Amplitude estimation via maximum likelihood on noisy quantum computer

Tomoki Tanaka[1,2,3] · Yohichi Suzuki[1] · Shumpei Uno[1,4] · Rudy Raymond[1,5] · Tamiya Onodera[1,5] · Naoki Yamamoto[1,6]

Abstract

Recently we find several candidates of quantum algorithms that may be implementable in near-term devices for estimating the amplitude of a given quantum state, which is a core subroutine in various computing tasks such as the Monte Carlo methods. One of those algorithms is based on the maximum likelihood estimate with parallelized quantum circuits. In this paper, we extend this method so that it incorporates the realistic noise effect, and then give an experimental demonstration on a superconducting IBM Quantum device. The maximum likelihood estimator is constructed based on the model assuming the depolarization noise. We then formulate the problem as a two-parameters estimation problem with respect to the target amplitude parameter and the noise parameter. In particular we show that there exist anomalous target values, where the Fisher information matrix becomes degenerate and consequently the estimation error cannot be improved even by increasing the number of amplitude amplifications. The experimental demonstration shows that the proposed maximum likelihood esti-

✉ Tomoki Tanaka
ttomoki@user.keio.ac.jp

Naoki Yamamoto
yamamoto@appi.keio.ac.jp

[1] Quantum Computing Center, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8522, Japan

[2] Mitsubishi UFJ Financial Group, Inc. and MUFG Bank, Ltd., 2-7-1 Marunouchi, Chiyoda-ku, Tokyo 100-8388, Japan

[3] Graduate School of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223- 8522, Japan

[4] Mizuho Research & Technologies, Ltd., 2-3 Kanda-Nishikicho, Chiyoda-ku, Tokyo 101-8443, Japan

[5] IBM Quantum, IBM Research - Tokyo, 19-21 Nihonbashi Hakozaki-cho, Chuo-ku, Tokyo 103-8510, Japan

[6] Department of Applied Physics and Physico-Informatics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8522, Japan

mator achieves quantum speedup in the number of queries, though the estimation error saturates due to the noise. This saturated value of estimation error is consistent to the theory, which implies the validity of the depolarization noise model and thereby enables us to predict the basic requirement on the hardware components (particularly the gate error) in quantum computers to realize the quantum speedup in the amplitude estimation task.

Keywords Quantum computing · Amplitude estimation · Maximum likelihood estimation · Depolarizing noise · IBM quantum systems

1 Introduction

Ideal and fault-tolerant quantum computers will provide us with game-changing platforms in various area such as security [1], chemical [2,3], and financial engineering [4–10]. Although quantum devices available now are all small and noisy [11], the recent progress in hardware [12] is certainly going to bridge this gap. At the same time various software methods, which are expected to lower the hurdle for realizing quantum computing, are now being developed as well [13–15]. With careful evaluation of both of these approaches, it is important to have a right prediction on what applications of quantum computing can be realized and when they might emerge in reality; for instance [16,17]. The quantum volume [18] is one reasonable measure along with discussing such predictions.

This paper focuses on the quantum amplitude estimation algorithm [19,20], which can be typically applied to speedup the classical Monte Carlo methods [4,5], etc; more precisely, in an ideal setup the quantum amplitude estimation algorithm can quadratically reduce the number of samples and thereby the computation time for Monte Carlo methods. Because the standard amplitude estimation algorithm is demanding to implement, due to several technical reasons including the use of many ancilla qubits for quantum Fourier transform and the use of many controlled gate operations, recently some new techniques that circumvent these challenges have been developed [21–24]. In particular, our approach [21] takes the maximum likelihood (ML) method to estimate the amplitude without both the ancilla qubits and the controlled operation conditioned on those qubits, thereby drastically reducing the number of quantum gates (particularly the controlled NOT gate) involved in this algorithm compared to the standard one; also the near quadratic speedup was demonstrated in a numerical simulation in the ideal setup.

This paper extends the ML method [21] so that it can be used for actual noisy quantum computers, and then gives an experimental demonstration on a superconducting quantum device. The point of the ML method is that, while the true probability distribution generating the data is in general unknown, the ML estimator is constructed based on a suitable model distribution. In our case, the true distribution is determined from the output state of a noisy quantum computer operated under several imperfections that are impossible to perfectly characterize. To approximate such unknown noise source for the purpose of constructing the ML estimator, in this work, we consider the depolarizing noise model; this is often chosen as the minimal or the worst-case noise

model [25–27] which is also used for quantifying the gate fidelity of a given quantum computing device, i.e., the randomized benchmarking [28–30]. We also add a fact that the depolarization noise process has a clear mathematical merit in that it commutes with an arbitrary unitary operation, which thereby enables analytic treatment for our estimation problem. We then formulate the problem as a two-parameters estimation problem with respect to the target amplitude parameter and the noise parameter. This problem formulation introduces a new important aspect to the near-term quantum computing field, in the sense that the unavoidable noise coming from the realistic imperfection has also to be estimated as a nuisance parameter. In our problem, thanks to the property of depolarizing noise, we have an explicit form of the Fisher information matrix for discussing the accuracy of estimation and thereby derive a formula for specifying the noise level so that near-quadratic speedup is achieved to reach a given estimation accuracy. Furthermore, the explicit formula of Fisher information matrix reveals the existence of anomalous case, where the target parameter cannot be efficiently estimated because the Fisher information matrix becomes degenerate. Note that, hence, such anomalous target parameter appears only in the multi-parameters estimation problem. Fortunately we can provide a simple way to circumvent this difficulty.

Below we show the organization of the paper, together with the summary of the results obtained in this paper

Section 2: We take the depolarizing noise model and then formulate the two-parameters estimation problem. With this noise model we can have the analytic expression of the Fisher information matrix, which gives an asymptotically achievable lower bound of the estimation error. This result is further used to derive a condition of the noise level required to have nearly quadratic speedup to reach a given estimation error. Furthermore, we show the anomalous case where the Fisher information matrix is degenerate and consequently the target parameter cannot be efficiently estimated; a simple strategy to circumvent this issue is provided.

Section 3: We give an experimental demonstration of our ML method for a simple Monte Carlo integration problem, using 2- and 3-qubits IBM Quantum devices [31,32]. The result is that, for the former case, a quantum speedup in the number of queries over the classical one is observed, while the estimation error saturates as the query becomes large. On the other hand, for the 3-qubit case, we found that the estimation error is always bigger than that via the classical method. For both cases the saturated value is consistent to the theoretical prediction, implying the validity of the depolarization noise model.

Section 4: We give the following two discussions. First, based on the condition obtained in Sect. 2 together with the experimental result shown in Sect. 3, we discuss the hardware requirements, e.g., the error rates of single-qubit and CNOT gates, to achieve a given precision for estimating the value of multidimensional integration. Second, the computational complexity of our proposed algorithm is discussed.

2 Analysis of estimation error under depolarizing noise

2.1 Preliminary

We briefly review the amplitude estimation algorithm that uses the ML method on parallelized circuits [21]. This algorithm can estimate the parameter (i.e., the unknown amplitude) quadratically faster than any classical sampling method to attain a given precision, even without the standard phase estimation subroutine. The algorithm mainly consists of two components: one is the amplitude amplification process [19,20], which is the generalized version of Grover's quantum search algorithm, and the other is the classical ML estimation part for the data obtained by the amplitude amplification. The essential idea of the algorithm is sketched below.

The amplitude amplification algorithm initially prepares the state given by $|\Psi\rangle_{n+1} = \mathcal{A}|0\rangle_{n+1}$, where $\mathcal{A}$ is a unitary operator acting on the $(n + 1)$ qubits. The operator $\mathcal{A}$ is designed to satisfy $\mathcal{A}|0\rangle_{n+1} = \sqrt{a}|\tilde{\Psi}_1\rangle_n|1\rangle + \sqrt{1-a}|\tilde{\Psi}_0\rangle_n|0\rangle$, where $a \in [0, 1]$ is an unknown parameter to be estimated, and $|\tilde{\Psi}_1\rangle_n$ and $|\tilde{\Psi}_0\rangle_n$ are the n-qubit normalized "good" and "bad" states. In terms of $\theta_a \in [0, \pi/2]$ satisfying $\sin^2 \theta_a = a$, the prepared state $|\Psi\rangle_{n+1}$ can also be expressed as

$$|\Psi\rangle_{n+1} = \sin \theta_a |\tilde{\Psi}_1\rangle_n|1\rangle + \cos \theta_a |\tilde{\Psi}_0\rangle_n|0\rangle. \tag{1}$$

The probability to measure the good state can be amplified by applying the unitary operator $\mathbf{Q} = -\mathcal{A}\mathbf{S}_0\mathcal{A}^{-1}\mathbf{S}_\chi$ on $|\Psi\rangle_{n+1}$, where $\mathbf{S}_\chi = \mathbf{I}_{n+1} - 2\mathbf{I}_n \otimes |1\rangle\langle 1|$ and $\mathbf{S}_0 = \mathbf{I}_{n+1} - 2|0\rangle_{n+1}\langle 0|$. Also $\mathbf{I}_k$ is an identity operator on k qubits. By applying $\mathbf{Q}$ on $|\Psi\rangle_{n+1}$ for m times, we have

$$\mathbf{Q}^m|\Psi\rangle_{n+1} = \sin((2m + 1)\theta_a)|\tilde{\Psi}_1\rangle_n|1\rangle + \cos((2m + 1)\theta_a)|\tilde{\Psi}_0\rangle_n|0\rangle. \tag{2}$$

This equation tells that the probability for measuring the good state is amplified depending on the number of repetitions of $\mathbf{Q}$ on $|\Psi\rangle_{n+1}$. Note that Eq. (2) is valid only in the absence of noise.

Next, we describe the ML part. The idea is to estimate the amplitude a using the number of measuring the good state after performing m_k (for $0 \le k \le M$) repetitions of $\mathbf{Q}$. Now for the ideal state $\mathbf{Q}^{m_k}|\Psi\rangle_{n+1}$, the probability measuring the good state is $P(m_k; a) = \sin^2((2m_k+1)\theta_a)$; then the probability to have the good state h_k times out of the total N_k measurement shots is proportional to $[P(m_k; a)]^{h_k}[1 - P(m_k; a)]^{N_k-h_k}$ and accordingly the likelihood function is

$$L(\mathbf{h}; a) = \prod_{k=0}^{M} [P(m_k; a)]^{h_k}[1 - P(m_k; a)]^{N_k-h_k}, \tag{3}$$

where $\mathbf{h} = (h_0, h_1, \ldots, h_M)$. The ML estimate of a is then given by $\text{argmax}_a L(\mathbf{h}; a)$, which was proven to achieve the theoretical lower bound in the estimation error (the detail is described just below). Note that, of course, the estimated precision depends on the choice of a sequence of integers $\{m_k\}$. In this paper, we consider the following

Linearly Incremental Sequence (LIS) and Exponential Incremental Sequence (EIS):

$$\text{(LIS)} \ m_k = k, \quad \text{(EIS)} \ m_k = \lfloor 2^{k-1} \rfloor, \ \text{for } 0 \le k \le M, \tag{4}$$

where in the hereafter we omit the floor notation for simplicity. To asymptotically achieve a given error ϵ, LIS and EIS need $O(1/\epsilon^{4/3})$ and $O(1/\epsilon)$ total query calls, respectively, while the classical case (i.e., $\forall k : m_k = 0$) needs $O(1/\epsilon^2)$ to achieve the same precision [21].

