

Title	On the bias in iterative quantum amplitude estimation
Author(s)	Miyamoto, Koichi
Citation	EPJ Quantum Technology. 2024, 11(1), p. 42
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/97840">https://hdl.handle.net/11094/97840</a>
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University



# On the bias in iterative quantum amplitude estimation

Koichi Miyamoto<sup>1\*</sup>

\*Correspondence:  
[miyamoto.kouichi.qiqb@osaka-u.ac.jp](mailto:miyamoto.kouichi.qiqb@osaka-u.ac.jp)

<sup>1</sup>Center for Quantum Information and Quantum Biology, Osaka University, Toyonaka, Osaka 560-0043, Japan

## Abstract

Quantum amplitude estimation (QAE) is a pivotal quantum algorithm to estimate the squared amplitude  $a$  of the target basis state in a quantum state  $|\Phi\rangle$ . Various improvements on the original quantum phase estimation-based QAE have been proposed for resource reduction. One of such improved versions is iterative quantum amplitude estimation (IQAE), which outputs an estimate  $\hat{a}$  of  $a$  through the iterated rounds of the measurements on the quantum states like  $G^k|\Phi\rangle$ , with the number  $k$  of operations of the Grover operator  $G$  (the Grover number) and the shot number determined adaptively. This paper investigates the bias in IQAE. Through the numerical experiments to simulate IQAE, we reveal that the estimate by IQAE is biased and the bias is enhanced for some specific values of  $a$ . We see that the termination criterion in IQAE that the estimated accuracy of  $\hat{a}$  falls below the threshold is a source of the bias. Besides, we observe that  $k_{\text{fin}}$ , the Grover number in the final round, and  $f_{\text{fin}}$ , a quantity affecting the probability distribution of measurement outcomes in the final round, are the key factors to determine the bias, and the bias enhancement for specific values of  $a$  is due to the skewed distribution of  $(k_{\text{fin}}, f_{\text{fin}})$ . We also present a bias mitigation method: just re-executing the final round with the Grover number and the shot number fixed.

**Keywords:** Quantum algorithm; Quantum amplitude estimation; Statistical bias

## 1 Introduction

Among various quantum algorithms, *quantum amplitude estimation (QAE)* is one of the prominent ones. It is originally proposed in [1] as a method to estimate the squared amplitude  $a$  of the target basis state  $|\phi_1\rangle$  in a quantum state  $|\Phi\rangle$ . If we have the oracle  $A$  to generate  $|\Phi\rangle$  and the reflection operator  $S$  with respect to  $|\phi_1\rangle$ , we can obtain an  $\epsilon$ -approximation of  $a$ ,<sup>1</sup> querying  $A$  and  $S$   $O(1/\epsilon)$  times. More concretely, the method in [1] is based on *quantum phase estimation (QPE)* [2]: using  $A$  and  $S$ , we construct the *Grover operator*  $G$  (defined later), which acts on  $|\Phi\rangle$  to amplify the amplitude of  $|\phi_1\rangle$ , and then operate the controlled version of  $G$   $O(1/\epsilon)$  times followed by an inverse quantum Fourier transform, which yields an approximation of  $a$ .

One reason why QAE is important is that it is the basis of other quantum algorithms. For example, it is used in the quantum algorithm for Monte Carlo integration (QMCI) [3],

<sup>1</sup>For  $x \in \mathbb{R}$ , we say that  $y \in \mathbb{R}$  is an  $\epsilon$ -approximation of  $x$  if  $|y - x| \leq \epsilon$ .

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

which estimates the expectation of a random variable quadratically faster than the classical counterpart. Furthermore, QMCI has many applications in industry, e.g., derivative pricing [4–9] in finance.

Partly because of such practical importance, many improvements to the original version of QAE have been proposed so far. In particular, some methods *without* QPE have been devised [10–21]. Since the controlled  $G$  requires a larger gate cost than the uncontrolled one, replacing the former with the latter leads to gate cost reduction. The first algorithm in such a direction is QAE based on the *maximum likelihood estimation (MLE)* [10], which we hereafter call *MLEQAE*. In this method, we apply  $G$  to  $|\Phi\rangle$   $k$  times and make  $N$  measurements on  $G^k|\Phi\rangle$  by which we distinguish  $|\phi_1\rangle$  and other basis states, increasing  $k$  according to some given schedule. We hereafter call this  $k$ , the number of operations of  $G$ , the *Grover number*. The outcomes of the measurements, namely the numbers of times we get  $|\phi_1\rangle$  for the various  $k$ , give us the information on  $a_k = \sin^2((2k+1) \times \arcsin(\sqrt{a}))$ , the squared amplitude of  $|\phi_1\rangle$  in  $G^k|\Phi\rangle$ , and thus the information on  $a$ . Then, we use this to construct the likelihood function of  $a$  and obtain an estimate of  $a$  as the maximum likelihood point.

Afterward, the paper [13] proposed *iterative quantum amplitude estimation (IQAE)*, which this paper focuses on. Like MLEQAE, in IQAE we use the outcomes of the measurements on  $G^k|\Phi\rangle$  with varying  $k$ , but in a different way. Starting from  $|\Phi\rangle$ , which corresponds to  $k=0$ , we obtain the confidence interval (CI) of  $a$  by the MLE based on the outcomes of the measurements on  $|\Phi\rangle$ . We then choose the next  $k$  adaptively in the way explained later, and make the measurements on  $G^k|\Phi\rangle$ . Via MLE, this yields the CI on  $a_k$ , which is translated into the CI of  $a$  narrower than the previous one. We repeat these steps, each of which is called a *round*, increasing the Grover number and narrowing the CI, until the CI width reaches the required accuracy  $\epsilon$ . In addition to this adaptive increment of the Grover number, another feature of IQAE is that  $N$  the number of the measurements (the *shot number*) in one round is increased gradually: if the CI of  $a$  derived from measurements with current  $k$  is so narrow that we can determine the next  $k$ , we stop the current round and go to the next round with the next  $k$ , and otherwise, we add more measurements with the current  $k$ . The advantage of such an adaptive increment of  $k$  and  $N$  is that it can lead to saving the total number of queries to  $G$  compared to fixing  $k$  and  $N$  in advance.<sup>2</sup>

In this paper, we focus on the *bias* in IQAE, which can be an issue in some situations but has not been focused on in previous studies. Note that the estimate  $\hat{a}$  of  $a$  by IQAE is stochastic since it is derived from outcomes of measurements on quantum states, which are intrinsically random. Although IQAE guarantees that the magnitude of the error  $\hat{a} - a$  is below the tolerance with high probability, it might be biased, that is, the expectation of the error might not be zero:  $b := \mathbb{E}[\hat{a} - a] \neq 0$ , where  $\mathbb{E}[\cdot]$  is the expectation with respect to the randomness of the measurement outcomes. We hereafter call  $b$  the bias and the residual  $\hat{a} - a - b$  as the random part.

The motivation to focus on the bias in QAE, including IQAE, is the possibility that it may matter more than random part in some cases. That is, when we want some quantities given as a combination of many outputs of different QAE runs, biases can accumulate and, even if small in each output, become significant in total. In worst cases, biases in

<sup>2</sup>Also in the framework of MLEQAE, some recent studies proposed setting the shot number adaptively [17].

the sum of  $N$  estimates scales as  $O(N)$ , whereas according to the central limit theorem, the random parts cancel each other and scale as  $O(\sqrt{N})$ . An example of situations where we combine many QAE outputs is calculating the total value of a portfolio of derivative contracts, where each contract is priced by individual runs of QAE. Another example is calculating Gibbs partition functions in statistical mechanics [22]. In the quantum algorithm in [22], a partition function is expressed as the product of expectations of certain random variables, each of which is estimated by QAE.

As far as the author knows, there has been no study on bias in IQAE, while bias in other types of QAE has been studied so far [19, 21, 22]. Thus, in this paper, we investigate the bias in IQAE. We conduct numerical experiments to reveal the nature of the bias in IQAE. One of our key findings from the numerical experiments is that IQAE is in fact biased, and its *termination criterion* induces the bias. That is, the procedure that the algorithm ends when the CI width of  $a$  reaches the required accuracy leads to the bias. This is because the CI is statistically inferred and its width depends on the realized value of the estimate  $\hat{a}$  of  $a$ . The algorithm tends to end when  $\hat{a}$  accidentally takes a value that yields a narrow CI. This effect affects the expectation of  $\hat{a}$  and then induces the bias.

In particular, we observe that the bias is enhanced for some specific values of  $a$ . This phenomenon is explained by the distribution of  $(k_{\text{fin}}, f_{\text{fin}})$ . Here,  $k_{\text{fin}}$  is the Grover number in the final round, and  $f_{\text{fin}} \in [0, 1]$  is a quantity determined by  $k_{\text{fin}}$  (see the definition later), which rapidly varies by a slight change of  $k_{\text{fin}}$  and largely affects the bias. Note that  $k_{\text{fin}}$  is also a random variable depending on the measurement outcomes, and thus so is  $(k_{\text{fin}}, f_{\text{fin}})$ . For a value of  $a$  other than the specific ones,  $f_{\text{fin}}$  takes the various values distributed widely in the range  $[0, 1]$  when  $k_{\text{fin}}$  varies. Then, over the wide distribution of realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  in the 2D plane, the various values of the bias for the various values of  $(k_{\text{fin}}, f_{\text{fin}})$  are canceled out on average, which yields a small bias in total. On the other hand, for the specific values of  $a$ , the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  are not distributed widely but concentrated in a small part of the 2D plane, in fact, on a few curves. Thus, the bias cancellation does not occur and the resultant bias remains considerable.

