Quantum-enhanced mean value estimation via adaptive measurement

Kaito Wada[1,a)] Kazuma Fukuchi[1] Naoki Yamamoto[1,2]

Abstract: Estimating the mean values of quantum observables is a fundamental task in quantum computing. In particular, efficient estimation in a noisy environment requires us to develop a sophisticated measurement strategy. Here, we propose a quantum-enhanced estimation method for the mean values, that adaptively optimizes the measurement (POVM) for each circuit; as a result of optimization, the estimation precision gets close to the quantum Cramér-Rao lower bound, that is, inverse of the quantum Fisher information. We provide a rigorous analysis for the statistical properties of the proposed adaptive estimation method such as consistency and asymptotic normality. Furthermore, several numerical simulations with large system dimension are provided to show that the estimator needs only a reasonable number of measurements to almost saturate the quantum Cramér-Rao bound.

1. Introduction

The mean value estimation of quantum observables is an important subroutine in many quantum algorithms for both near-term noisy and future fault-tolerant quantum devices. Especially, variational quantum algorithms, which have been studied extensively in recent years [5], require estimating mean values to train the variational parameters on a parameterized quantum circuit. However, typically when the target observable is a molecular Hamiltonian as in the variational quantum eigensolver [20, 25], the standard estimation method needs astronomically many measurements [25, 29, 9, 16], and thus any possible quantum advantage may be lost. In addition, real quantum devices suffer from noise induced by the interaction of system and environment, and this noise deteriorates the efficiency for reading out high-precision calculation results. Therefore, developing efficient mean-value estimation methods in a noisy environment is highly important [28, 16].

For this purpose, the statistical estimation theory [6, 21, 14] should of course be useful. Actually its quantum extension, the statistical quantum estimation theory, has been developed [14, 11] and successfully applied in many problems such as quantum sensing [7]. In particular, the seminal Cramér-Rao inequality and Fisher information, which can be used to characterize the limit of estimation precision, has been extended to quantum settings [13, 12, 11, 2]. This result, known as the quantum Cramér-Rao (QCR) inequality and quantum Fisher information, plays a practically important role in many applications and, further, reveals some fundamental limitation on information extractable from quantum mechanical systems. Hence the statistical quantum estimation theory could be effectively applied to improve estimation efficiency of subroutines in quantum computing. However, such exploration has just started. For instance, some less-demanding implementation methods for quantum-enhanced (i.e., less query complexity compared to any classical means) amplitude estimation algorithm have been proposed and their performance were investigated in depth using statistical estimation theory, yet without studying QCR bound [1, 10, 19, 4, 3, 8]. To our best knowledge, only Ref. [26] provides a quantum-enhanced amplitude estimation method that asymptotically achieves the QCR bound when several conditions are satisfied in a noisy environment.

Moreover, we recently find the quantum-enhanced mean value estimation method [28] that utilizes techniques from the quantum signal processing (QSP) [18]. This method employs QSP-inspired parameterized quantum circuits and executes the Bayesian inference for the target value based on the result of a fixed measurement on the parameterized quantum circuits. The key point for efficient estimation is that the variance of the posterior distribution can be rapidly decreased by adjusting the parameters of the quantum circuit in each measurement. However, this method does not consider optimizing the measurement for better estimation and, as a nature of Bayes method,

1 Department of Applied Physics and Physico-Informatics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, 223-8522, Japan
2 Quantum Computing Center, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, 223-8522, Japan
a) wkai1013keio840@keio.jp

no statistical guarantee of the estimator to hit the true mean value was provided. Importantly, Ref. [16] provided a benchmark using the same Bayesian inference, which clarifies the need for more efficient estimation method to achieve quantum advantage. Hence it is of increasing attraction to develop an algorithm that possibly achieves the QCR bound.

In this paper, we propose a new quantum-enhanced mean-estimation method that asymptotically achieves QCR bound with some provable statistical properties. This method uses a modified maximum likelihood estimation that incorporates an adaptive optimization based on the Fisher information. More precisely, this scheme adjusts the number of amplitude amplification operations and the measurement (more precisely POVM), which as a result enables the estimator to almost achieve the QCR bound as the number of qubits increases. As in the previous study [28, 26, 23], we employ the depolarizing channel as a noise model. As pointed out in the previous study, the quantum mean value estimation under the depolarizing noise may significantly deteriorate depending on the target value to be estimated, but we will numerically demonstrate that the proposed method can also get around the problem due to the optimization. In addition, we prove that the estimator satisfies the following statistical properties.

- Consistency: As the number of measurements increases, the estimation accuracy becomes better. More precisely, the probability that the estimated value is identical to the true value approaches 1, which is clearly a desirable statistical property.
- Asymptotic normality: For the target value, the estimator regularized by its Fisher information follows the standard normal distribution and as a result the estimation error asymptotically achieves the Cramér-Rao lower bound, as the number of measurements increases. Note that this property corresponds to the central limit theorem that holds for usual methods for estimating the quantum mean value in quantum computing.

These asymptotic properties guarantee the quality of our estimates, and we will also demonstrate the properties hold for practical number of measurements.

This article is organized as follows. In Section 2, we review a maximum likelihood amplitude estimation algorithm and show the estimation precision of this algorithm deteriorates in the presence of depolarizing noise. In Section 3, we propose a new quantum mean value estimation method with adaptive measurements. We first introduce two-type measurements (POVMs) and construct our algorithm with optimization for efficient estimation. Next, we show some Theorems for the statistical properties of our method. Section 4 is devoted to show the results of numerical simulations and discussion. Finally, we conclude this paper in Section 5.

2. Preliminary

The (quantum-enhanced) amplitude estimation algorithm of [22, 23, 26, 24], which does not use the hard-to-implement phase estimation algorithm, consists of two major steps. The first step is to perform amplitude amplification on the quantum state whose amplitude is to be estimated, and then make a measurement. The second step is to estimate the amplitude by classical post-processing for the measurement result, i.e., the maximum likelihood estimation.