Lastly we give a general framework of the above-described method for the vector of multiple parameters $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_K]^\top$ under noisy environment, which is indeed the scenario studied in this paper ($\bullet^\top$ denotes the transpose). Let $P(m_k; \boldsymbol{\theta})$ be the probability of measuring the good state for a density matrix obtained by applying $\mathbf{Q}^{m_k}$ on the state (1) under noisy environment. Then the likelihood function corresponding to the probability having the good state h_k times in total N_k measurement shots is, similar to the above, given by

$$L(\mathbf{h}; \boldsymbol{\theta}) = \prod_{k=0}^{M} [P(m_k; \boldsymbol{\theta})]^{h_k} [1 - P(m_k; \boldsymbol{\theta})]^{N_k - h_k}, \tag{5}$$

where again $\mathbf{h} = (h_0, h_1, \ldots, h_M)$. The ML estimate $\hat{\boldsymbol{\theta}}_{\text{ML}}$ is defined as

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \text{argmax}_{\boldsymbol{\theta}} L(\mathbf{h}; \boldsymbol{\theta}) = \text{argmax}_{\boldsymbol{\theta}} \ln L(\mathbf{h}; \boldsymbol{\theta}). \tag{6}$$

In general, the estimation error covariance matrix $\text{Cov}(\hat{\boldsymbol{\theta}}) = \mathbb{E}[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^{\text{T}}]$, with $\hat{\boldsymbol{\theta}}$ any unbiased estimate, satisfies the Cramér–Rao inequality:

$$\text{Cov}(\hat{\boldsymbol{\theta}}) \ge \mathcal{I}^{-1}(\boldsymbol{\theta}), \tag{7}$$

where $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information matrix defined as

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}\left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln L(\mathbf{h}; \boldsymbol{\theta}) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln L(\mathbf{h}; \boldsymbol{\theta}) \right)^\top \right]. \tag{8}$$

The expectation value $\mathbb{E}[\bullet]$ in Eq. (8) is defined as

$$\mathbb{E}[X(\mathbf{h})] = \sum_{h_0=0}^{N_0} \sum_{h_1=0}^{N_1} \cdots \sum_{h_M=0}^{N_M} X(\mathbf{h}) \left[\prod_{k=0}^{M} \binom{N_k}{h_k} P(m_k; \boldsymbol{\theta})^{h_k} (1 - P(m_k; \boldsymbol{\theta}))^{N_k - h_k} \right], \tag{9}$$

where $X(\mathbf{h})$ is a function of random variables $\mathbf{h}$. It is known that the ML estimate attains the lower bound of $\text{Cov}(\hat{\boldsymbol{\theta}})$ in Eq. (7), asymptotically in the limit of large samples. The elements of the Fisher information matrix are given as

$$[\mathcal{I}(\boldsymbol{\theta})]_{i,j} = \mathbb{E}\left[\left(\frac{\partial}{\partial \theta_i} \ln L(\mathbf{h}; \boldsymbol{\theta}) \right) \left(\frac{\partial}{\partial \theta_j} \ln L(\mathbf{h}; \boldsymbol{\theta}) \right) \right]. \tag{10}$$

In the case of the multi-parameter estimation problem, the ML estimation of the i-th and j-th ($i \neq j$) parameters can be performed independently, if the (i, j) element of the Fisher information matrix is zero. However, otherwise, the i-th and j-th parameters are correlated, and the estimation of parameter of interest may be adversely affected by other nuisance parameters, which is indeed the case in this work as shown later.

2.2 Fisher information in the presence of depolarizing noise

Quantum states are usually disturbed by noise that comes from the system-environment interaction. Then the ideal pure state (2) is replace by a mixed state, which is yet impossible to perfectly identify. On the other hand, the ML estimator is based on a model state which may well approximate such an unknown true mixed state. In this work, we assume the depolarization channel defined by [33,34]:

$$\mathcal{D}(\rho) = p\rho + (1 - p)\frac{\mathbf{I}_{n+1}}{d}. \tag{11}$$

Here ρ is a density matrix, $\mathcal{D}(\rho)$ is a completely positive trace-preserving (CPTP) map which represents the depolarization of qubits, $1 - p$ is the error probability that qubits are depolarized, and d is the dimension of the quantum system, i.e., $d = 2^{n+1}$. Note that p should also be treated as an unknown parameter, in addition to a; hence we are studying the two-parameter estimation problem, which is essentially harder compared to the one-parameter problem studied in [21].

Now, the CPTP map of the ideal amplitude amplification channel, in terms of the density matrix, is represented as

$$\mathcal{Q}(\rho) = \mathbf{Q}\rho\mathbf{Q}^\dagger. \tag{12}$$

From Eqs. (11) and (12), the amplitude amplification process in the presence of noise is thus given by

$$\rho_{\text{noise}} = \mathcal{Q}\mathcal{D}(\rho) = \mathcal{D}\mathcal{Q}(\rho) = p\mathbf{Q}\rho\mathbf{Q}^\dagger + (1 - p)\frac{\mathbf{I}_{n+1}}{d}. \tag{13}$$

Moreover, it is easy to see that m times repetition of this noisy amplitude amplification process end up with

$$\rho_{\text{noise}}^{(m)} = (\mathcal{Q}\mathcal{D})^m(\rho) = p^m\mathbf{Q}^m\rho\mathbf{Q}^{\dagger m} + (1 - p^m)\frac{\mathbf{I}_{n+1}}{d}. \tag{14}$$

Now the initial state is chosen as $\rho = |\Psi\rangle_{n+1}\langle\Psi|$, where $|\Psi\rangle_{n+1}$ is given in Eq. (1). As in the ideal case, we are interested in the probability of measuring the good state with which the last qubit is found to be $|1\rangle$, i.e., $P(m; \boldsymbol{\theta}) = \text{Tr}(\rho_{\text{noise}}^{(m)} E_1)$, where $E_1 = \mathbf{I}_n \otimes |1\rangle\langle 1|$; also $\boldsymbol{\theta} = [a, \kappa]^\top$ is the vector of unknown parameters, where κ is defined as $\kappa = -\ln p$ which we refer as the noise level of the amplitude amplification

process. By using Eq. (2) and $\mathrm{Tr}(\mathbf{I}_{n+1}E_1) = 2^n$, this probability is calculated as

$$P(m; \boldsymbol{\theta}) = P(m; a, \kappa) = \frac{1}{2} - \frac{1}{2}e^{-\kappa m}\cos(2(2m+1)\theta_a). \tag{15}$$

Then the likelihood function (5) is represented as

$$L(\mathbf{h}; \boldsymbol{\theta}) = L(\mathbf{h}; a, \kappa) = \prod_{k=0}^{M} [P(m_k; a, \kappa)]^{h_k} [1 - P(m_k; a, \kappa)]^{N_k - h_k}. \tag{16}$$

Our task is to estimate $\boldsymbol{\theta} = [a, \kappa]^\top$ via the ML estimate (6), i.e., $\boldsymbol{\theta}_{\mathrm{ML}} = \mathrm{argmax}_{\boldsymbol{\theta}} L(\mathbf{h}; \boldsymbol{\theta})$, where $\mathbf{h} = (h_0, h_1, \ldots, h_M)$ is the set of data obtained in the experiment. As mentioned in the previous subsection, $\boldsymbol{\theta}_{\mathrm{ML}}$ asymptotically achieves the lower bound in the Cramér–Rao inequality (7) if the data is generated from the model distribution; the Fisher information matrix (8) in this case can now be calculated as

$$\mathcal{I}_{1,1}(a, \kappa) = \mathbb{E}\left[\left(\frac{\partial}{\partial a}\ln L(\mathbf{h}; a, \kappa)\right)^2\right]$$

$$= \sum_{k=0}^{M} \frac{N_k(2m_k + 1)^2}{\sin^2(2\theta_a)} \frac{4\sin^2\left(2\left(2m_k + 1\right)\theta_a\right)}{e^{2\kappa m_k} - \cos^2\left(2\left(2m_k + 1\right)\theta_a\right)},$$

$$\mathcal{I}_{1,2}(a, \kappa) = \mathcal{I}_{2,1}(a, \kappa) = \mathbb{E}\left[\left(\frac{\partial}{\partial a}\ln L(\mathbf{h}; a, \kappa)\right)\left(\frac{\partial}{\partial \kappa}\ln L(\mathbf{h}; a, \kappa)\right)\right]$$

$$= \sum_{k=0}^{M} \frac{N_k m_k(2m_k + 1)}{\sin(2\theta_a)} \frac{\sin\left(4\left(2m_k + 1\right)\theta_a\right)}{e^{2\kappa m_k} - \cos^2\left(2\left(2m_k + 1\right)\theta_a\right)},$$

$$\mathcal{I}_{2,2}(a, \kappa) = \mathbb{E}\left[\left(\frac{\partial}{\partial \kappa}\ln L(\mathbf{h}; a, \kappa)\right)^2\right]$$

$$= \sum_{k=0}^{M} N_k m_k^2 \frac{\cos^2\left(2\left(2m_k + 1\right)\theta_a\right)}{e^{2\kappa m_k} - \cos^2\left(2\left(2m_k + 1\right)\theta_a\right)}. \tag{17}$$

Recall that a is the parameter of our true interest; from Eq. (7), the estimation error of a is lower bounded by the (1, 1) element of the inverse of the above Fisher information matrix as

$$\epsilon = \sqrt{\mathbb{E}\left[(a - \hat{a})^2\right]} \geq \sqrt{\left(\mathcal{I}(a, \kappa)^{-1}\right)_{1,1}} =: \epsilon_{\min}(a, \kappa). \tag{18}$$

We use $\epsilon_{\min}(a, \kappa)$ to discuss the condition on the noise level κ to satisfy a required estimation precision. These topics are studied in detail in the next subsection.

Lastly we remark on other noise models. The most general (Markovian) noise model, with multiple parameters, can be represented by the Kraus superoperator. However, in general this does not commute with Grover operator, and as a result we cannot obtain an analytic expression of Cramér–Rao lower bound. To discuss the estimation

Springer
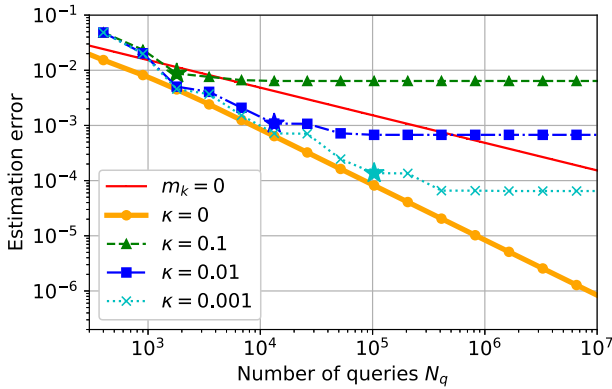
Fig. 1 Cramér–Rao lower bound $\epsilon_{\min}(a, \kappa)$ on the estimation error of a, versus the total number of queries. The red line corresponds to the classical case $\forall k : m_k = 0$, while the other lines are the quantum cases of EIS with several values of noise level κ. See the first paragraph of Sect. 2.3 for the details (Color figure online)

accuracy in such a case, several numerical methods have been developed in the field of quantum metrology [28]. Note that even for a special type of noise channel other than depolarization, it is still difficult to have an analytic expression of Cramér–Rao lower bound. For instance, in Ref. [34] considering the amplitude damping and dephasing noise model in the same amplitude estimation problem yet with known noise strength, the estimation error was evaluated numerically.