We also propose a simple way to mitigate the bias: just re-executing the final round. If the algorithm ends at the final round with the Grover number  $k_{\text{fin}}$  and the shot number  $N_{\text{fin}}$ , we perform another round with the same Grover number and shot number and obtain an estimate of  $a$  from the measurement outcomes in this additional round. Now, the gradual increment of the shot number and the termination criterion on the CI width are no longer adopted: we perform just  $N_{\text{fin}}$  shots and stop. This largely diminishes the bias by the termination criterion, as confirmed by the numerical experiments. We also confirm that, although this re-executing solution definitely increases the total number of queries to  $G$ , the increase rate is modest – about 25% in our experiment. This is because the final round does not dominate the other rounds in terms of the query number.

The rest of this paper is organized as follows. Section 2 is a preliminary one, where we outline QAE and IQAE. Section 3 is the main part of this paper, where the results of our numerical experiments are presented. We first show the magnitude of the bias for various values of  $a$ , which is enhanced for some specific values of  $a$ , and then elaborate the aforementioned understanding of such a phenomenon. We finally propose the bias mitigation method by re-executing the final round along with the numerical result. Section 4 summarizes this paper.

## 2 Preliminaries

### 2.1 Quantum amplitude estimation

In this paper, QAE is a generic term that means quantum algorithms to estimate the amplitude of a target basis state in a superposition state. Concretely, our aim is described as the following problem.

**Problem 1** Let  $\epsilon, \alpha \in (0, 1)$ . Suppose that we are given access to the following two quantum circuits  $A$  and  $S$  on a  $n$ -qubit register.  $A$  acts as

$$A|0\rangle = |\Phi\rangle := \sqrt{a}|\phi_1\rangle + \sqrt{1-a}|\phi_0\rangle, \quad (1)$$

where  $a \in [0, 1]$ ,  $|0\rangle$  is the computational basis state in which all the  $n$  qubits take  $|0\rangle$ , and  $|\phi_0\rangle$  and  $|\phi_1\rangle$  are quantum states orthogonal to each other.  $S$  acts as

$$\begin{cases} S|\phi_0\rangle &= |\phi_0\rangle \\ S|\phi_1\rangle &= -|\phi_1\rangle \end{cases}. \quad (2)$$

Also suppose that we can measure an observable corresponding to a Hermitian  $H$  on the same system, for which  $|\phi_0\rangle$  and  $|\phi_1\rangle$  are eigenstates with different eigenvalues. Then, we want to get an estimate of  $a$  with accuracy  $\epsilon$  with probability at least  $1 - \alpha$ .

Although the setup of Problem 1 seems quite simple, previous studies [1, 10–21] have generally considered this setup, and in fact, many applications can be boiled down to this form. For example, in QMCI [3], the expected value of a random variable is encoded into a quantum state like Eq. (1) as the squared amplitude  $a$  of a basis state  $|\phi_1\rangle$ , and through estimating this amplitude, we get an approximation of the expectation.

In many use cases of QAE,  $|\phi_1\rangle$  and  $|\phi_0\rangle$  are distinguished by whether a specific qubit takes  $|1\rangle$  or  $|0\rangle$ . In this case,  $S$  is the  $Z$  gate on the qubit and  $H$  is the projective measurement in the computational basis on the qubit.

[1] posed this problem and presented an algorithm for it based on QPE. We have the following theorem on its query complexity.

**Theorem 1** ([1], Theorem 12, modified) *Suppose that we are given access to the oracles  $A$  and  $S$  in Eqs. (1) and (2). Then, for any  $\epsilon, \alpha \in (0, 1)$ , there exists a quantum algorithm that outputs  $\hat{a} \in (0, 1)$  such that  $|\hat{a} - a| \leq \epsilon$  with probability at least  $1 - \alpha$  calling  $A$  and  $S$*

$$O\left(\frac{1}{\epsilon} \log\left(\frac{1}{\alpha}\right)\right) \quad (3)$$

*times.*

Although the success probability in the original algorithm in [1] is lower bounded by a constant  $8/\pi^2$ , it can be enhanced to an arbitrary value  $1 - \alpha$  at the expense of an  $O(\log(1/\alpha))$  overhead in the query complexity. This is done by a trick of taking the median of the results in the multiple runs of the algorithm [3], which is based on the powering lemma in [23].

We do not give the full details of this QPE-based QAE but present its outline briefly. The key ingredient is the Grover operator  $G$ , which is defined as

$$G := -AS_0A^\dagger S. \quad (4)$$

Here,  $S_0$  is a unitary such that

$$S_0|\phi\rangle = \begin{cases} -|0\rangle & ; \text{ if } |\phi\rangle = |0\rangle \\ |\phi\rangle & ; \text{ if } \langle\phi|0\rangle = 0 \end{cases}, \quad (5)$$

which can be implemented as a combination of  $X$  gates and a multi-controlled  $Z$  gate. The key property of  $G$  is that for any  $k \in \mathbb{N}$  it acts as

$$G^k|\Phi\rangle = \sin((2k+1)\theta_a)|\phi_1\rangle + \cos((2k+1)\theta_a)|\phi_0\rangle, \quad (6)$$

where  $\theta_a := \arcsin(\sqrt{a})$ . That is,  $G$  rotates the statevector by an angle  $2\theta_a$  in the 2-dimensional Hilbert space spanned by  $|\phi_1\rangle$  and  $|\phi_0\rangle$ . Because of this property, we have the following QPE-based approach to estimate  $a$ . We prepare a register  $R_G$  on which  $G$  acts and an ancillary  $m$ -qubit register  $R_{\text{anc}}$ . Then, using  $G, G^2, \dots, G^{2^{m-1}}$  controlled by the first, second,  $\dots$ ,  $m$ -th qubits in  $R_{\text{anc}}$ , respectively, we generate the state  $\frac{1}{2^{m/2}} \sum_{k=0}^{2^m-1} |k\rangle G^k |\Phi\rangle$ . Finally, we operate the inverse quantum Fourier transform on  $R_{\text{anc}}$ , which, thanks to the property in Eq. (6), yields an  $m$ -bit precision estimate of  $\theta_a$  and thus that of  $a = \sin^2 \theta_a$ . In this process, the number of uses of (controlled)  $G$  is  $1 + 2 + \dots + 2^{m-1} = 2^m$ , and thus  $A$  and  $S$  are queried  $O(2^m)$  times, which implies the  $O(1/\epsilon)$  query complexity for accuracy  $\epsilon$  as shown in Theorem 1. See [1] for more details.

Compared to a naive way for estimating  $a$  by repeating generations of  $|\Phi\rangle$  and measurements on it and letting the frequency of obtaining  $|\phi_1\rangle$  be an estimate of  $a$ , in which the number of queries to  $A$  scales as  $\tilde{O}(1/\epsilon^2)$ , QAE achieves the quadratic speedup with respect to  $\epsilon$ . This is the origin of the quadratic speedups in quantum algorithms built upon QAE, e.g., QMCI in comparison to the classical Monte Carlo method.

## 2.2 Iterative quantum amplitude estimation

After the original QAE was proposed, some variants have been proposed so far, aiming at the reduction of the resource. Avoiding the use of QPE is a common approach since QPE requires the controlled version of  $G$ , for which the resource for the implementation increases compared to the uncontrolled one.

IQAE [13] is in such a direction. The basic idea is as follows. By iterating the generation of  $|\Phi\rangle$  and the measurement on it, we get an estimate of  $a$  as the frequency of obtaining  $|\phi_1\rangle$  in the measurements. This is in fact a kind of MLE, since for  $\text{Be}(p)$ , the Bernoulli distribution with probability of 1 equal to  $p$ , the maximum likelihood estimate of  $p$  from multiple trials is nothing but the realized frequency of 1. We also have the CI of  $a$ , which contains the true value of  $a$  with high probability. Next, for some  $k \in \mathbb{N}$ , we repeat generating  $G^k|\Phi\rangle$  and the measurement on it, and from the measurement outcomes we get an estimate of

$$a_k := \sin^2((2k+1)\theta_a) \quad (7)$$

and then that of  $a$ . Here, due to the periodicity of  $a_k$  as the function of  $a$ , the  $2k + 1$  different values in  $[0, 1]$  can be the maximum likelihood estimates of  $a$ . However, since we already have the CI from the previous round consisting of the measurements on  $|\Phi\rangle$ , combining it with the measurements on  $G^k|\Phi\rangle$  yields the new CI as a single interval in  $[0, 1]$ , whose width is narrower than the previous one. We repeat this procedure for a sufficient number of rounds, making the width of the CI narrower. When the CI width reaches the required accuracy  $\epsilon$ , we have an estimate of  $a$  with an error below  $\epsilon$ . The CI width in each round is taken so that the probability that in every round the obtained CI successfully encloses the true value  $a$  is at least some required success probability  $1 - \alpha$ .

Concretely, we present this algorithm as Algorithm 1. This is a modification of the algorithm [20], which is already a modified version of the original IQAE algorithm in [13].

Leaving the full details of this algorithm to [20], we just present a theorem on the query complexity and some comments.