The problem is to estimate the amplitude $\sin(\phi^*)$, with unknown $\phi^* \in [0, \pi/2]$, of the following $(n+1)$-qubit quantum state $\mathcal{A}\left|0\right\rangle_{n+1}$:

$$\mathcal{A}\left|0\right\rangle_{n+1} := \cos\left(\phi^*\right)\left|\psi_0\right\rangle_n\left|0\right\rangle_1 + \sin\left(\phi^*\right)\left|\psi_1\right\rangle_n\left|1\right\rangle_1, \tag{1}$$

where $\left|0\right\rangle_{n+1}$ is the computational basis of $(n+1)$-qubit. Here, $\mathcal{A}$ is a unitary operator whose action is defined as Eq. (1), and $\left|\psi_0\right\rangle_n$ and $\left|\psi_1\right\rangle_n$ are normalized n-qubit quantum states. The first step is to amplify the amplitude via the operator $\mathcal{Q}$ defined as

$$\mathcal{Q} := \mathcal{A}\left(2\left|0\right\rangle_{n+1}\left\langle 0\right|_{n+1} - I_{n+1}\right)\mathcal{A}^\dagger\left(I_n \otimes Z\right), \tag{2}$$

where I_n is the identity operator on the n-qubit system, and Z denotes the 2×2 Pauli Z matrix. When $\mathcal{Q}$ is applied m times to the state (1), we obtain

$$\begin{aligned}\mathcal{Q}^m\mathcal{A}\left|0\right\rangle_{n+1} = \cos\left[(2m+1)\phi^*\right]\left|\psi_0\right\rangle_n\left|0\right\rangle_1 \\ + \sin\left[(2m+1)\phi^*\right]\left|\psi_1\right\rangle_n\left|1\right\rangle_1.\end{aligned} \tag{3}$$

We now measure the last qubit of Eq. (3) by the computational basis $\left|0\right\rangle_1$ and $\left|1\right\rangle_1$. Then the probability of obtaining "1" is given by

$$\tilde{\mathbf{P}}\left(m;\phi^*\right) := \frac{1}{2} - \frac{1}{2}\cos\left[2(2m+1)\phi^*\right]. \tag{4}$$

Note that this is equivalent to the Bernoulli trials with success probability $\tilde{\mathbf{P}}\left(m;\phi^*\right)$.

In the second step we estimate the amplitude based on the measurement results obtained in the first step by the maximum likelihood estimation method. For each of predefined odd numbers $2m_k+1$ $(k = 1, 2, \cdots, M)$, we prepare the state (3) and measure it N times with the computational basis independently. We write $x^{(k)} \in \{0, 1, \cdots, N\}$ as the number of hitting "1" for the state $\mathcal{Q}^{m_k}\mathcal{A}\left|0\right\rangle_{n+1}$. The measurement results for M different states are put together as $\boldsymbol{x}_M := (x^{(1)}, x^{(2)}, \cdots, x^{(M)})$. Then the likelihood function to have $\boldsymbol{x}_M$ is given by

$$\tilde{\mathcal{L}}_M(\boldsymbol{x}_M;\phi^*) := \prod_{k=1}^{M}\tilde{F}_k(x^{(k)};m_k,\phi^*), \tag{5}$$

where $\tilde{F}_k(x^{(k)};m_k,\phi^*)$ is the probability of obtaining $x^{(k)}$. Given the measurements $\boldsymbol{x}_M$, the maximum likelihood estimation assumes that the value of ϕ that maximizes

$\tilde{\mathcal{L}}_M(\boldsymbol{x}_M;\phi)$ is a plausible estimate of the true value ϕ^*. In other words, the maximum likelihood estimate $\hat{\phi}_M$ for ϕ^* from the M series of measurements is defined as

$$\hat{\phi}_M := \underset{\phi\in[0,\pi/2]}{\mathrm{argmax}}\ \tilde{\mathcal{L}}_M(\boldsymbol{x}_M;\phi). \tag{6}$$

We now consider the n-qubit depolarizing channel:

$$\mathcal{D}[\rho] := p\rho + \frac{1-p}{d} I_n, \quad d := 2^n, \tag{7}$$

where ρ is an arbitrary density operator and $1-p$ is the probability that depolarization occurs. In this paper, as in [28, 26, 23], we consider a noise model in which depolarization occurs with probability $1-p_s$ for state preparation and $1-p_q$ for each amplitude amplification. Then the probability of obtaining "1" for the final state under this noise model is given by

$$\tilde{\mathbf{P}}_{\mathcal{D}}\left(m;\phi^*\right) := \frac{1}{2} - \frac{p_s p_q^m}{2}\cos\left[2(2m+1)\phi^*\right]. \tag{8}$$

The subscript $\mathcal{D}$ means that the quantum state to be measured has passed through the above-defined depolarizing channels. Then the classical Fisher information associated with the probability (8) for ϕ^* is calculated as

$$\tilde{\mathcal{I}}_{\mathrm{c}}(m;\phi^*) := \frac{(2m+1)^2(p_s p_q^m)^2\sin^2\left[2(2m+1)\phi^*\right]}{\tilde{\mathbf{P}}_{\mathcal{D}}(m;\phi^*)\left(1-\tilde{\mathbf{P}}_{\mathcal{D}}(m;\phi^*)\right)}. \tag{9}$$

Equation (9) indicates that the estimation may become ineffective to ϕ^* if m satisfies $\sin\left[2(2m+1)\phi^*\right] \simeq 0$. More precisely, if m and ϕ^* satisfy this condition, the Cramér-Rao lower bound of the estimation error for ϕ^* significantly deteriorates [28, 23]. In addition, even if we use the above-described maximum likelihood estimation based on the M series of measurements for m_k $(k=1,2,\cdots,M)$, which are predefined independently to ϕ^*, the same deterioration may occur [23]. Importantly, this phenomena are also seen in the quantum mean value estimation problem discussed in the next section.

3. Quantum-enhanced mean value estimation

3.1 Dual amplitude amplification

The amplitude estimation method described in the previous section can be applied to the problem of estimating the mean value of a Hermitian operator $\mathcal{O}$ with eigenvalues ±1. Let us consider estimating the mean value $\langle\mathcal{O}\rangle := \langle A|\,\mathcal{O}\,|A\rangle$, where $|A\rangle := A\,|0\rangle_n$ is an n-qubit quantum state with a unitary operator A. Suppose $|A\rangle$ is not an eigenstate of $\mathcal{O}$, i.e., $|\langle\mathcal{O}\rangle| \neq 1$ and that the two quantum states $|A\rangle$ and $\mathcal{O}\,|A\rangle$ are linearly independent and form the following subspace [28, 17]

$$\mathcal{S} := \mathrm{Span}\left\{|A\rangle, \mathcal{O}|A\rangle\right\} = \mathrm{Span}\left\{|A\rangle, |A^{\perp}\rangle\right\}, \tag{10}$$

where $|A^{\perp}\rangle$ is an normalized state orthogonal to $|A\rangle$ obtained by Gram-Schmidt procedure to $\mathcal{O}\,|A\rangle$. For simplicity, we write $|A\rangle$ and $|A^{\perp}\rangle$ as $|\bar{0}\rangle$ and $|\bar{1}\rangle$, respectively, and identify $\mathcal{S}$ with the 1-qubit Hilbert space. We then define an n-qubit operator Q on $\mathcal{S}$, which corresponds to the amplitude amplification operator in Eq. (2), as

$$Q := A\left(2\,|0\rangle_n\,\langle 0|_n - I_n\right)A^{\dagger}\mathcal{O}. \tag{11}$$