2.3 Achievable estimation error and required depolarizing noise level

Figure 1 shows the Cramér–Rao lower bound of the estimation error, $\epsilon_{\min}(a, \kappa)$, versus the total number of query calls, $N_q = \sum_{k=0}^{M} N_k(2m_k + 1)$. In particular in the classical case, it is given by $N_q = \sum_{k=0}^{M} N_k$. The target value is chosen to be $a = \sin^2 \theta_a = 0.375$, and $N_k = 100$ for all k. The (red) thin solid line represents the lower bound in the classical case. The other lines represent the lower bounds obtained when using the amplitude amplification with the EIS for several noise level κ; the (yellow) thick solid line is the bound without noise ($\kappa = 0$); the (green) dashed line with triangles, the (blue) dash-dotted line with squares, and the (light blue) dotted line with crosses are the lower bound under depolarizing noise $\kappa = 10^{-1}$, 10^{-2}, and 10^{-3}, respectively. Recall that, in the ideal case $\kappa = 0$, the number of query calls needed to reach a specified value of ϵ is $N_q \sim O(1/\epsilon)$, i.e., the Heisenberg-scaling. We note that a similar dependence of $\epsilon_{\min}(a, \kappa)$ on the magnitude of κ is observed for many cases of a, but there exist cases such that $\epsilon_{\min}(a, \kappa)$ takes much bigger values than those shown in Fig. 1 even when κ is sufficiently small; see Sect. 2.4 about this special case.

A notable point observed in Fig. 1 is that, even under the depolarizing noise model, the estimation error $\epsilon_{\min}(a, \kappa)$ decreases in nearly the Heisenberg-scaling law up to $N_q \sim 10^4$ and $N_q \sim 10^5$ for the cases $\kappa = 0.01$ and $\kappa = 0.001$, respectively. However, the error does not decrease anymore, even by using more queries. In other words, $\epsilon_{\min}(a, \kappa)$ gets saturated at those points of N_q, and thus the algorithm has to
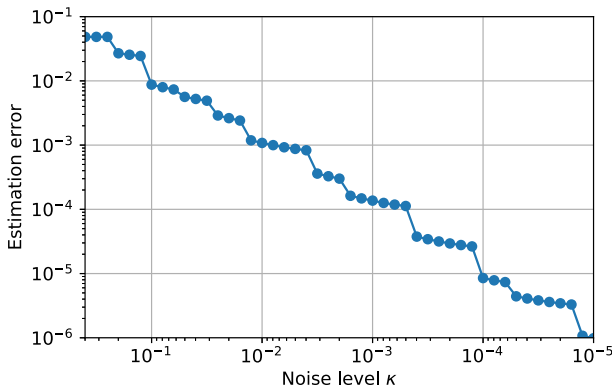
Fig. 2 Relationship between the noise level κ and the achievable error ϵ, in the case $a = 0.375$. The ML algorithm is with $\forall k : N_k = 100$ and EIS

be stopped[1]. The maximum number of queries within which the Heisenberg-scaling is guaranteed can be formally characterized as follows. That is, even under depolarizing noise model with κ, the Heisenberg-scaling is nearly preserved if the number of operations of $\mathcal{Q}$, m_k, is smaller than $\bar{m}$ defined as the maximum integer satisfying the following inequality [28]:

$$(2\bar{m} + 1)(1 - e^{-\kappa}) \leq 1. \tag{19}$$

This condition is derived as a sufficient condition to guarantee that the probability to measure the good state is not affected by the noise in the limit of large samples. The star marks in Fig. 1 are the total query calls $\bar{N}_q$ corresponding to $\bar{m}$ for given κ, showing that in fact the estimation error does not obey the Heisenberg-scaling law even by calling more queries than $\bar{N}_q$. Note that $\bar{m}$ given in Eq. (19) was originally derived in the one-parameter setting (that is, the case where κ is known), while $\epsilon_{\min}(a, \kappa)$ is a function of the two-parameter Fisher information matrix; despite of this gap $\bar{m}$ certainly captures the point of maximum number of query calls up to which the estimation error decreases according to the Heisenberg-scaling.

Based on the above-described fact, we obtain the condition on the noise level κ so that the ML algorithm reaches a specified estimation error ϵ with the number of query calls of the order $O(1/\epsilon)$, i.e., the Heisenberg-scaling, even under the noisy environment. Fig. 2 yields such condition; this is the relation between κ and the error at $\bar{N}_q$, for the case $a = \sin^2 \theta_a = 0.375$, $\forall k : N_k = 100$, and EIS. For instance, if we want to reach the estimation error $\epsilon = 10^{-4}$ using $O(1/\epsilon)$ query calls, then we need κ to be smaller than $\sim 10^{-3}$. Importantly, Fig. 2 indicates the "quasi linear relation" between a specified ϵ and the required value of κ; that is, if we need to decrease ϵ to $\epsilon/10$ then κ should be improved to simply $\kappa/10$. Note that this quasi linear relation is expected to hold from Eq. (19), which leads to $2\bar{m} + 1 = 1/\kappa$ when κ is small, together with the Heisenberg-scaling $\bar{N}_q = O(1/\epsilon)$, although Eq. (19)

---

[1] This saturation always occurs if $\kappa \neq 0$, which can be proven using the d'Alembert's ratio test together with the inequality $\epsilon_{\min}(a, \kappa) \geq \left(\mathcal{I}_{1,1}(a, \kappa)\right)^{-1/2} \geq \left(\sum_{k=0}^{M} \frac{N_k 4(2m_k+1)^2}{\sin^2 2\theta_a} \frac{e^{-2\kappa m_k}}{1 - e^{-2\kappa m_k}}\right)^{-1/2}$.
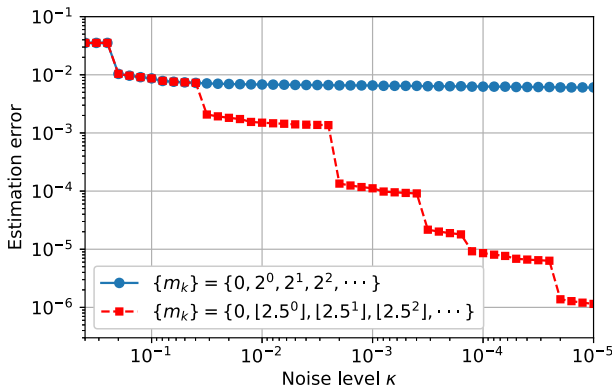
Fig. 3 Lower bound of the estimation error $\epsilon_{\min} = \sqrt{(\mathcal{I}^{-1})_{1,1}}$ versus the noise level κ. The target value is chosen as $a = \sin^2(\pi/8)$. The solid (blue) and dotted (red) lines are obtained when using $\{m_k\} = \{0, 2^0, 2^1, 2^2, \ldots\}$ and $\{m_k\} = \{0, \lfloor 2.5^0 \rfloor, \lfloor 2.5^1 \rfloor, \lfloor 2.5^2 \rfloor, \ldots\}$, respectively. $\lfloor \cdot \rfloor$ is the floor function (Color figure online)

was proven only in the one-parameter setting. The condition on κ can be further converted to that on the gate fidelity of elementary gates constructing the quantum circuit. That is, Fig. 2 represents the minimum hardware specification required to apply the quantum amplitude estimation method to solve a concrete problem such as a Monte Carlo integration task demonstrated later; a more detailed discussion on the hardware requirement will be given in Sect. 4.1.

The above discussion as well as Figs. 1 and 2 surely depend completely on the mathematical model described in Sect. 2.2. The real quantum system must not perfectly coincide with this model. The resulting ML estimate is then not guaranteed to reach the Cramér–Rao lower bound discussed here; also the quasi linear relation observed in Fig. 2 could be changed. That is, only with the materials posed up to now, we still could not say that Fig. 2 serves as a guide for discussing the condition on κ to reach a specified estimation error with Heisenberg-scaling law in the real world. This fully motivates us to execute some detailed experiment to see if the theory described above is consistent to the experimental result and thereby verify its usability to the real quantum computing applications; this topic will be discussed in Sect. 3.

Note that there is still a room for improvement on update strategy of m_k. For instance, if we employ $m_k = \lfloor r^{k-1} \rfloor$ where r is a real number which satisfies $r > 1$, the value of $\bar{N}_q$ changes depending on the value of r. This indicates that the achievable estimation error in the presence of noise can be reduced by changing the update strategy. More importantly, especially for the two-parameters problem considered in this paper, the estimation precision can be severely limited depending on the choice of r, which will be discussed in the next subsection.

2.4 Anomalous target value that induces a large estimation error

The previous discussion is based on the following two conditions; that is, the quasi linear relation between κ and $\epsilon_{\min}(a, \kappa)$, and the Heisenberg scaling of the total number
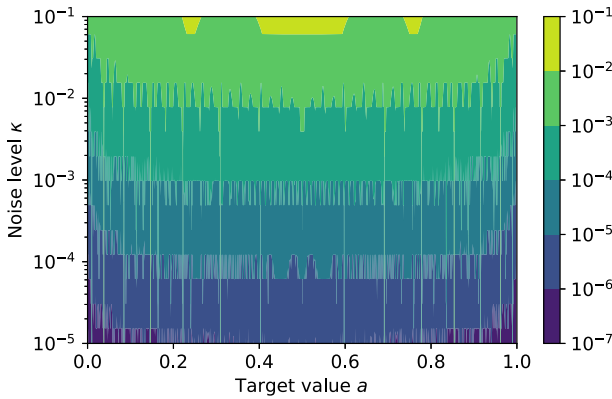
Fig. 4 Contour plot of the lower bound of the estimation error, $\epsilon_{\min} = \sqrt{(\mathcal{I}^{-1})_{1,1}}$, as a function of the target value a and the noise level κ

of queries $N_{\mathrm{q}} = \sum_{k=0}^{M} N_k (2m_k + 1)$ with respect to a given estimation error ϵ in the range $m_k \lesssim \bar{m}$. However, these conditions does not always hold. For instance, when the target value a is $a = \sin^2(\pi/8)$, the estimation error $\epsilon_{\min}(a, \kappa)$ does not obey the quasi linear relation with respect to κ, as shown by the blue solid line in Fig. 3. That is, even when the noise κ is sufficiently small, a precise estimate of a is not possible. This is due to the multi-parameter estimation setting, where in general the estimation error covariance matrix of the parameters $\boldsymbol{\theta}$ satisfies the Cramér–Rao inequality (7), i.e., $\mathrm{Cov}(\hat{\boldsymbol{\theta}}) \geq \mathcal{I}^{-1}(\boldsymbol{\theta})$, where $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information matrix. Clearly, if $\det \mathcal{I}(\boldsymbol{\theta})$ is nearly zero at a certain point of $\boldsymbol{\theta}$, then the estimation of those parameters has to be inefficient. Actually in the above-described example, our Fisher information matrix $\mathcal{I}(a, \kappa)$ is nearly degenerate at around $a = \sin^2(\pi/8)$; this is the reason why the quasi linear decrease of ϵ with respect to κ does not hold in this case. In this section, we investigate this "anomalous target" point of a in detail. But before moving forward, we would like to emphasize that this analysis never happen in the 1-parameter setting. That is, as stated in Sect. 1, to achieve quantum advantage on available noisy quantum devices, the noise parameter has to be incorporated into the parameters to be estimated and accordingly such a singular point analysis needs to be carried out.

First, to see how likely such an anomalous target of a exists, we plot the lower bound of the estimation error $\epsilon_{\min} = \sqrt{(\mathcal{I}^{-1})_{1,1}}$, as a function of a and κ in Fig. 4. This contour plot shows that, for almost all target values, the estimation error $\epsilon_{\min}$ almost linearly decreases with respect to κ; that is, $\epsilon_{\min}$ decreases approximately by one order of magnitude, as κ decreases by one order of magnitude. However, there exist anomalous target values of which estimation errors are insensitive to the value of κ; for instance, at around $a = \sin^2(\pi/8) = 0.146$, we observe a long spike where $\epsilon_{\min}$ takes almost the same value in the range $[1 \times 10^{-2}, 1 \times 10^{-3}]$ regardless of κ.