**Theorem 2** (Theorem 3.1 in [20]) *Let  $\epsilon, \alpha \in (0, 1)$ . Suppose that we are given access to the oracles  $A$  in Eq. (1) and  $S$  in Eq. (2). Then, Algorithm 1 outputs an  $\epsilon$ -approximation  $\hat{a}$  of  $a$  with probability at least  $1 - \alpha$ , making  $O(\frac{1}{\epsilon} \log \frac{1}{\alpha})$  queries to  $A$  and  $S$  in total.*

Let us make some comments that help us to understand the outline of the algorithm. In Algorithm 1,  $[a_{k_i}^l, a_{k_i}^u]$  and  $[\theta_{k_i}^l, \theta_{k_i}^u]$  are the CIs of  $a$  and the angle  $\theta_a$ , respectively. In the  $i$ th round, we set the Grover number to  $k_i$ , and we calculate the maximum likelihood estimate  $\hat{a}_{k_i}$  of  $a_{k_i}$  and its CI  $[a_{k_i}^l, a_{k_i}^u]$  from the outcomes of the measurements on  $G^{k_i}|\Phi\rangle$ . They are translated to the maximum likelihood estimate  $\hat{a}_{i,j}$  of  $a$  and the CIs  $[\theta_{i,j}^l, \theta_{i,j}^u]$  and  $[a_{i,j}^l, a_{i,j}^u]$ . The  $i$ th round ends if we find the Grover number in the next round by the procedure `FindNextK` shown as Algorithm 2. In this procedure, we search the next  $k$  greedily. Namely, we set it as large as possible, requiring that  $[K\theta_{i,j}^l, K\theta_{i,j}^u]$  the CI of  $\theta_a$  multiplied by  $K = 2k + 1$  lies in a single quadrant, which enables us to determine the CIs of  $a$  and  $\theta_a$  as single intervals. If we cannot find such  $k$  in the region that  $K \geq r_{\min}(2k_i + 1)$ , we continue the  $i$ th round. When  $\Delta a_{i,j}$ , the accuracy of the estimate of  $a$  from the measurement outcomes so far, goes below the predetermined accuracy level  $\epsilon$ , the algorithm stops and outputs the estimate at that time.

In the rounds of this process, we operate  $G^{k_1}, G^{k_2}, \dots$  with  $k_i$  increasing exponentially as  $K_{i+1} \geq r_{\min}K_i$ , and its value in the final round is of order  $O(1/\epsilon)$ . This means that  $G$  is queried  $O(1/\epsilon)$  times in total, and so are  $A$  and  $S$ , as stated in Theorem 2.

Note that the parameters  $N_{\text{shot}}$  and  $r_{\min}$  in Algorithm 1 are not mentioned in the statement of Theorem 2. Although there may be various settings on these that make the algorithm work, we adopt the following setting in this paper. According to [20],  $N_{\text{shot}}$  can be set to 1, which means that we search the next  $k$  every time we make one measurement on  $G^{k_i}|\Phi\rangle$ , and it reduces the query complexity keeping the accuracy. We thus set  $N_{\text{shot}} = 1$  hereafter. On  $r_{\min}$ , there may be some choices such as 2 in [13] and 3 in [20], and we adopt the former hereafter.

We also note that a slight modification in Algorithm 1 from the algorithm in [20]: the former outputs the maximum likelihood estimate  $\hat{a}$  when  $\Delta a_{i,j}$  becomes smaller than the required accuracy  $\epsilon$ , while the latter outputs the midpoint of  $[a_{i,j}^l, a_{i,j}^u]$  when  $\theta_{i,j}^u - \theta_{i,j}^l$ , the width of the CI of  $\theta_a$ , becomes smaller than  $2\epsilon$ . The easily checked relationship  $|a_{i,j}^u - a_{i,j}^l| \leq |\theta_{i,j}^u - \theta_{i,j}^l|$  implies that, if we impose the termination criterion on  $\theta_a$  even though we want

**Algorithm 1** Modified IQAE**Input:**  $\epsilon, \alpha \in (0, 1), N_{\text{shot}} \in \mathbb{N}, r_{\text{min}} > 1$ 


---

```

1:  $K_{\text{max}} \leftarrow \frac{\pi}{4\epsilon}$ 
2:  $k_1 \leftarrow 0$ 
3:  $\theta_{\text{last}}^l \leftarrow 0$ 
4: for  $i = 1, 2, \dots$  do
5:    $N_{i,0} \leftarrow 0, N_{i,0}^1 \leftarrow 0$ 
6:    $K_i \leftarrow 2k_i + 1$ 
7:    $\alpha_i \leftarrow \frac{2\alpha}{3} \frac{K_i}{K_{\text{max}}}$ 
8:    $N_i^{\text{max}} \leftarrow \frac{2}{\sin^2(\pi/21) \sin^2(8\pi/21)} \ln\left(\frac{2}{\alpha_i}\right)$ 
9:    $R_i \leftarrow \lfloor \frac{K_i \theta_{\text{last}}^l}{\pi/2} \rfloor$  ▷ indicates the quadrant enclosing  $[K_i \theta_{i,j}^l, K_i \theta_{i,j}^u]$ 
10:  for  $j = 1, 2, \dots$  do
11:     $n_{\text{shot}} \leftarrow \min\{N_{\text{shot}}, N_i^{\text{max}} - N_{i,j-1}\}$ 
12:    Iterate generating  $G^{k_i}|\Phi\rangle$  and measuring it  $n_{\text{shot}}$  times. Let the number of times  $|\phi_1\rangle$  is obtained be  $n_1$ .
13:     $N_{i,j} \leftarrow N_{i,j-1} + n_{\text{shot}}, N_{i,j}^1 \leftarrow N_{i,j-1}^1 + n_1$ 
14:     $\hat{a}_{k_{i,j}} \leftarrow \frac{N_{i,j}^1}{N_{i,j}}$  ▷ Maximum likelihood estimate of  $a_{k_i}$ 
15:     $\epsilon_{i,j} \leftarrow \sqrt{\frac{1}{2N_{i,j}} \ln\left(\frac{2}{\alpha_i}\right)}$ 
16:     $[a_{k_{i,j}}^l, a_{k_{i,j}}^u] \leftarrow [\max(0, \hat{a}_{k_{i,j}} - \epsilon_{i,j}), \min(1, \hat{a}_{k_{i,j}} + \epsilon_{i,j})]$  (8)
▷ CI of  $a_{k_i}$ 
17:     $\hat{\gamma}_{i,j} \leftarrow \gamma(\hat{a}_{k_{i,j}}, R_i), \gamma_{i,j}^l \leftarrow \gamma(a_{k_{i,j}}^l, R_i), \gamma_{i,j}^u \leftarrow \gamma(a_{k_{i,j}}^u, R_i)$  (9)
    where
    
$$\gamma(a', r) := \begin{cases} \arcsin(\sqrt{a'}) & ; \text{ if } r \text{ is even} \\ \frac{\pi}{2} - \arcsin(\sqrt{a'}) & ; \text{ if } r \text{ is odd} \end{cases}$$
 (10)
    for  $a' \in [0, 1]$  and  $r \in \mathbb{N}$ .
18:     $\hat{\theta}_{i,j} \leftarrow \frac{R_i \frac{\pi}{2} + \hat{\gamma}_{i,j}}{K_i}$  ▷ Maximum likelihood estimate of  $\theta_a$ 
19:     $[\theta_{i,j}^l, \theta_{i,j}^u] \leftarrow \left[ \frac{R_i \frac{\pi}{2} + \gamma_{i,j}^l}{K_i}, \frac{R_i \frac{\pi}{2} + \gamma_{i,j}^u}{K_i} \right]$  ▷ CI of  $\theta_a$ 
20:     $\hat{a}_{i,j} \leftarrow \sin^2 \hat{\theta}_{i,j}$  ▷ Maximum likelihood estimate of  $a$ 
21:     $[a_{i,j}^l, a_{i,j}^u] \leftarrow [\sin^2 \theta_{i,j}^l, \sin^2 \theta_{i,j}^u]$  ▷ CI of  $a$ 
22:     $\Delta a_{i,j} \leftarrow \max\{\hat{a}_{i,j} - a_{i,j}^l, a_{i,j}^u - \hat{a}_{i,j}\}$  (11)
▷ Estimated accuracy of  $\hat{a}_{i,j}$ 
23:    if  $\Delta a_{i,j} \leq \epsilon$  then
24:      return  $\hat{a}_{i,j}$  as  $\hat{a}$ 
25:    end if
26:     $k_{\text{temp}} \leftarrow \text{FindNextK}(k_i, \theta_{i,j}^l, \theta_{i,j}^u, r_{\text{min}})$ 
27:    if  $k_{\text{temp}} > k_i$  then
28:       $k_{i+1} \leftarrow k_{\text{temp}}$ 
29:       $\theta_{\text{last}}^l \leftarrow \theta_{i,j}^l$ 
30:    break the inner for-loop
31:    end if
32:  end for
33: end for

```

---

to guarantee the accuracy of  $a$ , we may take unnecessarily many iterations. We thus adopt the termination criterion on  $a$ , and in fact, we confirmed that this leads to a smaller query number than the criterion on  $\theta_a$  in the later numerical experiment, although we will not show the result in the latter setting.



**Algorithm 2** FindNextK**Input:**  $k_i, \theta_l, \theta_u, r_{\min}$ 


---

```

1:  $K_i \leftarrow 2k_i + 1$ 
2:  $K \leftarrow \left\lfloor \frac{\pi/2}{\theta_u - \theta_l} \right\rfloor$ 
3: if  $K$  is even then
4:    $K \leftarrow K - 1$ 
5: end if
6: while  $K \geq r_{\min} K_i$  do
7:   if  $\left\lfloor \frac{K\theta_l}{\pi/2} \right\rfloor = \left\lceil \frac{K\theta_u}{\pi/2} \right\rceil - 1$  then
8:     return  $(K - 1)/2$ 
9:   end if
10:   $K \leftarrow K - 2$ 
11: end while
12: return  $k_i$ 

```

---

**3 Numerical experiments on the bias in IQAE****3.1 Biases for various  $a$** 

Hereafter, in order to understand the bias in IQAE and how it arises, we conduct some numerical experiments.