This operator keeps the subspace $\mathcal{S}$ invariant. Now the representation of Q on $\mathcal{S}$ is expressed as $Q|_{\mathcal{S}} = R_{\bar{Y}}\left(-2\theta^*\right)$, where $\bar{Y}$ is the Pauli Y matrix in the basis $|\bar{0}\rangle$ and $|\bar{1}\rangle$, and $\theta^* := \arccos\left(\langle\mathcal{O}\rangle\right)$. Note that, for the target mean value $\langle\mathcal{O}\rangle = \cos\theta^*$, the domain of θ^* is given by $\theta^* \in (0,\pi)$, which is twice as that of ϕ^* in the previous section. Acting Q on $|A\rangle$ m times yields

$$\begin{aligned} Q^m\,|A\rangle &= \left(Q|_{\mathcal{S}}\right)^m|\bar{0}\rangle \\ &= \cos\left(m\theta^*\right)|\bar{0}\rangle - \sin\left(m\theta^*\right)|\bar{1}\rangle. \end{aligned} \tag{12}$$

Here, as in the previous case, we assume that the state is subjected to the depolarization noise through the amplitude amplification process; that is, the output state after m repetition of Q is given by

$$\rho_{\mathcal{D}}(m;\theta^*) := p_s p_q^m \rho(m;\theta^*) + \frac{1-p_s p_q^m}{d} I_n, \tag{13}$$

where $\rho(m;\theta^*) := Q^m\,|A\rangle\,\langle A|\,(Q^m)^{\dagger}$. Here, $1-p_q$ (>0) denotes the probability of depolarization noise occurring along the single query of the operator Q.

One strategy to efficiently read out quantum mean values from the amplitude-amplified quantum state Eq. (13) is to use a sophisticated POVM whose classical Fisher information achievable to the quantum Fisher information for the state. More details are provided in [27]. Although we found a 3-valued POVM that satisfies the above feature with respect to certain θ^*, it is difficult to implement because the 3-valued POVM includes the θ^*-dependent quantum states [27]. Therefore, considering a marginalisation of the POVM, we define the following 2-valued POVM

$$\begin{aligned} M_1^{(\mathrm{even})} &:= I_n - M_0^{(\mathrm{even})}, \\ M_0^{(\mathrm{even})} &:= \mathcal{O}^{\dagger}\,|A\rangle\,\langle A|\,\mathcal{O}, \end{aligned} \tag{14}$$

where the meaning of *even* is explained later. Measuring Eq. (13) with the POVM, we obtain the result corresponding to $M_1^{(\mathrm{even})}$ in the probability

$$\begin{aligned} &\mathbf{P}_{\mathcal{D}}^{(\mathrm{even})}(m;\theta^*) \\ &:= \frac{d-1}{d} + p_s p_q^{m+1/2}\left[\sin^2\left[(m+1)\theta^*\right] - \frac{d-1}{d}\right], \end{aligned} \tag{15}$$

where $p_q^{1/2}$ denotes the additional noise associated with the implementation of $A^{\dagger}$ in $M_0^{(\mathrm{even})}$. This measurement has powerful estimation capabilities in the sense of its Fisher information. The classical Fisher information per shot with Eq. (15) can be calculated as

$$\begin{aligned} &\mathcal{I}_{\mathrm{c}}^{(\mathrm{even})}(m;\theta^*) \\ &:= \frac{(2m+2)^2 p_s^2 p_q^{2m+1}\sin^2\left[(2m+2)\theta^*\right]}{4\mathbf{P}_{\mathcal{D}}^{(\mathrm{even})}(m;\theta^*)\left(1-\mathbf{P}_{\mathcal{D}}^{(\mathrm{even})}(m;\theta^*)\right)}. \end{aligned} \tag{16}$$

As for the quantum Fisher information, we can calculate it as [15, 30]

$$\mathcal{I}_{\mathrm{q}}^{(\mathrm{even})}(m) := \frac{(p_s p_q^{m+1/2})^2}{\frac{2}{d} + \left(1-\frac{2}{d}\right) p_s p_q^{m+1/2}}(2m+2)^2. \quad (17)$$

When $\cos\left[(m+1)\theta^*\right] \neq 0$, the classical Fisher information asymptotically approaches the quantum Fisher information, in the large dimension limit $d = 2^n \to \infty$:

$$\mathcal{I}_{\mathrm{c}}^{(\mathrm{even})}(m;\theta^*) \to \mathcal{I}_{\mathrm{q}}^{(\mathrm{even})}(m). \quad (18)$$

This means that, to estimate a *particular* θ^*, the POVM Eq. (14) becomes optimal with respect to the Fisher information, as the number of qubits increases; more precise analysis is provided in [27]. In the next subsection, for estimating θ^*, i.e., $\langle\mathcal{O}\rangle$, we introduce an new algorithm that is sufficiently optimal for *almost all* θ^* in the large system dimension via the optimization of the number of query m.

Although the even POVM Eq. (14) can be optimal with its Fisher information, it is important to evaluate the performance of estimators in terms of not only the amount of its Fisher information, but also whether it can correctly estimate the true target value. Since the probability Eq. (15) is in fact an even function when $\theta = \pi/2$ is the center, the maximum likelihood estimation associated with this measurements cannot distinguish the sign of quantum mean values. We therefore introduce the following second POVM broken this symmetry:

$$M_1^{(\mathrm{odd})} := I_n - M_0^{(\mathrm{odd})},$$
$$M_0^{(\mathrm{odd})} := \frac{I_n + \mathcal{O}}{2}. \quad (19)$$

This POVM corresponds to Eq. (8) in the standard amplitude estimation algorithm because the probability of obtaining "1" by measureing the state Eq. (13) with the POVM is given as

$$\mathbf{P}_{\mathcal{D}}^{(\mathrm{odd})}(m;\theta^*) = \frac{1}{2} - \frac{p_s p_q^m}{2}\cos\left[(2m+1)\theta^*\right]. \quad (20)$$

Note that the coefficient of θ^* differs by 2 compared to that of ϕ^* in Eq. (8). This is because the presence of sign for quantum mean values doubles the domain of the target value. Since the probability is an odd function unlike Eq. (15), the measurement can distinguish the sign of the target mean value. Note that we add the superscript *even* or *odd* for the POVM Eqs. (14), (19) based on this symmetry. In contrast to the even POVM, the corresponding classical Fisher information $\mathcal{I}_{\mathrm{c}}^{(\mathrm{odd})}(m;\theta^*)$ deviates from the quantum Fisher information even if the number of qubits increases [27].

