubit ($m = 1$), this column vector is the minimum eigenvector when $L_1 \leq 0$, which is satisfied by any even nondecreasing loss function. For optimality at a given register size m , one requires that $L_j = L_{2^m-j}$ for a positive integer j . For this condition to hold for every m , all coefficients L_j for $j > 0$ must be equal. However, a finite loss function has $L_N \rightarrow 0$ for large N due to the decay of Fourier coefficients. Thus, $L_j = 0$ for all $j > 0$. Loss functions that satisfy this condition are of the form $L(\delta) = L_0$, a trivial constant. With this loss function, Eq. (9) suggests that $\mathcal{R}(\Psi) = L_0$ for any initial-register state. In fact, the risk is L_0 for any estimation procedure, even a random guess of θ . Therefore, one concludes that estimation with

the uniform state is optimal for any m only in pathological situations where the QPEA does not do better in determining θ than random guesses.

IV. MULTIPLE MEASUREMENTS OF A SINGLE CIRCUIT

The performance of the QPEA is limited by the fact that resources can only be added to the algorithm by increasing the size of the input register and thus circuit depth. We now consider a simple modification to the QPEA. In particular, we consider making multiple evaluations of the same QPEA circuit (with a fixed register size m) and then analyzing the results. This provides a method to increase resources without increasing the level of depolarization present in the algorithm. Consider carrying out a number, M , of measurements of the QPEA circuit in Fig. 1. In total, $M(2^m - 1)$ applications of $U(\theta)$ are used during this protocol. The likelihood of observing the outcome $\mathbf{y} = (y_1, y_2, \dots, y_M)$ is

$$p(\mathbf{y}|\theta) = \prod_{i=1}^M p(y_i|\theta). \quad (25)$$

This new likelihood function $p(\mathbf{y}|\theta)$ can be substituted for $p(y|\theta)$ in Eqs. (5)–(9) to find the measurement risk. These calculations could be used to derive the optimal number of qubits m and measurements M to use in the QPEA in order to minimize risk.

Figure 5 plots the Holevo and 1-0 risks as a function of the resources used for $1 \leq M \leq 5$ when the QPEA register is prepared in the optimal cosine state. We have calculated these risks numerically because no closed-form solution for them exists. These calculations are computationally intensive due to the 2^{Mm} possible measurement outcomes upon which the risk depends, so we only present data for tractable values of Mm . For noiseless circuits (top row), the Holevo risk scales as $\mathcal{R}_Q$ for any fixed M . The minimum risk per resource occurs when $M = 1$. Therefore, the QPEA circuit should be run once in this situation. The same is true for the absolute and squared risks in noiseless systems (not plotted here). For the 1-0 loss in noiseless circuits, better risk scaling is seen for larger M . Risk below $\mathcal{R}_C$ is achieved for $M \geq 3$. One concludes that the QPEA circuit with a register prepared in the optimal cosine state should be measured at least three times. When depolarizing noise is present (bottom row of Fig. 5), both risks tend to L_0 as register size increases for any M . The minimum risk obtained over all register sizes decreases as M increases. For large enough M , risks below the shot-noise value $\mathcal{R}_C$ can be achieved by fixing register size and M appropriately. Unfortunately, the exact optimal values that m and M should be fixed to in order to minimize risk cannot be derived from this small dataset. Future work with significant computational power may address this problem.

V. CONCLUSIONS

We have derived the optimal input state for the QPEA with respect to Bayesian risk functions. The state, we showed, is equivalent to the eigenvector of a Toeplitz matrix with the smallest eigenvalue. The elements of this matrix are the