This insensitivity of $\epsilon_{\min}$ at a certain point of a is originated from the fact that, as mentioned in the first paragraph of this section, the Fisher information matrix $\mathcal{I}(a, \kappa)$ is nearly degenerate at around a. To closely look into the relation between the estimation error and the degeneracy of $\mathcal{I}(a, \kappa)$, in Fig. 5 we plot $\epsilon_{\min}$ with the solid blue line and
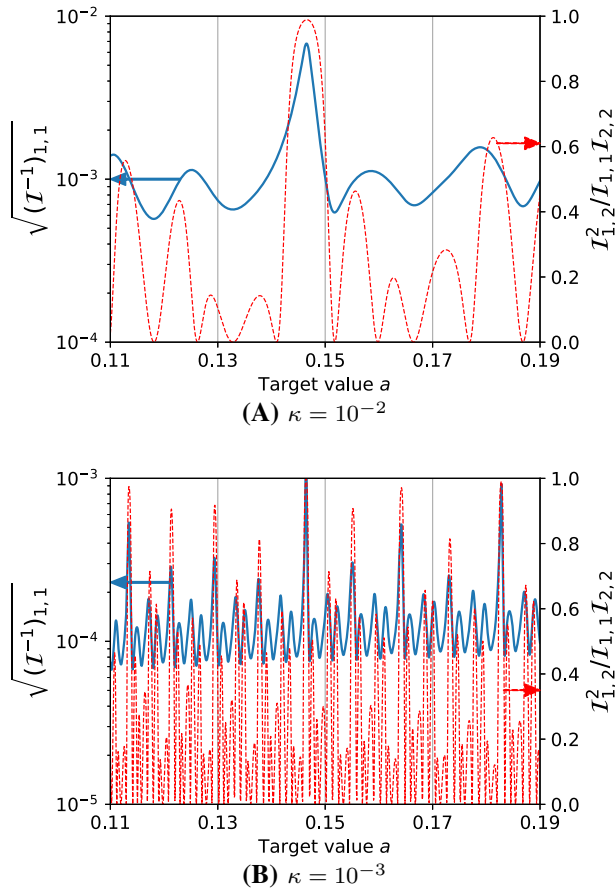
Fig. 5 Estimation error $\epsilon_{\min} = \sqrt{(\mathcal{I}^{-1})_{1,1}}$ represented with the solid (blue) line and the anomality $\beta = \mathcal{I}_{1,2}^2/\mathcal{I}_{1,1}\mathcal{I}_{2,2}$ represented with the dotted (red) line, as a function of the target value a. The noise parameter κ is fixed to (A) $\kappa = 10^{-2}$ and (B) $\kappa = 10^{-3}$ (Color figure online)

the "anomality" $\beta = \mathcal{I}_{1,2}^2/\mathcal{I}_{1,1}\mathcal{I}_{2,2}$ with the dotted red line, as a function of the target value a. The noise parameter κ is fixed to (A) $\kappa = 10^{-2}$ and (B) $\kappa = 10^{-3}$. Note that $0 < \beta \leq 1$ and $\beta = 1$ is equivalent to $\det \mathcal{I}(a, \kappa) = 0$; hence, $\beta \sim 1$ means that such (a, κ) are difficult to estimate precisely. Both of (A) and (B) of Fig. 5 indeed show that, at around the anomalous target of a where $\beta = 1$, the estimation error takes a relatively large value.

The existence ratio of anomalous target values can be quantitatively analyzed in terms of the *linear density* defined as follows. Here, $N = 10^5$ samples of a are randomly chosen from the uniform distribution on $[0, 1]$, and and the ratio of a satisfying $\beta > 0.9$ is computed; the linear density is given by this ratio. Table 1 shows the linear density for several values of κ. Importantly, the linear density is almost independent of the value of κ; it is about $1\% \sim 2\%$ regardless of κ. This is due to the composite of

Table 1 The linear density of anomalous targets in the range [0, 1]

κ	The linear density of anomalous targets a
10^{-10}	1.56 ± 0.04 %
10^{-9}	1.84 ± 0.04 %
10^{-8}	1.90 ± 0.04 %
10^{-7}	1.56 ± 0.04 %
10^{-6}	2.01 ± 0.04 %
10^{-5}	1.70 ± 0.04 %
10^{-4}	1.67 ± 0.04 %
10^{-3}	2.02 ± 0.04 %
10^{-2}	1.28 ± 0.04 %
10^{-1}	0.00 ± 0.00 %

The statistical error of the linear density is calculated under the assumption that the linear density obeys a binomial distribution

the following two properties of the anomalous targets: (i) the number of anomalous segments satisfying $\beta > 0.9$ is inversely proportional to κ, and (ii) the length of each anomalous segment is proportional to κ. Since the linear density of anomalous targets is approximately the product of the number of the anomalous regions with the average length of anomalous regions, it is almost insensitive to κ.

It should be noted that the linear density of anomalous targets takes finite value in the limit of $\kappa \to 0$, while the anomalous targets themselves are not present in the case of $\kappa = 0$. This is essentially originated from whether one has the information of the noise or not, i.e. the Fisher information $\mathcal{I}_{1,1}$ can be employed to obtain the lower bound of the estimation error in the absence of the noise, however $(\mathcal{I}^{-1})_{1,1}$ should be used if the noise level is unknown.

Finally, we propose two approaches to avoid the anomalous case. The first one is based on the observation that the determinant of the Fisher information matrix changes depending on the choice of the sequence $\{m_k\}$ of the amplitude amplification. Therefore, if the underlying target value is detected to be anomalous, then we can try another sequence $\{m_k\}$ to avoid the anomaly. For instance, when $\{m_k\} = \{0, \lfloor 2.5^0 \rfloor, \lfloor 2.5^1 \rfloor, \lfloor 2.5^2 \rfloor, \ldots\}$, the quasi linear relation between κ and $\epsilon_{\min}$ is recovered even when $a = \sin^2(\pi/8)$, as shown with the dotted red line in Fig. 3. This is a clear evidence showing that by suitably choosing $\{m_k\}$ the determinant of the Fisher information matrix does not get smaller. Our view is that it might be possible to detect the anomalous target by calculating the empirical Fisher information, which eventually allows us to tune the sequence and thereby avoid the anomaly.

The second approach is by modifying the amplitude to be estimated. After detecting the anomalous target, we introduce an extra ancilla qubit; then $R_y(\phi)$ rotation (i.e., the single qubit rotation around the y-axis with a fixed parameter ϕ) is applied to the ancilla qubit as follows:

$$
\begin{aligned}
\mathcal{A}|0\rangle_{n+2} =& \sqrt{a}|\tilde{\Psi}_1\rangle_n|1\rangle|0\rangle + \sqrt{1-a}|\tilde{\Psi}_0\rangle_n|0\rangle|0\rangle \\
\xrightarrow{Ry(\phi)}& \cos(\phi)\sqrt{a}|\tilde{\Psi}_1\rangle_n|1\rangle|0\rangle + \sin(\phi)\sqrt{a}|\tilde{\Psi}_1\rangle_n|1\rangle|1\rangle \\
&+ \cos(\phi)\sqrt{1-a}|\tilde{\Psi}_0\rangle_n|0\rangle|0\rangle + \sin(\phi)\sqrt{1-a}|\tilde{\Psi}_0\rangle_n|0\rangle|1\rangle.
\end{aligned}
\tag{20}
$$

By estimating the probability that the last two-qubit state is $|1\rangle|1\rangle$, we could avoid the anomalous target problem, because the amplitude to be estimated is modified from a to $a \sin^2(\phi)$.

3 Experiment with a real quantum computing device

This section is devoted to show experimental result using the real-backend device of IBM Quantum Systems called *ibmq_valencia* [25], to evaluate the estimation performance of the proposed ML estimate and thereby the validity of the employed depolarization model. In particular we consider the Monte Carlo integration problem, whose computational (sampling) cost can be quadratically reduced via the amplitude estimation algorithm [4–10]. In this section, we begin with a brief explanation on the target integration problem, followed by showing the execution results of the real device for two-qubit and three-qubit cases.

3.1 Monte Carlo integration via amplitude estimation

Let us consider a general 1-dimensional integration $\mathbb{E}[f(x)] = \int_0^1 dx q(x) f(x)$, where $f : [0, 1] \rightarrow [0, 1]$ is any real-valued smooth function and $q : [0, 1] \rightarrow [0, 1]$ is the probability distribution function which satisfies $\int_0^1 dx q(x) = 1$. This quantity can be in practice obtained by calculating the approximation

$$
S(f) = \sum_{j=0}^{2^n-1} p(x_j) f(x_j),
\tag{21}
$$

where x_j is defined as $x_j = (j+1/2)/2^n$ and p is the probability mass function defined as $p(x_j) = \int_{x_j-1/2^{n+1}}^{x_j+1/2^{n+1}} q(x) dx$. It should be noted that there is an approximation error due to the discretization of $f(x)$, i.e., $S(f) \neq \mathbb{E}[f(x)]$. In our analysis, however, we evaluate the error between $S(f)$ and the value obtained by Monte Carlo integration in order to assess the performance of our algorithm on a real quantum device. The amplitude estimation algorithm is run via the following operators:

$$
\mathcal{P}|0\rangle_n = \sum_{j=0}^{2^n-1} \sqrt{p(x_j)}|j\rangle_n, \quad \mathcal{R}|j\rangle_n|0\rangle = |j\rangle_n \left(\sqrt{f(x_j)}|1\rangle + \sqrt{1-f(x_j)}|0\rangle\right).
\tag{22}
$$
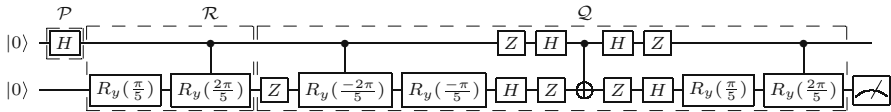
Fig. 6 The 2-qubit circuit of the unitary operators $\mathcal{Q}$ and $\mathcal{R}(\mathcal{P} \otimes I)$, for computing the probability $P(m; \boldsymbol{\theta})$ where the target value $S(f)$ is given in Eq. (25) with $n = 1$, and $b = 2\pi/5$

By applying these operators to the $(n + 1)$-qubit initial state, $|0\rangle_n |0\rangle$, we obtain

$$
\begin{aligned}
\mathcal{R}(\mathcal{P} \otimes \mathbf{I}_1)|0\rangle_n |0\rangle &= \sum_{j=0}^{2^n-1} \sqrt{p(x_j)}|j\rangle_n \left(\sqrt{f(x_j)}|1\rangle + \sqrt{1 - f(x_j)}|0\rangle \right) \\
&= \sqrt{S(f)}|\tilde{\Psi}_1\rangle|1\rangle + \sqrt{1 - S(f)}|\tilde{\Psi}_0\rangle|0\rangle,
\end{aligned}
\tag{23}
$$

where $|\tilde{\Psi}_1\rangle$ and $|\tilde{\Psi}_0\rangle$ are defined as

$$
|\tilde{\Psi}_1\rangle = \frac{1}{\sqrt{S(f)}} \sum_{j=0}^{2^n-1} \sqrt{p(x_j)f(x_j)}|j\rangle_n, \quad |\tilde{\Psi}_0\rangle = \frac{1}{\sqrt{1 - S(f)}} \sum_{j=0}^{2^n-1} \sqrt{p(x_j)(1 - f(x_j))}|j\rangle_n.
\tag{24}
$$

This is exactly the state of the form (1). Thus, the ideal amplitude estimation algorithm gives an approximation of $S(f)$ with the precision ϵ, with only $O(1/\epsilon)$ queries.