Combining Eqs. (15), (20), we summarize the probabilities of obtaining a measurement corresponding to "1" with the two POVMs as follows.

$$\mathbf{P}_{\mathcal{D}}(\alpha;\theta^*) := \begin{cases} \frac{1}{2} - \frac{\eta_\alpha}{2}\cos\left(\alpha\theta^*\right), & \alpha \text{ is odd} \\ \frac{d-1}{d} + \eta_\alpha\left[\sin^2\left(\frac{\alpha}{2}\theta^*\right) - \frac{d-1}{d}\right], & \alpha \text{ is even} \end{cases}, \quad (21)$$

where η_α denotes $p_s p_q^{\frac{\alpha-1}{2}}$ and $\alpha \in \mathbb{N}$ corresponds to the number of queries to the operator A or $A^\dagger$. We perform the measurement by Eq. (19) when α is odd and by Eq. (14) when α is even. In the following, we call α amplified level. Also, the classical/quantum Fisher information corresponding to each measurement is written as

$$\mathcal{I}_{\mathrm{c}}(\alpha;\theta^*) := \frac{\alpha^2\eta_\alpha^2\sin^2\left(\alpha\theta^*\right)}{4\mathbf{P}_{\mathcal{D}}(\alpha;\theta^*)(1-\mathbf{P}_{\mathcal{D}}(\alpha;\theta^*))}, \quad (22)$$

$$\mathcal{I}_{\mathrm{q}}(\alpha) := \frac{\alpha^2\eta_\alpha^2}{\frac{2}{d} + \left(1-\frac{2}{d}\right)\eta_\alpha}. \quad (23)$$

By combining the above measurements with the maximum likelihood estimation framework described in the previous section, we can estimate the quantum mean value $\cos\theta^*$ as detailed in the next subsection. Note that there are also ineffective points in the classical Fisher information Eq. (22) in the presence of noise. Therefore, we are motivated to avoid α that satisfies $\sin\left(\alpha\theta^*\right) \simeq 0$ for both Eqs. (14) and (19) as much as possible. In addition, we would like to employ measurements whose classical Fisher information approaches the quantum Fisher information and enhance the estimation power. Thus, we propose a new quantum mean value estimation method that satisfies these requirements by adaptively adjusting amplified levels α, that is, POVMs. Importantly, this method inherits the good properties of the maximum likelihood estimation.

3.2 Our algorithm

The general framework of our algorithm is a maximum likelihood estimation method using random variables following the statistics of Eq. (21). An overview is shown in Fig. 1. In the following, we explain the procedures in detail.

(i) First, before starting the estimation, determine the probabilities of depolarization $1-p_s$ and $1-p_q$ based on the quantum device to be used. For the first measurements, we also set the amplified level α_1 to 1.

(ii) We measure the quantum state (13) N times independently with either POVM Eq. (14) or Eq. (19), which is determined based on the amplified level α_k. We write the total number of hitting "1" as $x^{(k)} \in \{0,1,\cdots,N\}$. Here, $x^{(k)}$ can be considered as the realization of the Binomial random variable $X^{(k)} \sim \mathrm{Bin}(N, \mathbf{P}_{\mathcal{D}}(\alpha_k;\theta^*))$.

(iii) Calculate the maximum likelihood estimate $\hat{\theta}_k$ based on the total k measurement results $\boldsymbol{x}_k := (x^{(1)}, x^{(2)}, \cdots, x^{(k)})$. Here, the estimate maximizes the following likelihood function

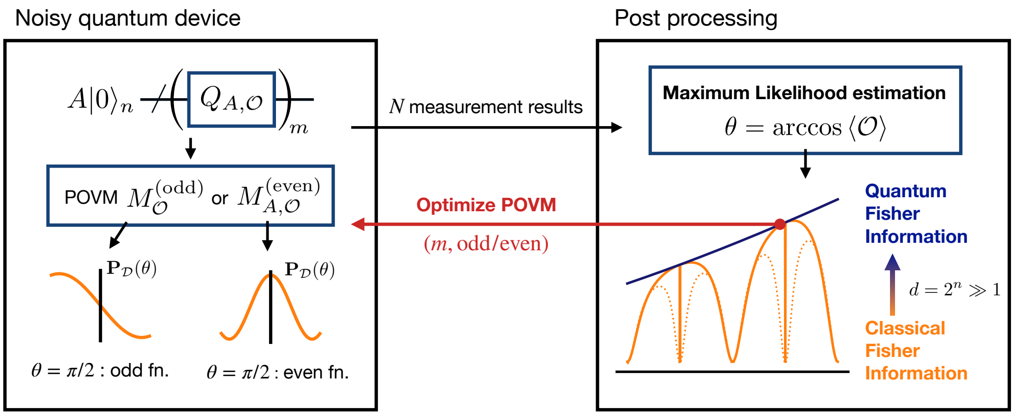


Fig. 1 Overview of our algorithm for estimating $\langle A|\mathcal{O}|A\rangle$. Here, $|A\rangle$ is a target state prepared by an n-qubit quantum circuit A and the computational basis $|0\rangle_n$, and $\mathcal{O}$ is a target Hermitian operator with eigenvalues ±1. The general framework is a modified maximum likelihood estimation method using the specific quantum circuits with the amplitude amplification operator $Q_{A,\mathcal{O}}$ and the two POVMs $M_{A,\mathcal{O}}^{(\mathrm{even})}, M_{\mathcal{O}}^{(\mathrm{odd})}$. Here, we assume the depolarization noise is induced by preparing $|A\rangle$ and querying $Q_{A,\mathcal{O}}$. Even under this noisy condition, our algorithm can estimate the quantum mean value $\langle A|\mathcal{O}|A\rangle$ efficiently in the sense of its Fisher information via optimization of the number of querying $Q_{A,\mathcal{O}}$ and the POVM. In particular, the optimization every N measurements improves the classical Cramér-Rao lower bound, and besides it gets close to the quantum Cramér-Rao lower bound sufficiently in large system dimension d.

$$\mathcal{L}_k(\boldsymbol{x}_k;\theta) := \prod_{m=1}^{k} F_m\left(x^{(m)};\alpha_m(\boldsymbol{x}_{m-1}),\theta\right), \tag{24}$$

where $F_m\left(x^{(m)};\alpha_m(\boldsymbol{x}_{m-1}),\theta\right)$ is the probability function of random variable $X^{(m)}$.

Note that the measurements $\boldsymbol{x}_k$ collect the already obtained results via measurements with α_m $(m=1,2,\cdots,k)$. Since the amplified level α_k depends on the previous results $\boldsymbol{x}_{k-1}$ as described below, the likelihood function $\mathcal{L}_k(\boldsymbol{x}_k;\theta)$ has a hierarchical structure, which does not find in the conventional one Eq. (5).