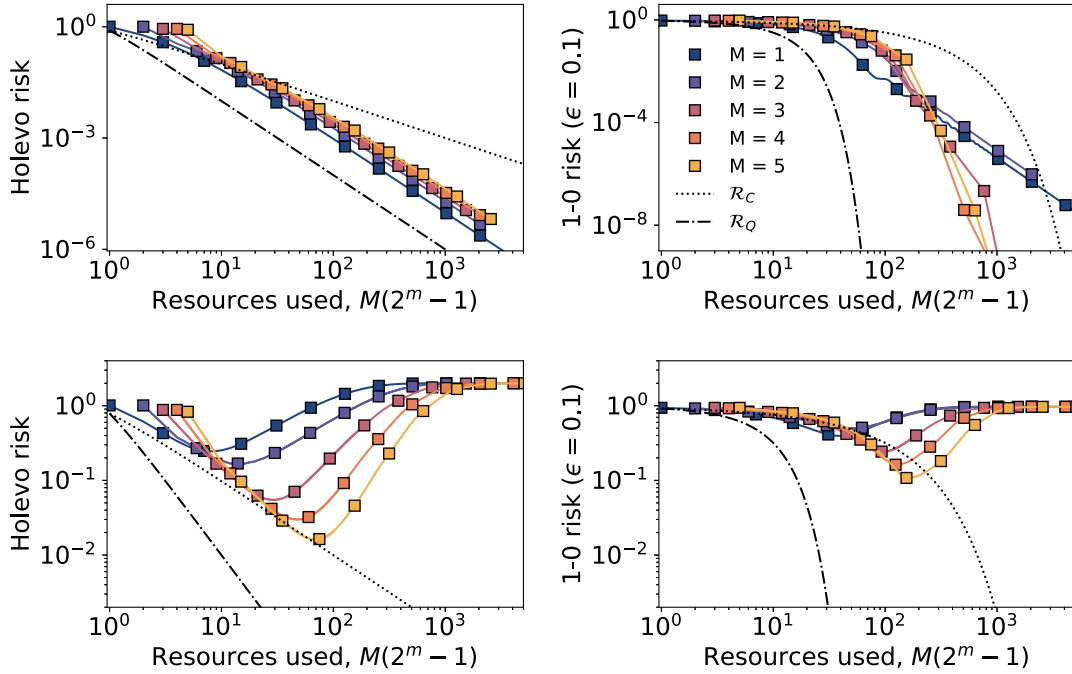


FIG. 5. The risk of estimates of θ made by M measurements of the QPEA with an m -qubit register prepared in the optimal cosine state against the total resource use $M(2^m - 1)$. The top row shows risks when using noiseless circuits, whilst the bottom row shows risks when using circuits with $\lambda = 1\%$.

Fourier coefficients of the target loss function that is used to calculate the risk of interest. Although a closed form of these states may not exist in general, we find that a good approximation exists in the form of a cosine state. Optimizing the frequency of the cosine state enables theoretically optimal noiseless risk scalings for three common risks. In addition, we demonstrate that the traditional QPEA state never minimizes risk for any nonconstant loss function. We have also shown that if any depolarizing noise exists, the risk will tend to a constant in the large-resource limit. This is regardless of the preparation of the register state. Nevertheless, we show that systems with a strength of depolarization less than some critical value can obtain quantum advantage for finite register sizes. Furthermore, the asymptotic risk scalings can be improved by repeating measurements of a circuit with a fixed register size m a number of times M . Data for small Mm shows

that performance can be improved beyond classical shot-noise risk limits.

Whilst this article has shown that suitable modifications to the QPEA can result in a quantum advantage, the requirement of entanglement between multiple qubits would limit implementation on current NISQ hardware. If entanglement between qubits is difficult, one may consider alternative approaches that use statistical inference techniques on measurements of single-qubit circuits, such as those outlined in Refs. [17,35,36]. A full comparison between the risks of inference-based approaches and the improved QPEA presented here is the subject of future work.

DATA AVAILABILITY

The data supporting this study's findings are available within the article.

