scientific reports

OPEN

Quantum error mitigation via quantum-noise-effect circuit groups

Yusuke Hama[1]✉ & Hirofumi Nishi[1,2]

Near-term quantum computers have been built as intermediate-scale quantum devices and are fragile against quantum noise effects, namely, NISQ devices. Traditional quantum-error-correcting codes are not implemented on such devices and to perform quantum computation in good accuracy with these machines we need to develop alternative approaches for mitigating quantum computational errors. In this work, we propose quantum error mitigation (QEM) scheme for quantum computational errors which occur due to couplings with environments during gate operations, i.e., decoherence. To establish our QEM scheme, first we estimate the quantum noise effects on single-qubit states and represent them as groups of quantum circuits, namely, quantum-noise-effect circuit groups. Then our QEM scheme is conducted by subtracting expectation values generated by the quantum-noise-effect circuit groups from those obtained by the quantum circuits for the quantum algorithms under consideration. As a result, the quantum noise effects are reduced, and we obtain approximately the ideal expectation values via the quantum-noise-effect circuit groups and the numbers of elementary quantum circuits composing them scale polynomial with respect to the products of the depths of quantum algorithms and the numbers of register bits. To numerically demonstrate the validity of our QEM scheme, we run noisy quantum simulations of qubits under amplitude damping effects for four types of quantum algorithms. Furthermore, we implement our QEM scheme on IBM Q Experience processors and examine its efficacy. Consequently, the validity of our scheme is verified via both the quantum simulations and the quantum computations on the real quantum devices. Our QEM scheme is solely composed of quantum-computational operations (quantum gates and measurements), and thus, it can be conducted by any type of quantum device. In addition, it can be applied to error mitigation for many other types of quantum noise effects as well as noisy quantum computing of long-depth quantum algorithms.

The research and development of quantum computers are currently an important and active field of quantum information science and technology[1–13]. On the one side, quantum computer devices have been engineered with state-of-the-art technologies using various kinds of elements including superconducting circuits[7–11,14,15,16,17] and trapped ions[7,12,13,18–22]. On the other side, toward the application to, for example, material science, quantum chemistry, optimization problems, and quantum machine learning, many new kinds of quantum algorithms have been recently developed such as Variational Quantum Eigensolver (VQE)[23–27], Quantum Approximate Optimization Algorithm (QAOA)[27–33], and hybrid quantum-classical machine learning algorithms[27,34–36]. These algorithms have characteristics such that they are constructed by the hybridization between quantum and classical computational procedures. Recently, in the task of sampling random quantum circuits, quantum supremacy has been demonstrated using the superconducting circuit device[37]. All these facts are implying important milestones for the advancement of the research and development of the quantum computers and the broadening of quantum-computing applications to many fields of science and engineering.

While the above successful results of the research and development of quantum computers have been reported, near-term quantum computers based on circuit models have been built as intermediate-scale quantum devices yet and are fragile against quantum noise effects: they are called noisy intermediate-scale quantum (NISQ) devices[10,12,38]. Quantum noise effects (decoherence) are major obstacles for performing quantum computation and historically many great efforts have been made on reducing such effects[39,40]. One of the traditional and representative schemes for this is the quantum-error-correction (QEC) coding[5,8,10,11,17,41–44]. Another important one is

[1]Quemix Inc., 2-11-2 Nihombashi, Chuo-ku, Tokyo 103-0027, Japan. [2]Laboratory for Materials and Structures, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8503, Japan. ✉email: yhama@quemix.com

1

the dynamical decoupling which plays fundamental role in extending coherence times of qubits[9,12,17,43,45–49]. The QEC codes are, however, not implemented on NISQ devices and to obtain quantum computational results in good accuracy with NISQ devices we need to search for alternative approaches for mitigating quantum noise effects. This research field is called quantum error mitigation (QEM), and these days, it is one of the important themes of the research and development of quantum computation[26,27,50–76]. The difficulty of the treatment of quantum noises (e.g., amplitude damping, phase damping (dephasing), depolarizing channel) is that we cannot directly construct their inverse processes by quantum gates due to their non-unitarity. On the other hand, it is possible to formulate quantum noise effects as quantum circuits by using ancilla bits and measurements on them[5,77–84]. By utilizing the quantum circuits representing the quantum noise effects under consideration, we expect that we can establish QEM schemes for reducing such effects. If this is established, we become able to mitigate the quantum noise effects by the gate operations and measurements, i.e., QEM conducted by all-quantum-computational operations. In other words, we become able to programmably run quantum algorithms with mitigating the quantum noise effects solely by the quantum computational operations and realize high-accurate quantum computation.

In this work, we propose our QEM schemes for quantum computational errors which occur owing to couplings with environments (decoherence) during gate operations: errors of state preparation (initialization) and measurement, imperfections of quantum gates, and cross talks among qubits are not taken into account. In particular, we make detailed analysis on quantum computational errors generated by amplitude damping (AD) of single-qubit states. We show the schematic representation of our QEM scheme in Fig. 1 and it consists of two elements, the quantum circuit for the quantum algorithm under consideration (original circuit) represented by the blue rectangle and the ensemble of quantum circuits which yields the theoretical value of the quantum computational error due to the quantum noise effect, namely, quantum-noise-effect circuit group and is represented by the orange rectangles. By utilizing the quantum-noise-effect circuit groups, we formulate our QEM scheme as a perturbation theory with respect to a strength(s) of quantum noise(s) and perform it by subtracting the expectation values given by the quantum-noise-effect circuit groups from those generated by the quantum circuits for the quantum algorithms under consideration as expressed by the formula in the green rectangle; see also the right-hand side of the first line in Eq. (8). As a result, the quantum noise effects are mitigated and we approximately obtain the ideal expectation values. Then, we discuss the numbers of elementary quantum circuits which compose the quantum-noise-effect circuit groups and show that they scale polynomial (linear) with respect to the products of the numbers of register bits and the depths of the quantum algorithms (circuit depths or the numbers of unitary gates composing the quantum algorithm). Finally, we numerically demonstrate the validity of our QEM scheme by running noisy quantum simulators of qubits under the AD effects for four types of quantum algorithms in the linear-order perturbation regime. Furthermore, we examine the effectiveness of our QEM scheme by using IBM Q Experience processors[86]. The detailed explanation on how to extend our QEM scheme to other kinds of quantum noise effects including phase damping, generalized amplitude damping (thermalization), and depolarizing channel, and extension of our QEM scheme to higher-order quantum noise effects are given in Supplementary Information.

The structure of this paper is given as follows. It begins by "QEM schemes" with our modeling of the quantum computation under the influence of the quantum noise effect. After then we explain the formalisms of our QEM scheme. In "Numerical simulations", which presents our main results, we demonstrate numerically our QEM
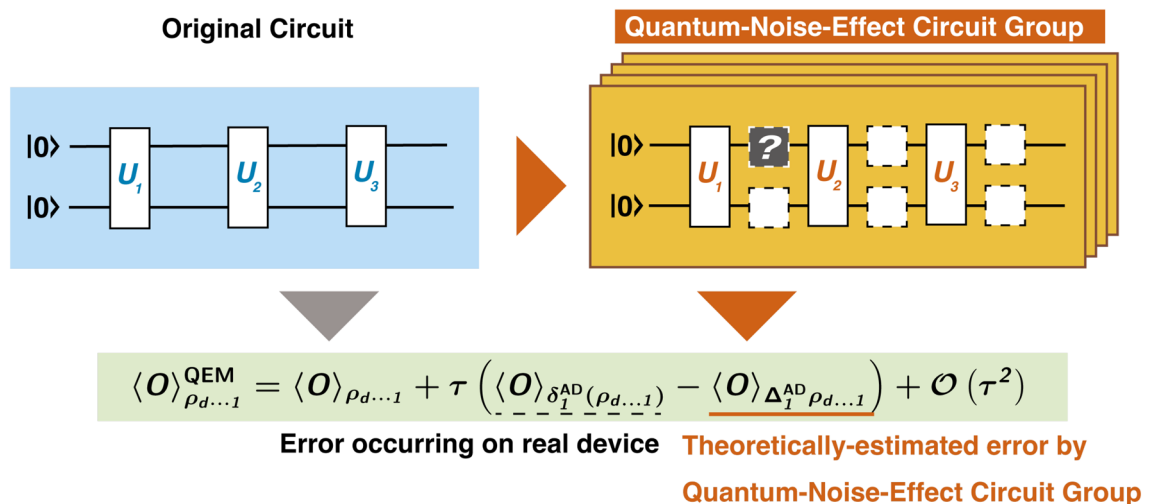

Figure 1. Schematic of our proposed QEM method. The original circuit represented by the blue rectangle (left side) describes the quantum circuit for the quantum algorithm to be run and is composed of the unitary operations U_k with $k = 1, \ldots, d$ and d denotes the depth of the quantum algorithm. It yields the expectation value $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}}$. On the other hand, the quantum-noise-effect circuit group, which is represented by the orange rectangles (right side), is constructed from the original circuit by inserting an additional operation between U_k and U_{k+1} (gray box). It yields the theoretically-estimated quantum computational error $\langle \hat{O} \rangle_{(\Delta_1^{\text{AD}} \rho_{d\cdots1})^{\text{real}}}$. By using these two expectation values, we obtain the equation for our QEM scheme in the green rectangle. Here we have taken $d = 3$.

schemes for the noisy quantum simulations for four types of quantum algorithms. These simulations are done by both our original numerical code and Qiskit[85]. In "QEM scheme implementation", we discuss our quantum computation results for our QEM scheme run on the IBM Q Experience processors[86]. In "Comparison with other methods", we make comparisons between our scheme and other QEM methods. Section "Conclusion and outlook" is devoted to the conclusion of this paper.

QEM schemes
Modeling and formulation

Let us explain our modeling of quantum computation under the influence of quantum noise effects[26,27,50,60,61,85]. In the following, we focus on the amplitude damping (AD) effect: generalized-amplitude-damping (GAD) effect at zero temperature. As discussed later, it is straightforward to generalize the argument for the AD effect to other quantum noise effects such as phase damping (PD) and stochastic Pauli noises. We schematically represent such a circumstance as a quantum circuit and show it in Fig. 2.

There are N_q register bits and the quantum algorithm to be run is represented by the unitary transformation U^{QC}. It is comprised of d unitary transformations described by $U^{\mathrm{QC}} = \prod_{k=1}^{d} U_k = U_d \cdot U_{d-1} \cdots U_2 \cdot U_1$. The unitary transformation $U_k (k = 1, 2, \ldots, d)$ is composed of single- and two-qubit gates. We assume that the duration time (gate operation time) of the unitary transformation U_k is Δt for any k. During the time interval Δt, the register bits are influenced by the AD effects due to couplings with environments, e.g., electromagnetic field in the vacuum, phonons in solids, etc. The quantum master equation describing the AD process in the interaction picture is given by[5,87–89]

$$\frac{\partial \rho(t)}{\partial t} = \gamma \mathcal{L}_{\mathrm{AD}}[\rho(t)]$$
$$= \gamma \sum_{j=0}^{N_q-1} \left[\tilde{\sigma}_j^- \rho(t) \tilde{\sigma}_j^+ - \frac{1}{2} \{ \tilde{\sigma}_j^+ \tilde{\sigma}_j^-, \rho(t) \} \right], \qquad (1)$$

where $\rho(t)$ is the density matrix of the N_q register bits at time t and γ is the decay rate. The symbol $\mathcal{L}_{\mathrm{AD}}$ denotes the Lindblad superoperator of the AD process and the operators $\tilde{\sigma}_j^\pm = \frac{X_j \mp i Y_j}{2}$ are the ladder (raising and lowering) operators acting on the register bit $Q_{\mathrm{r}j}$. X_j and Y_j are X and Y gates acting on $Q_{\mathrm{r}j}$, respectively. $\{A, B\}$ is the anti-commutator between the operators A and B. In our model, we assume that the N_q register bits experience homogeneously the AD effect of single-qubit state given by the decay rate γ. At the initial time $t = 0$, all the register bits are in $|0\rangle$ state (ground state), namely, $\rho(0) = |0\rangle\langle 0|^{\otimes N_q}$. Let us write the total amount of quantum computational time (running time of the quantum algorithm under consideration) by $T (= d \cdot \Delta t)$ while we introduce the dimensionless time $\tau = \gamma \Delta t$. By assuming $\tau \ll 1$, in the following let us evaluate the density matrix at the time T, $\rho(T)$, by using the quantum master equation (1) and express it as a perturbation series with respect to τ given by

$$\rho(T) = \sum_{p=0}^{\infty} \frac{\tau^p}{p!} \cdot \Delta_p^{\mathrm{AD}} \rho_{d\cdots 1}$$
$$= \rho_{d\cdots 1} + \tau \cdot \Delta_1^{\mathrm{AD}} \rho_{d\cdots 1} + \mathcal{O}(\tau^2). \qquad (2)$$

here $\rho_{d\cdots 1} = U^{\mathrm{QC}} \cdot \rho(0) \cdot \left(U^{\mathrm{QC}}\right)^\dagger$ describes the noise-free (ideal) quantum state of the register bits. In other words, it is the ideal output quantum state generated by the quantum algorithm given by U^{QC}. The quantity $\Delta_p^{\mathrm{AD}} \rho_{d\cdots 1} (p \geq 1)$ is the theoretically-evaluated p-th-order AD effect. Let us focus on the first-order AD effect $\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}$ which has the form

$$\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1} = \sum_{k=1}^{d} \Delta_{1,k}^{\mathrm{AD}} \rho_{d\cdots 1},$$
$$\Delta_{1,k}^{\mathrm{AD}} \rho_{d\cdots 1} = \left(\prod_{l=k+1}^{d} U_l \right) \cdot \tilde{\rho}_{k\cdots 1}^{\mathrm{AD}} \cdot \left(\prod_{l=k+1}^{d} U_l \right)^\dagger, \qquad (3)$$
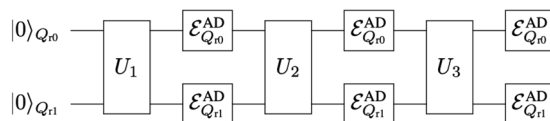
where

Figure 2. Illustration of quantum computation under AD effects represented as a quantum circuit. Here we show it for $d = 3$ and $N_q = 2$. The symbol $\mathcal{E}_{Q_{\mathrm{r}j}}^{\mathrm{AD}}$ expresses the occurrence of AD effect on the register bit $Q_{\mathrm{r}j}$ (the j-th register bit).

$$\tilde{\rho}_{k\cdots1}^{\mathrm{AD}} = \mathcal{L}_{\mathrm{AD}}\left[\rho_{k\cdots1}\right] = \sum_{j=0}^{N_q-1}\left[\tilde{\sigma}_j^- \rho_{k\cdots1}\tilde{\sigma}_j^+ - \frac{1}{2}\left\{P_j^1, \rho_{k\cdots1}\right\}\right],$$

$$\rho_{k\cdots1} = \left(\prod_{l=1}^{k} U_l\right)\cdot\rho(0)\cdot\left(\prod_{l=1}^{k} U_l\right)^\dagger, \tag{4}$$

with $\prod_{l=k+1}^{d} U_l = U_d \cdot U_{d-1}\cdots U_{k+2}\cdot U_{k+1}$ and $\prod_{l=1}^{k} U_l = U_k \cdot U_{k-1}\cdots U_2 \cdot U_1$. In the above equation we have used $\prod_{l=d+1}^{d} U_l = \mathbf{1}$, where $\mathbf{1}$ denotes the identity operator. The operator $P_j^1 = \tilde{\sigma}_j^+ \tilde{\sigma}_j^- = \frac{\mathbf{1}_j - Z_j}{2}$ describes the projection onto the quantum state $|1\rangle_j$ with $\mathbf{1}_j$ and Z_j denoting the identity operator and the Z gate acting on $Q_{\mathrm{r}j}$, respectively: On the other hand, the projection operator of the quantum state $|0\rangle_j$ is given by $P_j^0 = \tilde{\sigma}_j^- \tilde{\sigma}_j^+ = \frac{\mathbf{1}_j + Z_j}{2}$.