In this paper, we consider the simple integration $\int_0^1 \sin^2(bx) \, dx$ with b a constant, which is approximated as

$$
S(f) = \sum_{j=0}^{2^n-1} \frac{1}{2^n} \sin^2\left(bx_j\right).
\tag{25}
$$

In this case the operators $\mathcal{P}$ and $\mathcal{R}$ in Eq. (22) can be implemented only with Hadamard gates and controlled Y-rotation gates, as shown in Fig. 6.

3.2 Experimental result for the two-qubits case

We now show the experimental result of applying the ML algorithm in the real device, to compute Eq. (25). In this subsection, we consider the simple case $n = 1$, meaning that the integration is approximated via the discrete sum $S(f)$ having only two domain values $x = 0$ or $x = 1$, in which case Eq. (25) takes $b = \pi/20$, that is, the value $S(f) = a = \sin^2 \theta_a = 0.0077$. Also this setting means that we need only two qubits; in the experiment we chose the 0-th and 1-st qubit of *ibmq_valencia*.
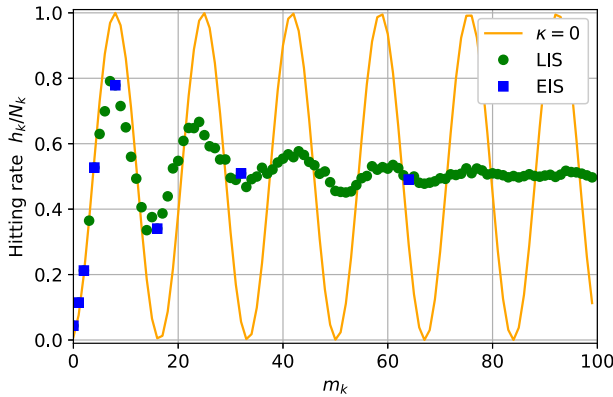
Fig. 7 The hitting rate of ancilla qubit being 1 (i.e., h_k/N_k), versus the number of $\mathcal{Q}$ in amplitude amplifications, m_k. The (green) round and (blue) square points show the experimental result obtained by running the ML algorithm with LIS and EIS, respectively. The (orange) line is the analytic result of the ideal probability without noise. To mark each point, we performed $N_k = 8192$ measurements for all k (Color figure online)

First, we show the experimental result of the quantum algorithm to compute the probability $P(m_k; a, \kappa)$ given in Eq. (15); recall that

$$
\begin{aligned}
P(m_k; a, \kappa) = \mathrm{Tr}(E_1 \rho_{\mathrm{noise}}^{(m_k)}) &= \mathrm{Tr}\left[(\mathbf{I}_1 \otimes |1\rangle\langle 1|)(\mathcal{QD})^{m_k}(|\Psi\rangle_2\langle\Psi|) \right] \\
&= \frac{1}{2} - \frac{1}{2}e^{-\kappa m_k}\cos(2(2m_k + 1)\theta_a),
\end{aligned}
\tag{26}
$$

and this is used to construct the likelihood function (16).

The quantum circuit to execute the unitary operators $\mathcal{Q}$ and $\mathcal{R}(\mathcal{P} \otimes I)$ is shown in Fig. 6. Here $|\Psi\rangle_2 = \mathcal{R}(\mathcal{P} \otimes I)|0\rangle|0\rangle$ is given by Eq. (23). In Fig 7, the green round points are plotted by computing the hitting rate of ancilla qubit being 1 (i.e., the hitting rate of measuring the good state, which corresponds to $P(m_k; a, \kappa)$), for the LIS setting. Note that these points cover the points of the EIS setting which are marked with the blue square points depicted in Fig. 7. Importantly, the figures show that the hitting rate has a trend of exponentially-decaying oscillation and approaches to 0.5 as the number of $\mathcal{Q}$ in amplitude amplification, m_k, increases. As a minimal model, we take the depolarizing noise (11) to model this decayed oscillation; actually the resulting probability distribution (26) well interpolates all the points obtained in the experiment, as shown in Fig. 7. Also, for reference, the analytic result of the ideal probability in the absence of noise, i.e., the case $\kappa = 0$, is depicted with the orange line in the Fig. 7.

We next experimentally executed the ML algorithm based on the model (26), for estimating $a = S(f)$ and κ. Recall that the best ML estimate of (a, κ) is given by the maximum of the likelihood function (16) with $P(m_k; a, \kappa)$ the model (26) and h_k the experimental result of the number of hit for a fixed number of $\mathcal{Q}$ in amplitude amplification, m_k ($k = 0, \ldots, M$); here we tested 6 patterns $M = 1, \ldots, 6$. In particular, we used EIS, fixed $N_k = 100$ and $b = 2\pi/5$ therefore $a = S(f) = 0.375$. The result is given in Fig. 8, which shows the relation between the estimation error of $S(f)$ and the
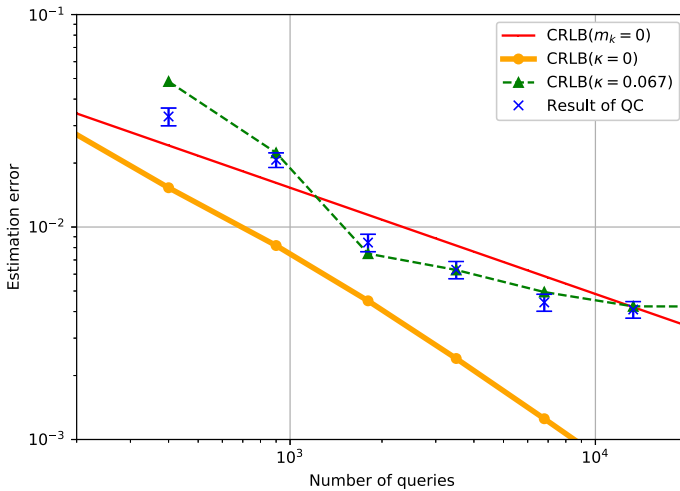
Fig. 8 Estimation error of a versus the total number of queries for the 2-qubit case. The thin red and thick yellow lines are the theoretical Cramér–Rao lower bound, obtained via the classical method and the ideal quantum ML method without noise ($\kappa = 0$), respectively. The green dotted line shows the Cramér–Rao lower bound with noise $\kappa = 0.067$. The blue cross marks show the standard deviation between the true value $S(f) = a = \sin^2 \theta_a = 0.375$ and the estimated values of a obtained via the ML method of experiments on *ibmq_valencia* (Color figure online)

total number of query calls, $N_q = \sum_{k=0}^{M} N_k(2m_k + 1)$. The solid thin red and thick yellow lines are the theoretical Cramér–Rao lower bound $\epsilon_{\min}(a, \kappa)$ given in Eq. (18), obtained via the classical method and the ideal quantum ML method without noise ($\kappa = 0$), respectively. The blue cross marks are the standard deviation of the estimated values of $S(f)$ obtained via the ML method, from the true value $S(f) = 0.375$. Note that, for instance to mark the blue point at $M = 4$ (or equivalently $N_q \sim 3.5 \times 10^3$), in which case the estimation error is about 0.65×10^{-2}, the ML algorithm uses the likelihood function constructed from the amplitude amplification processes with different operation length m_k ($k = 0, \ldots, 4$). Further, to reduce the fluctuation of those points, we repeated the same experiment 1064 times and averaged out to determine each point; the three-times standard errors are indicated by the error bars. The green dotted line shows the Cramér–Rao lower bound with noise level $\kappa = 0.067$, which is the single ML estimate of κ based on the $1064 \times N_k = 1064 \times 100$ data at $M = 6$ (that is, roughly speaking, the *best* estimate of κ over the whole execution of the algorithm).

As seen from Fig. 8, the estimation error experimentally obtained using the ML method (the blue points) is in good agreement with the theoretical Cramér–Rao lower bound (the green dotted line). A few slight deviation, particularly the points where the ML estimate is below the Cramér–Rao lower bound, might be due to some imperfections other than depolarizing noise, such as the rotation error of the gate operation and unnecessary coupling to neighboring qubits on the device. We would also like to emphasize that, in this example, there are several points where the experiment achieves the more precise estimate than that obtained via the classical method (the red line). This is an important evidence that even a noisy quantum computer can be beneficial
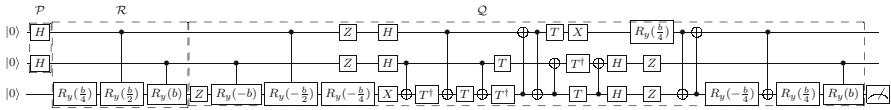
Fig. 9 The 3-qubit circuit for computing the probability $P(m; a, \kappa)$ where the target value $S(f)$ is given in Eq. (25) with $n = 2$ and $b = 2\pi/5$
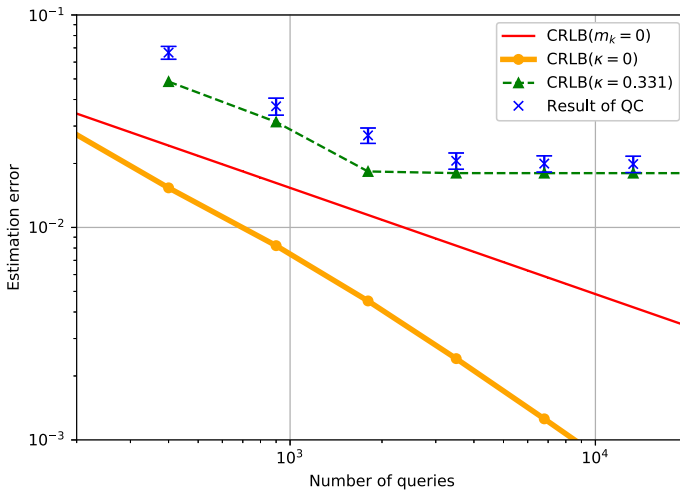

Fig. 10 Estimation error of a versus the total number of queries for the 3-qubit case. The thin red and thick yellow lines are the theoretical Cramér–Rao lower bound, obtained via the classical method and the ideal quantum ML method without noise ($\kappa = 0$), respectively. The green dotted line shows the Cramér–Rao lower bound with noise $\kappa = 0.331$. The blue cross marks show the standard deviation between the true value $S(f) = a = \sin^2 \theta_a = 0.381$ and the estimated values of a obtained via the ML method of experiments on *ibmq_valencia* (Color figure online)

over the classical one, in the measure of query complexity. Finally we remark that a similar behavior was observed, in other settings that use different two qubits in the device and a different target value $S(f)$; see Appendix C.

3.3 Experimental result for the three-qubit case

Here we present the experimental result for the case where the target integration $\int_0^1 \sin^2(bx) \, dx$ with $b = 2\pi/5$ is to be approximated by $S(f)$ in Eq. (25) with $2^n = 2^2$ segments. In this case, $S(f) = a = \sin^2 \theta_a = 0.381$. Also then we need $n + 1 = 3$ qubits to implement the amplitude amplification operation.

The quantum circuit to execute the ML algorithm is shown in Fig. 9. Because *ibmq_valencia* does not allow direct coupling for arbitrary pair of qubits, we chose three qubits to form a chain structure. In the experiment, the qubit in the middle of chain was chosen as the ancilla qubit, and it is placed as the third qubit from the top in the circuit at Fig. 9.