-
- [1] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, *Contemp. Math.* **305**, 53 (2002).
 - [2] Y. Suzuki, S. Uno, R. Raymond, T. Tanaka, T. Onodera, and N. Yamamoto, *Quantum Inf. Process.* **19**, 75 (2020).
 - [3] S. Lloyd, S. Bosch, G. D. Palma, B. Kiani, Z.-W. Liu, M. Marvian, P. Rebentrost, and D. M. Arvidsson-Shukur, [arXiv:2006.00841](https://arxiv.org/abs/2006.00841).
 - [4] P. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, in *Proceedings of the 35th FOCS* (IEEE, New York, 1994), pp. 124–134.
 - [5] S. Lloyd, B. T. Kiani, D. R. M. Arvidsson-Shukur, S. Bosch, G. D. Palma, W. M. Kaminsky, Z.-W. Liu, and M. Marvian, [arXiv:2104.01410](https://arxiv.org/abs/2104.01410).
 - [6] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, *Chem. Rev.* **120**, 12685 (2020).
 - [7] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, *Mol. Phys.* **109**, 735 (2011).
 - [8] L. Lin and Y. Tong, *Phys. Rev. X Quantum* **3**, 010318 (2022).
 - [9] R. Demkowicz-Dobrzanski, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, *Phys. Rev. A* **80**, 013825 (2009).
 - [10] D. Oblak, P. G. Petrov, C. L. G. Alzar, W. Tittel, A. K. Vershovski, J. K. Mikkelsen, J. L. Sørensen, and E. S. Polzik, *Phys. Rev. A* **71**, 043807 (2005).
 - [11] B. C. Barish and R. Weiss, *Phys. Today* **52**(10), 44 (1999).

- [12] V. Giovannetti, S. Lloyd, and L. Maccone, *Nat. Photon.* **5**, 222 (2011).
- [13] V. Giovannetti, S. Lloyd, and L. Maccone, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [14] J. G. Smith, C. H. W. Barnes, and D. R. M. Arvidsson-Shukur, *Phys. Rev. A* **106**, 062615 (2022).
- [15] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, *Nat. Commun.* **3**, 1063 (2012).
- [16] F. Albarelli and R. Demkowicz-Dobrzański, *Phys. Rev. X* **12**, 011039 (2022).
- [17] J. G. Smith, C. H. Barnes, and D. R. M. Arvidsson-Shukur, *Phys. Rev. A* **109**, 042412 (2024).
- [18] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, *Phys. Rev. A* **80**, 052114 (2009).
- [19] K. M. Svore, M. Hastings, and M. Freedman, *Quantum Inf. Comput.* **14**, 306 (2014).
- [20] T. E. O'Brien, B. Tarasinski, and B. M. Terhal, *New J. Phys.* **21**, 023022 (2019).
- [21] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, *Phys. Rev. Lett.* **111**, 070403 (2013).
- [22] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University, Cambridge, England, 2000).
- [23] R. B. Griffiths and C.-S. Niu, *Phys. Rev. Lett.* **76**, 3228 (1996).
- [24] J. a. Chiaverini, J. Britton, D. Leibfried, E. Knill, M. D. Barrett, R. Blakestad, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, T. Schaetz, and D. J. Wineland, *Science* **308**, 997 (2005).
- [25] A. Y. Kitaev, [arXiv:quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026).
- [26] X. Lu and H. Lin, [arXiv:2210.00231](https://arxiv.org/abs/2210.00231).
- [27] F. Chapeau-Blondeau and E. Belin, *Ann. Telecommun.* **75**, 641 (2020).
- [28] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, *Nature (London)* **450**, 393 (2007).
- [29] W. van Dam, G. M. D'Ariano, A. Ekert, C. Macchiavello, and M. Mosca, *Phys. Rev. Lett.* **98**, 090501 (2007).
- [30] G. Rendon, T. Izubuchi, and Y. Kikuchi, *Phys. Rev. D* **106**, 034503 (2022).
- [31] S.-E. Ekström, C. Garoni, and S. Serra-Capizzano, *Exp. Math.* **27**, 478 (2018).
- [32] S.-E. Ekström and S. Serra-Capizzano, *Numer. Linear Algebra Appl.* **25**, e2137 (2018).
- [33] A. W. Harrow, A. Hassidim, and S. Lloyd, *Phys. Rev. Lett.* **103**, 150502 (2009).
- [34] D. Patel, S. J. S. Tan, Y. Subasi, and A. T. Sornborger, [arXiv:2403.18927](https://arxiv.org/abs/2403.18927).
- [35] 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).
- [36] S. Kimmel, G. H. Low, and T. J. Yoder, *Phys. Rev. A* **92**, 062315 (2015).