QEM scheme

Since we have evaluated the single-qubit-state AD effect, next we discuss our quantum error mitigation (QEM) scheme. We denote the operator of which we are aiming to take an expectation value by $\hat{O}$. When we implement the quantum state ρ on a real device what we actually obtain is a quantum state which is different from ρ due to quantum noise effects: note again that hereinafter we only consider the AD effect. Let us write it by ρ^{real}. We represent the density matrix ρ^{real} in terms of ρ (ideal state) as $\rho^{\mathrm{real}} = \rho + \delta^{\mathrm{AD}}\rho$, where $\delta^{\mathrm{AD}}\rho$ represents the deviation from ρ owing to the AD effect on a real device. Note that we use the symbol δ^{AD} to describe the AD effect on a real device while we use Δ^{AD} to describe the theoretically-estimated AD effect like Eq. (2). Namely, a quantum computational error occurs due to the deviation $\delta^{\mathrm{AD}}\rho$. QEM is a prescription for mitigating the error coming from the deviation $\delta^{\mathrm{AD}}\rho$. Mathematically, this is a task to make the value of $\mathrm{Tr}(\hat{O}\delta^{\mathrm{AD}}\rho)$ as small as possible. In our scheme, we mitigate the error $\mathrm{Tr}(\hat{O}\delta^{\mathrm{AD}}\rho)$ by perturbatively treating the deviation $\delta^{\mathrm{AD}}\rho$ with respect to τ and using the theoretically-estimated AD effect $\Delta_P^{\mathrm{AD}}\rho$. In the following we show such a perturbative analysis up to the first order in τ. The extension of our QEM scheme to higher-order AD effect is discussed in Sect. I in the Supplementary Information. The key procedure of our QEM scheme is to construct quantum circuits for computing the quantity $\mathrm{Tr}(\hat{O}\Delta_1^{\mathrm{AD}}\rho_{d\cdots1})$, which describes the theoretically-estimated quantum computational error of the expectation value $\mathrm{Tr}(\hat{O}\rho)$ in the first order of τ. For doing this, there are two difficulties: (1) the generation of the anti-commutator term $\{P_j^1, \rho_{k\cdots1}\}$ in Eq. (4) and (2) the implementation of the non-unitary operators $\tilde{\sigma}_j^-$ and P_j^1. Let us discuss from our solution to the difficulty (1). We denote some sort of quantum-computational operation (gate operation or measurement) by $\mathcal{A}$. When the operation $\mathcal{A}$ acts on the quantum state $\rho_{k\cdots1}$ the output state we have is $\rho_{k\cdots1} \rightarrow \mathcal{A}\rho_{k\cdots1}\mathcal{A}^\dagger$. The anti-commutator term $\{P_j^1, \rho_{k\cdots1}\}$, in contrast, is not represented in this way, and thus, it is not clear how to generate such a term by the quantum-computational operations. We solve this in the following way. To make our argument simple, here let us focus on the single-register-bit system ($N_q = 1$); the generalization to $N_q \geq 2$ is straightforward and is discussed later. First, we rewrite $\tilde{\rho}_{k\cdots1}^{\mathrm{AD}}$ in Eq. (4) as

$$\tilde{\rho}_{k\cdots1}^{\mathrm{AD}} = -\frac{\rho_{k\cdots1}}{4} + \frac{Z\rho_{k\cdots1}Z}{4} + \tilde{\sigma}^- \rho_{k\cdots1}\tilde{\sigma}^+ - P^1 \rho_{k\cdots1}P^1. \tag{5}$$

In the above way, all the four terms in Eq. (5) are written in the form $\mathcal{A}\rho_{k\cdots1}\mathcal{A}^\dagger$, and thus, we have solved the difficulty (1). Let us analyze the mathematical structure of the right-hand side of Eq. (5). The quantum circuit for creating the first term is straightforward because it is obtained by the quantum circuit composed of U^{QC} (the quantum algorithm under consideration). The implementation of the quantum circuit for the second term $\frac{Z\rho_{k\cdots1}Z}{4}$ is also straightforward because we just apply the Z gate after the operation of U_k. The unclear part is to find ways to construct the quantum circuits for generating the third and fourth terms given by the non-unitary operators $\tilde{\sigma}^-$ and P^1 and this is nothing but the difficulty (2). We solve this by using an ancilla bit and a measurement on
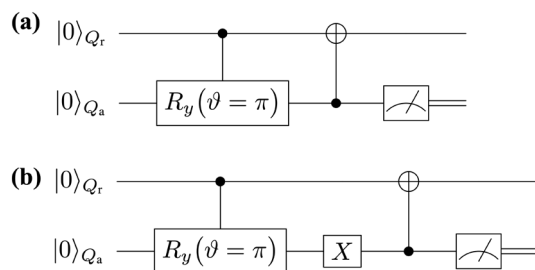

Figure 3. (**a**) Schematic of AD-effect circuit A. When we set $\vartheta = \pi$ and post-select the measurement result of the quantum state of Q_{a} to be $|1\rangle_{Q_{\mathrm{a}}}$, we realize the operation of $\tilde{\sigma}^-$ on Q_{r0}. (**b**) Schematic of AD-effect circuit B. By setting $\vartheta = \pi$ and post-selecting the measurement result of the quantum state of Q_{a} to be $|0\rangle_{Q_{\mathrm{a}}}$, we the operation of P^1 on Q_{r0} is created.

it[5]. For the creating the operation $\tilde{\sigma}^-$, we use the quantum circuit presented in Fig. 3a (AD-effect circuit A) while for the operation of P^1 we use the one in Fig. 3b (AD-effect circuit B).

In both quantum circuits, we regard the ancilla bit Q_a as the environment which induces the AD effect on the register bit Q_{r0}. The interactions between these two qubits are represented by the controlled-rotational gate $U_{CR_y}[Q_{r0}; Q_a](\vartheta)$ and the controlled-not gate $U_{CX}[Q_a; Q_{r0}]$. The controlled-rotational gate $U_{CR_y}[Q_{r0}; Q_a](\vartheta)$ describes the rotation about y axis by the rotational angle ϑ and it is composed of the control bit Q_{r0} and the target bit Q_a. On the other hand, for the gate operation $U_{CX}[Q_a; Q_{r0}]$ the ancilla bit Q_a is the control bit while the register bit Q_{r0} is the target bit. We have used the notation such that the control bit(s) comes before the semicolon while the target bit(s) comes after. Let us explain the output states generated by the AD-effect circuits A and B. For both quantum circuits, the initial quantum states of Q_{r0} and Q_a are the same and it is $\rho_{Q_{r0}Q_a}(0) = \rho_{Q_{r0}}(0) \otimes \rho_{Q_a}(0)$ with $\rho_{Q_{r0}}(0) = |0\rangle_{Q_{r0}}\langle 0|$ and $\rho_{Q_a}(0) = |0\rangle_{Q_a}\langle 0|$. The AD-effect circuit A is given by the unitary operation $U_{ADA} = U_{CX}[Q_a; Q_{r0}] \cdot U_{CR_y}[Q_{r0}; Q_a](\vartheta)$ while the AD-effect circuit B is given by $U_{ADB} = U_{CX}[Q_a; Q_{r0}] \cdot (1_{2 \times 2} \otimes X_{Q_a}) \cdot U_{CR_y}[Q_{r0}; Q_a](\vartheta)$, where $1_{2 \times 2}$ is the two by two identity operator. Owing to these unitary operations, the quantum state generated by the AD-effect circuit A is given by $\rho_{ADA,Q_{r0}Q_a}(\vartheta) = U_{ADA} \cdot \rho_{Q_{r0}Q_a}(0) \cdot (U_{ADA})^\dagger$ while the quantum state created by the AD-effect circuit B is $\rho_{ADB,Q_{r0}}(\vartheta) = U_{ADB} \cdot \rho_{Q_{r0}Q_a}(0) \cdot (U_{ADB})^\dagger$. At the end, we measure the ancilla bit Q_a. Then the quantum states of Q_{r0} (reduced density matrices) are described by the Kraus operators[5,90].

$$\mathcal{K}_0^{ADA} = {}_{Q_a}\langle 0|U_{ADA}|0\rangle_{Q_a} = P^0 + \cos\left(\frac{\vartheta}{2}\right)P^1$$

$$\mathcal{K}_1^{ADA} = {}_{Q_a}\langle 1|U_{ADA}|0\rangle_{Q_a} = \sin\left(\frac{\vartheta}{2}\right)\tilde{\sigma}^-,$$

$$\mathcal{K}_0^{ADB} = {}_{Q_a}\langle 0|U_{ADB}|0\rangle_{Q_a} = \sin\left(\frac{\vartheta}{2}\right)P^1,$$ (6)

$$\mathcal{K}_1^{ADB} = {}_{Q_a}\langle 1|U_{ADB}|0\rangle_{Q_a} = \tilde{\sigma}^+ + \cos\left(\frac{\vartheta}{2}\right)\tilde{\sigma}^-.$$

For the AD-effect circuit A (B) the Kraus operators $\mathcal{K}_0^{ADA}$ ($\mathcal{K}_0^{ADB}$) acts on the register bit Q_{r0} when the measurement outcome of the quantum state of the ancilla bit Q_a is $|0\rangle_{Q_a}$ while $\mathcal{K}_1^{ADA}$ ($\mathcal{K}_1^{ADB}$) operates when the measurement outcome is $|1\rangle_{Q_a}$. When we average these two outcomes, the quantum state of Q_{r0} created by the AD-effect circuit A is given by $\rho_{ADA,Q_{r0}}(\vartheta) = \text{Tr}_{Q_a}\left[\rho_{ADA,Q_{r0}Q_a}(\vartheta)\right] = \sum_{\mu=0,1} \mathcal{K}_\mu^{ADA} \cdot \rho_{Q_{r0}}(0) \cdot (\mathcal{K}_\mu^{ADA})^\dagger$, where the symbol Tr_{Q_a} denotes the partial trace with respect to Q_a degrees of freedom. Similarly, for the AD-effect circuit B we have $\rho_{ADB,Q_{r0}}(\vartheta) = \sum_{\mu=0,1} \mathcal{K}_\mu^{ADB} \cdot \rho_{Q_{r0}}(0) \cdot (\mathcal{K}_\mu^{ADB})^\dagger$. In particular, for the AD-effect circuit A when we take ϑ to be ϑ_t such that $\cos^2\left(\frac{\vartheta_t}{2}\right) = e^{-\gamma t}$[5], the matrix representation of $\rho_{ADA,Q_{r0}}(\vartheta_t)$ is given by

$$\rho_{ADA,Q_{r0}}(\vartheta_t) = \begin{pmatrix} \left[\rho_{ADA,Q_{r0}}(0)\right]_{00} + \left[\rho_{ADA,Q_{r0}}(0)\right]_{11}(1 - e^{-\gamma t}) & \left[\rho_{ADA,Q_{r0}}(0)\right]_{01}e^{-\frac{\gamma t}{2}} \\ \left[\rho_{ADA,Q_{r0}}(0)\right]_{10} \cdot e^{-\frac{\gamma t}{2}} & \left[\rho_{ADA,Q_{r0}}(0)\right]_{11}e^{-\gamma t} \end{pmatrix}. \quad (7)$$

The matrix element $\left[\rho_{ADA,Q_{r0}}(0)\right]_{nn'}$ ($n, n' = 0, 1$) denotes the (n, n')-element of $\rho_{ADA,Q_{r0}}(0)$. The reduced density matrix $\rho_{ADA,Q_{r0}}(\vartheta_t)$ in Eq. (7) is nothing but the solution of the quantum master equation (1). Further, when we take $\vartheta_t \to \pi$, the Kraus operators in Eq. (6) becomes $\{\mathcal{K}_0^{ADA}, \mathcal{K}_1^{ADA}\} \to \{P^0, \tilde{\sigma}^-\}$, $\{\mathcal{K}_0^{ADB}, \mathcal{K}_1^{ADB}\} \to \{P^1, \tilde{\sigma}^+\}$. Therefore, for the case of the AD-effect circuit A by using the measurement result of Q_a such that we post-select the output state of Q_a to be $|1\rangle_{Q_a}$ we can realize the operation of $\tilde{\sigma}^-$ on Q_{r0}. On the other hand, for the AD-effect circuit B by post-selecting the output state of
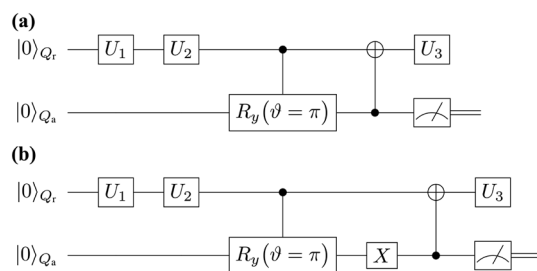

Figure 4. (**a**) Schematic of an elementary quantum circuit in the AD-effect circuit group given by the AD-effect circuit A which generates $\left(\prod_{l=k+1}^{d} U_l\right) \cdot \left(\tilde{\sigma}^- \rho_{k\cdots1} \tilde{\sigma}^+\right) \cdot \left(\prod_{l=k+1}^{d} U_l\right)^\dagger$. For doing so, we set $\vartheta = \pi$ and post-select the measurement result of the quantum state of Q_a to be $|1\rangle_{Q_a}$. (**b**) Schematic of an elementary quantum circuit in the AD-effect circuit group given by the AD-effect circuit B. When we post-select the measurement outcome of Q_a to be $|0\rangle_{Q_a}$, we have $\left(\prod_{l=k+1}^{d} U_l\right) \cdot \left(P^1 \rho_{k\cdots1} P^1\right) \cdot \left(\prod_{l=k+1}^{d} U_l\right)^\dagger$. Both of these quantum circuits in (**a**) and (**b**) are for $k = 2, d = 3$ with $\vartheta = \pi$.

Q_a to be $|0\rangle_{Q_a}$ we realize the operation of P^1 on Q_{r0}. To show the above things concretely, let us present the examples of the quantum circuits for the generation of $\left(\prod_{l=k+1}^{d} U_l\right) \cdot \left(\tilde{\sigma}^- \rho_{k\cdots 1} \tilde{\sigma}^+\right) \cdot \left(\prod_{l=k+1}^{d} U_l\right)^\dagger$ and $\left(\prod_{l=k+1}^{d} U_l\right) \cdot \left(P^1 \rho_{k\cdots 1} P^1\right) \cdot \left(\prod_{l=k+1}^{d} U_l\right)^\dagger$ for $k = 2, d = 3$, and we show them in Fig. 4a,b, respectively. As a result, by using the AD-effect circuits A and B we can perform the actions of $\tilde{\sigma}^-$ and P^1 as described by the third and four terms in Eq. (5), and thus, we have solved the second difficulty (2).

Since the difficulties (1) and (2) have been solved, we are now ready to establish our QEM scheme. To compute the quantity $\mathrm{Tr}(\hat{O} \cdot \Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})$ we need four types of quantum circuits: the original circuit given by U^{QC}, the quantum circuit where the additional Z gate is performed, and the AD-effect circuits A and B. The latter three quantum-circuit ensembles composed of the additional Z-gate, $\tilde{\sigma}^-$, and P^1 operations form the quantum-noise-effect circuit group for the AD effect, i.e, AD-effect circuit group. Hereinafter, let us write the trace of the product between the operator $\hat{O}$ and ρ by $\mathrm{Tr}(\hat{O}\rho) = \langle \hat{O} \rangle_\rho$. We perturbatively express the quantum states $\rho_{d\cdots 1}^{\mathrm{real}}$ with respect to τ as $\rho_{d\cdots 1}^{\mathrm{real}} = \rho_{d\cdots 1} + \delta^{\mathrm{AD}}(\rho_{d\cdots 1})$ with $\delta^{\mathrm{AD}}(\rho_{d\cdots 1}) = \sum_{p=1}^{\infty} \frac{\tau^p}{p!} \cdot \delta_p^{\mathrm{AD}}(\rho_{d\cdots 1})$. Furthermore, we write the quantity which is obtained by the implementation of $\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}$ on a real device by $(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})^{\mathrm{real}}$. With similar to $\rho_{d\cdots 1}^{\mathrm{real}}$, we perturbatively express $(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})^{\mathrm{real}}$ in terms of $\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}$ and τ as $(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})^{\mathrm{real}} = \Delta_1^{\mathrm{AD}} \rho_{d\cdots 1} + \delta^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})$ with $\delta^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}) = \sum_{p=1}^{\infty} \frac{\tau^p}{p!} \cdot \delta_p^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})$. Then, by using Eqs. (2)–(5) we obtain the quantum-error-mitigated expectation value of $\hat{O}$ given by

$$
\begin{aligned}
\langle \hat{O} \rangle_{\rho_{d\cdots 1}}^{\mathrm{QEM}} &\equiv \langle \hat{O} \rangle_{\rho_{d\cdots 1}^{\mathrm{real}}} - \tau \langle \hat{O} \rangle_{(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})^{\mathrm{real}}} \\
&= \langle \hat{O} \rangle_{\rho_{d\cdots 1}} + \tau \left(\langle \hat{O} \rangle_{\delta_1^{\mathrm{AD}}(\rho_{d\cdots 1})} - \langle \hat{O} \rangle_{\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}} \right) \\
&\quad + \mathcal{O}(\tau^2).
\end{aligned}
\tag{8}
$$