First of all, since the current problem is characterized by  $a$ , the amplitude we want to estimate, we run IQAE and see the bias for the various values of  $a$ . Here, “run” does not mean running Algorithm 1 on a real quantum computer or a quantum circuit simulator but a classical simulation of the algorithm based on the probability distribution of the outcomes of the measurements made in the algorithm. Concretely, we replace the step 32 in Algorithm 1 with “Draw  $n_{\text{shot}}$  samples from  $\text{Be}(a_{k_i})$  and let the number of 1’s be  $n_1$ ”. Note that, if we know the value of  $a$ , we know the probability distribution of the outcome in measuring  $G^{k_i}|\Phi\rangle$ , that is,  $|\phi_1\rangle$  with probability  $a_{k_i}$  and  $|\phi_0\rangle$  with probability  $1 - a_{k_i}$ , which is equivalent to  $\text{Be}(a_{k_i})$ . Therefore, with the above replacement, we can produce the output of Algorithm 1 under the probability distribution it obeys. We hereafter call this simulation procedure Algorithm 1’.

For the estimator  $\hat{a}$  of  $a$ , we define its bias as

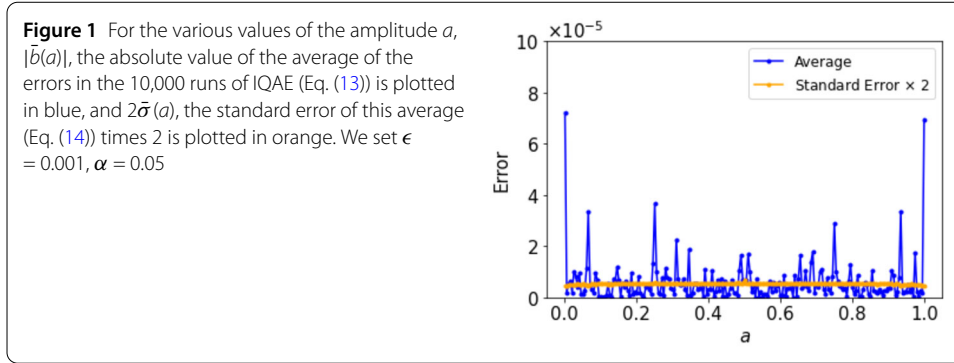
$$b(a) := \mathbb{E}[\hat{a} - a], \quad (12)$$

where  $\mathbb{E}[\cdot]$  denotes the expectation with respect to the randomness of the measurement outcomes in Algorithm 1. In order to estimate the magnitude of the bias for various values of  $a$ , we ran Algorithm 1’  $N_{\text{run}}$  times, and calculate the value of the average of the errors in the runs

$$\bar{b}(a) := \frac{1}{N_{\text{run}}} \sum_{n=1}^{N_{\text{run}}} (\hat{a}^{(n)} - a) \quad (13)$$

and the standard error of  $\bar{b}(a)$

$$\bar{\sigma}(a) := \frac{1}{\sqrt{N_{\text{run}}}} \left[ \frac{1}{N_{\text{run}}} \sum_{n=1}^{N_{\text{run}}} (\hat{a}^{(n)} - a)^2 \right]^{1/2}, \quad (14)$$



where  $\hat{a}^{(n)}$  is the output in the  $n$ th run. We can grasp the magnitude of the bias by comparing  $\bar{b}(a)$  and  $2\bar{\sigma}(a)$ , since, if  $b(a) = 0$ ,  $\bar{b}(a)$  is comparable to  $2\bar{\sigma}(a)$  with high probability, e.g.,  $|\bar{b}(a)| \leq 2\bar{\sigma}(a)$  with about 95%, for sufficiently large  $N_{\text{run}}$ . The result is shown in Fig. 1. In this figure, we set  $a$  to the 201 equally spaced points in  $[0.001, 0.999]$  including the ends. The other parameters are set as  $N_{\text{run}} = 10^4$ ,  $\epsilon = 0.001$ ,  $\alpha = 0.05$ , which also applies hereafter. From this figure, we see that, for a considerable fraction of the examined values of  $a$ ,  $|\bar{b}(a)|$  exceeds  $2\bar{\sigma}(a)$ , which implies  $\hat{a}$  is biased. In particular,  $\bar{b}(a)$  takes much larger values for specific values of  $a$  than other values. Namely, the bias is enhanced for some specific values of  $a$ .

### 3.2 Reason why the bias is enhanced for specific values of $a$

We then investigate the reason why the bias is enhanced for specific values of  $a$ , fixing  $a$  to 0.2505, one of such values. What makes the situation that  $a = 0.2505$  different from the others?

To make the investigation as simple as possible, we should note some points. First, we note that it is sufficient to focus on the final round in Algorithm 1. The output of Algorithm 1 is determined by the result of the MLE in the final round. Therefore, as long as the CI obtained in the round just before the final one successfully encloses  $a$ , which occurs with high probability at least  $1 - \alpha$ , the error is determined by the final round only. Second, we note that the final round is characterized by only  $k_{\text{fin}}$ , the Grover number in that round: other quantities that define the procedure in the final round are automatically determined when  $k_{\text{fin}}$  is fixed, including  $R_{\text{fin}}$ , as long as the CI of  $a$  encloses its true value. We finally make a note on notation: here and hereafter, like  $k_{\text{fin}}$  and  $R_{\text{fin}}$ , when we write quantities that have  $i_{\text{fin}}$ , the index indicating the final round, in the subscript, we replace  $i_{\text{fin}}$  with  $\text{fin}$  for conciseness.

Based on the above discussion and noting that  $k_{\text{fin}}$  is also a random variable affected by the measurement outcomes in the previous rounds, we write the bias as

$$b(a) = \mathbb{E}[\hat{a} - a] = \sum_{k'_{\text{fin}}=0}^{\infty} p_{k'_{\text{fin}}} b_{k'_{\text{fin}}}(a) \quad (15)$$

$$b_{k'_{\text{fin}}}(a) := \mathbb{E}[\hat{a} - a | \mathcal{G}_{k'_{\text{fin}}}], \quad (16)$$

where  $\mathcal{G}_{k'_{\text{fin}}}$  is the event that  $k_{\text{fin}}$  takes a value  $k'_{\text{fin}}$ , and  $p_{k'_{\text{fin}}}$  is its probability.

### 3.2.1 Bias conditioned on the final Grover number

We then consider how  $k_{\text{fin}}$  affects  $b_{k_{\text{fin}}}(a)$ , the bias conditioned on  $k_{\text{fin}}$ . First, we rewrite  $\hat{a}$  the output of Algorithm 1 as

$$\hat{a} = \sin^2 \left( \frac{R_{\text{fin}} \frac{\pi}{2} + \hat{\gamma}_{\text{fin}}}{K_{\text{fin}}} \right), \quad (17)$$

where  $\hat{\gamma}_{\text{fin}}$  is the value of  $\hat{\gamma}_{\text{fin},j}$  when

$$\begin{aligned} a_{\text{fin},j}^u - \hat{a}_{\text{fin},j} &= \sin^2 \left( \frac{R_{\text{fin}} \frac{\pi}{2} + \gamma_{\text{fin},j}^u}{K_{\text{fin}}} \right) - \sin^2 \left( \frac{R_{\text{fin}} \frac{\pi}{2} + \hat{\gamma}_{\text{fin},j}}{K_{\text{fin}}} \right) \leq \epsilon, \\ \hat{a}_{\text{fin},j} - a_{\text{fin},j}^l &= \sin^2 \left( \frac{R_{\text{fin}} \frac{\pi}{2} + \hat{\gamma}_{\text{fin},j}}{K_{\text{fin}}} \right) - \sin^2 \left( \frac{R_{\text{fin}} \frac{\pi}{2} + \gamma_{\text{fin},j}^l}{K_{\text{fin}}} \right) \leq \epsilon \end{aligned} \quad (18)$$

holds for the first time. Note that  $\gamma_{\text{fin},j}^l$ ,  $\gamma_{\text{fin},j}^u$  and  $\hat{\gamma}_{\text{fin},j}$  are random variables determined by the measurement outcomes in Algorithm 1, and so is  $\hat{\gamma}_{\text{fin}}$ . When  $k_{\text{fin}} \gg 1$ , which is a typical situation for small  $\epsilon$ , the most sensitive dependence of  $\hat{a}$ 's distribution on  $k_{\text{fin}}$  is through the distributions of  $\gamma_{\text{fin},j}^l$ ,  $\gamma_{\text{fin},j}^u$  and  $\hat{\gamma}_{\text{fin},j}$ , while  $k_{\text{fin}}$  also affects  $\hat{a}$  through  $K_{\text{fin}}$  and  $R_{\text{fin}}$  in Eqs. (17) and (18). This is because the distributions of  $\hat{\gamma}_{\text{fin},j}$  etc. largely change even when  $k_{\text{fin}}$  changes slightly. Under the definition in Eq. (9), the distributions of  $\hat{\gamma}_{\text{fin},j}$  etc. are determined by the distribution of  $\hat{a}_{k_{\text{fin},j}}$  and  $\mathfrak{p}(R_{\text{fin}})$ , the parity of  $R_{\text{fin}}$ . The distribution of  $\hat{a}_{k_{\text{fin},j}}$  is determined by that of the outcome of the measurement on  $G^{k_{\text{fin}}}|\Phi\rangle$ , which is equivalent to  $\text{Be}(a_{k_{\text{fin}}})$ , and  $a_{k_{\text{fin}}}$  can change largely even by shifting  $k_{\text{fin}}$  by 1. When  $\mathfrak{p}(R_{\text{fin}})$  flips, the form of  $\hat{\gamma}_{\text{fin},j}$  as a function of  $\hat{a}_{k_{\text{fin},j}}$  also flips and so does the distribution of  $\hat{\gamma}_{\text{fin},j}$ . These observations imply that we should focus on  $a_{k_{\text{fin}}}$  and  $\mathfrak{p}(R_{\text{fin}})$  as key factors for  $b_{k_{\text{fin}}}(a)$ . However, as an equivalent to this, we instead focus on how  $b_{k_{\text{fin}}}(a)$  is affected by

$$f_{\text{fin}} := \text{frac} \left( \frac{(2k_{\text{fin}} + 1)\theta_a}{\pi} \right), \quad (19)$$

where  $\text{frac}(x) := x - \lfloor x \rfloor$  is the fractional part of  $x \in \mathbb{R}$ .  $f_{\text{fin}}$  has a one-to-one correspondence to the pair  $(a_{k_{\text{fin}}}, \mathfrak{p}(R_{\text{fin}}))$ , and taking a single quantity  $f_{\text{fin}}$  as a key factor rather than the pair  $(a_{k_{\text{fin}}}, \mathfrak{p}(R_{\text{fin}}))$  makes the following discussion simpler.