Fig. 10 is the three-qubit version of that at Fig. 8, and shows the relationship between the number of queries and the estimation error of $a = S(f)$. The Cramér–Rao lower bound under noise (the green dotted line with triangle) was plotted with $\kappa = 0.331$. Due to such a high noise level, as expected from Fig. 1, this Cramér–Rao lower bound is above the classical one (the red line), meaning that the quantum computation has no advantage in the amplitude estimation task. However, what is important in our context is that the estimation error obtained from the experiment (the blue points) lie near the green line; that is, the ML estimate computed with the 3-qubit real quantum device asymptotically reaches the theoretical Cramér–Rao lower bound.

Therefore, together with the result for the 2-qubit case, we now would like to conclude that Eq. (11) is a good model of the noise process, at least for a small size qubit device. This means that the theoretical predictions illustrated in Fig. 1 and thereby Fig. 2 might be usable as a practical guide for discussing the condition on the noise level κ to realize the Heisenberg-scaling in the quantum amplitude estimation task.

4 Discussion

This section is divided to two topics. In Sect. 4.1, based on the result obtained in Sect. 2.3, we show a procedure to assess the gate errors required to achieve a given task as well as the expected execution time of the algorithm, and discuss the limitations such as the gate error and the coherence time when using IBM Quantum devices. Next in Sect. 4.2, we discuss the computational complexity that takes into account the classical optimization procedure to compute the ML estimate.

4.1 Hardware specification for the amplitude estimation task

Recall that Fig. 2 is used to predict the maximum allowed noise level $\bar{\kappa}$ for achieving a given estimation error ϵ with the query calls obeying the Heisenberg-scaling law. Here we connect this value of $\bar{\kappa}$ to the error of elementary gates in the quantum circuit; also the execution time of the algorithm is assessed. The entire procedure to compute these quantities is composed of the following four steps.

1. An amplitude estimation problem, together with a target estimation error ϵ, is given to us. Then the amplitude amplification operator $\mathcal{Q}$ (plus $\mathcal{P}$ and $\mathcal{R}$ in the Monte Carlo case), and accordingly the elementary gates constituting those operators are identified; for instance, see Fig. 6 or Fig. 9.
2. Given the target estimation error ϵ, we use Fig. 2 to compute the maximum allowed noise level $\bar{\kappa}$ when N_k for all k is fixed. We can then compare $\bar{\kappa}$ to the ML estimate $\hat{\kappa}$ obtained through the experiment with a real device, to see whether the device can produce the desired estimate $\hat{a}$. Also Eq. (19) enables us to have the maximum number of operations of $\mathcal{Q}$, i.e., $\bar{m}$, within which the Heisenberg-scaling nearly holds in estimating the parameter. More precisely it is given by $\bar{m} = 0.5/(e^{\bar{\kappa}} - 1)$; for example, $\bar{m}$ can take only $\bar{m} = 5$ when $\bar{\kappa} = 0.1$, but $\bar{m} = 500$ when $\bar{\kappa} = 0.001$.
3. Furthermore, from $\bar{m}$ and the number of gates constituting the operator $\mathcal{Q}$, denoted by L, we can have a rough estimate of the execution time of each quantum circuit

with length $m_1, m_2, \ldots, \bar{m}$. Then the execution time in $\bar{m}$ and the total execution time in $m_1, m_2, \ldots, \bar{m}$ can also be estimated, which are then compared to the coherence time of the available real device, and then the feasibility of this algorithm can be assessed respectively.

4. Under the assumption that all qubits are subjected to only depolarizing noise, we approximate the $p = e^{-\bar{\kappa}}$, the success probability of the operator $\mathcal{Q}$, as $\prod_{i=1}^{L}(1 - \epsilon_i)$, where L is the number of gates that constitute the operator $\mathcal{Q}$ and ϵ_i is the depolarizing error probability of the ith qubit. Using this equation we may also have a rough estimate of the gate error, using $\bar{\kappa}$ and L. The gate error is compared to that of real device (which might be identified via the standard randomized benchmarking test) to see the feasibility of the algorithm.

A detailed procedure for computing the above quantities, especially the gate errors and the total execution time, is given in Appendix B.

As an example, let us consider a multiple integration over 5 variables, which assumes direct correlations between any of two variables, e.g., $\int (x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_5) d\mathbf{x}$. The reason of this choice is that, to compute such a multiple integration with more than five variables, the Monte-Carlo method is usually employed rather than grid-based numerical integration approaches. Here we follow the above four steps one by one. (Step 1) Suppose that we are required to achieve the target estimation error $\epsilon = 0.001$ with the Heisenberg-scaling query calls. (Step 2) Then Fig. 2 tells us that we need $\bar{\kappa} = 0.005$ when $N_k = 100$ for all k and 99 qubits wherein 48 qubits are used for the ancilla to gate decomposition. Also we have $\bar{m} = 99$. (Step 3) For some quantum devices, the operating time of single gates and the CNOT gate are publicly available; in the case of *ibmq_valencia*, they are 7.1×10^{-8} s and 2.8×10^{-7} s, respectively. Also the measurement time is 3.5×10^{-6} s. Moreover we assume that the interval time between the measurement and the next execution of the algorithm is 10 times longer than the execution time of the algorithm with length $\bar{m}$. Then, we find that the execution time of the algorithm with length $\bar{m}$ is 0.54 s and the total computation time of the entire algorithm is about 1082 s; see Table 2 in Appendix B for the detailed calculation. (Step 4) We can also assess the gate error required for the algorithm to follow the Heisenberg-scaling to reach the given estimation error. Under the assumption that the error of CNOT gate, ϵ_d, is 10 times bigger than that of any single gate error, ϵ_s, we end up with $\epsilon_d = 2.8 \times 10^{-7}$ and $\epsilon_s = 2.8 \times 10^{-8}$. Now, in the case of *ibmq_valencia*, they are given by $\epsilon_d \sim 1.0 \times 10^{-2}$ and $\epsilon_s \sim 1.0 \times 10^{-3}$.

An important finding is in quantifying the difference between the gate error of the currently available devices and that of the near ideal machine which realizes the Heisenberg-scaling to reach the given estimation error. The execution time (0.54 s) is also clarified, which may be much longer than the coherence time of the currently available devices. In fact, even though these gaps seem to be large, they are informative for us because we now know how much improvement in the hardware development is necessary to fill these gaps. In addition, algorithm improvements would help close these gaps. For example, recent research [29] suggested a method for reducing the circuit depth and the number of qubits.

Finally, we note that the estimated noise level κ is comparable with the value calculated based on the publicly available information on the gate error of the IBM

Quantum device. In our case, the 2-qubits Grover circuit contains 5 CNOT gates (the Q operator part of Fig. 6), and the 3-qubits one contains 16 CNOT gates (the Q operator part of Fig. 9); note that R_y gate is composed of 2 CNOT gates. Then, for the 2 qubits case, the CNOT gate error was 0.00565, meaning that $\kappa \sim -\ln((1 - 0.00565)^5) \sim 0.0283$. For the 3 qubits case, we used two types of CNOT gate with error rate 0.008923 and 0.01119, and they were used 8 times; hence we have $\kappa \sim -\ln((1 - 0.008923)^8 * (1 - 0.01119)^8) \sim 0.162$. By taking into account some other impact of noise such as single gate error, we consider that these values becomes closer to the estimated values $\kappa = 0.067$ and $\kappa = 0.331$.

4.2 Computational time complexity of maximum likelihood estimation

We have argued about the possibility of less number of queries achieved by the ML method against classical cases under noise influences, and how to deal with anomalous target values. Here, assuming anomalous targets are detectable, we show the classical post-processing for maximum likelihood estimation of the target value $\hat{a}$ and noise level $\hat{\kappa}$ can be performed in $O(\ln^{5/2}(1/\epsilon))$, which is much less than that of the quantum part.

Because under depolarizing noise model the success probability now has two parameters, the target value $\hat{a}$ and noise level $\hat{\kappa}$ as shown in Sect. 2.2, it is necessary to perform two-dimensional maximum likelihood to obtain the estimated values of the parameters. A straightforward random search or grid search requires $O(1/\epsilon^2)$ evaluations of likelihood function to achieve the error ϵ [30]. This completely ruins the advantage obtained from the quantum method because the total time complexity now becomes $O(1/\epsilon^2)$ which is at least the same as the classical Monte-Carlo approach.

In order to avoid the problem of computational complexity, we employed an adaptive constant grid search at each stage k for $0 \le k \le M$ in the experiment of Sect. 3. Namely, at the k-th stage, we performed the grid search with a constant number of divisions only in the range of confidence interval defined as C_ϵ times larger than the error estimated from the Fisher information at the $(k - 1)$-th stage. By the Chebysev inequality, the error estimated at each stage is guaranteed to be within C_ϵ times the confidence interval with probability at least $1 - 1/C_\epsilon^2$. To obtain an estimation within ϵ error, the number of stages M is $O(\ln(1/\epsilon))$. Thus, the total probability of obtaining good parameter estimation is at least $(1 - 1/C_\epsilon^2)^M$, which is $\Omega(1)$ when $C_\epsilon = \Theta(\ln^{1/2}(1/\epsilon))$. The total time complexity of the maximum likelihood computation is thus $C_\epsilon \cdot M \cdot O(\ln(1/\epsilon))$, where the last term is that for evaluating the likelihood function. This gives the bound $O(\ln^{5/2}(1/\epsilon))$ time complexity for the classical post processing.

Except for the anomalous cases, we should note the estimation error of the target value a is not affected even if the estimation error of the noise level κ is large. The details are shown in Appendix A. This may allow us to reduce the two-dimensional ML estimation to a one-dimensional one. In this case, the estimated values can be obtained with lower computational cost; roughly estimate the noise level κ with large grid size, and then perform one-dimensional ML of the target value a with high accuracy. Furthermore, for sufficiently large number of shots the two-dimensional

ML estimation can be approximated with the weighted least square estimation. The latter can be solved much more efficiently in practice thanks to the specific algorithms, such as the Levenberg-Marquardt's [31,32].

5 Conclusion

All quantum algorithms running on currently available quantum devices must consider the effects of noise. Hence, to perform an appropriate evaluation of such algorithms, particularly those aimed at possible quantum advantage, it is necessary to carefully model the noise, analyze its effects to the algorithm, and make a validation via experiments. This paper provides a demonstration of this evaluation, yielding Fig. 2, a relationship between the target estimation error and the required noise threshold for the amplitude estimation problem, which is validated by experiments on quantum devices.

Lastly we again point out the importance of multi-parameters estimation problem in the quantum computing framework. In our case, the nuisance noise parameter in addition to the target amplitude parameter must be estimated. We have seen that the Fisher information matrix can be degenerate, in which case the target amplitude parameter cannot be efficiently estimated even by increasing the number of amplitude amplifications. This singular problem arises only in multi-parameters estimation problems, and a similar challenge may easily appear in other scenarios such as the phase estimation problem under noisy environment. In this sense, while this paper employed the ML estimate approach, we can have various options to design an efficient estimator for such multi-parameter estimation problems in noisy quantum devices. For instance Ref. [33] showed the method to tune the likelihood function for Bayesian inference of the amplitude parameter, and they have derived the run time estimation to achieve a target accuracy under noisy environment.

Appendix A: Effect of estimation error of noise level on the target value

In this paper the problem is formulated as a two-parameters estimation problem with respect to (a, κ), where a is our main interest while κ is the nuisance parameter. Hence here we briefly sketch how much the estimation error of κ may affect on the target value a.