The idea of our QEM scheme is clearly represented in the second line of Eq. (8). The first term is the ideal expectation value while the second term represents the conduction of our QEM scheme. It is described as the subtraction between $\langle \hat{O} \rangle_{\delta_1^{\mathrm{AD}}(\rho_{d\cdots 1})}$ (the quantum computational error occurring on a real device) and $\langle \hat{O} \rangle_{\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}}$ (theoretically-evaluated quantum computational error), which is computed by the AD-effect circuit group. The heart of the idea for doing this is that we have considered that the real noise effect $\delta_1^{\mathrm{AD}}(\rho_{d\cdots 1})$ is (approximately) equivalent to the theoretically-estimated noise effect $\Delta_1^{\mathrm{AD}}(\rho_{d\cdots 1})$. When the second term in the second line of Eq. (8) becomes small enough, we consider that we have accomplished in mitigating the error of quantum computation on a real device. Note that the quantum noise effects coming from $\delta^{\mathrm{AD}}(\Delta^{\mathrm{AD}} \rho_{d\cdots 1})$ are suppressed by multiplying $\langle \hat{O} \rangle_{(\Delta^{\mathrm{AD}} \rho_{d\cdots 1})^{\mathrm{real}}}$ by τ (the second term in the right-hand side of the first line of Eq. (8)). This is because the lowest-order error AD effect on the implementation of $\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}$, which is $\delta_1^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})$, becomes $\mathcal{O}(\tau^2)$ due to the multiplication by τ: $\delta^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}) \rightarrow \tau \cdot \delta^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1}) = \sum_{p=1}^{\infty} \tau^{p+1} \cdot \delta_p^{\mathrm{AD}}(\Delta_1^{\mathrm{AD}} \rho_{d\cdots 1})$. Consequently, in the first-order perturbation theory with respect to τ we have established our QEM scheme by the usage of the AD-effect circuit group and is expressed by the formula given by Eq. (8). The above argument on QEM-scheme derivation can be straightforwardly generalized to register-bit systems for $N_q \geq 2$. In this case, $\tilde{\rho}_{k\cdots 1}^{\mathrm{AD}}$ in Eq. (5) is represented as

$$
\begin{aligned}
\tilde{\rho}_{k\cdots 1}^{\mathrm{AD}} &= \mathcal{L}_{\mathrm{AD}}[\rho_{k\cdots 1}] \\
&= \sum_{j=0}^{N_q-1} -\frac{\rho_{k\cdots 1}}{4} + \frac{Z_j \rho_{k\cdots 1} Z_j}{4} + \tilde{\sigma}_j^- \rho_{k\cdots 1} \tilde{\sigma}_j^+ \\
&\quad - P_j^1 \rho_{k\cdots 1} P_j^1.
\end{aligned}
\tag{9}
$$

We can apply our QEM scheme to the N_q register-bit system in the following way. We prepare N_q register bits and one ancilla bit $\left\{ Q_{r0}, Q_{r1}, \ldots, Q_{rN_q-1}, Q_a \right\}$. Then we create an ensemble of quantum circuits composed of the j-th register bit $Q_{rj}(j = 0, 1, \ldots, N_q - 1)$ and the ancilla bit Q_a which describes that Q_{rj} is subject to the AD effect induced by the ancilla bit Q_a. Namely, we create the ensemble of four types of quantum circuits composed of Q_{rj} and Q_a, the original quantum circuits given by U^{QC}, the quantum circuits with additional Z-gate operations, and the AD-effect circuits A and B. By summing up all these quantum circuits, we obtain the AD-effect circuit group which enables us to perform QEM for N_q-register-bit system under the AD effect. The total number of quantum circuits which compose the AD-effect circuit group is $3dN_q + 1$, and thus, it scales polynomial in dN_q, which is not so high-cost computational performance.

Before ending this section, let us explain two ways to extend our QEM formalism. Firstly, we can extend into cases of other quantum noise channels including generalized amplitude damping (GAD), phase damping (PD), their composite channels, and stochastic Pauli noise models such as bit flip, phase flip, bit-phase flip, and depolarizing channel. Secondly, we can create quantum-noise-effect circuit groups which enable us to perform QEM for higher-order quantum noise effects. We present arguments on these two cases in Sect. I in the Supplementary Information.

Numerical simulations

In this section, we numerically demonstrate our QEM scheme for four types of algorithms. For the quantum noise effect we choose AD effect. In "Preliminary", as a preliminary of our QEM demonstration, we present the results of two algorithms: the algorithm composed of the initial X-gate operation and the repetition of the Hadamard gate acting on a single register bit and that composed of the initial $X \otimes X$ operation and the controlled-Hadamard gate acting on two register bits. In "Quantum amplitude amplification", we show the results of QEM for a long-term quantum algorithm and here we choose quantum amplitude amplification (QAA) algorithm (Grover's search algorithm). In "QAOA". we show the results of recently developed NISQ quantum algorithms, Quantum Approximate Optimization Algorithm (QAOA).

In the following, let us explain the formalism of our noisy quantum simulations (numerical simulations of running quantum algorithms with real quantum devices performed by classical computers). As shown in Fig. 2, every time we apply an unitary (gate) operation, the N_q register bits experience AD effects. Suppose that at time t_0 the quantum states of the register bits were given by the density matrix $\rho(t_0)$. According to the quantum master equation (1), when the unitary gate U has been applied to the register bits within the time interval Δt the quantum state of the register bits at $t = t_0 + \Delta t$ is expressed by the density matrix

$$\rho^{\mathrm{AD}}(t_0 + \Delta t) = \mathcal{E}^{\mathrm{AD}}[U\rho(t_0)U^{\dagger}], \tag{10}$$

where $\mathcal{E}^{\mathrm{AD}}[\cdots]$ is the superoperator which describes the AD effect on the N_q register bits and it is given by[5,85]

$$\mathcal{E}^{\mathrm{AD}}[\rho] = \sum_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}}} \mathcal{M}_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}}} \cdot \rho \cdot \mathcal{M}_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}}}^{\dagger},$$

$$\mathcal{M}_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}}} = \mathcal{M}_{n_{Q_{r0}}}^{\mathrm{AD}} \otimes \cdots \otimes \mathcal{M}_{n_{Q_{rN_q-1}}}^{\mathrm{AD}}, \tag{11}$$

$$\mathcal{M}_0^{\mathrm{AD}} = \begin{bmatrix} 1 & 0 \\ 0 & \cos\left(\frac{\vartheta_{\tau}}{2}\right) \end{bmatrix}, \quad \mathcal{M}_1^{\mathrm{AD}} = \begin{bmatrix} 0 & \sin\left(\frac{\vartheta_{\tau}}{2}\right) \\ 0 & 0 \end{bmatrix},$$

where $n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}} = 0, 1$, and $\cos^2\left(\frac{\vartheta_{\tau}}{2}\right) = e^{-\tau}$. The operators $\mathcal{M}_0^{\mathrm{AD}}$ and $\mathcal{M}_1^{\mathrm{AD}}$ are the Kraus operators acting on single-qubit states and describe the influence of the AD effect on a single register bit during the time interval Δt. Here we assume that the N_q register bits homogeneously experience the single-qubit-state AD effect as described in Eq. (11). For later convenience, let us introduce the notation which describes the operation of the unitary operator U on the quantum state ρ by $\mathcal{T}[\rho, U](= U\rho U^{\dagger})$. Let us write the quantum state generated by the unitary transformation U^{QC} under the AD effect by $\rho_{d\cdots 1}^{\mathrm{AD}}$. By using the superoperator $\mathcal{E}^{\mathrm{AD}}[\cdots]$, the output state $\rho_{d\cdots 1}^{\mathrm{AD}}$ is represented as

$$\rho_{d\cdots 1}^{\mathrm{AD}} = \mathcal{E}^{\mathrm{AD}}\big[\mathcal{T}[\cdots \mathcal{E}^{\mathrm{AD}}\big[\mathcal{T}[\mathcal{E}^{\mathrm{AD}}\big[\mathcal{T}[\rho(0), U_1]\big], U_2]\big], \ldots U_d]\big]. \tag{12}$$

Equations (11) and (12) are the basic equations of our noisy quantum simulations. Namely, we perform our noisy quantum simulations by identifying the quantum state $\rho_{d\cdots 1}^{\mathrm{AD}}$ in the above equation with $\rho_{d\cdots 1}^{\mathrm{real}}$, which is the AD-affected quantum state generated by the unitary transformation U^{QC} on real devices. We conduct QEM described by Eq. (8) for various values of τ by tuning the value of ϑ_{τ}. When we execute the AD circutis A and B, we run quantum simulations of the $N_q + 1$ qubit systems. The right-hand side of Eq. (11) becomes $\sum_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}},n_{Q_a}} \mathcal{M}_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}},n_{Q_a}} \cdot \rho \cdot \mathcal{M}_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}},n_{Q_a}}^{\dagger}$, where $\mathcal{M}_{n_{Q_{r0}},\ldots,n_{Q_{rN_q-1}},n_{Q_a}} = \mathcal{M}_{n_{Q_{r0}}}^{\mathrm{AD}} \otimes \cdots \otimes \mathcal{M}_{n_{Q_{rN_q-1}}}^{\mathrm{AD}} \otimes \mathcal{M}_{n_{Q_a}}^{\mathrm{AD}}$. Note that the ancilla bit Q_a is treated similarly as the other N_q register bits such that Q_a is subject to the same quantum noise effect described by the Kraus operators in Eq. (11) as the other N_q register bits do. Correspondingly, we mitigate the quantum noise effects influencing both Q_a and the N_q register bits via Eq. (8).

For performing our noisy quantum simulations, we have created two types of numerical codes. The first one is our original numerical code and the second one is the numerical code created by Qiskit[85]. In our numerical code, we simply compute the matrix products of the density matrices, the unitary operations U_k, the Kraus operators, and the additional operations such as $Z, \tilde{\sigma}^{\pm}, P^1$. Furthermore, we compute the trace operations between the density matrices and the physical operators O. Namely, our original numerical code is a density-matrix simulator. On the other hand, the Qiskit code is programmed by two numbers, N_{QC} and N_{samp}. To understand them concretely, first let us discuss an example of quantum computing of a single-qubit system. We execute the given quantum circuit and we obtain an output state which is either $|0\rangle$ or $|1\rangle$. We repeat this process N_{QC} times and say that we obtained $|0\rangle$ for $N_{|0\rangle}$ times as the output state while we obtained $|1\rangle$ for $N_{|1\rangle}$ times. Then, the probability weight of $|0\rangle$ is $\frac{N_{|0\rangle}}{N_{\mathrm{QC}}}$ while that of $|1\rangle$ is $\frac{N_{|1\rangle}}{N_{\mathrm{QC}}}$. Namely, the number N_{QC} is the repetition number of quantum computing executed by the given quantum circuit, and the numerical simulations in our original code describes the quantum simulations in the limit $N_{\mathrm{QC}} \to \infty$. Next let us explain what N_{samp} is. It is the repetition number of "the execution of the quantum computation for N_{QC} times under the given (fixed) quantum circuit". By introducing such a number, we effectively perform the quantum computation under the given quantum circuit with the repetition number $N_{\mathrm{QC}} \times N_{\mathrm{samp}}$ in our Qiskit code. In contrast to N_{QC}, the repetition number N_{samp} is not coming from foundations of quantum mechanics and we have introduced this quantity owing to the following two reasons. First, due to our survey there is an upper limit on N_{QC} in the Qiskit program and is 10^7. By introducing N_{samp}, we become able to effectively perform quantum computations with repetition numbers greater

than the upper limit of $N_{QC} (= 10^7)$ in the Qiskit program. Second, we consider that it is not enough to just show a single data point of "the quantum computational result obtained by the fixed quantum circuit and N_{QC}" to see how largely it deviates from the true quantum computational result (quantum computation in the limit $N_{QC} \to \infty$). To present how largely the quantum computational results for fixed and finite N_{QC} deviate (finite-size effects of N_{QC}) from the true ones, we introduce N_{samp} and show visually such deviations. In order to describe the second reason more mathematically, let write a binary which labels a quantum state of qubit Q_α ($\alpha = 0, \ldots, N_q - 1$) by $n_{Q_\alpha}: |n_{Q_\alpha}\rangle$ with $n_{Q_\alpha} = 0, 1$ and an output state is described by the computational basis states $|n_{Q_{N_q-1}}\rangle_{Q_{N_q-1}} \otimes \cdots \otimes |n_{Q_0}\rangle_{Q_0} = |n_{Q_{N_q-1}} \cdots n_{Q_0}\rangle$. Let us say that we focus on the specific quantum state $|\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}\rangle$ and consider how many times it is obtained as the output state for given N_{QC}. By saying that $|\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}\rangle$ has been obtained as the output state for $N_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (N_{QC})$ times, the probability of an output state being $|\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}\rangle$ is $w_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (N_{QC}) = \frac{N_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (N_{QC})}{N_{QC}}$. Furthermore, we write the probability such that $|\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}\rangle$ is going to become observed in quantum computing under $N_{QC} \to \infty$ as $p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}}$. We redescribe such a circumstance as the binomial distribution denoted by $B(p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}}, N_{QC})$ and introduce the random variable X_i such that in the ith round of quantum computation we take $X_i = 1$ when the output state is $|\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}\rangle$, otherwise $X_i = 0$. Next, we introduce another random variable $\bar{X} = \frac{\sum_{j=0}^{N_{QC}-1} X_j}{N_{QC}}$, which is equivalent to $w_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (N_{QC})$. Owing to the central limit theorem, we obtain $\lim_{N_{QC} \to \infty} w_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (N_{QC}) = p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}}$. In other words, the binomial distribution $B(p, N_{QC})$ approaches to the Gaussian distribution function given by the mean $p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}}$ and the standard deviation $\sqrt{\frac{p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (1 - p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}})}{N_{QC}}}$, namely $N\left(p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}}, \frac{p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (1 - p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}})}{N_{QC}}\right)$. Let us say that we perform quantum computing with accuracy ϵ. From the standard deviation of $N\left(p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}}, \frac{p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (1 - p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}})}{N_{QC}}\right)$, we can estimate the lower bound of N_{QC} for doing this and is equal to $\frac{p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}} (1 - p_{\hat{n}_{Q_{N_q-1}} \cdots \hat{n}_{Q_0}})}{\epsilon^2}$. Note that in order to perform QEM with the accuracy ϵ the total repetition number of quantum computing gets larger with respect to d and N_q due to the quantum-noise-effect circuit group and this issue is addressed later. The number N_{QC} describes the repetition number of quantum computation owing to the given quantum circuit whereas the number N_{samp} describes how many times you conduct the sampling for the expectation values of physical operators obtained by the N_{QC}-repeated quantum computation. Owing to this sampling, the repetitive number of the quantum computations effectively becomes $N_{QC} \times N_{samp}$, and our simulation results become more trustable. In our simulations we take $N_{QC} = 2^{10}$ and $N_{samp} = 100$ for both the original quantum circuit and each elementary circuit of quantum-noise-effect circuit group. Our original numerical code, on the other hand, is the code for a noisy quantum simulation in the limit $N_{QC} \to \infty$, and basically, it performs pure linear algebraical computations such as matrix-product and trace operations. We note that when we create the numerical codes with Qiskit, we need to be careful with how controlled-unitary operators are implemented. We have examined that on Qiskit program the controlled-unitary operators are implemented as the decomposition of U_{CX} gates and single-qubit unitary gates. For example, the control-R_y gate $U_{CR_y(\vartheta)}[Q_{r0}; Q_{r1}]$ is decomposed as $U_{CR_y(\vartheta)}[Q_{r0}; Q_{r1}] = (\mathbf{1}_{2 \times 2_{Q_{r0}}} \otimes R_y(\frac{\vartheta}{2})_{Q_{r1}}) \cdot U_{CX}[Q_{r1}; Q_{r0}] \cdot (\mathbf{1}_{2 \times 2_{Q_{r0}}} \otimes R_y(-\frac{\vartheta}{2})_{Q_{r1}}) \cdot U_{CX}[Q_{r0}; Q_{r1}]$. Therefore, when we simulate QAA for three-qubit systems and QAOA with our original code we implement $U_{CR_y(\vartheta)}[Q_{r0}; Q_{r1}]$ in the same way as Qiskit program does.