Since  $f_{\text{fin}}$  is also a quantity rapidly changing with respect to  $k_{\text{fin}}$ , in order to understand how it affect  $b_{k_{\text{fin}}}(a)$ , we temporarily deal with  $k_{\text{fin}}$  and  $f_{\text{fin}}$  as independent variables, even though  $k_{\text{fin}}$  determines  $f_{\text{fin}}$ . We set  $k_{\text{fin}}$  and  $f_{\text{fin}}$  separately, and, as an estimate of  $b_{k_{\text{fin}}}(a)$ , calculate the quantity  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  by the procedure in Algorithm 3. That is, we run one round of IQAE many times for fixed  $k_{\text{fin}}$  and  $f_{\text{fin}}$ , and take the average of the resultant errors in the runs in which IQAE itself is terminated, neglecting the other runs, which move to the next round. We calculate the average only when at least 1000 runs out of the 10,000 total runs yield outputs other than NaN, since averaging a small number of random results leads to an inaccurate estimate. Note that, in compensation for setting  $k_{\text{fin}}$  and  $f_{\text{fin}}$  independently, the true amplitude  $a$  (now, 0.2505) is slightly adjusted to  $\tilde{a}$  so that Eq. (19)

---

**Algorithm 3** Estimate the bias with  $k_{\text{fin}}$  and  $f_{\text{fin}}$  fixed

---

**Input:**  $k_{\text{fin}} \in \mathbb{N}, f_{\text{fin}} \in [0, 1], a \in (0, 1), N_{\text{run}} \in \mathbb{N}, \epsilon, \alpha \in (0, 1), N_{\text{shot}} \in \mathbb{N}, r_{\text{min}} > 1$

1:

$$\begin{aligned} \tilde{R} &\leftarrow \left\lfloor \frac{(2k_{\text{fin}} + 1)\theta_a}{\pi} \right\rfloor \\ \tilde{a} &\leftarrow \sin^2 \left( \frac{(\tilde{R} + f_{\text{fin}})\pi}{2k_{\text{fin}} + 1} \right) \end{aligned} \quad (20)$$

2: **for**  $c = 1, \dots, N_{\text{run}}$  **do**

3: Run one round of IQAE (Steps 32-32 in Algorithm 1), setting  $k_i \leftarrow k_{\text{fin}}$  and

$$R_i \leftarrow \left\lfloor \frac{(2k_{\text{fin}} + 1)\theta_{\tilde{a}}}{\pi/2} \right\rfloor, \quad (21)$$

and adjusting  $a$  to  $\tilde{a}$ .

4: **if** the round ends with  $\Delta a_{i,j} \leq \epsilon$  **then**

5: Let  $\tilde{a}_c$  be the value of  $\hat{a}_{i,j}$  at the end of the round.

6:  $I_c^{\text{end}} \leftarrow 1$

7: **else**

8:  $\tilde{a}_c \leftarrow \text{NaN}$  (not a number)

9:  $I_c^{\text{end}} \leftarrow 0$

10: **end if**

11: **end for**

12:  $N_{\text{end}} \leftarrow \sum_{c=1}^{N_{\text{run}}} I_c^{\text{end}}$

13: Output

$$\begin{aligned} &\tilde{b}(k_{\text{fin}}, f_{\text{fin}}) \\ &\leftarrow \begin{cases} \frac{1}{N_{\text{end}}} \sum_{\substack{c=1, \dots, N_{\text{run}} \\ I_c^{\text{end}}=1}} (\tilde{a}_c - \tilde{a}) & ; \text{ if } N_{\text{end}} \geq 0.1 \times N_{\text{run}} \\ \text{NaN} & ; \text{ otherwise} \end{cases} \end{aligned} \quad (22)$$

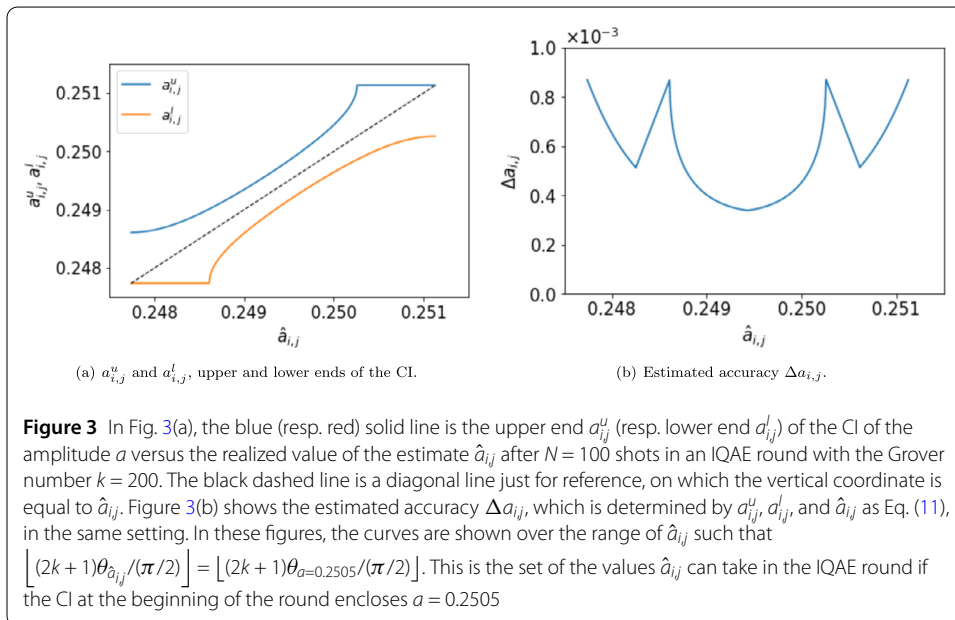
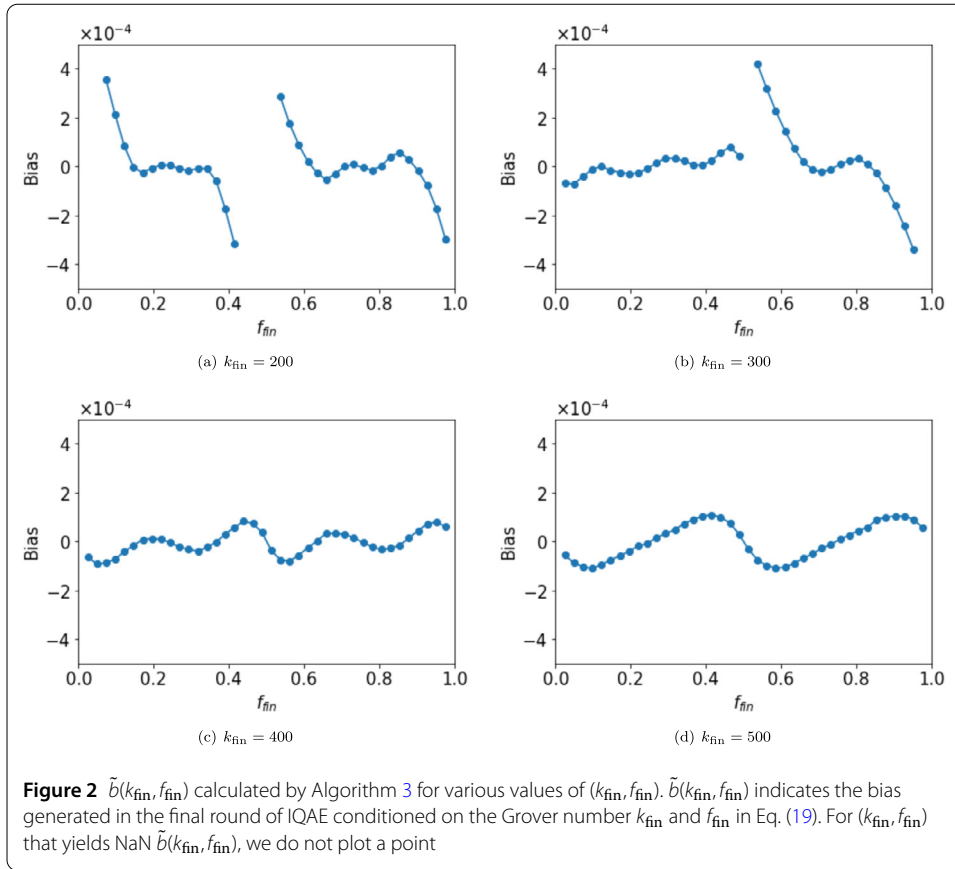

---

holds. Nevertheless, we expect that  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  gives us a good illustration of the behavior of  $b_{k_{\text{fin}}}(a)$ . Also note that the setting in Eq. (21) corresponds to the assumption that the true amplitude  $\tilde{a}$  is enclosed in the CI in the previous round.

In Fig. 2, we show the results for the various values of  $f_{\text{fin}}$  with  $k_{\text{fin}}$  set to several values. For  $(k_{\text{fin}}, f_{\text{fin}})$  that yields NaN  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$ , we do not plot a point.

We can intuitively understand how the bias is generated by noting the following two points.