For this purpose, let us consider one-dimensional ML estimation for the target a, with a fixed noise level $\hat{\kappa}$, which is supposed to be roughly estimated in advance in some way. The true values of parameters, a and κ, are denoted as a_0 and κ_0 only in this section, and the differences between the true values and the estimation errors are denoted as $\delta a = \hat{a} - a_0$ and $\delta \kappa = \hat{\kappa} - \kappa_0$, respectively. By definition, the one-dimensional ML estimate $\hat{a}$ is given by

$$\hat{a} = \mathrm{argmax}_a \ln L(\mathbf{h}; \{a, \hat{\kappa}\}), \tag{27}$$

where L is the likelihood function introduced in Eq. (5). Assuming the smoothness of the likelihood function with respect to a, we have the following equation:

$$\frac{\partial}{\partial a} \ln L(\mathbf{h}; \{\hat{a}, \hat{\kappa}\}) = 0. \tag{28}$$

Hence, the first order Taylor expansion of the left-hand side of this equation, at around the true parameters a_0 and κ_0, yields

$$\frac{\partial}{\partial a} \ln L(\mathbf{h}; \{a_0, \kappa_0\}) = -\delta a \frac{\partial^2}{\partial a^2} \ln L(\mathbf{h}; \{a_0, \kappa_0\}) - \delta \kappa \frac{\partial^2}{\partial a \partial \kappa} \ln L(\mathbf{h}; \{a_0, \kappa_0\}). \tag{29}$$

As a consequence of the central limit theorem, the left-hand side of this equation asymptotically follows $\mathbb{N}(0, \mathcal{I}_{1,1})$, i.e., the normal distribution with mean 0 and variance $\mathcal{I}_{1,1}$, in the limit with a large number of shots. Also, the first term on the right-hand side asymptotically converges to $\mathcal{I}_{1,1} \delta a$ and the second term converges to $\mathcal{I}_{1,2} \delta \kappa$. Thus in the asymptotic regime, Eq. (29) implies that δa obeys the normal distribution with mean $-\mathcal{I}_{1,2} \delta \kappa / \mathcal{I}_{1,1}$ and variance $1/\mathcal{I}_{1,1}$, i.e.,

$$\delta a \sim \mathbb{N}\left(-\frac{\mathcal{I}_{1,2}}{\mathcal{I}_{1,1}} \delta \kappa, \frac{1}{\mathcal{I}_{1,1}} \right). \tag{30}$$

Then the mean squared error of $\hat{a}$ can be calculated as

$$\mathbb{E}\left[\delta a^2 \right] = \delta \kappa^2 \frac{\mathcal{I}_{1,2}^2}{\mathcal{I}_{1,1}^2} + \frac{1}{\mathcal{I}_{1,1}}. \tag{31}$$

If the squared error of $\hat{\kappa}$ is c times bigger than the Cramér–Rao lower bound of the variance, $\mathcal{I}_{1,1}/(\mathcal{I}_{1,1}\mathcal{I}_{2,2} - \mathcal{I}_{1,2}^2)$, this equation yields

$$\mathbb{E}\left[\delta a^2\right] = c\frac{\mathcal{I}_{1,1}}{(\mathcal{I}_{1,1}\mathcal{I}_{2,2} - \mathcal{I}_{1,2}^2)}\frac{\mathcal{I}_{1,2}^2}{\mathcal{I}_{1,1}^2} + \frac{1}{\mathcal{I}_{1,1}} = (\mathcal{I}^{-1})_{1,1}\left(1 + (c-1)\frac{\mathcal{I}_{1,2}^2}{\mathcal{I}_{1,1}\mathcal{I}_{2,2}}\right). \tag{32}$$

The first term of the rightmost side of this equation is just the Cramér–Rao lower bound and the second term is an additional term due to the effect from the error of noise parameter κ. This expression indicates that the estimation error of κ does not significantly affect on the estimation error of the target value a except for the anomalous target values, because $\mathcal{I}_{1,2}^2/\mathcal{I}_{1,1}\mathcal{I}_{2,2}$ is not large for such typical target values.

Appendix B: Detailed procedure for determining the hardware components

In Sect. 4.1 we have discussed the method to evaluate the hardware components required to achieve the Heisenberg-scaling estimation under the influence of noise. Here we show a detailed procedure to determine those quantities for an N_{int}-dimensional integration problem, which is summarized in Table 2. Our task is to use the quantum ML method to estimate the value of integral, within the target estimation error ϵ. Note that then the amplitude amplification (AA) operator $\mathcal{Q}$ as well as $\mathcal{R}$ and $\mathcal{P}$ are determined. Also we set the following assumptions; the gate error of CNOT, ϵ_d, is 10 times bigger than that of any single gate error, ϵ_s; the qubits are fully connected in the device; the operating time of single and CNOT gates are 7.1×10^{-8} s and 2.8×10^{-7} s, respectively; the measurement duration time is 3.5×10^{-6} s; the interval time between the measurements is 10 times bigger than the execution time; the number measurement of $N_k = 100$ and the sequence is EIS. The main focused quantities are the gate errors (ϵ_d, ϵ_s) and the executing time of the algorithm.

## Appendix C: Two qubit experimental result with other settings

In Sect. 3.2, we demonstrated the results using the real quantum computer with specific parameters. Here we present additional results in different experimental settings, to see how the basic quantities such as the estimation error would be affected by those changes. Fig. 11(a) shows the result when we used the third and the fourth qubits on *ibmq_valencia* to represent the target value $S(f) = a = 0.375$, whereas the other settings are the same as before; e.g., EIS ($m_0 = 0, m_k = 2^{k-1}$) is taken and the number of shots is $N_k = 100 \ \forall k$. Next, Fig. 11(b) shows the case where the target value is chosen as $S(f) = a = \sin^2 \pi/40 + \sin^2 3\pi/40 \sim 0.03$, meaning that $b = \pi/10$; the other settings are the same as in Fig. 11(a).

These figures show that the estimation error obtained with the experiments (the blue cross marks) lies near the green dotted line which is the Cramér–Rao lower bound with depolarizing noise, supporting our depolarizing noise model discussed in

Table 2 List of quantities to be determined

Element	Calculated by	Value	Supplement
N_{nq} (# qubits for an integral)	$\log_2 1/\epsilon$	10	$\epsilon = 0.001$
N_{tnq} (total number of qubits)	$2N_{nq}N_{int} - 1$	99	$N_{nq}N_{int} + 1$ uses ML method $N_{nq}N_{int} - 2$ comes from the ancilla qubits for decomposition of $\mathbf{S}_0$
$\bar{\kappa}$ (noise level)	$5\epsilon/\sqrt{N_k}$	0.005	From Fig. 2.
N_y (number of CCR_y)	$N_{nq}(N_{nq}-1)N_{int}^2/2$	1000	From $\mathcal{R}$
N_s (# single gates in $\mathcal{Q}$)	$2(N_{nq}N_{int} + 1 + 6N_y) + 12N_{nq}N_{int} - 15$	12, 687	$2(N_{nq}N_{int} + 1)$ is from $\mathcal{P}$ and $\mathcal{P}^{-1}$ $2(6N_y*2)$ is from $\mathcal{R}$ and $\mathcal{R}^{-1}$ $12N_{nq}N_{int} - 15$ is from decomposition of $\mathbf{S}_0$
N_d (# CNOT gates in $\mathcal{Q}$)	$8N_y * 2 + 6N_{nq}N_{int} - 5$	16, 295	$8N_y \times 2$ is from $\mathcal{R}$ and $\mathcal{R}^{-1}$ $6N_{nq}N_{int} - 5$ is from decomposition of $\mathbf{S}_0$ ϵ_s (single gate error)
ϵ_d (CNOT gate error)	$e^{-\bar{\kappa}} = (1 - \epsilon_s)^{N_s}(1 - \epsilon_d)^{N_d}$	2.8×10^{-8} 2.8×10^{-7}	Assuming $\epsilon_s = 0.1\epsilon_d$ and including approximation calculations.
$\bar{m}$ (maximum number of $\mathcal{Q}$)	$(2\bar{m} + 1)(1 - e^{-\bar{\kappa}}) < 1$	99	From Eq. (19)
t_{AA} (the time of a $\mathcal{Q}$)	$t_{AA} = t_s N_s + t_d N_d$	5.4×10^{-3}	t_s, t_d are the gate time of single and CNOT gate respectively which are 7.1×10^{-8}, 2.8×10^{-7} s from IBM Q.
$t_{\bar{m}}$ (executing time with $\bar{m}$)	$t_{AA} \times \bar{m} + t_m$	0.54	t_m is the measurement time, 3.5×10^{-6} s.
t_t (total executing time)	$\sum_{k=1}^{N}((t_{AA}m_k + t_m + t_i) * N_k)$	1082	N is from $m_N = \bar{m}$. t_i is an interval time between shots calculated as 10 times the execution time. $N_k = 100$ is the number of shots, for all k
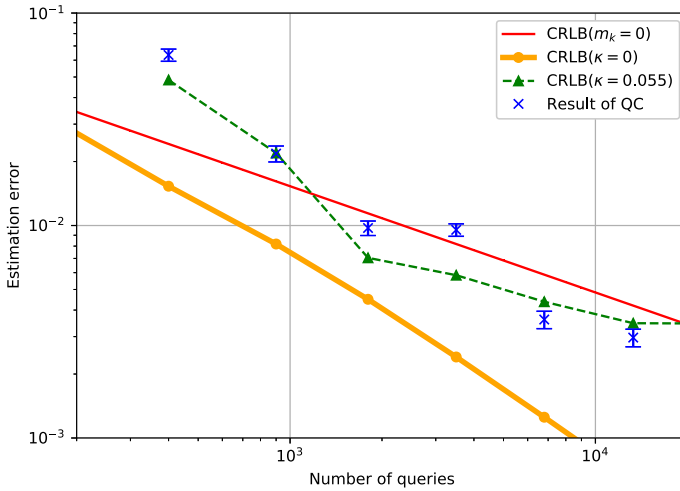
(a) The case of target value $S(f) = a = 0.375$
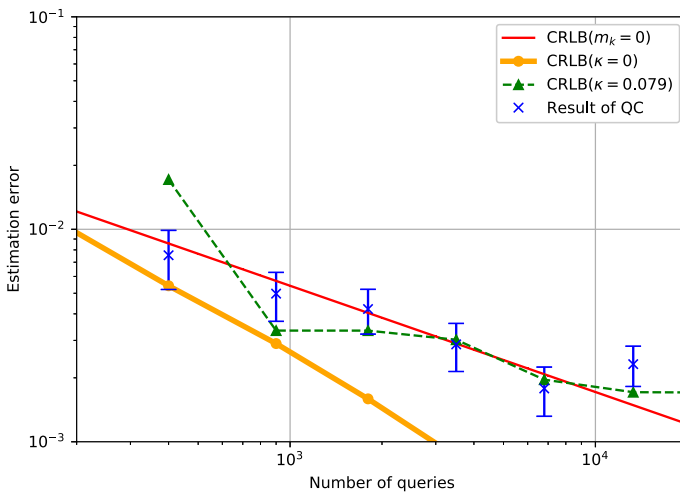
(b) The case of target value $S(f) = a \sim 0.03$

Fig. 11 Estimation error of a versus the total number of queries. These figures have different target values, while the other settings are the same. The thin red, thick yellow, and green dotted lines are the theoretical Cramér–Rao lower bounds, while the blue error bars represent the three-times standard errors. For the detail, see Sect. 3.2 (Color figure online)

Sect. 2.3. We note that the noise level κ ended up with different values, 0.055 and 0.079 respectively, depending on the different target values a; hence it is important to perform the multi-parameter estimation each time even when using the same qubits. In addition, the experimental results of Figs. 8 and 11a show that the noise level κ

differs depending on the qubits; hence, a careful qubit mapping (i.e., assigning the location of qubits) is important to get to lower noise level.