In order to quantitatively describe the validity of our QEM scheme we introduce the measure defined by

$$\mathrm{RT_{QEM}} = \frac{|\langle \hat{O} \rangle_{\rho_{d \cdots 1}^{real}} - \langle \hat{O} \rangle_{\rho_{d \cdots 1}}|}{|\langle \hat{O} \rangle_{\rho_{d \cdots 1}} - \langle \hat{O} \rangle_{\rho_{d \cdots 1}^{QEM}}|}. \tag{13}$$

The numerator of $\mathrm{RT_{QEM}}$ in Eq. (13) describes the absolute of the difference between the expectation value owing to the noisy quantum simulation $\langle \hat{O} \rangle_{\rho_{d \cdots 1}^{real}}$ (no QEM) and the ideal expectation value $\langle \hat{O} \rangle_{\rho_{d \cdots 1}^{ideal}}$. On the other hand, the denominator represents the absolute of the difference between the expectation value obtained by our QEM scheme $\langle \hat{O} \rangle_{\rho_{d \cdots 1}^{QEM}}$ (see Eq. (8)) and the ideal expectation value. In other words, the measure $\mathrm{RT_{QEM}}$ in Eq. (13) is the ratio between the absolute of the error without QEM and the one with QEM. Thus, when $\mathrm{RT_{QEM}} > 1$ the expectation value $\langle \hat{O} \rangle_{\rho_{d \cdots 1}^{QEM}}$ is closer to the ideal value than $\langle \hat{O} \rangle_{\rho_{d \cdots 1}^{real}}$, which means that our QEM scheme is working. In addition to the ratio $\mathrm{RT_{QEM}}$ in Eq. (13), we display the results of the expectation values obtained by the ideal simulations, the noisy simulations without QEM, and the noisy simulations with QEM, and show explicitly the validity of our QEM scheme. Note that for our original code we take the noise-strength parameter to be $\vartheta_\tau = i_o \times 0.01$ with $i_o = 0, 1, \ldots, 50$ while for Qiskit code we take $\vartheta_\tau = i_Q \times 0.05$ with $i_Q = 0, 1, \ldots, 10$. For computing the expectation values we include the case $i_o = 0$ and $i_Q = 0$ whereas for computing the ratio $\mathrm{RT_{QEM}}$ we omit $i_o = 0$ and $i_Q = 0$. This is because in this case we have $\langle \hat{O} \rangle_{\rho_{d \cdots 1}} = \langle \hat{O} \rangle_{\rho_{d \cdots 1}^{real}} = \langle \hat{O} \rangle_{\rho_{d \cdots 1}^{QEM}}$ and we encounter in the indefinite $\frac{0}{0}$. In the following, we create a subsection for each quantum algorithm and discuss the results in detail.
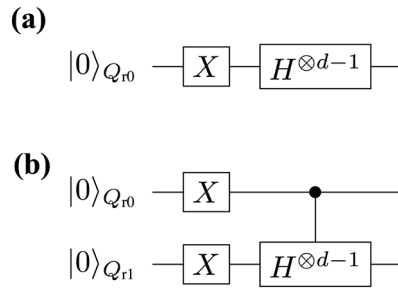
(a)

(b)

Figure 5. Quantum circuits for (**a**) $U_{\text{pre1}}^{\text{QC}} = H^{\otimes d-1} \cdot X$ and (**b**) $U_{\text{pre2}}^{\text{QC}} = (U_{CH}[Q_{r0}; Q_{r1}])^{\otimes d-1}) \cdot X_{Q_{r0}} \otimes X_{Q_{r1}}$.
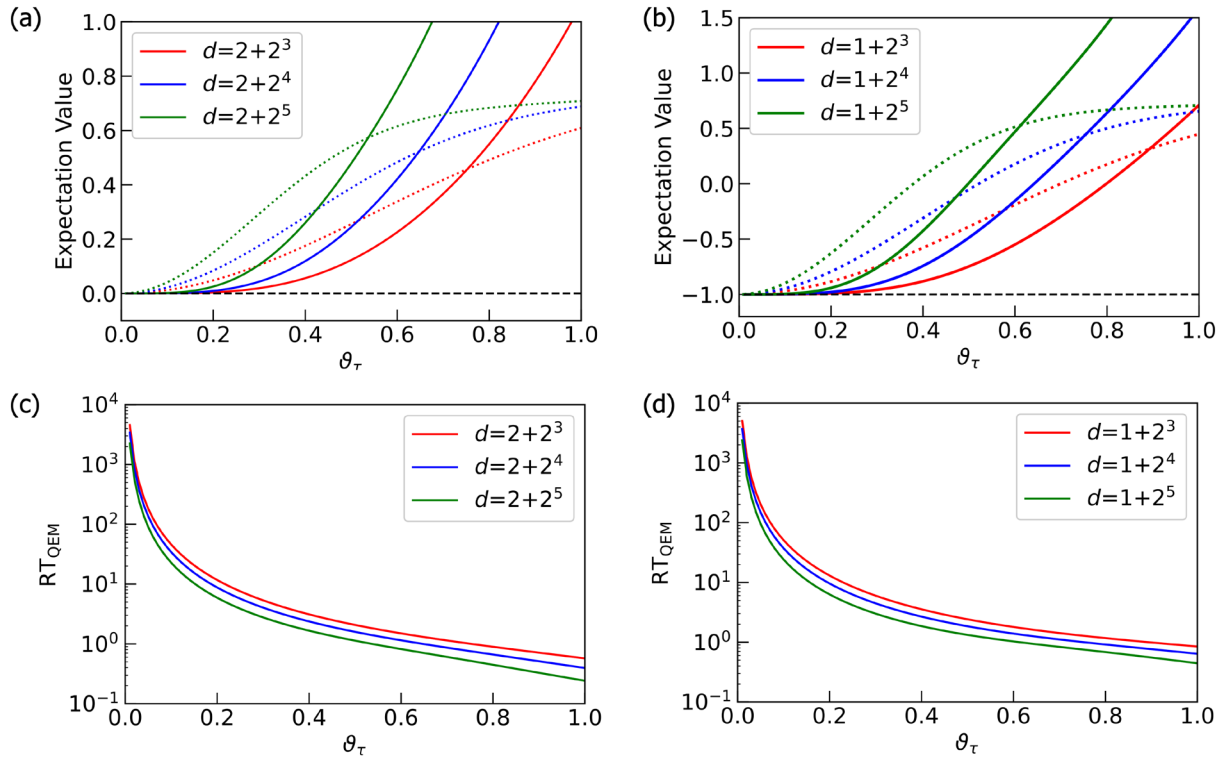

Figure 6. Quantum simulations for the quantum algorithm $U_{\text{pre1}}^{\text{QC}} = H^{\otimes d-1} \cdot X$. Plots in (**a**) and (**b**) are the results of QEM for the expectation value of X and Z, respectively. The dotted lines are the expectation values without QEM while the solid lines represent the expectation values with QEM. The black dotted lines are the results of the ideal simulations. In (**c**) and (**d**), we show the ratio RT_{QEM} for the expectation values of X and Z respectively. The red, blue, and green plots are the results for $d = 1 + 2^3, d = 1 + 2^4$, and $d = 1 + 2^5$, respectively. Note that in (**c**) we have $d = 2 + 2^n$ with $n = 3, 4, 5$ since the Hadamard gate is applied to Q_{r0} at the end.

Preliminary

As a preliminary, let us conduct the noisy quantum simulations for two simple algorithms. The first algorithm is given by the unitary operation $U_{\text{pre1}}^{\text{QC}} = H^{\otimes d-1} \cdot X$, where H denotes the Hadamard gate. The second one is given by $U_{\text{pre2}}^{\text{QC}} = (U_{CH}[Q_{r0}; Q_{r1}])^{\otimes d-1}) \cdot X_{Q_{r0}} \otimes X_{Q_{r1}}$. The unitary operator $U_{CH}[Q_{r0}; Q_{r1}]$ is the controlled-Hadamard gate composed of the control bit Q_{r0} and the target bit Q_{r1}. Here we take the circuit depth d to be $d = 1 + 2^n$ with n being positive integers. Let us show the quantum circuits for these two algorithms in Fig. 5. By changing the values of d and τ, we numerically analyze how well our QEM scheme works in terms of these two parameters. Let us discuss from the results of noisy quantum simulations conducted by the quantum circuit in Fig. 5a and show them in Fig. 6. We have taken the physical operators $\hat{O}$ as $\hat{O} = X, Z$. The horizontal axis represents ϑ_τ, which can be regarded as the strength of AD effect. On the other hand, the vertical axis in Fig. 6a,b denote the expectation values of $\hat{O}$ while those in Fig. 6c,d describe the ratio RT_{QEM} given in Eq. (13): Fig. 6a,c are the results for $\hat{O} = X$ while Fig. 6b,d are those for $\hat{O} = Z$. The dotted lines in Fig. 6a,b describe the

expectation values without QEM being performed whereas the solid lines represent the expectation values with QEM being performed. For both the dotted and solid lines the red, blue, and green plots in Fig. 6a,b are the computational results of $\langle X \rangle$ and $\langle Z \rangle$ for $d = 2^3 + 1$, $2^4 + 1$, and $2^5 + 1$, respectively. The black dotted lines are the results of the ideal simulations. We have computed $\langle \hat{X} \rangle_{\rho_{d\cdots1}}$, $\langle \hat{X} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, and $\langle \hat{X} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ not as the expectation values of X with respect to the quantum states generated by $U_{\text{pre1}}^{\text{QC}}$ but as those of Z with respect to the quantum states generated by $U_{\text{pre1}}^{\text{QC}}$ and the subsequent operation of H, i.e., we have switched the basis vectors for the measurement from the computational basis to $\{|+\rangle, |-\rangle\}$, where $|+\rangle = H|0\rangle$, $|-\rangle = H|1\rangle$. Correspondingly, we have $d = 2 + 2^n$: see Fig. 6c. Let us analyze our simulation results by comparing the behaviors of the expectation values and the ratio RT_{QEM} in Eq. (13) as functions of ϑ_τ. In this way, we can clearly see whether our QEM scheme is working or not, and for this purpose in the following we rewrite RT_{QEM} as $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ to emphasize that they are the functions of ϑ_τ. Furthermore, we introduce the angle ϑ_τ^c such that $\text{RT}_{\text{QEM}}(\vartheta_\tau = \vartheta_\tau^c) = 1$, which indicates that the point $\vartheta_\tau = \vartheta_\tau^c$ is the critical point of our QEM to become failed. Let us look from the results shown in Fig. 6a,c by focusing on how the behaviors of $\langle \hat{O} \rangle_{\rho_{d\cdots1}}$, $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$, and RT_{QEM} change by increasing ϑ_τ. In Fig. 6a, as the definition of ϑ_τ^c we certainly see that in the range $0 < \vartheta_\tau \leq \vartheta_\tau^c$ the absolute $|\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}}|$ is bigger than $|\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}}|$, which implies that $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ is numerically closer to $\langle \hat{O} \rangle_{\rho_{d\cdots1}}$ than $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, and correspondingly, in Fig. 6c we see $\text{RT}_{\text{QEM}}(\vartheta_\tau) \geq 1$. As we increase the value of ϑ_τ from ϑ_τ^c, the absolute $|\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}}|$ becomes smaller than $|\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}}|$, and correspondingly, the ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ decreases monotonically from one. For the region $\vartheta_\tau \geq \vartheta_\tau^c$ to improve the quality of our QEM we need take into account higher-order AD effects and establish QEM schemes for mitigating them and we expect the value of ϑ_τ^c to become larger. Next, let us analyze how the quality of our QEM becomes when we vary the circuit depth d. We see that for every ϑ_τ both $|\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}}|$ and $|\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}}|$ become bigger and $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ decreases as we increase d. This is reasonable because when d gets larger the amount of error gets bigger. For the results in Fig. 6b,d, basically we see that both the expectation values of Z and $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ show the similar behaviors as those for $\hat{O} = X$: (I) the validity of QEM ($\text{RT}_{\text{QEM}}(\vartheta_\tau) \geq 1$) in the range $0 < \vartheta_\tau \leq \vartheta_\tau^c$ and monotonic decrease of $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ for $\vartheta_\tau > \vartheta_\tau^c$, and (II) worsening of the quality of our QEM for large d. In contrast to the above characteristics of $\langle X \rangle$ and $\langle Z \rangle$, we have numerically verified that the expectation value of Y takes zero for any ϑ_τ. This is because when the density matrix ρ is a real matrix the expectation value $\langle Y \rangle$ is zero. Since both
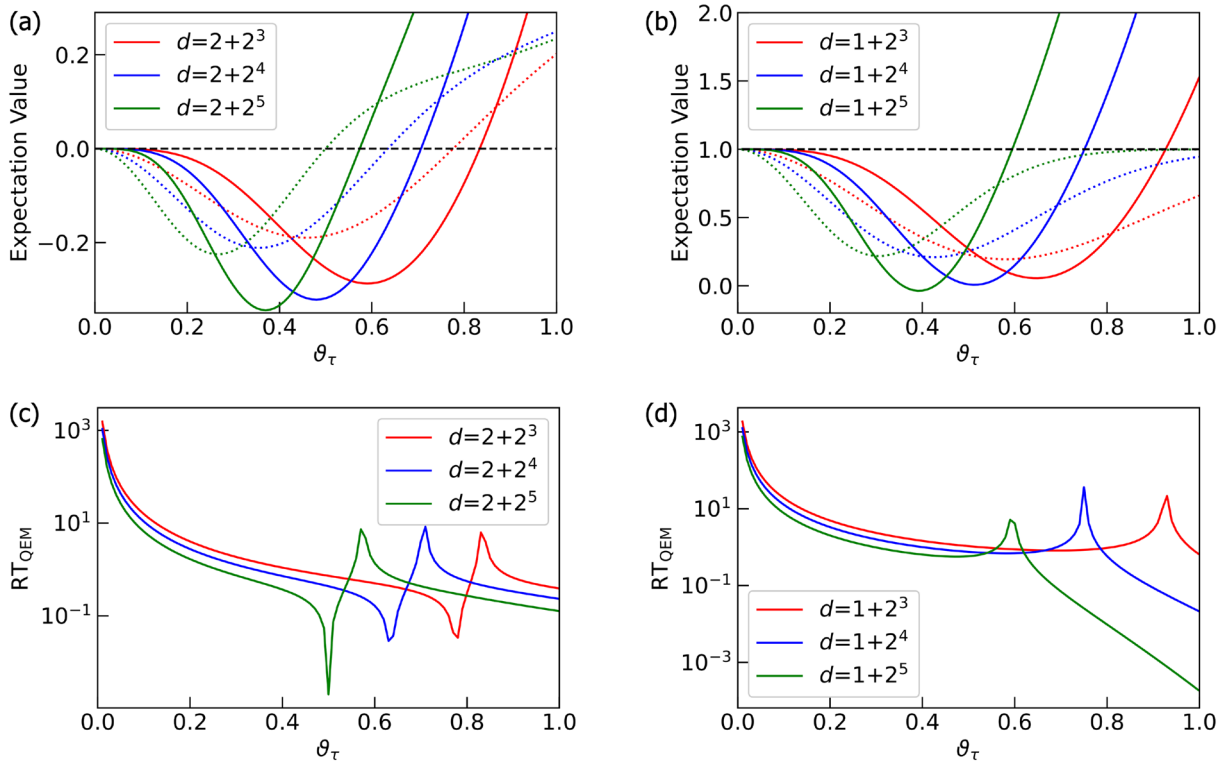

Figure 7. Quantum simulations for the quantum algorithm $U_{\text{pre2}}^{\text{QC}} = (U_{CH}[Q_{r0}; Q_{r1}])^{\otimes d-1}) \cdot X_{Q_{r0}} \otimes X_{Q_{r1}}$. Plots in (**a**) and (**b**) are the results of QEM for the expectation value of ZX and ZZ, respectively. The dotted lines are the expectation values without QEM while the solid lines represent the expectation values with QEM. The black dotted lines are the ideal simulation results. We plot the ratio RT_{QEM} for expectation value of ZX and ZZ in (**c**) and (**d**), respectively. The red, blue, and green plots are the results for $d = 1 + 2^3$, $d = 1 + 2^4$, and $d = 1 + 2^5$, respectively. Note that in (**c**) d is given as $d = 2 + 2^n$ ($n = 3, 4, 5$) since the Hadamard gate operates on Q_{r1} at the end.

the quantum algorithm and the AD effect are described by real numbers (see also Eq. (7) or the Kraus operators in Eq. (11)) the density matrix generated by these two things is real and we have $\langle Y \rangle = 0$.