(iv) To enhance the classical Fisher information, we determine α_{k+1} for the next measurement based on the current estimate $\hat{\theta}_k$ by numerically solving the following optimization problem

$$\alpha_{k+1} := \underset{\alpha\in D_k}{\mathrm{argmax}}\ \mathcal{I}_{\mathrm{c}}(\alpha;\hat{\theta}_k)\left|\sin\left(\alpha\hat{\theta}_k\right)\right|^{\delta}, \tag{25}$$

where $D_k \subset \mathbb{N}$ is the range of α_{k+1} to be optimized.

Once $(k+1)$-th amplified level α_{k+1} is determined, return to (ii) and perform the measurement again.

Here, $|\sin(\alpha\hat{\theta}_k)|^{\delta}$ is a regularization term to avoid instability of the numerical optimization, which comes from the incontinuity of $\mathcal{I}_{\mathrm{c}}(\alpha;\theta)$ [27]. In this paper, we used $\delta=0.10$. Note that since the elements of D_k are integer, the optimization problem can be solved fast. Importantly, this optimization takes a role of not only selecting an optimal α for large Fisher information but also controlling the following trade-off relation by setting the optimization range of each α_{k+1}. A large α is preferred to enhance the Fisher information per 1 shot, but if we choose an α which dramatically (e.g., super-exponentially) increases with respect to m, the estimation will be failed in practical N because the likelihood function $\mathcal{L}_k(\boldsymbol{x}_k;\theta)$ has many peaks around the true target value.

3.3 Statistical properties of our estimator

Although in our algorithm the random variables describing the measurement results are dependent each other and have the hierarchical structure in Eq. (24), we can prove the consistency of our estimator based on the convergence of α_k, which intuitively means that the random variables become independent each other in large N.

Theorem 1. (Informal version) *There exists an unique maximum likelihood estimator $\hat{\theta}_M$ of $\mathcal{L}_M(\boldsymbol{X}_M;\theta)$ such that*

$$\cos\hat{\theta}_M \to \langle\mathcal{O}\rangle, \tag{26}$$

as $N\to\infty$. In addition, $\alpha_k(\boldsymbol{X}_{k-1})$ converges to a constant α_k^* in the sense of probability as follows

$$\lim_{N\to\infty} P\left[\alpha_k(\boldsymbol{X}_{k-1})=\alpha_k^*\right]=1, \quad k=2,\cdots,M \tag{27}$$

We show the formal version of this theorem and proof in [27]. This theorem shows that our estimator is reasonable because more data gives more accurate estimation.

Furthermore, the asymptotic variance of our estimator can achieve the Cramér-Rao lower bound when N is large. To state it more precisely, we first introduce the following total classical/quantum Fisher information of our estimator:

$$\mathcal{I}_{\mathrm{c/q,tot}}\left(\theta^*\right) = N\sum_{m=1}^{M}\mathbf{E}_{\boldsymbol{X}_{m-1}}\left[\mathcal{I}_{\mathrm{c/q}}\left(\alpha_m\left(\boldsymbol{X}_{m-1}\right);\theta^*\right)\right], \tag{28}$$

where $\alpha_m(\boldsymbol{x}_{m-1})$ denotes the optimized amplified level based on the measurement results $\boldsymbol{x}_{m-1}$. The derivation is provided in [27]. Since Theorem 1 states that $\alpha_m(\boldsymbol{X}_{m-1})$ becomes a constant with high probability as N increases, the total Fisher information also converges as follows

$$\frac{\mathcal{I}_{\mathrm{c/q,tot}}(\theta^*)}{N} \to \frac{\mathcal{I}^*_{\mathrm{c/q,tot}}(\theta^*)}{N}, \tag{29}$$

where $\mathcal{I}^*_{\mathrm{c/q,tot}}(\theta^*)$ is the asymptotic value of the total classical/quantum Fisher information defined as

$$\mathcal{I}^*_{\mathrm{c/q,tot}}\left(\theta^*\right) := N\sum_{m=1}^{M}\mathcal{I}_{\mathrm{c/q}}\left(\alpha^*_m;\theta^*\right). \tag{30}$$

Here, we provide the convergence theorem for the distribution of our estimator.

Theorem 2. (Informal version) *If $\hat{\theta}_M$ is an maximum likelihood estimator of $\mathcal{L}_M(\boldsymbol{X}_M;\theta)$, then the following convergence holds*

$$\sqrt{\mathcal{I}^*_{\mathrm{c,tot}}(\theta^*)}\left(\cos\hat{\theta}_M - \langle\mathcal{O}\rangle\right) \to \mathcal{N}(0,\sin^2\theta^*), \tag{31}$$

where $\to$ means the convergence in distribution as N increases.

The formal version of this theorem and proof are shown in [27]. This theorem shows that the asymptotic variance of our estimator is equivalent to inverse of the asymptotic value of the total classical Fisher information, and thus, our estimator can asymptotically achieve the classical Cramér-Rao lower bound. In the next section, we demonstrate that the asymptotic property holds in practical N.

Recalling the optimality of the even POVM (14), the total classical Fisher information is likely to be sufficiently close to the total quantum Fisher information as the number of qubits increases:

$$\mathcal{I}^*_{\mathrm{c,tot}}(\theta^*) \approx \mathcal{I}^*_{\mathrm{q,tot}}(\theta^*). \tag{32}$$

Since the optimization in Eq. (25) decreases this deviation, it depends on the selection of $\{D_k\}_k$. In result section, we demonstrate that the deviation is sufficiently small for *almost all* θ^*, i.e., the total classical Fisher information of our algorithm can be sufficiently close to its total quantum Fisher information with use of a certain set $\{D_k\}_k$.

4. Results and Discussion

In this section, we numerically verify the performance of our algorithm with noisy quantum devices. First, in practical N, we demonstrate that the asymptotic properties of our estimator hold. Next, evaluating the asymptotic value of the total Fisher information, we confirm that our algorithm retains large Fisher information regardless of the target value θ^*. We also show that the total classical Fisher information of our algorithm are sufficiently close to its quantum Fisher information in the large system dimension.