First, we note that  $\Delta a_{i,j}$ , the estimated accuracy in the intermediate step in Algorithm 1, is also a random variable. It is determined by  $\hat{a}_{k_{i,j}}$ , the realized frequency of  $|\phi_1\rangle$  in the repeated measurements on  $G^{k_i}|\Phi\rangle$ , and  $N_{i,j}$ , the number of the measurements. How  $\Delta a_{i,j}$  depends on  $\hat{a}_{k_{i,j}}$  is understood as follows. Because of the upper bound 1 and lower bound 0 of  $\hat{a}_{k_{i,j}}$ , the width of the CI  $[a_{k_{i,j}}^l, a_{k_{i,j}}^u]$  of  $a_{k_i}$  given as Eq. (8) already depends on not only



$N_{i,j}$  but also  $\hat{a}_{k_{i,j}}$ : if  $\hat{a}_{k_{i,j}}$  is closer to 0 or 1, the CI width becomes smaller. In addition to this, the derivation of  $\Delta a_{i,j}$  from  $\hat{a}_{k_{i,j}}$  and  $[a_{k_{i,j}}^l, a_{k_{i,j}}^u]$  by the nonlinear relationship introduce the dependence of  $\Delta a_{i,j}$  on  $\hat{a}_{k_{i,j}}$ . Besides, since  $\hat{a}_{k_{i,j}}$  has the one-to-one correspondence to  $\hat{a}_{i,j}$ , we can regard  $\Delta a_{i,j}$  as a function of  $\hat{a}_{i,j}$ . Figure 3 illustrates this. Figure 3(a) shows  $a_{i,j}^u$  and

$a_{i,j}^l$ , the upper and lower ends of the CI of  $a$  versus the realized value of  $\hat{a}_{i,j}$  after  $N = 100$  shots in a round with the Grover number  $k = 200$ , and Fig. 3(b) shows  $\Delta a_{i,j}$  versus  $\hat{a}_{i,j}$ .

The second point is the termination criterion  $\Delta a_{i,j} \leq \epsilon$  in Step 32 in Algorithm 1, which means that IQAE ends when the estimated accuracy  $\Delta a_{i,j}$  becomes smaller than  $\epsilon$ . This criterion, along with the aforementioned dependence of  $\Delta a_{i,j}$  on  $\hat{a}_{i,j}$ , leads to the bias. If  $\hat{a}_{i,j}$  goes to the region where  $\Delta a_{i,j}$  is relatively small, Algorithm 1 tends to end early, and, if  $\hat{a}_{i,j}$  goes to the region with larger  $\Delta a_{i,j}$ , the algorithm tends to take more shots by the end, or even go to the next round. This makes the probability that the algorithm ends with  $\hat{a}_{i,j}$  corresponding to smaller  $\Delta a_{i,j}$  and outputs such  $\hat{a}_{i,j}$  as  $\hat{a}$  higher. Such an effect yields the probability distribution of  $\hat{a}$  asymmetric about the point  $\hat{a} = a$ , and the nonzero bias  $b_{k_{\text{fin}}} \neq 0$ .

Although this understanding on how the bias arises is an intuitive one, we will indirectly confirm its validity by seeing in the numerical experiment in Sect. 3.3 that re-executing the final round without the termination criterion mitigates the bias.

With the above understanding, we can also see why the bias  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  depends on  $f_{\text{fin}}$  with  $k_{\text{fin}}$  fixed. As explained above, changing  $f_{\text{fin}}$  corresponds to changing  $a$ . Assuming that  $\hat{a}$  mainly distributes in the neighborhood of  $a$ , the shape of  $\Delta a_{i,j}$  as a function of  $\hat{a}_{i,j}$  around the point  $\hat{a}_{i,j} = a$  affects the bias, which causes the dependence of the bias on  $a$ , and then  $f_{\text{fin}}$ . For example, if  $\Delta a_{i,j}$  is increasing in the neighborhood of  $\hat{a}_{i,j} = a$ , the preference of  $\hat{a}_{i,j}$  to values corresponding to small  $\Delta a_{i,j}$  leads to the negative bias.

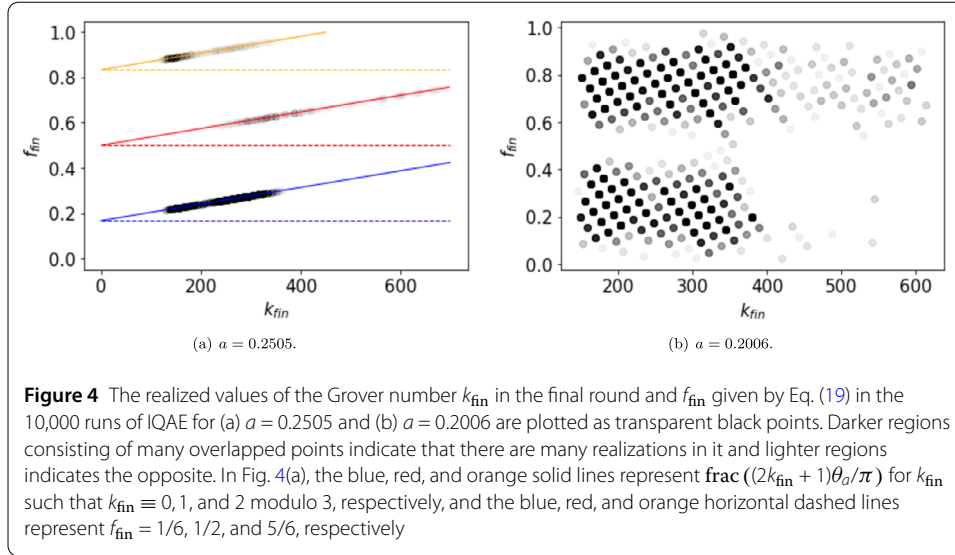
We also note that, in Fig. 2, the graph of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  is nearly antisymmetric with respect to reflection about the vertical line  $f_{\text{fin}} = \frac{1}{2}$ , that is,

$$\tilde{b}(k_{\text{fin}}, f_{\text{fin}}) \approx -\tilde{b}(k_{\text{fin}}, 1 - f_{\text{fin}}), \quad (23)$$

except the case of  $k_{\text{fin}} = 300$ . This antisymmetry is understood as follows. We temporarily ignore the process for the transition to the next round. Then, note that, with  $k_{\text{fin}}$  fixed, the transform  $f_{\text{fin}} \rightarrow 1 - f_{\text{fin}}$  conserves  $a_{k_{\text{fin}}}$  and flips  $p(R_{\text{fin}})$ . Thus, by this transform, the distribution of  $\hat{a}_{k_{\text{fin}},j}$  is unchanged and the relationship between  $\hat{y}_{\text{fin},j}$  and  $\hat{a}_{k_{\text{fin}},j}$  in Eq. (10) is switched. Therefore, assuming that  $k_{\text{fin}} \gg 1$  and neglecting the slight change of  $a$  by this transform, we see that this transform flips the distribution of the error  $\hat{a} - a$ :  $p(\hat{a} - a; k_{\text{fin}}, f_{\text{fin}}) \approx p(-(\hat{a} - a); k_{\text{fin}}, 1 - f_{\text{fin}})$ , where  $p(\cdot; k_{\text{fin}}, f_{\text{fin}})$  is the probability density of  $\hat{a} - a$  conditioned by  $k_{\text{fin}}$  and  $f_{\text{fin}}$ . This means that the transform  $f_{\text{fin}} \rightarrow 1 - f_{\text{fin}}$  also flips the bias as Eq. (23). In fact, this discussion is not rigorous, since the process for the transition to the next round breaks the symmetry with respect to  $f_{\text{fin}} \rightarrow 1 - f_{\text{fin}}$ . This makes the antisymmetry disappear in Fig. 2(b) for  $k_{\text{fin}} = 300$ . Nevertheless, this antisymmetry holds for the wide region of  $(k_{\text{fin}}, f_{\text{fin}})$ , and becomes a key for the phenomenon that the bias is enhanced only for the specific values of  $a$ , as seen below.

### 3.2.2 Distribution of $(k_{\text{fin}}, f_{\text{fin}})$

Next, we consider the distribution of  $(k_{\text{fin}}, f_{\text{fin}})$ . Since  $f_{\text{fin}}$  is determined by  $k_{\text{fin}}$  with  $a$  fixed and  $k_{\text{fin}}$  takes natural numbers,  $(k_{\text{fin}}, f_{\text{fin}})$  distributes in the 2D plane not continuously but as discrete points. In Fig. 4(a), we show the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  for  $a = 0.2505$  in the 10,000 runs of Algorithm 1'. We see that the plotted points are concentrated only on the three lines. This is not a phenomenon observed for general  $a$ . For example, in Fig. 4(b), the similar plot for  $a = 0.2006$ , for which the observed bias  $\tilde{b}(a)$  is much smaller than  $a = 0.2505$ , the points are distributed more broadly than those for  $a = 0.2505$ .