Let us discuss the results in Fig. 7. They are the noisy simulation results of the quantum algorithm given by $U_{\text{pre2}}^{\text{QC}}$ (see the quantum circuit in Fig. 5b) and we have taken $d = 1 + 2^n$ as in the case of simulations for $U_{\text{pre1}}^{\text{QC}}$. Here we have simulated $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ for the expectation values of the operators $\hat{O} = ZX, ZZ$. As similar to the computations of $\langle \hat{X} \rangle_{\rho_{d\cdots1}}$, $\langle \hat{X} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, and $\langle \hat{X} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$, we have computed $\langle \hat{ZX} \rangle_{\rho_{d\cdots1}}$, $\langle \hat{ZX} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, and $\langle \hat{ZX} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ not as the expectation values of ZX with respect to the quantum states generated by $U_{\text{pre2}}^{\text{QC}}$ but as those of ZZ with respect to the quantum states generated by $U_{\text{pre2}}^{\text{QC}}$ and the subsequent operation of H on $Q_{\text{r1}} : d = 2 + 2^n$ as indicated in Fig. 7c. Overall, we see the same characteristics with the cases of $\hat{O} = X, Z$: the characteristics (I) and (II) mentioned above. For any ϑ_τ, the ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ for the noisy simulations of $U_{\text{pre2}}^{\text{QC}}$ are smaller than those of $U_{\text{pre1}}^{\text{QC}}$. This is because $U_{\text{pre1}}^{\text{QC}}$ is solely comprised of the single-qubit gates (X and H) while $U_{\text{pre2}}^{\text{QC}}$ is constructed by n-operation of the controlled-Hadamard gate (two-qubit gate), and thus, the bigger amount of errors are accumulated in the latter case. The difference between the characteristics of noisy simulations for $U_{\text{pre1}}^{\text{QC}}$ and those for $U_{\text{pre2}}^{\text{QC}}$, although it is not an essential point for the validity of our QEM, is that we see both one minima and one maxima in each plot for $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ in Fig. 7c while only one maxima appears in Fig. 7d. Let us denote the point where $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ takes the minimum (maximum) by $\vartheta_\tau^{\text{min}}$ ($\vartheta_\tau^{\text{max}}$): note that these values depend on d. We can understand why these points emerge by looking at Fig. 7a,b. Let us explain from the origins of the minima and the maxima in Fig. 7c by looking at the plots in Fig. 7a. In the range $0 < \vartheta_\tau \leq \vartheta_\tau^{\text{min}}$ we have $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} < 0$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} < 0$ whereas in the range $\vartheta_\tau^{\text{min}} < \vartheta_\tau \leq \vartheta_\tau^{\text{max}}$ we have $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} > 0$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} < 0$. Then, in the range $\vartheta_\tau^{\text{max}} < \vartheta_\tau$ we have $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} > 0$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} > 0$. As a result, the minima appears at $\vartheta_\tau = \vartheta_\tau^{\text{min}}$ whereas the maxima emerges at $\vartheta_\tau = \vartheta_\tau^{\text{max}}$ in Fig. 7c. The origin of the maxima in Fig. 7d can be similarly explained by looking at the plots in Fig. 7b. In the range $0 < \vartheta_\tau \leq \vartheta_\tau^{\text{max}}$ we have $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} < 0$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} < 0$, while in $\vartheta_\tau > \vartheta_\tau^{\text{max}}$ we have $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} < 0$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}} > 0$, and consequently, the maxima appears at $\vartheta_\tau = \vartheta_\tau^{\text{max}}$. Like the case of the noisy simulations of $U_{\text{pre1}}^{\text{QC}}$, the density matrices are generated as real matrices (the unitary transformation $U_{\text{pre2}}^{\text{QC}}$ as well as the AD effects are described by real numbers), and the expectation values of the Pauli operators, IY, XY, YI, YX, YZ, ZY are zero for both ideal and noisy simulations. Here we have rewritten $\mathbf{1}_{2\times2}$ as I for convenience. Note that the expectation value of the identity operator ($= \mathbf{1}_{4\times4}$: four by four identity operator) is one for any quantum state including noise-affected quantum states since the trace of density matrix is one for any quantum state. In other words, it is unnecessary to do QEM for the expectation value of the identity operator. Note, however, that when leakage occurs the trace preservation is not held anymore and we need to consider QEM for the error induced by the leakage.

Quantum amplitude amplification

By taking account of the previous analysis, let us apply our QEM scheme to Quantum Amplitude Amplification (QAA)[33] for three-qubit systems: two-register bits and one oracle bit. One of the important application of QAA is the database retrieval and the quantum algorithms for this is called the Grover's search algorithm[33,91–94]. Let us denote the (classical) oracle function by f and a binary by x which takes "00", "01", "10", "11". We consider that we have only one solution of f and write it by x^*, which satisfy $f(x^*) = 1$, and assume $x^* = $ "11": for $x = $ "00", "01", "10" we have $f(x) = 0$. The oracle operator O_{QAA} can be implemented on a quantum circuit by using one oracle bit Q_{o} such that $O_{\text{QAA}} \left[|x\rangle_{Q_{\text{r0}}Q_{\text{r1}}} \otimes \left(\frac{|0\rangle_{Q_{\text{o}}} - |1\rangle_{Q_{\text{o}}}}{\sqrt{2}} \right) \right] = (-1)^{f(x)} \left[|x\rangle_{Q_{\text{r0}}Q_{\text{r1}}} \otimes \left(\frac{|0\rangle_{Q_{\text{o}}} - |1\rangle_{Q_{\text{o}}}}{\sqrt{2}} \right) \right]$, where the superposition state $\frac{|0\rangle_{Q_{\text{o}}} - |1\rangle_{Q_{\text{o}}}}{\sqrt{2}}$ is created by applying $H \cdot X$ on the oracle-bit state $|0\rangle_{Q_{\text{o}}}$. In our case, $(-1)^{f(x)} = -1$ when $|x\rangle_{Q_{\text{r0}}Q_{\text{r1}}} = |11\rangle_{Q_{\text{r0}}Q_{\text{r1}}}$, and the oracle operator O_{QAA} is equivalent to the Toffoli gate comprised of the two controlled bits Q_{r0} and Q_{r1} and the target bit Q_{o}[33], and write it by $U_{CX}[Q_{\text{r0}}Q_{\text{r1}}; Q_{\text{o}}]$. To construct QAA we need one more unitary transformation and that is $U_\psi = (\mathbf{1}_{4\times4} - 2|\psi\rangle\langle\psi|) \otimes \mathbf{1}_{2\times2}$, where $|\psi\rangle = H^{\otimes2}|00\rangle_{Q_{\text{r0}}Q_{\text{r1}}}$. By introducing $U_{\text{init}} = H^{\otimes2} \otimes (H \cdot X)_{Q_{\text{o}}}$, QAA is given by the unitary operation[33]
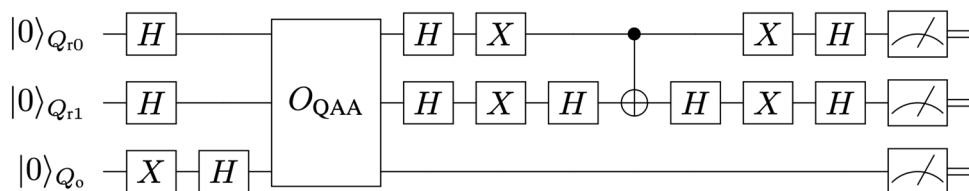

Figure 8. Quantum circuits for U^{QAA} in Eq. (14).
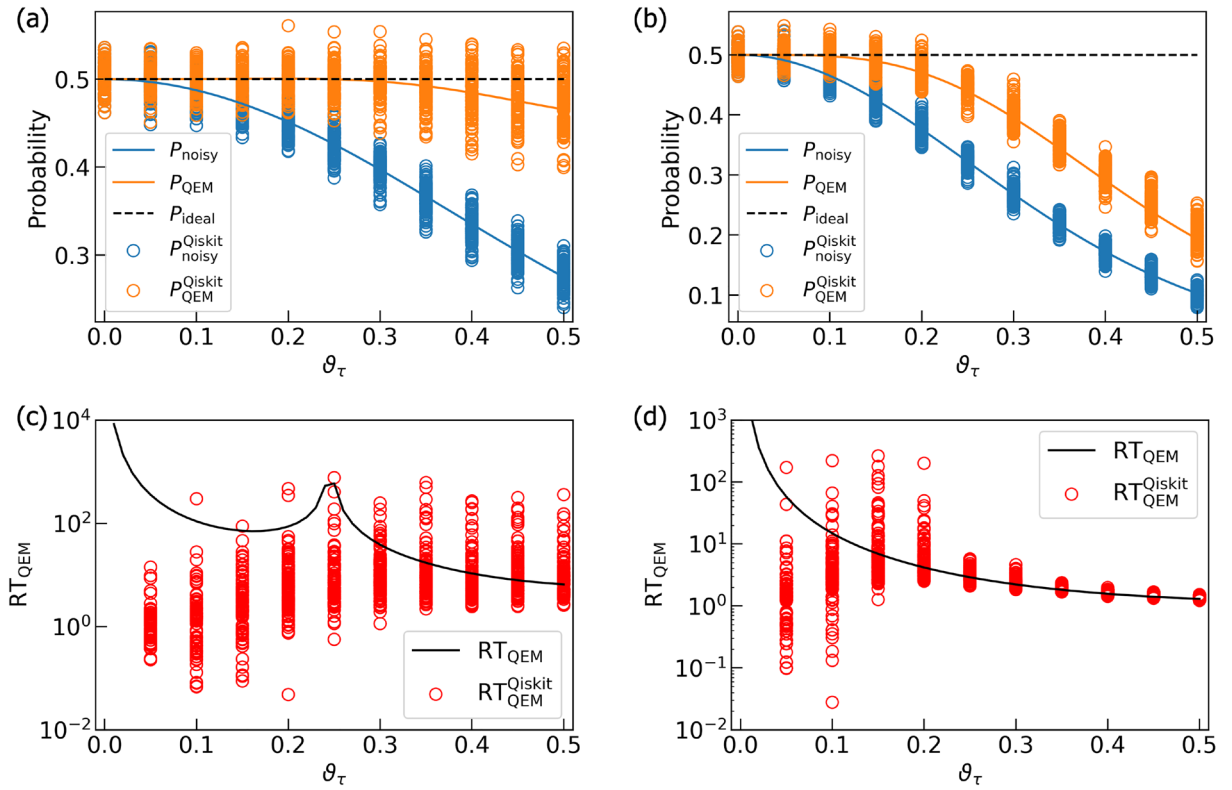
Figure 9. Quantum simulation results for the quantum algorithm U^{QAA} in Eq. (14). Plots in (**a**) and (**b**) are the results of the probabilities P_{110} (probability of $|110\rangle$) and P_{111} (probability of $|111\rangle$), respectively. The dotted black lines are the ideal simulation results whereas the blue and orange curves are the noisy simulation results without QEM and the ones with QEM, respectively. All of them are obtained by our original code. The blue and orange circles are the noisy simulation results without QEM and the ones with QEM, respectively, and they are obtained by our Qiskit code. Plots in (**c**) and (**d**) are the results of the ratio RT_{QEM} for P_{110} and P_{111}, respectively. The black curves are obtained by our original code while the red circles by our Qiskit code. For each ϑ_τ we have plotted 100 circles in (**a**)–(**d**), i.e., $N_{\text{samp}} = 100$..

$$
\begin{aligned}
U^{\text{QAA}} &= (U_{\text{G}})^k \cdot U_{\text{init}}, \\
U^{\text{G}} &= -U_\psi \cdot O_{\text{QAA}}, \\
U_\psi &= \left((H \cdot X)^{\otimes 2} \otimes \mathbf{1}_{2\times 2, Q_o}\right) \cdot \left(\mathbf{1}_{2\times 2, Q_{r0}} \otimes H_{Q_{r1}} \otimes \mathbf{1}_{2\times 2, Q_o}\right) \cdot \left(U_{CX}[Q_{r0}; Q_{r1}] \otimes \mathbf{1}_{2\times 2, Q_o}\right) \\
&\quad \cdot \left(\mathbf{1}_{2\times 2, Q_{r0}} \otimes H_{Q_{r1}} \otimes \mathbf{1}_{2\times 2, Q_o}\right) \cdot \left((X \cdot H)^{\otimes 2} \otimes \mathbf{1}_{2\times 2, Q_o}\right)
\end{aligned}
\tag{14}
$$

We show the quantum circuit for the unitary operation U^{QAA} in Fig. 8[33]: the quantum circuit for the oracle operator O_{QAA} ($= U_{CX}[Q_{r0}Q_{r1}; Q_o]$) is shown in Sect. II in the Supplementary Information. We note that on the quantum circuit in Fig. 8, what is actually implemented is $-U^{\text{G}} = U_\psi \cdot O_{\text{QAA}}$ and the global phase factor (-1) does not affect our result. The unitary operation U_{G} is called the Grover operator and k is the repetitive number number of its operation and here we have $k = 1$. After the operation of U^{QAA} in Eq. (14), ideally both of the probability weights of $|110\rangle$ and $|111\rangle$ are $\frac{1}{2}$, and thus, the probability of obtaining the quantum state $|11\rangle$ as the output state is $\frac{1}{2} \times 2 = 1$, which implies the success of searching the solution $x^* = $ "11". By taking account of the above theoretical framework, we examine whether our QEM scheme works or not for QAA given by U^{QAA} in Eq. (14) by computing the probability weights of $|110\rangle$ and $|111\rangle$ which we name as P_{110} and P_{111}, respectively, and show these results in Fig. 9. Solid lines in Fig. 9a,b describe the probability weights obtained by our original numerical code and we have denoted $\langle P_{110(111)}\rangle_{\rho_{d\cdots 1}^{\text{ideal}}}$, $\langle P_{110(111)}\rangle_{\rho_{d\cdots 1}^{\text{real}}}$, and $\langle P_{110(111)}\rangle_{\rho_{d\cdots 1}^{\text{QEM}}}$ by P_{ideal}, P_{real}, and P_{QEM}, respectively. On the other hand, the blue and orange circles are calculated by our Qiskit code and we have denoted $\langle P_{110(111)}\rangle_{\rho_{d\cdots 1}^{\text{real}}}$ and $\langle P_{110(111)}\rangle_{\rho_{d\cdots 1}^{\text{QEM}}}$ by $P_{\text{noisy}}^{\text{Qiskit}}$ and $P_{\text{QEM}}^{\text{Qiskit}}$, respectively. Similarly, in Fig. 9c,d, we have denoted the ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ calculated by our Qiskit code by $\text{RT}_{\text{QEM}}^{\text{Qiskit}}$: for the results obtained by our original code we have just used the notation RT_{QEM} for describing them. Let us look from the simulation results of P_{110} and the associated ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ given by Fig. 9a,c, respectively. In the range $0 \leq \vartheta_\tau \leq 0.5$, overall the simulation results with QEM are numerically closer to the ideal values than the noisy simulation results without QEM, and correspondingly, we have $\text{RT}_{\text{QEM}}(\vartheta_\tau) > 1$. The similar characteristics can be seen in Fig. 9b (simulation results of P_{111}) and 9d ($\text{RT}_{\text{QEM}}(\vartheta_\tau)$ for P_{111}). For the results obtained by our Qiskit code, the deviation between RT_{QEM}
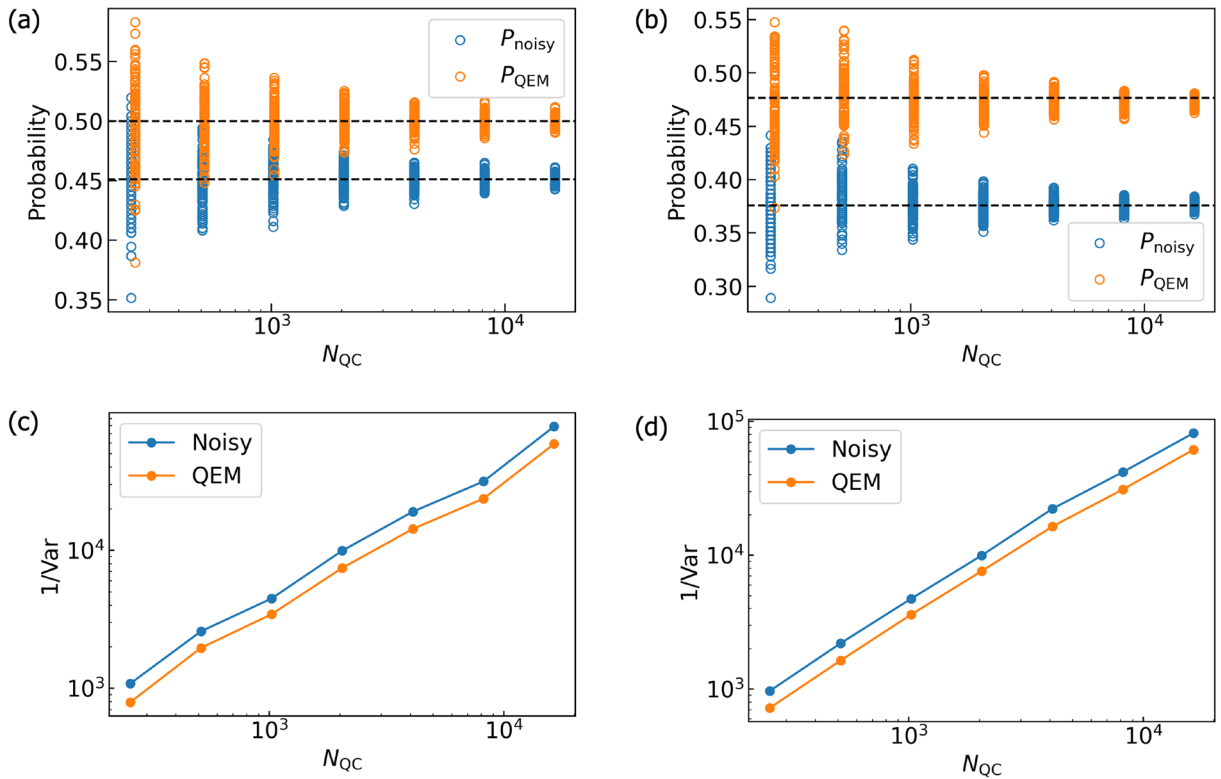
Figure 10. Numerical results of $\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ and $\langle O \rangle_{\mathrm{QEM}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ for (**a**) P_{110} and (**b**) P_{111}. The plots in (**c**) and (**d**) are $\left(\sigma_{\mathrm{noisy,Qiskit}}^2[O, N_{\mathrm{QC}}, N_{\mathrm{samp}}]\right)^{-1}$ (blue) and $\left[\sigma_{\mathrm{QEM,Qiskit}}^2[O, N_{\mathrm{QC}}, N_{\mathrm{samp}}, N_q, d]\right]^{-1}$ (orange) for P_{110} and P_{111}, respectively. We take $N_{\mathrm{QC}} = 2^{n_{\mathrm{QC}}}$ with $n_{\mathrm{QC}} = 8, 9, \ldots, 14$, $N_{\mathrm{samp}} = 100$, and $\vartheta_\tau = 0.2$.