4.1 Asymptotic properties in practical N

To demonstrate the properties of our algorithm, we numerically evaluated the Root Mean Squared Error (RMSE) of $\cos\hat{\theta}$ defined as

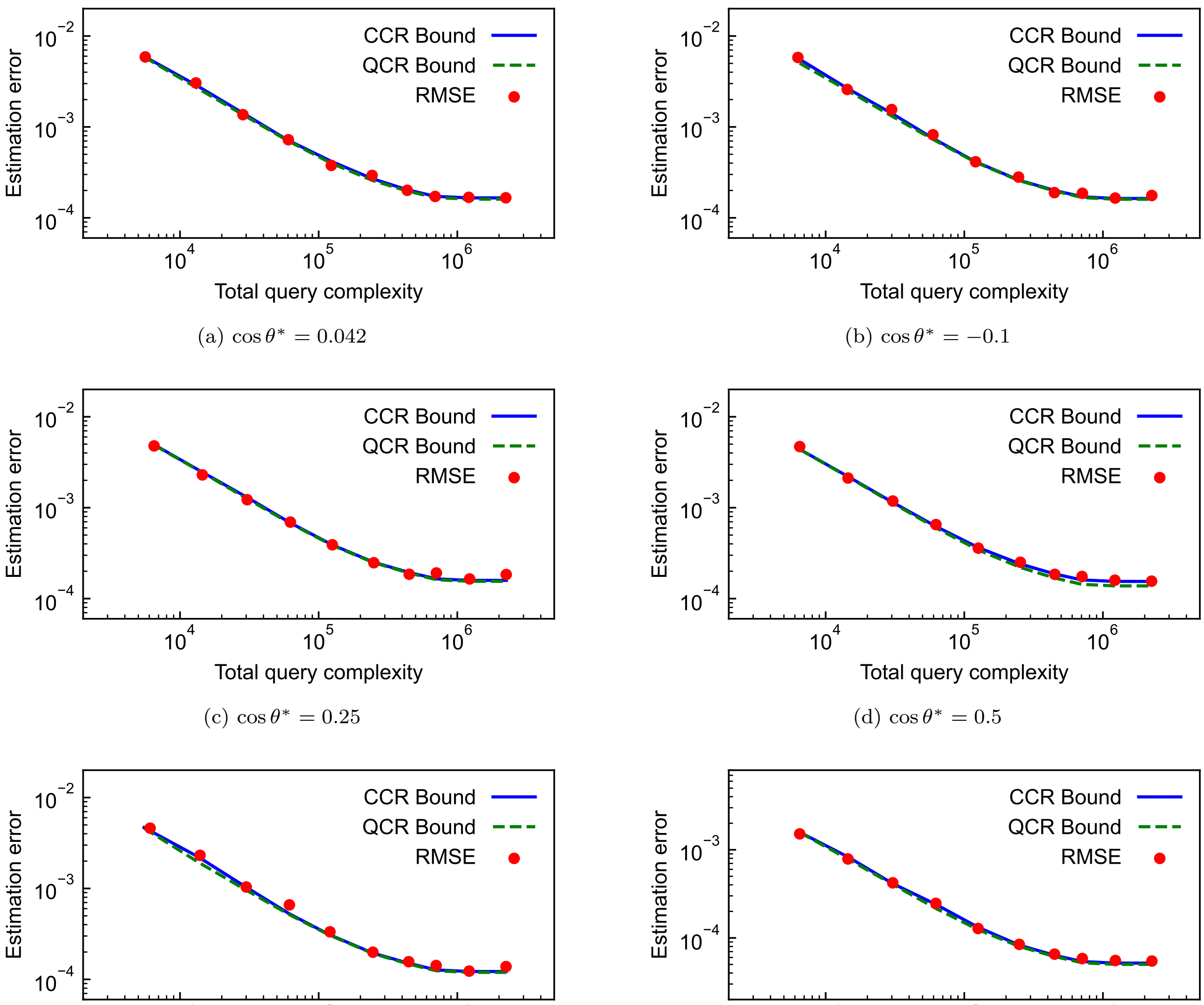
$$\mathrm{RMSE}\left[\cos\hat{\theta}\right] := \sqrt{\mathbf{E}_{\hat{\theta}}\left[\left(\cos\hat{\theta}-\cos\theta^*\right)^2\right]}$$
$$\simeq \sqrt{\frac{1}{\#}\sum_{i=1}^{\#}\left(\cos\hat{\theta}[i]-\cos\theta^*\right)^2}, \tag{33}$$

where # is the total number of trials, and $\hat{\theta}[i]$ denotes the estimate of ith trial. In the following, we used $\# = 200$ samples to evaluate the RMSE. We assumed the depolarization noise with the parameters $p_s = 1$ and $p_q = 0.99$. The number of qubits was set to 20, which corresponds to $d = 2^{20}$. We also set the number of measurements $N = 500$ for each circuit. To obtain maximum likelihood estimates based on the measurements, we used a modified brute force, in which the search domain is narrowed as the measurement process proceeds. The amplified level α_k of the kth measurement process ($k = 2, ..., M$) was determined from the maximum likelihood estimate $\hat{\theta}_{k-1}$ and the optimization range D_{k-1}. Here, the discrete set D_{k-1} was set to

$$D_{k-1} := \{2^{k-1}, 2^{k-1}+1, \cdots, 2^k\}. \tag{34}$$

Note that the exponential increase of the number of elements was inspired by the fact: the previous research [22] unveiled that when there is no noise, the exponential increment sequence $m_k = 2^{k-1}$ can achieve the Heisenberg limit in the context of standard amplitude estimation method introduced in Section 2.

Figure 2 shows the relationship between total query complexity and estimation errors. Here, the total query complexity is defined by $N\sum_{k=1}^{M}\alpha_k$, and the plots correspond to $M = 3, 4, \cdots, 12$ from left to right. Note that since the total query complexity is a random variable due to the randomness of α_k, we plotted the RMSE as a function of the arithmetic mean of the total query complexity realized in $\# = 200$ trials. We calculated the classical/quantum Cramér-Rao lower bound: CCR bound and QCR bound, respectively, based on the true value θ^* and the asymptotic sequence $\{\alpha^*_k\}_k$ as in Eq. (30). It is worth noting that we confirmed that the difference between the asymptotic CCR/QCR bounds and the bounds from the $\# = 200$ average of the realized total Fisher information can be negligible. For the six target values, the RMSE achieves the CCR bound sufficiently, and the ratio of the RMSE to the bound is at most 1.25. Thus, we consider $N = 500$ is sufficient for the asymptotic property Eq. (31). Note that if we employ larger N such as the maximum number of shots available for real quantum devices, Eq. (31) guarantees that the ratio get closer to one. On the other hand, as the total query complexity increases, the CCR/QCR bounds are saturated around $M = 10$. This is



(a) $\cos\theta^* = 0.042$ (b) $\cos\theta^* = -0.1$

(c) $\cos\theta^* = 0.25$ (d) $\cos\theta^* = 0.5$

(e) $\cos\theta^* = -0.67$ (f) $\cos\theta^* = 0.95$

Fig. 2 The estimation error and the total query complexity for several target values. The root mean squared error (RMSE) of the estimator $\cos\hat{\theta}$ was evaluated by 200 trials. Since the total query complexity is a random variable in our method, the RMSE was plotted as a function of the averaged total query complexity. The blue solid line and green dashed line represent the asymptotic values of CCR/QCR bounds obtained from the corresponding classical/quantum Fisher information $\mathcal{I}^*_{\mathrm{c/q,tot}}(\theta^*)$, respectively.

because the Fisher information from the circuit employed in $M > 10$ is buried in depolarization noise.