Note also that, in Fig. 11a, b, the error bars (the three-times standard errors) seem to have different values, but this is merely because the vertical axis is in the log scale and each difference is only within a factor of two.

References

1. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Rev. **41**(2), 303–332 (1999). https://doi.org/10.1137/S0036144598347011
2. Knill, E., Ortiz, G., Somma, R.D.: Optimal quantum measurements of expectation values of observables. Phys. Rev. A **75**(1), 012328 (2007). https://doi.org/10.1103/PhysRevA.75.012328
3. Kassal, I., Jordan, S.P., Love, P.J., Mohseni, M., Aspuru-Guzik, A.: Polynomial-time quantum algorithm for the simulation of chemical dynamics. Proc. Natl. Acad. Sci. **105**(48), 18681–18686 (2008). https://doi.org/10.1073/pnas.0808245105
4. Montanaro, A.: Quantum speedup of Monte Carlo methods. Proc. R. Soc. A Math. Phys. Eng. Sci. **471**(2181), 20150301 (2015). https://doi.org/10.1098/rspa.2015.0301
5. Rebentrost, P., Gupt, B., Bromley, T.R.: Quantum computational finance: Monte Carlo pricing of financial derivatives. Phys. Rev. A **98**(2), 022321 (2018). https://doi.org/10.1103/PhysRevA.98.022321
6. Woerner, S., Egger, D.J.: Quantum risk analysis. NPJ Quant. Inform. **5**(1), 1–8 (2019). https://doi.org/10.1038/s41534-019-0130-6
7. Stamatopoulos, N., Egger, D.J., Sun, Y., Zoufal, C., Iten, R., Shen, N., Woerner, S.: Option pricing using quantum computers. arXiv preprint arXiv:1905.02666 (2019)
8. Martin, A., Candelas, B., Rodríguez-Rozas, A., Martín-Guerrero, J.D., Chen, X., Lamata, L., Orús, R., Solano, E., Sanz, M.: Towards pricing financial derivatives with an IBM quantum computer. arXiv preprint arXiv:1904.05803 (2019)
9. Egger, D.J., Gutiérrez, R.G., Mestre, J.C., Woerner, S.: Credit risk analysis using quantum computers. arXiv preprint arXiv:1907.03044 (2019)
10. Miyamoto, K., Shiohara, K.: Reduction of qubits in quantum algorithm for Monte Carlo simulation by pseudo-random number generator. arXiv preprint arXiv:1911.12469 (2019)
11. Preskill, J.: Quantum computing in the NISQ era and beyond. Quantum **2**, 79 (2018). https://doi.org/10.22331/q-2018-08-06-79
12. Córcoles, A.D., Kandala, A., Javadi-Abhari, A., McClure, D.T., Cross, A.W., Temme, K., Nation, P.D., Steffen, M., Gambetta, J.M.: Challenges and opportunities of near-term quantum computing systems. Proc. IEEE (2019). https://doi.org/10.1109/JPROC.2019.2954005
13. Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.H., Zhou, X.Q., Love, P.J., Aspuru-Guzik, A., Obrien, J.L.: A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. (2014). https://doi.org/10.1038/ncomms5213
14. Farhi, E., Goldstone, J., Gutmann, S.: A quantum approximate optimization algorithm. arXiv preprint arXiv:1411.4028 (2014)
15. Temme, K., Bravyi, S., Gambetta, J.M.: Error mitigation for short-depth quantum circuits. Phys. Rev. Lett. **119**(18), 180509 (2017). https://doi.org/10.1103/PhysRevLett.119.180509
16. Cody Jones, N., Van Meter, R., Fowler, A.G., McMahon, P.L., Kim, J., Ladd, T.D., Yamamoto, Y.: Layered architecture for quantum computing. Phys. Rev. X **2**(3), 031007 (2012). https://doi.org/10.1103/PhysRevX.2.031007
17. Gidney, C., Ekerå, M.: How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. arXiv preprint arXiv:1905.09749 (2019)
18. Cross, A.W., Bishop, L.S., Sheldon, S., Nation, P.D., Gambetta, J.M.: Validating quantum computers using randomized model circuits. Phys. Rev. A **100**(3), 032328 (2019). https://doi.org/10.1103/PhysRevA.100.032328
19. Grover, L.K.: Quantum computers can search rapidly by using almost any transformation. Phys. Rev. Lett. **80**(19), 4329 (1998). https://doi.org/10.1103/PhysRevLett.80.4329
20. Brassard, G., Hoyer, P., Mosca, M., Tapp, A.: Quantum amplitude amplification and estimation. Contemp. Math. **305**, 53–74 (2002). https://doi.org/10.1090/conm/305/05215

21. Suzuki, Y., Uno, S., Raymond, R., Tanaka, T., Onodera, T., Yamamoto, N.: Amplitude estimation without phase estimation. Quant. Inf. Process. **19**(2), 75 (2020). https://doi.org/10.1007/s11128-019-2565-2

22. Aaronson, S., Rall, P.: Quantum approximate counting, simplified. In Symposium on Simplicity in Algorithms, pp. 24–32. SIAM (2020). https://doi.org/10.1137/1.9781611976014.5

23. Grinko, D., Gacon, J., Zoufal, C., Woerner, S.: Iterative quantum amplitude estimation. arXiv preprint arXiv:1912.05559 (2019)

24. Nakaji, K.: Faster amplitude estimation. arXiv preprint arXiv:2003.02417 (2020)

25. IBM Quantum Experience. https://quantum-computing.ibm.com

26. Abraham, H., Akhalwaya, I.Y., Aleksandrowicz, G., Alexander, T., Alexandrowics, G., Arbel, E., Asfaw, A., Azaustre, C., AzizNgoueya, Barkoutsos, P., Barron, G., Bello, L., Ben-Haim, Y., Bevenius, D., Bishop, L.S., Bosch, S., Bravyi, S., Bucher, D., Cabrera, F., Calpin, P., Capelluto, L., Carballo, J., Carrascal, G., Chen, A., Chen, C.F., Chen, R., Chow J.M., Claus, C., Clauss, C., Cross, A.J., Cross, A.W., Cross, S., Cruz-Benito, J., Culverm, C., Córcoles-Gonzales, A.D., Dague, S., El Dandachi, T., Dartiailh, M., DavideFrr, Davila, A.R., Ding, D., Doi, J., Drechsler, E., Drew, Dumitrescu, E., Dumon, K., Duran, I., EL-Safty, K., Eastman, E., Eendebak, P., Egger, D., Everitt, M., Fernández, P.M., Ferrera, A.H., Frisch, A., Fuhrer, A., GEORGE, M., Gacon, J., Gadi, Gago, B.G., Gambetta, J.M., Gammanpila, A., Garcia, L., Garion., S., Gomez-Mosquera, J., de la Puente González, S., Gould, I., Greenberg, D., Grinko, D., Guan, W., Gunnels, J.A., Haide, I., Hamamura, I., Havlicek, V., Hellmers, J., Herok, L., Hillmich, S., Horii, H., Howington, C., Hu, S., Hu, W., Imai, H., Imamichi, T., Ishizaki, K., Iten, R., Itoko, T., Javadi-Abhari, A., Jessica, Johns, K., Kachmann, T., Kanazawa, N., Kang-Bae, Karazeev, A., Kassebaum, P., King, K., Knabberjoe, Kovyrshin, A., Krishnan, V., Krsulich, K., Kus, G., LaRose, R., Lambert, R., Latone, J., Lawrence, S., Liu, D., Liu, P., Maeng, Y., Malyshev, A., Marecek, J., Marques, M., Mathews, D., Matsuo, A., McClure, D.T., McGarry, C., McKay, D., McPherson, D., Meesala, S., Mevissen, M., Mezzacapo, A., Midha, R., Minev, Z., Mitchell, A., Moll, N., Mooring, M.D., Morales, R., Moran, N., Murali, P., Müggenburg, P., Nadlinger, D., Nannicini, G., Nation, P., Naveh, Y., Neuweiler, P., Niroula, P., Norlen, H., O'Riordan, L.J., Ogunbayo, O., Ollitrault, P., Oud, S., Padilha, D., Paik, H., Perriello, S., Phan, A., Pistoia, M., Pozas-iKerstjens, A., Prutyanov, V., Puzzuoli, D., Pérez, J., Quintiii, Raymond, R., Martín-Cuevas Redondo, R., Reuter, M., Rice, J., Rodríguez, D.M., Rossmannek, M., Ryu, M., SAPV, T., SamFerracin, Sandberg, M., Sathaye, N., Schmitt, B., Schnabel, C., Schoenfeld, Z., Scholten, T.L., Schoute, E., Schwarm, J., Sertage, I.F., Setia, K., Shammah, N., Shi, Y., Silva, A., Simonetto, A., Singstock, N., Siraichi, Y., Sitdikov, I., Sivarajah, S., Sletfjerding, M.B., Smolin, J.A., Soeken, M., Sokolov, I.O., SooluThomas, Steenken, D., Stypulkoski, M., Suen, J., Takahashi, H., Tavernelli, I., Taylor, C., Taylour, P., Thomas, S., Tillet, M., Tod, M., de la Torre, E., Trabing, K., Treinish, M., TrishaPe, Turner, W., Vaknin, Y., Valcarce, C.R., Varchon, F., Vazquez, A.C., Vogt-Lee, D., Vuillot, C., Weaver, J., Wieczorek, R., Wildstrom, J.A., Wille, R., Winston, E., Woehr, J.J., Woerner, S., Woo, R., Wood, C.J., Wood, R., Wood, S., Wootton, J., Yeralin, D., Young, R., Yu, J., Zachow, C., Zdanski, L., Zoufal, C., Zoufalc, azulehner, bcamorrison, brandhsn, chlorophyll zz, dan1pal, dime10, drholmie, elfrocampeador, faisaldebouni, fanizzamarco, gruu, kanejess, klinvill, kurarrr, lerongil, ma5x, merav aharoni, ordmoj, sethmerkel, strickroman, sumitpuri, tigerjack, toural, vvilpas, welien, willhbang, yang.luh, yelojakit, and yotamvakninibm. Qiskit: An open-source framework for quantum computing. 2019. https://doi.org/10.5281/zenodo.2562110

27. Nielsen, M.A., Chuang, I.: Quantum computation and quantum information (2002)

28. Ji, Z., Wang, G., Duan, R., Feng, Y., Ying, M.: Parameter estimation of quantum channels. IEEE Trans. Inf. Theory **54**(11), 5172–5185 (2008). https://doi.org/10.1109/TIT.2008.929940

29. Vazquez, A.C., Woerner, S.: Efficient state preparation for quantum amplitude estimation. arXiv preprint arXiv:2005.07711 (2020)

30. Zabinsky, Z.B.: Stochastic adaptive search for global optimization, vol. 72. Springer, New York (2013). https://doi.org/10.1007/978-1-4419-9182-9

31. Nocedal, J., Wright, S.: Numerical optimization. Springer, New York (2006). https://doi.org/10.1007/b98874

32. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963). https://doi.org/10.1137/0111030

33. Wang, G., Koh, D.E., Johnson, P.D., Cao, Y.: Bayesian inference with engineered likelihood functions for robust amplitude estimation. arXiv preprint arXiv:2006.09350 (2020)

34. Brown, E.G., Goktas, O., Tham, W.K.: Quantum amplitude estimation in the presence of noise. arXiv preprint arXiv:2006.14145 (2020)