and $\mathrm{RT}_{\mathrm{QEM}}^{\mathrm{Qiskit}}$ becomes prominent in the small ϑ_τ region and we consider this as follows. When noise strength ϑ_τ is weak enough, on the Qiskit code the difference between the noisy value and the ideal value is very tiny such that our QEM becomes invalid and $\mathrm{RT}_{\mathrm{QEM}}^{\mathrm{Qiskit}}$ gets lower than one. On the other hand, we see that some red points are above $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_\tau) = 1$. We consider that by greatly increasing N_{QC}, we expect that $\mathrm{RT}_{\mathrm{QEM}}^{\mathrm{Qiskit}}$ approaches to $\mathrm{RT}_{\mathrm{QEM}}$. As a result, our QEM works for the noisy simulations of both P_{110} and P_{111}. To show clearly that the simulation results obtained by our Qiskit code approach to those obtained by our original code by increasing N_{QC}, in Fig. 10a,b we show the N_{QC} dependencies of P_{110} and P_{111}, respectively. In these plots the horizontal axes represent N_{QC} whereas the vertical axis in Fig. 10a represents the numerical values of P_{110} and that in Fig. 10b represents those of P_{111}. In the following let us write a noisy expectation value obtained in the ith round of quantum computing by $\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ ($O = P_{110}, P_{111}$ and $i = 1, \ldots, N_{\mathrm{samp}}$) and similarly an expectation value with QEM by $\langle O \rangle_{\mathrm{QEM}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ with taking $N_{\mathrm{QC}} = 2^{n_{\mathrm{QC}}}$ with $n_{\mathrm{QC}} = 8, 9, \ldots, 14$, $N_{\mathrm{samp}} = 100$, and $\vartheta_\tau = 0.2$. From these two figures, we clearly see that both $\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ and $\langle O \rangle_{\mathrm{QEM}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ approach to $\langle O \rangle_{\mathrm{noisy}}$ and $\langle O \rangle_{\mathrm{QEM}}$, respectively, i.e., the deviations of $\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ from $\langle O \rangle_{\mathrm{noisy}}$ and those of $\langle O \rangle_{\mathrm{QEM}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ from $\langle O \rangle_{\mathrm{QEM}}$ get smaller for larger $N_{\mathrm{QC}} = 2^{n_{\mathrm{QC}}}$. To evaluate these deviations numerically, in Fig. 10c ($O = P_{110}$) and 10d ($O = P_{111}$) we plot inverse variances defined by $\left(\sigma_{noisy,\mathrm{Qiskit}}^2[O, N_{\mathrm{QC}}, N_{\mathrm{samp}}]\right)^{-1} = \left[\sum_{i=1}^{N_{\mathrm{samp}}} \frac{1}{N_{\mathrm{samp}}} \left(\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}}) - \overline{\langle O \rangle}_{\mathrm{noisy}}^{\mathrm{Qiskit}}(N_{\mathrm{QC}})\right)^2\right]^{-1}$, where $\langle \bar{O} \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(N_{\mathrm{QC}}, N_{\mathrm{samp}}) = \sum_{i=1}^{N_{\mathrm{samp}}} \frac{\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})}{N_{\mathrm{samp}}}$. We approximate $\left(\sigma_{noisy,\mathrm{Qiskit}}^2[O, N_{\mathrm{QC}}, N_{\mathrm{samp}}]\right)^{-1}$ as a linear function of N_{QC} as $\left(\sigma_{noisy,\mathrm{Qiskit}}^2[O, N_{\mathrm{QC}}, N_{\mathrm{samp}}]\right)^{-1} = \alpha_{\mathrm{noisy}}^{\mathrm{Qiskit}}[O, N_{\mathrm{samp}}] N_{\mathrm{QC}}$ and we have $\alpha_{\mathrm{noisy}}^{\mathrm{Qiskit}}[O, N_{\mathrm{samp}}] = 4.62$ and 5.05 for P_{110} and P_{111}, respectively. On the other hand, the variances of $\langle O \rangle_{\mathrm{noisy}}^{\mathrm{Qiskit}}(i, N_{\mathrm{QC}})$ can be analytically evaluated (see also the description in pages 7 and 8) as $\left(\sigma_{\mathrm{noisy}}^2[O, N_{\mathrm{QC}}]\right)^{-1} = \left(\frac{\langle O \rangle_{\mathrm{noisy}}(1 - \langle O \rangle_{\mathrm{noisy}})}{N_{\mathrm{QC}}}\right)^{-1} = \alpha_{\mathrm{noisy}}[O] N_{\mathrm{QC}}$, and we have $\alpha_{\mathrm{noisy}}[O] \simeq 4.04$ for P_{110} and $\alpha_{\mathrm{noisy}}[O] \simeq 4.26$ for P_{111}. Therefore, the two variances $\sigma_{\mathrm{noisy}}^2[O, N_{\mathrm{QC}}]$ and $\sigma_{\mathrm{noisy,Qiskit}}^2[O, N_{\mathrm{QC}} N_{\mathrm{samp}}]$ have good numerical agreements. In addition to these two variances, we also plot the quantity defined by
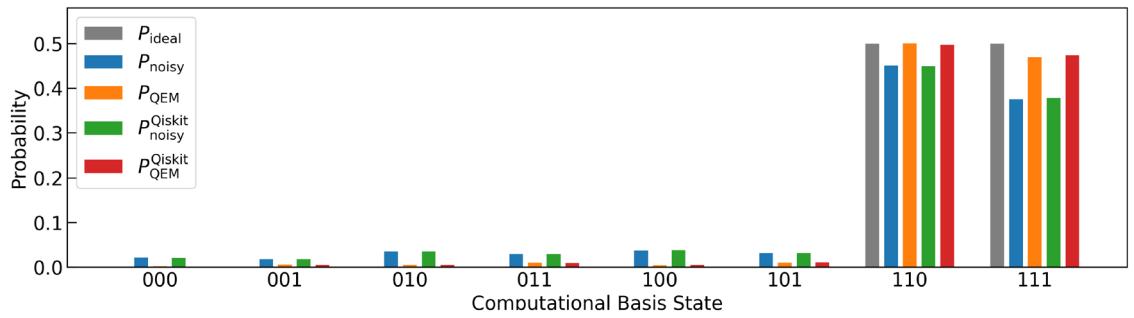
Figure 11. Histogram of the probability distribution of the computational basis states for QAA simulations given by U^{QAA} in Eq. (14). Here we have set $\vartheta_\tau = 0.2$.
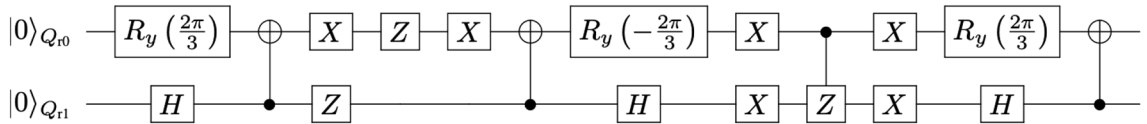

Figure 12. Quantum circuit for QAA given by V^{QAA} with $k = 1$.

$$\left[\sigma^2_{\text{QEM,Qiskit}}[O, N_{\text{QC}}, N_{\text{samp}}, N_q, d]\right]^{-1} = \left[\sum_{i=1}^{N_{\text{samp}}} \frac{1}{N_{\text{samp}}} \left(\langle O \rangle^{\text{Qiskit}}_{\text{QEM}}(i, N_{\text{QC}}) - \langle \bar{O} \rangle^{\text{Qiskit}}_{\text{QEM}}(N_{\text{QC}}, N_{\text{samp}})\right)^2\right]$$

$$= \alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d)N_{\text{QC}}, \text{ where } \langle \bar{O} \rangle^{\text{Qiskit}}_{\text{QEM}}(N_{\text{QC}}, N_{\text{samp}}) = \sum_{i=1}^{N_{\text{samp}}} \frac{\langle O \rangle^{\text{Qiskit}}_{\text{QEM}}(i, N_{\text{QC}})}{N_{\text{samp}}}.$$ Such a quantity describes

the deviations of $\langle O \rangle^{\text{Qiskit}}_{\text{QEM}}(i, N_{\text{QC}})$ from $\langle O \rangle_{\text{QEM}}$ owing to the finite effect of N_{QC} and plays the role of variance. Like $\left(\sigma^2_{\text{noisy}}[O, N_{\text{QC}}]\right)^{-1}$ and $\left[\sigma^2_{\text{noisy,Qiskit}}[O, N_{\text{QC}}, N_{\text{samp}}]]\right]^{-1}$, we take it as the linear function of N_{QC} given by the coefficient $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d)$. Note that the dependency of the coefficient $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d)$ in terms of N_q and d originates in the size of the quantum-noise-effect circuit group $3N_qd$. Owing to our simulation, we obtain $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d) = 3.45$ for $O = P_{110}$ and $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d) = 3.77$ for $O = P_{111}$. The coefficients $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d)$ are smaller than $\alpha^{\text{Qiskit}}_{\text{noisy}}[O, N_{\text{samp}}]$, as expected, since the deviations get larger owing to the usage of the quantum-noise-effect circuit group (additional computational resource for QEM). The ratio $\frac{\alpha^{\text{Qiskit}}_{\text{noisy}}[O, N_{\text{samp}}]}{\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d)}$, however, is about 1.34 for both cases which implies that the broadening of the deviations is not so big. We leave the detailed mathematical analysis on $\left(\sigma^2_{\text{QEM}}[O, N_{\text{QC}}]\right)^{-1}$ and $\left[\sigma^2_{\text{QEM,Qiskit}}[O, N_{\text{QC}}, N_{\text{samp}}, N_q, d]\right]^{-1}$ as well as the coefficient $\alpha_{\text{QEM}}(O, N_{\text{samp}}, N_q, d)$ as our future work. In addition to $N_{\text{samp}} = 100$, we also perform the simulations for $N_{\text{samp}} = 1000$, and as a result, we obtain $\alpha^{\text{Qiskit}}_{\text{noisy}}[O, N_{\text{samp}}] = 4.25$ and $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d) = 3.18$ for P_{110} and $\alpha^{\text{Qiskit}}_{\text{noisy}}[O, N_{\text{samp}}] = 4.47$ and $\alpha^{\text{Qiskit}}_{\text{QEM}}(O, N_{\text{samp}}, N_q, d) = 3.34$ for P_{111}. As a result, by increasing N_{QC} both $\langle O \rangle^{\text{Qiskit}}_{\text{noisy}}(i, N_{\text{QC}})$ and $\langle O \rangle^{\text{Qiskit}}_{\text{QEM}}(i, N_{\text{QC}})$ approach to $\langle O \rangle_{\text{noisy}}$ and $\langle O \rangle_{\text{QEM}}$, respectively, owing to the law of large numbers. Let us also show the simulation results of the rest of the six probabilities of the computational basis states for $\vartheta_\tau = 0.2$ as the histogram in Fig. 11, which also includes P_{110} and P_{111}. The ideal values of these six probabilities are all zero and we see that the noisy simulation results with QEM are numerically close to them compared with those without QEM, which indicates that our QEM scheme also works for the other six probabilities.

In addition to the above simulation, let us present the simplified version of QAA for the two-qubit systems[95]. In this problem setting, we consider the effective two-dimensional space spanned by the two Bell states $|\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$ and $|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$, and write their superposition state by $|\Sigma\rangle = c_{\Psi^+}|\Psi^+\rangle + c_{\Phi^+}|\Phi^+\rangle$, where the complex coefficients c_{Ψ^+}, c_{Φ^+} satisfy $|c_{\Psi^+}|^2 + |c_{\Phi^+}|^2 = 1$. Our goal is to amplify the probability amplitude c_{Φ^+}. Here we take the initialization operator to be $V_{\text{init}} = U_{CX}[Q_{r1}; Q_{r0}] \cdot (R_y\left(\frac{2\pi}{3}\right)_{Q_{r0}} \otimes H_{Q_{r1}})$ while we take the Grover operator to be $V_G = V_s V_\omega$, where $V_\omega = (X \cdot Z \cdot X)_{Q_{r0}} \otimes Z_{Q_{r1}}$ and $V_s = V_{\text{init}} \cdot (2|0\rangle^{\otimes 2}\langle 0| - \mathbf{1}_{4 \times 4}) \cdot V^\dagger_{\text{init}} \equiv -\tilde{V}_s$ with $\tilde{V}_s = V_{\text{init}} \cdot (X \otimes X \cdot U_{CZ}[Q_{r0}; Q_{r1}] \cdot X \otimes X) \cdot V^\dagger_{\text{init}}$. k is the number of V_G to be applied. In total, the unitary operation for running QAA is given by $V^{\text{QAA}} = (V_G)^k \cdot V_{\text{init}}$. We present the corresponding quantum circuit in Fig. 12. As similar to the above case, on the quantum circuit we implement $\tilde{V}_s$ instead of V_s. In the following, let us take a look at the meanings of the three unitary operations $V_{\text{init}}, V_s,$ and V_ω. First, the unitary operation V_{init} generates the superposition of $|\Phi^+\rangle$ and $|\Psi^+\rangle$ as $|00\rangle \to |s\rangle = V_{\text{init}}|00\rangle = \cos\frac{\vartheta_V}{2}|\Psi^+\rangle + \sin\frac{\vartheta_V}{2}|\Phi^+\rangle$ with $\vartheta_V = \pi/3$. Second, the unitary operation V_ω is the oracle operator and when it is applied to the initial state $|s\rangle$ we have $V_\omega|s\rangle = \cos\frac{\vartheta_V}{2}|\Psi^+\rangle - \sin\frac{\vartheta_V}{2}|\Phi^+\rangle$, i.e., the
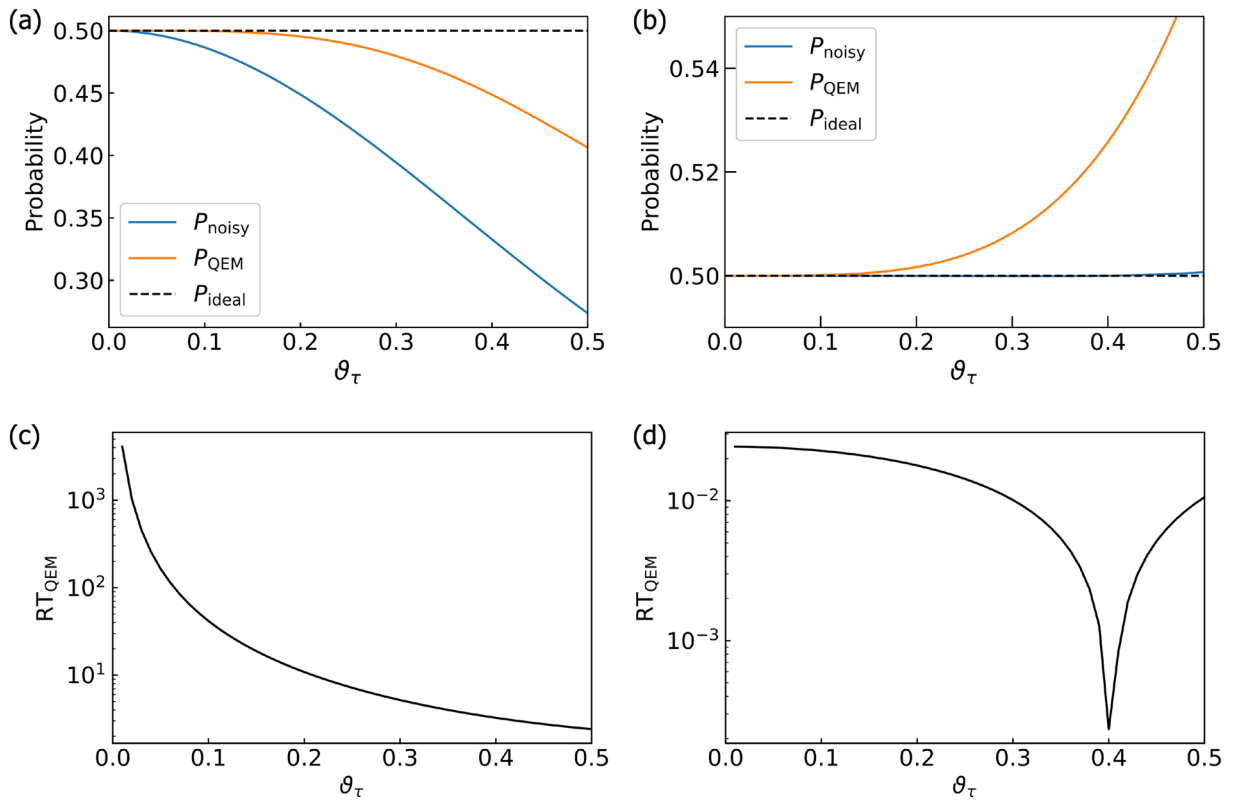
Figure 13. Noisy quantum simulations for QAA given by V^{QAA} for $\hat{O} = P_{00}, P_{11}$. All these results are obtained by our numerical code. (**a**) Plots of the results of $P_{\text{ideal}} = \langle P_{11} \rangle_{\rho_{d\cdots1}}$, $P_{\text{noisy}} = \langle P_{11} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, and $P_{\text{QEM}} = \langle P_{11} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ presented by the black dashed line, the blue solid line, and the orange solid line, respectively. (**b**) Plots of the results of $P_{\text{ideal}} = \langle P_{00} \rangle_{\rho_{d\cdots1}}$, $P_{\text{noisy}} = \langle P_{00} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, and $P_{\text{QEM}} = \langle P_{00} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ presented by the black dashed line, the blue solid line, and the orange solid line, respectively. (**c**) The ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ for P_{11}. (**d**) The ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$ for P_{00}.
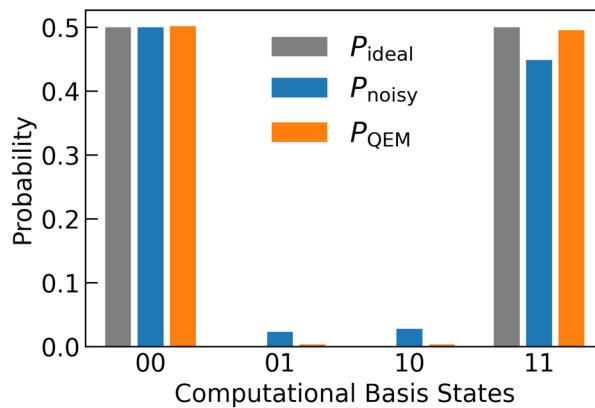