In addition, it should be emphasized that the CCR bound is nearly identical to the QCR bound. As seen in the next subsetion, this holds for almost all θ^*.

4.2 Efficiency of our estimator

As described in Section 3, the classical Fisher information associated with measurements Eqs. (14), (19) vanishes at certain points. However, our algorithm can avoid the points and retain large Fisher information due to the optimization of amplified levels. To confirm that, we calculated the asymptotic total Fisher information $\mathcal{I}^*_{\mathrm{c/q,tot}}(\theta^*)$ for every $\theta^* \in (0, \pi)$. In the following, we considered the same experimental setting in the previous subsection. The number of measurement processes M was set to nine, which includes the maximum point of the quantum Fisher information under the current noise level.

Figure 3 shows target value dependency of the ratio between the (asymptotic) total classical Fisher information of our method and that of other methods. The classical Fisher information for the standard sampling method, which is usually employed for VQE [20] calculations, is defined by Eq. (30) with $\alpha^*_k = 1\ \forall k$, while that for the non-adaptive method [22, 23, 26] is defined by Eq. (30) with $\alpha^*_1 = 1,\ \alpha^*_k = 2^k\ \forall k \geq 2$. Since the classical Fisher information for the standard sampling is a constant regardless of the target value in the current setting, its plot reflects the behavior of our classical Fisher information to the target value. In contrast to our total classical Fisher

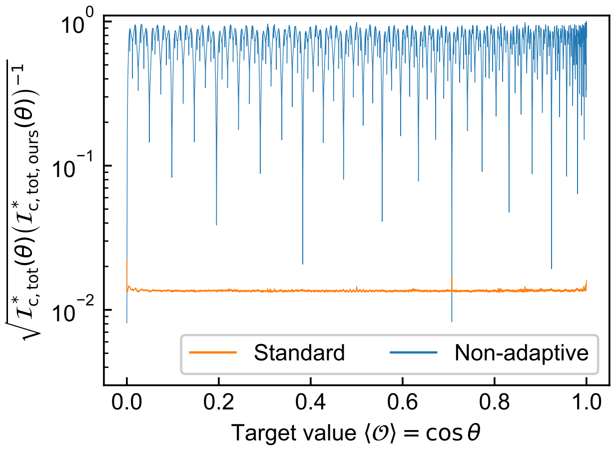


Fig. 3 Comparison of the total classical Fisher information of our method and that of other methods. Orange (bottom) and blue (top) lines represent the ratio of the Fisher information corresponding to standard sampling method $\alpha_k^* = 1$ and non-adaptive method $\alpha_1^* = 1,\ \alpha_k^* = 2^k\ (k \geq 2)$, respectively. As for the our classical Fisher information, we used its asymptotic value to plot this figure.

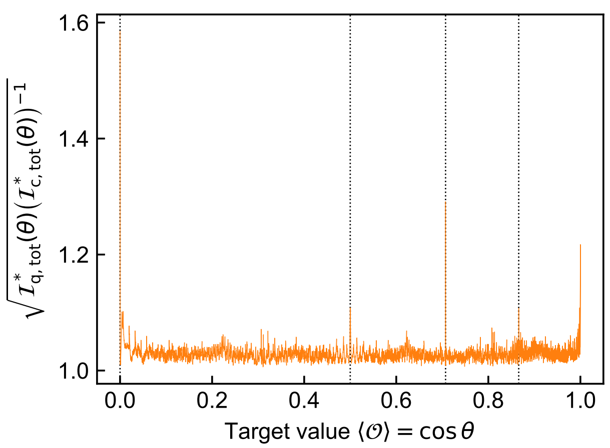


Fig. 4 Comparison of the asymptotic total classical and quantum Fisher information of our method in twenty qubits. The vertical dotted lines represent the target values corresponding to $\theta^* = \pi/j,\ j = 2, 3, 4, 6$ from left to right.

information, that for the non-adaptive method heavily depends on the target value. In addition, there are some target values for which the estimation error bound deviates by more than two order of magnitude from ours. Therefore, when the amplified level $\{\alpha_k\}_k$ is chosen non-adaptively, even if the asymptotic properties of maximum likelihood estimators hold, the estimation efficiency is significantly decreased. Comparing to the standard sampling method, we see an improvement of about two order of magnitude for all target values. The improvement depends on the noise level of the quantum devices, and therefore the estimation efficiency is further accelerated when there is less noise. Consequently, we confirm that our total classical Fisher information retains large Fisher information regardless of the target value.

Figure 4 shows the ratio of the asymptotic total classical/quantum Fisher information. For almost all target values, the difference between asymptotic values of CCR bound and the QCR bound is sufficiently small, which is in line with the results of the previous subsection. In general, the QCR bound, which means the ultimate limit of the estimation precision, can only be achieved when an optimal measurement tailored for a given quantum state performs. The results demonstrate that Eq. (32) holds, that is, our estimation method can select the almost optimal measurements (POVM) with respect to its Fisher information. For the target values $\theta^* = \pi/2, \pi/3, \pi/4, \pi/6$, the difference of the QCR bound and the CCR bound is slightly larger than that for other target values. This is because values of $(m+1)\theta^* \pmod{2\pi}$ obtained from the amplitude amplification are limited at these points, and therefore the classical Fisher information cannot get close to the quantum Fisher information, i.e., $(m+1)\theta^*$ does not get sufficiently close to $k\pi/2,\ k \in \mathbb{Z}$ (not identical to $k\pi/2$, see [27]). Fortunately, the peaks at these points in Fig. 4 are very sharp, and it may be neglected in usual. This is because when the target values shift slightly from these points the values of $(m+1)\theta^* \pmod{2\pi}$ will fill the domain $[0, 2\pi)$ exponentially in the optimization of amplified levels on the current $\{D_k\}_k$.

5. Conclusions

We have proposed a quantum mean value estimation method in a noisy environment for the ultimate precision based on the quantum Cramér-Rao lower bound. This method employs a modified maximum likelihood estimation with the adaptive measurements consisting of multiple POVMs and the amplitude amplifications, which are assumed to induce a depolarization noise. The key element in our method is an optimization of the POVMs and the number of querying the amplitude amplification operators in order to enhance the Fisher information. Importantly, employing the 2-type POVMs our method has not only large Fisher information but also the provable statistical properties such as consistency and asymptotic normality.

To demonstrate the asymptotic properties, we evaluated the root mean squared error (RMSE) of our method for several target values and confirmed that the RMSE saturated the asymptotic classical (also quantum) Cramér-Rao lower bound sufficiently in practical number of measurements. In addition, we clarified the dependance of our total Fisher information on the target value by evaluating the asymptotic classical/quantum Cramér-Rao lower bounds. Although the previous research [28, 23] implies that the classical Fisher information deteriorates significantly for the certain target values under depolarization noise, our method has confirmed to retain large Fisher information regardless of the target value due to the adaptive optimization. Moreover, we also confirmed that the asymptotic classical Cramér-Rao lower bound of our method almost saturate the ultimate bound in large system dimension $d = 2^{20}$, i.e., twenty-qubit system.