The distribution of the points for  $a = 0.2505$  is understood as follows. We note that, for  $\theta \in (0, \pi/2)$  written as  $\theta = l\pi/2m$  with positive integers  $l$  and  $m$  such that  $l < m$ ,  $f_\theta(k) := \text{frac}((2k + 1)\theta/\pi)$  is a periodic function of  $k \in \mathbb{Z}$  with period  $m$ . Thus, it takes at most  $m$  values, and, if we plot  $f_\theta(k)$  versus  $k$ , the points lie on at most  $m$  horizontal lines. If we slightly shift  $\theta$  from such a value,  $f_\theta(k)$  can take different values from the original ones, which means the horizontal lines transform into lines with small slopes.  $\theta_a$  for  $a = 0.2505$  applies to such a case. It can be written in the form of

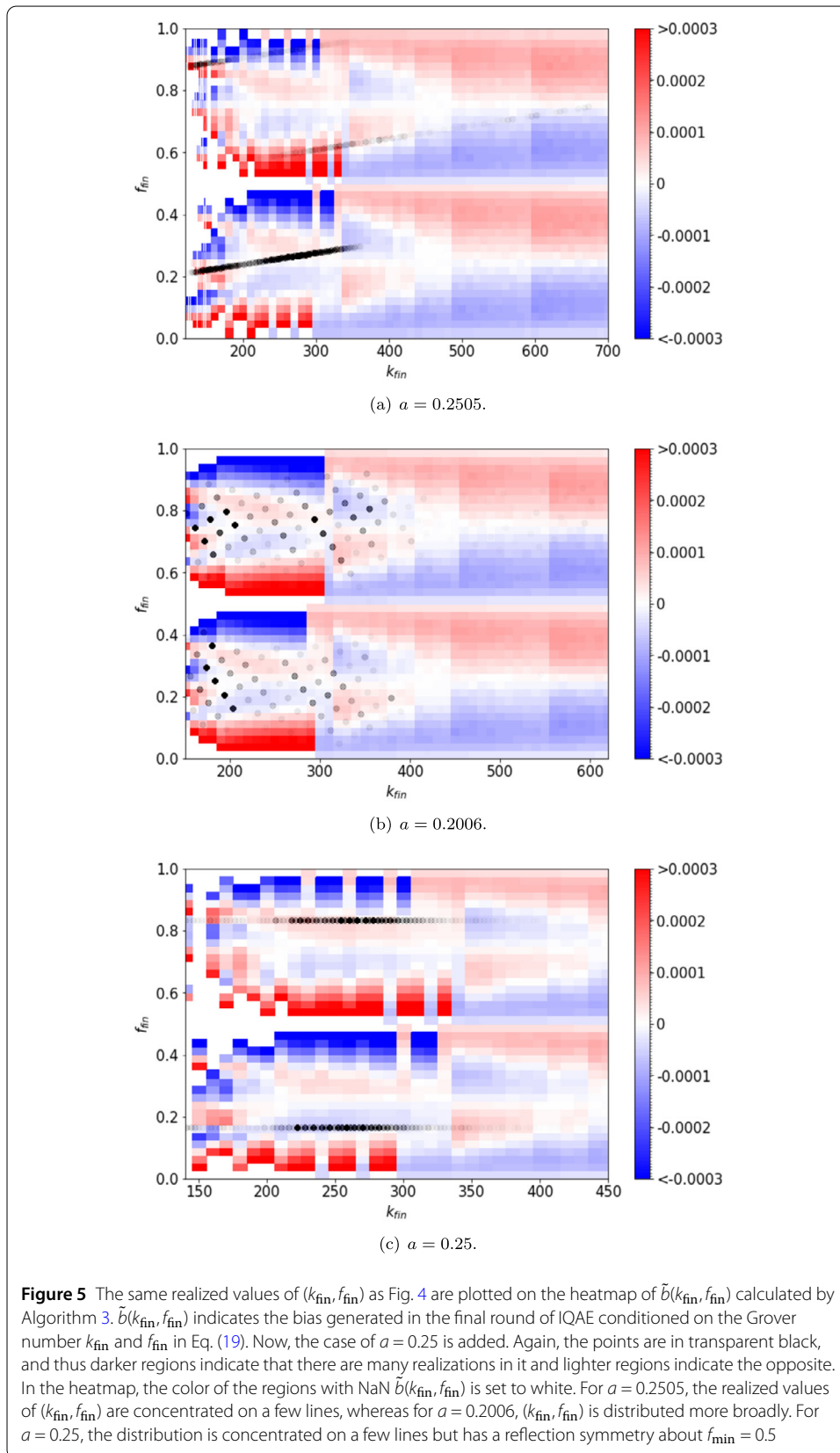
$$\theta_a = \frac{l\pi}{2m} + \delta \quad (24)$$

with small  $m \in \mathbb{N}$  and  $\delta \in \mathbb{R}$  such that  $|\delta| \ll 1$ . Concretely,  $l = 1$ ,  $m = 3$  and  $\delta \approx 0.000577$ . For this  $a$ ,  $f_{\theta_a}(k)$  consists of only three slightly tilted lines shown in Fig. 4(a) as solid lines, which were dashed horizontal lines in the same figure if  $\delta = 0$ .

We now plot the distribution of  $(k_{\text{fin}}, f_{\text{fin}})$  on the heatmap of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$ , which is highly illustrative for understanding why the bias is enhanced for specific values of  $a$ . In Fig. 5, we show the results for  $a = 0.2505, 0.2006,$  and, in addition,  $a = 0.25$ , for which the 10,000 runs of Algorithm 1' are conducted too.

We obtain  $b(a)$  by averaging the values of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  over the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$ , which is represented by the black points, with the weight proportional to the realization frequency, which is represented by the darkness of the points. In the case of  $a = 0.2505$ , as seen above, the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  are not broadly distributed but located on a few specific lines. Contrary to this, in the case of  $a = 0.2006$ , the points are distributed more broadly, and the values of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  on the points take various values, both positive and negative, canceling out each other when averaged. This cancellation tends not to occur in the case of  $a = 0.2505$  due to the concentration of the points. In fact, we see that the lines consisting of the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  mainly go through the regions where  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  is positive, which leads to positive  $b(a)$ .

Of course, even if Eq. (24) holds and thus the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  concentrate on a few lines, the values of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  on the points may accidentally cancel out each other, resulting in a relatively small  $b(a)$ . Conversely, even if the points are distributed broadly, the





distribution is not uniform and the heatmap of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  is not completely antisymmetric, which means that the bias cancellation is not perfect. Nevertheless, the distribution of realized  $(k_{\text{fin}}, f_{\text{fin}})$  and the cancellation of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  on the points roughly describe the phenomenon that the bias is enhanced for the specific values of  $a$ .

We remark on the case that  $\theta_a$  can be exactly written in the form of  $l\pi/2m$  with small  $m$ . For example,  $a = 0.25$  gives  $\theta_a = l\pi/2m$  with  $l = 1$  and  $m = 3$ . Although  $a = 0.25$  has only a slight difference from  $a = 0.2505$ ,  $a = 0.25$  does not yield the large bias: for it,  $b(a = 0.25) = 3.5 \times 10^{-6}$ , which is one order of magnitude smaller than  $b(a = 0.2505) = 3.7 \times 10^{-5}$ . Figure 5(c) shows the distribution of realized  $(k_{\text{fin}}, f_{\text{fin}})$  on the heatmap of  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  for  $a = 0.25$ . We see that the points are distributed on the two horizontal lines<sup>3</sup> located at the symmetric positions with respect to reflection about the line  $f_{\text{fin}} = 1/2$ . Since  $\tilde{b}(k_{\text{fin}}, f_{\text{fin}})$  is nearly antisymmetric about this line, its values at the realized points largely cancel out each other, yielding the relatively small  $b(a)$ .

### 3.3 Mitigation of the bias

Finally, we propose a simple method to mitigate the bias: just re-executing the final round. Namely, we modify Algorithm 1 as follows. Suppose that we get the CI of  $a$  such that  $\Delta a_{i,j} \leq \epsilon$  in the round with the Grover number  $k_{\text{fin}}$  and the total shot number  $N_{\text{fin}}$ . Then, we perform just one additional round using the same  $k_{\text{fin}}$  and  $N_{\text{fin}}$ , and let the resultant maximum likelihood estimate  $\hat{a}$  in the added round be the output of the algorithm. In this additional round, we do not impose the termination criterion  $\Delta a_{i,j} \leq \epsilon$ . We make  $N_{\text{fin}}$  measurements on  $G^{k_{\text{fin}}}|\Phi\rangle$  even if  $\Delta a_{i,j} \leq \epsilon$  is satisfied in the middle. We show the modified algorithm as Algorithm 4.

The reason why we expect that this re-execution-based bias mitigation works is as follows. Recall that, in Algorithm 1, the error is generated only in the final round as long as the second-to-last round yields the CI of  $a$  enclosing the true value, and the bias is induced by the termination criterion  $\Delta a_{i,j} \leq \epsilon$ . We thus consider that running an additional round without the criterion  $\Delta a_{i,j} \leq \epsilon$  and taking the result as the final output will mitigate the bias.

---

#### Algorithm 4 Modified IQAE with the final round re-executed

---

**Input:**  $\epsilon, \alpha \in (0, 1)$ ,  $N_{\text{shot}} \in \mathbb{N}$ ,  $r_{\text{min}} > 1$

- 1: Run Algorithm 1. Let the Grover number, the total shot number, and  $R_i$  in the final round be  $k_{\text{fin}}$ ,  $N_{\text{fin}}$ , and  $R_{\text{fin}}$ , respectively.
- 2: Iterate generating  $G^{k_{\text{fin}}}|\Phi\rangle$  and measuring it  $N_{\text{fin}}$  times. Let the number of times  $|\phi_1\rangle$  is obtained be  $N_1$ .
- 3: **return**

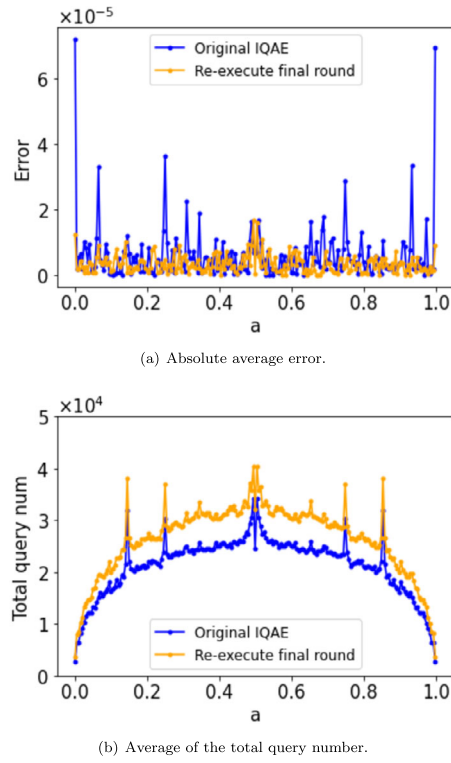
$$\sin^2 \left( \frac{R_{\text{fin}} \frac{\pi}{2} + \gamma \left( \sqrt{\frac{N_1}{N_{\text{fin}}}}, R_{\text{fin}} \right)}{2k_{\text{fin}} + 1} \right) \quad (25)$$

as  $\hat{a}$ .