Figure 14. Histogram of the probability distribution of the computational basis states for two-qubit-system QAA simulations. We have set $\vartheta_\tau = 0.2$.

oracle operator V_ω is the operator such that it reverses the sign of the Bell state $|\Phi^+\rangle$. Third, V_s is the reflection of the vector $V_\omega|s\rangle$ with respect to the vector $|s\rangle$. When we operate V_G on $|s\rangle$ for k times we have $(V_G)^k|s\rangle = \cos\frac{(2k+1)\vartheta_V}{2}|\Psi^+\rangle + \sin\frac{(2k+1)\vartheta_V}{2}|\Phi^+\rangle$, and since $\vartheta_V = \pi/3$ we have $k = 1$. We can understand the geometrical meaning of the operation V_G as follows. Let us consider the effective three-dimensional space spanned by the two Bell states $|\Phi^+\rangle$ and $|\Psi^+\rangle$ and the vector $|\xi\rangle$ which is perpendicular to both $|\Phi^+\rangle$ and $|\Psi^+\rangle$. Furthermore, we call the axis which is parallel to $|\xi\rangle$ (perpendicular to the two-dimensional plane spanned by $|\Phi^+\rangle$ and $|\Psi^+\rangle$) as ξ-axis. The unitary operation V_G is the rotation about ξ-axis by the angle ϑ_V in this effective three-dimensional space. We can verify whether $|\Phi^+\rangle$ has been generated as the output state or not by measuring

the probabilities of $|00\rangle$ state and $|11\rangle$ state, which are denoted by p_{00} and p_{11}, respectively. In other words, the probabilities p_{00} and p_{11} are the expectation values of the projection operators P_{00} and P_{11}, respectively, where P_{00} (P_{11}) is the projection operator of $|00\rangle$ ($|11\rangle$) state. Namely, we run noisy quantum simulation for $\hat{O} = P_{00}, P_{11}$ and perform QEM on them: Note that in the case of the ideal simulation we obtain $p_{00} = p_{11} = \frac{1}{2}$. We plot the numerical results of p_{11} and p_{00} for the range $0 \leq \vartheta_{\tau} \leq 0.5$ in Fig. 13a,b, respectively, and in Fig. 13c,d we plot $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ for p_{11} and p_{00}, respectively. All these results shown here are obtained by our original numerical code and we have denoted $\langle P_{00(11)}\rangle_{\rho_{d\cdots1}^{\mathrm{ideal}}}$, $\langle P_{00(11)}\rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$, and $\langle P_{00(11)}\rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}}$ by P_{ideal}, P_{real}, and P_{QEM}, respectively. Let us look from the results of the probability p_{11}. In Fig. 13a we see that for any ϑ_{τ} the absolute of the deviation $\left|\langle P_{11}\rangle_{\rho_{d\cdots1}^{\mathrm{real}}} - \langle P_{11}\rangle_{\rho_{d\cdots1}}\right|$ is bigger than $\left|\langle P_{11}\rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}} - \langle P_{11}\rangle_{\rho_{d\cdots1}}\right|$, and correspondingly, as we see in Fig. 13c the ratio $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ is greater than one. Therefore, our QEM scheme works well for the noisy simulation of p_{11}. In contrast, in Fig. 13b,d we see that the probability p_{00} shows essentially a different behavior. That is the expectation value without QEM $\langle P_{00}\rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$ is numerically close to the ideal value $\langle P_{00}\rangle_{\rho_{d\cdots1}}$ compared with the QEM-performed expectation value $\langle P_{00}\rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}}$, and correspondingly, the ratio $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ is lower than one. Such a characteristic is understood as follows. Firstly, we have analytically examined that the expectation value $\langle P_{00}\rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$ does not include first-order term in τ, i.e., $\langle P_{00}\rangle_{\Delta_{1}^{\mathrm{AD}}\rho_{d\cdots1}} = 0$. The lowest-order term included in the numerator of $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ is $\mathcal{O}(\tau^2)$. Secondly, due to our QEM the lowest order of the denominator of $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ is also $\mathcal{O}(\tau^2)$. As a result, the ratio $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ becomes lower than one, which indicates that it is not appropriate to adopt our QEM scheme. We consider that this is because our QEM scheme described by Eq. (8) is the scheme for mitigating the first-order AD effect. In the limit of $\vartheta_{\tau} \to 0$, the ratio $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ takes finite value, and analytically it is the ratio between the absolute of the coefficient of $\langle P_{00}\rangle_{\delta_2^{\mathrm{AD}}\rho_{d\cdots1}}$ and that of $\langle P_{00}\rangle_{\delta_1(\Delta_1^{\mathrm{AD}}\rho_{d\cdots1})}$. We note that we have performed two types of simulations with our original code. In the first one we have directly implemented $U_{CZ}[Q_{\mathrm{r1}}; Q_{\mathrm{r0}}]$ gate while in the second one we have implemented it via the decomposition $\left(\mathbf{1}_{2\times2\,Q_{\mathrm{r0}}} \otimes H_{Q_{\mathrm{r1}}}\right) \cdot U_{CX}[Q_{\mathrm{r1}}; Q_{\mathrm{r0}}] \cdot \left(\mathbf{1}_{2\times2\,Q_{\mathrm{r0}}} \otimes H_{Q_{\mathrm{r1}}}\right)$. According to the results of these two simulations, we have analyzed that on the Qiskit code $U_{CZ}[Q_{\mathrm{r1}}; Q_{\mathrm{r0}}]$ gate is automatically implemented by the above decomposition since the result of the second simulation with our original code shows better matching with that obtained by our Qiskit code. In such a case, the first-order term in τ appears for $\langle P_{00}\rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$ and our QEM works well. Besides p_{00} and p_{11}, let us briefly discuss the noisy simulation results of the probability weights of $|01\rangle$ and $|10\rangle$ and write them by p_{01} and p_{10}, respectively. We show them in the histogram in Fig. 14 which describes the probability distribution of the computational basis states of the two register qubits Q_{r0} and Q_{r1}. Here we have taken $\vartheta_{\tau} = 0.2$. We

Let us end this subsection by giving the following comment. In the previous subsection, we have seen that our QEM scheme becomes meaningless in the cases when the expectation values of the ideal simulations are equivalent to those of noisy simulations such as the simulation for the expectation value $\langle Y\rangle$. Besides these cases, our QEM scheme represented by the formula in Eq. (8) does not work when noisy expectation values do not include the first-order term in τ like the noisy simulation for the probability p_{00} discussed above. In other words, if we construct the QEM formula which describes the mitigation for a higher-order quantum noise effect, which is discussed in Sect. IB in the Supplementary Information, by using it we become able to accomplish the noisy quantum simulation obtaining $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau}) > 1$. In practice, however, when we run quantum algorithms on real quantum devices we cannot compute $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_{\tau})$ since we cannot compute ideal expectation values. We can check whether noisy expectation values include the first-order terms in τ or not, for example, in the following way. We perform two types of QEM, QEM of both the first- and second order quantum noise effects and the one of only the second-order effect. Let us denote the density matrices obtained by the former QEM and the latter one by $\rho_{2\mathrm{nd}}^{\mathrm{QEM}}$ and $\rho_{2\mathrm{nd,only}}^{\mathrm{QEM}}$, respectively. Next, we introduce the measure $M_{\mathrm{1st/2nd}} = \left|\mathrm{Tr}\left(\left(\rho_{2\mathrm{nd}}^{\mathrm{QEM}} - \rho_{2\mathrm{nd,only}}^{\mathrm{QEM}}\right)O\right) \times \tau^{-1}\right|$, where O is a physical operator. If the first-order QEM fails (noisy
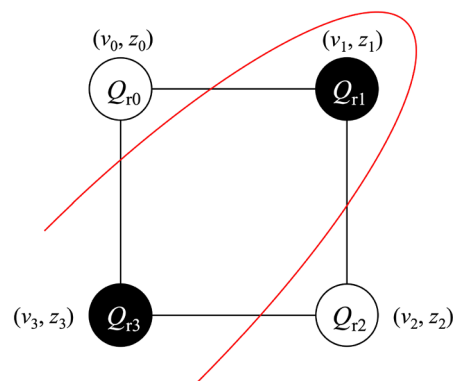

Figure 15. Structure of the graph $G = (V, E)$. It is the square composed of the four vertices $v_0, v_1, v_2,$ and v_3 and the four edges $\langle v_0 v_1\rangle$, $\langle v_1 v_2\rangle$, $\langle v_2 v_3\rangle$, and $\langle v_3 v_0\rangle$. For each vertex v_i ($i = 0, 1, 2, 3$) the binary value $z_i = \pm1$ is assigned and the set (v_i, z_i) is encoded in the qubit Q_{ri} in the QAOA simulation.

expectation values do not include the first-order terms in τ) then the measure $M_{1\text{st/2nd}}$ is $\mathcal{O}(\tau^2)$. On the other hand, if the first-order QEM succeeds (noisy expectation values include the first-order terms in τ and is mitigated) then $M_{1\text{st/2nd}}$ is $\mathcal{O}(1)$. By extending this approach we can examine whether noisy expectation values include higher-order quantum noise effects or not. We consider, however, that the failure of the first-order QEM for the noisy quantum computation when the linear order in τ does not appear is not so crucial compared with the case when we have failed in mitigating the linear-order quantum noise effects included in noisy expectation values, and indeed we can see this by looking at Fig. 14. For the probability weight of $|00>$ state, the numerical differences among the three expectation values, $\langle \hat{O} \rangle_{\rho_{d\cdots1}}$, $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}}$, $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$, are small. On the other hand, for the probability weight of $|11>$ state, $\langle \hat{O} \rangle_{\rho_{d\cdots1}}$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ is close enough while $\langle \hat{O} \rangle_{\rho_{d\cdots1}}$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\text{real}}}$ are quite separated.

QAOA

As a final example, let us apply our QEM scheme to the noisy simulation of the variational quantum algorithm called Quantum Approximate Optimization Algorithm (QAOA)[27–33]. In the following, we analyze QAOA for a max-cut problem which is to divide vertices (nodes) of a given graph into two groups so that the number of edges connecting two vertices belonging to the different groups is maximized and is a NP (Non-deterministic Polynomial time)-hard problem.

First, we discuss from a theoretical framework of a classical approximate optimization. We express the given graph G as $G = (V, E)$, where $V = \{v_0, \ldots, v_i, \ldots, v_{N_V-1}\}$ is the set of vertices with N_V denoting their total number and for each vertex v_i the binary value $z_i = \pm1$ is assigned. $E = \left\{ \{\langle v_0 v_1 \rangle, C_{0,1}\}, \ldots, \{\langle v_i v_{i+1} \rangle, C_{i,i+1}\}, \ldots, \{\langle v_{N_V-2} v_{N_V-1} \rangle, C_{N_V-2,N_V-1}\} \right\}$ is the set of the edges with $\langle v_i v_{i+1} \rangle$ denoting the edge connected by the vertices v_i and v_{i+1}. The quantity C_{ij} $(i, j = 0, \ldots, N_V - 1$ with $i \neq j)$ is the adjacency matrix element (weight) for the edge $\langle v_i v_j \rangle$ which is semi-positive. Let us write the N_V strings of z_i by $\boldsymbol{z} = (z_0, \ldots, z_{N_V-1})$. The goal of a classical approximate optimization is to minimize the cost function

$$C(\boldsymbol{z}) = \frac{1}{2} \sum_{(v_i, v_j)} C_{ij}(z_i z_j - 1), \tag{15}$$

or equivalently to maximize the ratio r_{CAO} (≤ 1) which satisfies

$$\frac{C(\boldsymbol{z})}{C_{\min}} \geq r_{\text{CAO}}, \tag{16}$$

where $C_{\min}$ is the minimum value of $C(\boldsymbol{z})$. In our simulation, as illustrated in Fig. 15 we adopt the square graph given by the four vertices v_0, v_1, v_2, and v_3, and the edges are $\langle v_0 v_1 \rangle$, $\langle v_1 v_2 \rangle$, $\langle v_2 v_3 \rangle$, and $\langle v_3 v_0 \rangle$, and take $C_{ij} = 1$ for any edge $\langle v_i v_j \rangle$. Next, we discuss the theoretical framework for QAOA. The four vertices v_0, v_1, v_2, and v_3 are encoded in four qubits Q_{r0}, Q_{r1}, Q_{r2}, and Q_{r3}, respectively, and the values $z_i z_j$ in the expectation values of the operators $Z_{Q_i} \otimes Z_{Q_j}$. The cost function $C(\boldsymbol{z})$ in Eq. (15) is given by the expectation of the Hamiltonian

$$H_{\text{C}} = \frac{1}{2} \sum_{(i,j)} \left(Z_{Q_{ri}} \otimes Z_{Q_{rj}} - 1 \right), \tag{17}$$

where the symbol (i, j) $(i, j = 0, 1, 2, 3)$ denotes the summation for the edges connected by the qubits Q_{ri} and Q_{rj} under the square-graph structure in Fig. 15. In this simulation the physical operator $\hat{O}$ is the Hamiltonian H_{C} in Eq. (17). The unitary operation for running QAOA, which we denote by U^{QAOA}, consists of three elements. The first one is the unitary operation for creating the reference state and is given by the Hadamard-gate operation on all four qubits, $U_{\text{int}} = H^{\otimes 4}$. The other two are the unitary operations $U_C(\vartheta_j^{\text{QAOA}})$ and $U_X(\varphi_j^{\text{QAOA}})$ which are generated by the Hamiltonian H_{C} in Eq. (17) with the angle $\vartheta_j^{\text{QAOA}}$ and the term $H_X = \sum_i X_{Q_{ri}}$ with the angle φ_j^{QAOA}, respectively: in QAOA H_X is called the transverse-field (mixing or driving) term. The two types of angles $\vartheta_j^{\text{QAOA}}$ and φ_j^{QAOA} $(j = 1, \ldots, p)$ are the variational parameters and p is the repetition number of applying the
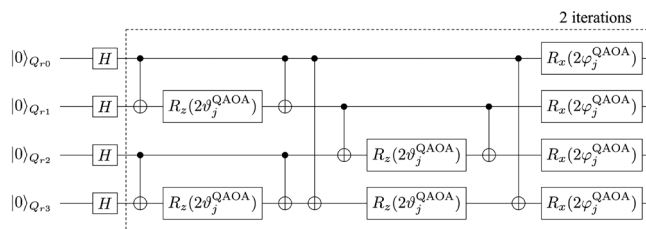

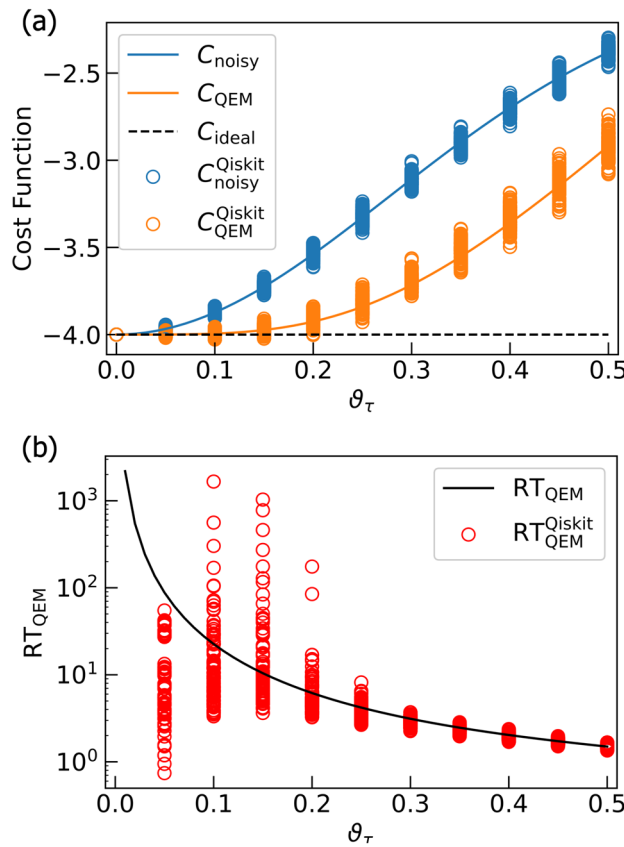**Figure 16.** Quantum circuit for QAOA.