6. Acknowledgements

K.W. and K.F. thank IPA for its support through MITOU target program. We would like to thank Dr. Yasunari Suzuki and Dr. Yuuki Tokunaga. This work was supported by MEXT Quantum Leap Flagship Program Grant Number JPMXS0118067285 and JPMXS0120319794.

References

[1] Aaronson, S. and Rall, P.: Quantum approximate counting, simplified, *Symposium on Simplicity in Algorithms*, SIAM, pp. 24–32 (2020).

[2] Braunstein, S. L. and Caves, C. M.: Statistical distance and the geometry of quantum states, *Physical Review Letters*, Vol. 72, No. 22, p. 3439 (1994).

[3] Brown, E. G., Goktas, O. and Tham, W.: Quantum amplitude estimation in the presence of noise, *arXiv preprint arXiv:2006.14145* (2020).

[4] Callison, A. and Browne, D.: Improved maximum-likelihood quantum amplitude estimation, *arXiv preprint arXiv:2209.03321* (2022).

[5] Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S. C., Endo, S., Fujii, K., McClean, J. R., Mitarai, K., Yuan, X., Cincio, L. et al.: Variational quantum algorithms, *Nature Reviews Physics*, Vol. 3, No. 9, pp. 625–644 (2021).

[6] Cramér, H.: Mathematical methods of statistics. (1946).

[7] Degen, C. L., Reinhard, F. and Cappellaro, P.: Quantum sensing, *Reviews of modern physics*, Vol. 89, No. 3, p. 035002 (2017).

[8] Giurgica-Tiron, T., Kerenidis, I., Labib, F., Prakash, A. and Zeng, W.: Low depth algorithms for quantum amplitude estimation, *Quantum*, Vol. 6, p. 745 (2022).

[9] Gonthier, J. F., Radin, M. D., Buda, C., Doskocil, E. J., Abuan, C. M. and Romero, J.: Identifying challenges towards practical quantum advantage through resource estimation: the measurement roadblock in the variational quantum eigensolver, *arXiv preprint arXiv:2012.04001* (2020).

[10] Grinko, D., Gacon, J., Zoufal, C. and Woerner, S.: Iterative quantum amplitude estimation, *npj Quantum Information*, Vol. 7, No. 1, pp. 1–6 (2021).

[11] Hayashi, M.: *Quantum information*, Springer (2006).

[12] Helstrom, C.: The minimum variance of estimates in quantum signal detection, *IEEE Transactions on information theory*, Vol. 14, No. 2, pp. 234–242 (1968).

[13] Helstrom, C. W.: Quantum detection and estimation theory, *Journal of Statistical Physics*, Vol. 1, No. 2, pp. 231–252 (1969).

[14] Holevo, A. S.: *Probabilistic and statistical aspects of quantum theory*, Vol. 1, Springer Science & Business Media (2011).

[15] Jiang, Z.: Quantum Fisher information for states in exponential form, *Physical Review A*, Vol. 89, No. 3, p. 032128 (2014).

[16] Johnson, P. D., Kunitsa, A. A., Gonthier, J. F., Radin, M. D., Buda, C., Doskocil, E. J., Abuan, C. M. and Romero, J.: Reducing the cost of energy estimation in the variational quantum eigensolver algorithm with robust amplitude estimation, *arXiv preprint arXiv:2203.07275* (2022).

[17] Koh, D. E., Wang, G., Johnson, P. D. and Cao, Y.: A framework for engineering quantum likelihood functions for expectation estimation, *arXiv preprint arXiv:2006.09349* (2020).

[18] Martyn, J. M., Rossi, Z. M., Tan, A. K. and Chuang, I. L.: Grand unification of quantum algorithms, *PRX Quantum*, Vol. 2, No. 4, p. 040203 (2021).

[19] Nakaji, K.: Faster amplitude estimation, *arXiv preprint arXiv:2003.02417* (2020).

[20] Peruzzo, A.McClean, J.Shadbolt, P.Yung, M.-H.Zhou, X.-Q.Love, P. J.Aspuru-Guzik, A.Jeremy LObrienA variational eigenvalue solver on a photonic quantum processor, *Nature communications*, Vol. 5, No. 1, pp. 1–7 (2014).

[21] Shao, J.: *Mathematical statistics*, Springer Science & Business Media (2003).

[22] Suzuki, Y., Uno, S., Raymond, R., Tanaka, T., Onodera, T. and Yamamoto, N.: Amplitude estimation without phase estimation, *Quantum Information Processing*, Vol. 19, No. 2, pp. 1–17 (2020).

[23] Tanaka, T., Suzuki, Y., Uno, S., Raymond, R., Onodera, T. and Yamamoto, N.: Amplitude estimation via maximum likelihood on noisy quantum computer, *Quantum Information Processing*, Vol. 20, No. 9, pp. 1–29 (2021).

[24] Tanaka, T., Uno, S., Onodera, T., Yamamoto, N. and Suzuki, Y.: Noisy quantum amplitude estimation without noise estimation, *Physical Review A*, Vol. 105, No. 1, p. 012411 (2022).

[25] Tilly, J., Chen, H., Cao, S., Picozzi, D., Setia, K., Li, Y., Grant, E., Wossnig, L., Rungger, I., Booth, G. H. et al.: The variational quantum eigensolver: a review of methods and best practices, *Physics Reports*, Vol. 986, pp. 1–128 (2022).

[26] Uno, S., Suzuki, Y., Hisanaga, K., Raymond, R., Tanaka, T., Onodera, T. and Yamamoto, N.: Modified Grover operator for quantum amplitude estimation, *New Journal of Physics*, Vol. 23, No. 8, p. 083031 (2021).

[27] Wada, K., Fukuchi, K. and Yamamoto, N.: Quantum-enhanced mean value estimation via adaptive measurement, *in preparation* (xxxx).

[28] Wang, G., Koh, D. E., Johnson, P. D. and Cao, Y.: Minimizing estimation runtime on noisy quantum computers, *PRX Quantum*, Vol. 2, No. 1, p. 010346 (2021).

[29] Wecker, D., Hastings, M. B. and Troyer, M.:

Progress towards practical quantum variational algorithms, *Physical Review A*, Vol. 92, No. 4, p. 042303 (2015).

[30] Yao, Y., Ge, L., Xiao, X., Wang, X. and Sun, C.: Multiple phase estimation for arbitrary pure states under white noise, *Physical Review A*, Vol. 90, No. 6, p. 062113 (2014).