---

<sup>3</sup>Although  $f_{\text{fin}}$  can take  $\frac{1}{2}$  for  $a = 0.25$ , none of the 10,000 runs ended with  $f_{\text{fin}} = \frac{1}{2}$ , and thus there is no point on the line  $f_{\text{fin}} = \frac{1}{2}$  in Fig. 5(c).

**Figure 6** For the various values of the amplitude  $a$ , we plot (a)  $|\bar{b}(a)|$ , the absolute value of the average of the errors in the 10,000 runs of IQAE (Eq. (13)) and (b) the average number of queries to  $G$  in the same runs of IQAE. The blue line denotes the result for Algorithm 1, which does not incorporate bias mitigation, and the orange one denotes that for Algorithm 4, which includes bias mitigation



In Fig. 6(a), we plot  $|\bar{b}(a)|$  in Algorithm 1, which is plotted also in Fig. 1, and that in Algorithm 4. As this figure shows, the enhanced biases for the specific values of  $a$  in Algorithm 1 are largely reduced in Algorithm 4, which indicates that re-executing the final round mitigates the bias. If we take, among the examined values of  $a$ , those for which  $|\bar{b}(a)| \geq 2\bar{\sigma}(a)$ , the average and maximum of the bias reduction rates for these values of  $a$  are 57.8% and 99.2%, respectively.

An obvious drawback of this bias mitigation method is the increase of the total number of queries to  $G$  by running an additional round. We have confirmed that, at least in our numerical experiments, this query number increase is mild. Figure 6(b) shows the average number of queries to  $G$  in the 10,000 run of Algorithm 1 and Algorithm 4.<sup>4</sup> The ratio of the average query number is about 1.25 for any examined value of  $a$ . This mild increase is because the final round in Algorithm 1 is not dominant among all the rounds with respect to query number. Namely, on average, the final round accounts for about 25% of the total number of queries to  $G$  across all the rounds.

#### 4 Summary

In this paper, we focused on the bias in IQAE, a widely studied version of QAE. We saw that the bias is enhanced for the specific values of the estimated amplitude  $a$ . The termination criterion that the estimated accuracy  $\Delta a_{i,j}$  of  $a$  falls below the predetermined accuracy  $\epsilon$

<sup>4</sup>The reason why the curves in Fig. 6(b) have several peaks is as follows. In each round in Algorithm 1, we search the next Grover number  $k_{i+1}$  by Algorithm 2, and depending on the current Grover number  $k_i$  and the true angle  $\theta_o$ , this search tends to require many shots in order to narrow down the CI  $[\theta_{i,j}^l, \theta_{i,j}^h]$  of  $\theta_o$  so much that  $[K_{i+1}\theta_{i,j}^l, K_{i+1}\theta_{i,j}^h]$  lies within a single quadrant. This can happen for any  $\theta_o$ , but for specific values of  $\theta_o$ , the probability of this phenomenon is relatively high, which leads to the larger total query number in expectation.

is a source of the bias. Decomposing the bias into the bias conditioned by  $k_{\text{fin}}$  the Grover number in the final round, we found a key factor that determines the magnitude of the bias: the distribution of the realized values of  $(k_{\text{fin}}, f_{\text{fin}})$  in the landscape of the conditional bias. Here,  $f_{\text{fin}}$  is defined as Eq. (19), and a main factor to determine the probability distribution of the IQAE estimate  $\hat{a}$  and thus its bias. We found that for  $a$  such that Eq. (24) holds with small  $m$  and tiny  $\delta$ , the points of realized  $(k_{\text{fin}}, f_{\text{fin}})$  are located only on a few lines in a 2D plane, and the conditional biases at the points do not tend to cancel each other so much, resulting in the large bias. We also proposed a simple bias mitigation method by re-executing the final round with the same Grover number  $k_{\text{fin}}$  and shot number  $N_{\text{fin}}$ . We saw that the increase of the total number of queries to  $G$  is mild, about 25%.

#### Abbreviations

QAE, quantum amplitude estimation; IQAE, iterative quantum amplitude estimation; QPE, quantum phase estimation; QMCI, quantum algorithm for Monte Carlo integration; MLE, maximum likelihood estimation; MLEQAE, maximum likelihood estimation-based quantum amplitude estimation; CI, confidence interval.

#### Acknowledgements

Not applicable.

#### Author contributions

KM as the sole author of the manuscript, conceived, designed, and performed the analysis; he also wrote and reviewed the paper. The author read and approved the final manuscript.

#### Funding

The author is supported by MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant no. JPMXS0120319794, JSPS KAKENHI Grant no. JP22K11924, and JST COI-NEXT Program Grant No. JPMJPF2014.

#### Data Availability

The data generated and analysed during the current study can be obtained from the corresponding author upon a reasonable request.

## Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

The author has approved the publication. The research in this work did not involve any human, animal or other participants.

#### Competing interests

The authors declare no competing interests.

Received: 13 December 2023 Accepted: 17 June 2024 Published online: 25 June 2024

## References

1. Brassard G, Hoyer P, Mosca M, Tapp A. Quantum amplitude amplification and estimation. *Contemp Math.* 2002;305:53.
2. Kitaev AY. Quantum measurements and the Abelian stabilizer problem. *Electron Colloq Comput Complex.* 1995;3:22.
3. Montanaro A. Quantum speedup of Monte Carlo methods. *Proc R Soc A.* 2015;471:20150301.
4. Rebstroff P, Gupt B, Bromley TR. Quantum computational finance: Monte Carlo pricing of financial derivatives. *Phys Rev A.* 2018;98:022321.
5. Stamatopoulos N, Egger DJ, Sun Y, Zoufal C, Iten R, Shen N, Woerner S. Option pricing using quantum computers. *Quantum.* 2020;4:291.
6. Chakrabarti S, Krishnakumar R, Mazzola G, Stamatopoulos N, Woerner S, Zeng WJ. A threshold for quantum advantage in derivative pricing. *Quantum.* 2021;5:463.
7. Miyamoto K. Bermudan option pricing by quantum amplitude estimation and Chebyshev interpolation. *EPJ Quantum Technol.* 2022;9:3.
8. Kaneko K, Miyamoto K, Takeda N, Yoshino K. Quantum pricing with a smile: implementation of local volatility model on quantum computer. *EPJ Quantum Technol.* 2022;9:7.
9. Doriguello JF, Luongo A, Bao J, Rebstroff P, Santha M. Quantum algorithm for stochastic optimal stopping problems with applications in finance. In: 17th conference on the theory of quantum computation, communication and cryptography (TQC 2022). 2022. p. 2:1–2:24.
10. Suzuki Y, Uno S, Raymond R, Tanaka T, Onodera T, Yamamoto N. Amplitude estimation without phase estimation. *Quantum Inf Process.* 2020;19:75.

11. Aaronson S, Rall P. Quantum approximate counting, simplified. In: Symposium on simplicity in algorithms (SOSA). 2020. p. 24–32.
12. Nakaji K. Faster amplitude estimation. *Quantum Inf Comput* 2020;20:1109–23. <https://doi.org/10.26421/QIC20.13-14-2>.
13. Grinko D, Gacon J, Zoufal C, Woerner S. Iterative quantum amplitude estimation. *npj Quantum Inf*. 2021;7:52.
14. Tanaka T, Suzuki Y, Uno S, Raymond R, Onodera T, Yamamoto N. Amplitude estimation via maximum likelihood on noisy quantum computer. *Quantum Inf Process*. 2021;20:293.
15. Giurgica-Tiron T, Kerenidis I, Labib F, Prakash A, Zeng W. Low depth algorithms for quantum amplitude estimation. *Quantum*. 2022;6:745.
16. Uno S, Suzuki Y, Hisanaga K, Raymond R, Tanaka T, Onodera T, Yamamoto N. Modified Grover operator for quantum amplitude estimation. *New J Phys*. 2021;23:083031.
17. Wada K, Fukuchi K, Yamamoto N. Quantum-enhanced mean value estimation via adaptive measurement. 2022. <https://doi.org/10.48550/arXiv.2210.15624>.
18. Tanaka T, Uno S, Onodera T, Yamamoto N, Suzuki Y. Noisy quantum amplitude estimation without noise estimation. *Phys Rev A*. 2022;105:012411.
19. Callison A, Browne DE. Improved maximum-likelihood quantum amplitude estimation. 2022. <https://doi.org/10.48550/arXiv.2209.03321>.
20. Fukuzawa S, Ho C, Irani S, Zion J. Modified iterative quantum amplitude estimation is asymptotically optimal. In: 2023 proceedings of the symposium on algorithm engineering and experiments (ALENEX). 2023. p. 135–47.
21. Lu X, Lin H. Random-depth quantum amplitude estimation. 2023. <https://doi.org/10.48550/arXiv.2301.00528>.
22. Cornelissen A, Hamoudi Y. A sublinear-time quantum algorithm for approximating partition functions. In: Proceedings of the 2023 annual ACM-Siam symposium on discrete algorithms (SODA). 2023. p. 1245–64.
23. Jerrum MR, Valiant LG, Vazirani VV. Random generation of combinatorial structures from a uniform distribution. *Theor Comput Sci*. 1986;43:169.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---