Figure 17. Quantum simulations for QAOA. In (**a**) we show the results of QEM for the cost function $C\left(\boldsymbol{\vartheta}^{\mathrm{QAOA}}, \boldsymbol{\varphi}^{\mathrm{QAOA}}\right)$. The blue and orange solid lines are $C_{\mathrm{noisy}} = \langle H_C \rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$ and $C_{\mathrm{QEM}} = \langle H_C \rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}}$, respectively, and they obtained by our original code. The blue and orange circles are the results of $\langle H_C \rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$ and $\langle H_C \rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}}$ obtained by our Qiskit code, respectively. For this simulation, we have described $\langle H_C \rangle_{\rho_{d\cdots1}^{\mathrm{real}}}$ and $\langle H_C \rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}}$ as $C_{\mathrm{noisy}}^{\mathrm{Qiskit}}$ and $C_{\mathrm{QEM}}^{\mathrm{Qiskit}}$, respectively. The dotted black line is the ideal expectation value $C_{\mathrm{ideal}} = \langle H_C \rangle_{\rho_{d\cdots1}}$. In (**b**) we have plotted the results of the ratio $\mathrm{RT}_{\mathrm{QEM}}(\vartheta_\tau)$. The black curve is the result obtained by our original code whereas the red circles are the one obtained by our Qiskit code. For each ϑ_τ, we have plotted 100 circles.

unitary operation $U_X(\varphi_j^{\mathrm{QAOA}}) \cdot U_C(\vartheta_j^{\mathrm{QAOA}})$ (the number of iteration), which determines the accuracy of QAOA. In total, U^{QAOA} is given by

$$
U^{\mathrm{QAOA}} = \left[\prod_{j=1}^{p} \left(U_X(\varphi_j^{\mathrm{QAOA}}) \cdot U_C(\vartheta_j^{\mathrm{QAOA}}) \right) \right] \cdot \left[\bigotimes_{a=0}^{3} H_{Q_{ra}} \right],
$$

$$
U_C(\vartheta_j^{\mathrm{QAOA}}) = e^{-i\vartheta_j^{\mathrm{QAOA}} H_C}, \quad U_X(\varphi_j^{\mathrm{QAOA}}) = e^{-i\varphi_j^{\mathrm{QAOA}} H_X}. \tag{18}
$$

The quantum circuit for U^{QAOA} in Eq. (18) is presented in Fig. 16. In our simulation we set $p = 2$ and the circuit depth is $d = 15$. The unitary operation $U_X(\varphi_j^{\mathrm{QAOA}})$ is implemented by the R_x gate (rotation about x axis) with the angle $2\varphi_j^{\mathrm{QAOA}}$. Meanwhile, the quantum circuit for $U_C(\vartheta_j^{\mathrm{QAOA}})$ is composed of the sets of the quantum gates $\left[U_{CX}[Q_{ri}; Q_{ri}], R_z(2\vartheta_j^{\mathrm{QAOA}}) \right]$, where R_z denotes the rotational gate about z axis and the associated angle is $2\vartheta_j^{\mathrm{QAOA}}$. Corresponding to Eq. (16), the goal of QAOA simulation is to compute and minimize the expectation value

$$
C\left(\boldsymbol{\vartheta}^{\mathrm{QAOA}}, \boldsymbol{\varphi}^{\mathrm{QAOA}}\right) = \langle \boldsymbol{\vartheta}^{\mathrm{QAOA}}, \boldsymbol{\varphi}^{\mathrm{QAOA}} | H_C | \boldsymbol{\vartheta}^{\mathrm{QAOA}}, \boldsymbol{\varphi}^{\mathrm{QAOA}} \rangle, \tag{19}
$$

where $|\boldsymbol{\vartheta}^{\mathrm{QAOA}}, \boldsymbol{\varphi}^{\mathrm{QAOA}}\rangle = U^{\mathrm{QAOA}}|0\rangle^{\otimes 4}$ with $\boldsymbol{\vartheta}^{\mathrm{QAOA}} = (\vartheta_1^{\mathrm{QAOA}}, \vartheta_2^{\mathrm{QAOA}})$ and $\boldsymbol{\varphi}^{\mathrm{QAOA}} = (\varphi_1^{\mathrm{QAOA}}, \varphi_2^{\mathrm{QAOA}})$. Let us also call the expectation value $C\left(\boldsymbol{\vartheta}^{\mathrm{QAOA}}, \boldsymbol{\varphi}^{\mathrm{QAOA}}\right)$ in Eq. (19) as the cost function and its minimization is equivalent to the optimization of the variational parameters $\vartheta_j^{\mathrm{QAOA}}, \varphi_j^{\mathrm{QAOA}}$, and write them by
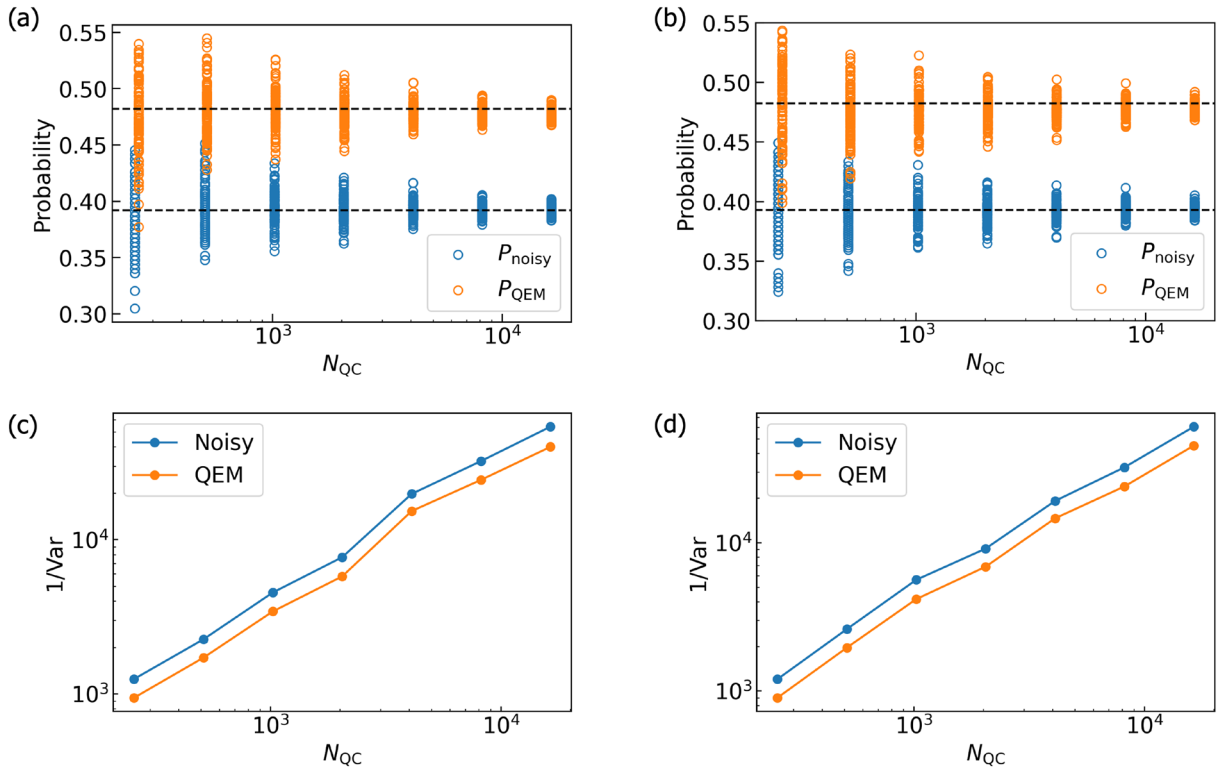
Figure 18. Numerical results of $\langle O\rangle_{\text{noisy}}^{\text{Qiskit}}(i, N_{\text{QC}})$ and $\langle O\rangle_{\text{QEM}}^{\text{Qiskit}}(i, N_{\text{QC}})$ for (**a**) P_{0101} and (**b**) P_{1010}. We plot $\left(\sigma_{\text{noisy,Qiskit}}^2[O, N_{\text{QC}}, N_{\text{samp}}]\right)^{-1}$ (blue) and $\left[\sigma_{\text{QEM,Qiskit}}^2[O, N_{\text{QC}}, N_{\text{samp}}, N_q, d]\right]^{-1}$ (orange) for P_{0101} and P_{1010} in (**c**) and (**d**), respectively. We take $N_{\text{QC}} = 2^{n_{\text{QC}}}$ with $n_{\text{QC}} = 8, 9, \ldots, 14$, $N_{\text{samp}} = 100$, and $\vartheta_\tau = 0.2$.
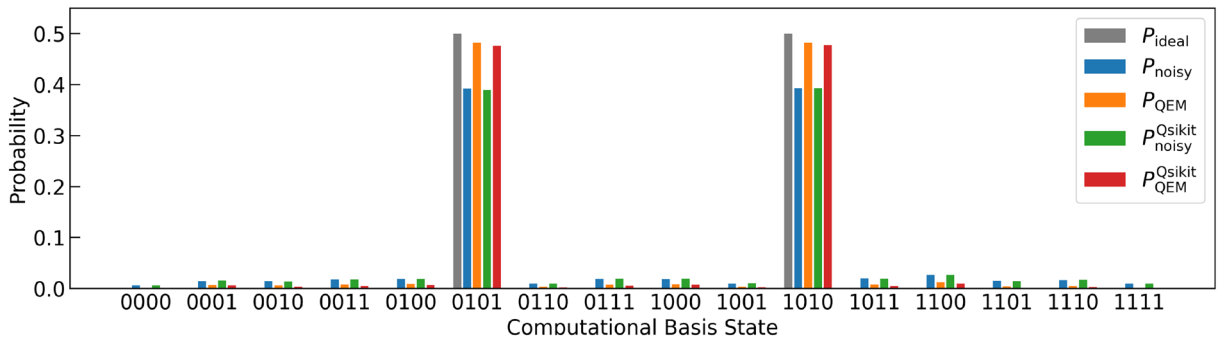
Figure 19. Histogram of the probability distribution of the computational basis states for QAOA simulations. We have set $\vartheta_\tau = 0.2$. P_{ideal}, P_{noisy}, and P_{QEM} are obtained by our original code while $P_{\text{noisy}}^{\text{Qiskit}}$ and $P_{\text{QEM}}^{\text{Qiskit}}$ are obtained by the Qiskit code.

$\vartheta_j^{\text{QAOA,opt}}$, $\varphi_j^{\text{QAOA,opt}}$. We show their values in Sect. III in the Supplementary Information. Let us now discuss our simulation results shown in Figs. 17, 18 and 19. All these results are obtained by computing the probability distribution of the computational basis states since the Hamiltonian H_C in Eq. (17) is given by the Z gate operations. First, let us take a look at the simulation results in Fig. 17. In Fig. 17a, we have plotted the results of the cost function $C(\vartheta^{\text{QAOA}}, \varphi^{\text{QAOA}})$ for the ideal simulation ($C_{\text{ideal}} = \langle H_C\rangle_{\rho_{d\cdots1}}$), the noisy simulation ($C_{\text{noisy}} = \langle H_C\rangle_{\rho_{d\cdots1}^{\text{real}}}$), and the simulation with QEM ($C_{\text{QEM}} = \langle H_C\rangle_{\rho_{d\cdots1}^{\text{QEM}}}$). The dashed black line, the blue solid line, and the orange solid line are C_{ideal}, C_{noisy}, and C_{QEM}, respectively, and they are all obtained by our original code. On the other hand, the blue and orange circles are $\langle H_C\rangle_{\rho_{d\cdots1}^{\text{real}}}$ and $\langle H_C\rangle_{\rho_{d\cdots1}^{\text{QEM}}}$, respectively, and they have been calculated by our Qiskit code. We have denoted $\langle H_C\rangle_{\rho_{d\cdots1}^{\text{real}}}$ and $\langle H_C\rangle_{\rho_{d\cdots1}^{\text{QEM}}}$ by $C_{\text{noisy}}^{\text{Qiskit}}$ and $C_{\text{QEM}}^{\text{Qiskit}}$, respectively. We have plotted 100 circles for each ϑ_τ. We see that in the range $0.1 \leq \vartheta_\tau \leq 0.5$ our QEM works well. In Fig. 17b, we have plotted the ratio $\text{RT}_{\text{QEM}}(\vartheta_\tau)$. The black curve is the result obtained by our original code while the red circles are those obtained by our Qiskit code and 100 circles are plotted for each ϑ_τ. Corresponding to the result

shown in Fig. 17a, the ratio satisfies $RT_{QEM}(\vartheta_\tau) > 1$. Like the results in Fig. 10, we present the N_{QC} dependencies of the probabilities P_{0101} and P_{1010} in Fig. 18a,b, respectively. We take $N_{QC} = 2^{n_{QC}}$ with $n_{QC} = 8, 9, \ldots, 14$, $N_{samp} = 100$, and $\vartheta_\tau = 0.2$. We see that as we increase N_{QC} both $\langle P_{0101} \rangle_{noisy}^{Qiskit}(i, N_{QC})$ and $\langle P_{1010} \rangle_{noisy}^{Qiskit}(i, N_{QC})$ approach to $\langle P_{0101} \rangle_{noisy}$ and $\langle P_{1010} \rangle_{noisy}$, respectively, and similarly, $\langle P_{0101} \rangle_{QEM}^{Qiskit}(i, N_{QC})$ and $\langle P_{1010} \rangle_{QEM}^{Qiskit}(i, N_{QC})$ approach to $\langle P_{0101} \rangle_{QEM}$ and $\langle P_{1010} \rangle_{QEM}$, respectively. In Fig. 18c,d we plot $\left(\sigma_{noisy,Qiskit}^2[O, N_{QC}, N_{samp}] \right)^{-1}$, and $\left[\sigma_{QEM,Qiskit}^2[O, N_{QC}, N_{samp}, N_q, d] \right]^{-1}$ for P_{0101} and P_{1010}, respectively. For P_{0101} we obtain $\alpha_{noisy}[O] \simeq 4.20, \alpha_{noisy}^{Qiskit}[O, N_{samp}] \simeq 3.50, \alpha_{QEM}^{Qiskit}[O, N_{samp}, N_q, d] \simeq 2.61$ and for P_{1010} $\alpha_{noisy}[O] \simeq 4.19, \alpha_{noisy}^{Qiskit}[O, N_{samp}] \simeq 3.80, \alpha_{QEM}^{Qiskit}[O, N_{samp}, N_q, d] \simeq 2.84$. For both P_{0101} and P_{1010} the ratio $\frac{\alpha_{noisy}^{Qiskit}[O, N_{samp}]}{\alpha_{QEM}^{Qiskit}(O, N_{samp}, N_q, d)}$ is 1.34, which indicates that the broadening of the deviations of $\langle O \rangle_{QEM}^{Qiskit}(i, N_{QC})$ from $\langle O \rangle_{QEM}$ due to our QEM method is not so large (not so high computational cost). We have also performed our simulations for $N_{samp} = 1000$ and we obtain $\alpha_{noisy}^{Qiskit}[O, N_{samp}] \simeq 4.04, \alpha_{QEM}^{Qiskit}[O, N_{samp}, N_q, d] \simeq 3.02$ for P_{0101} and $\alpha_{noisy}^{Qiskit}[O, N_{samp}] \simeq 4.33, \alpha_{QEM}^{Qiskit}[O, N_{samp}, N_q, d] \simeq 3.27$ for P_{1010}. Therefore, compared with the results for $N_{samp} = 100$ the coefficients $\alpha_{noisy}^{Qiskit}[O, N_{samp}]$ get bigger and become closer to $\alpha_{noisy}[O]$ and $\alpha_{QEM}^{Qiskit}[O, N_{samp}, N_q, d]$. Finally, let us explain the results in Fig. 19. Here we have presented the histogram of the probability distribution of the computational basis states for the ideal case (P_{ideal}), the noisy case ($P_{noisy}, P_{noisy}^{Qiskit}$), and the case with QEM being performed ($P_{QEM}, P_{QEM}^{Qiskit}$), with setting $\vartheta_\tau = 0.2$. We see that ideally the probabilities of the two quantum states $|0101\rangle$ and $|1010\rangle$ are both equal to 0.5. This implies that under the optimized variational parameters the cost function $C(\boldsymbol{\vartheta}^{QAOA}, \boldsymbol{\varphi}^{QAOA})$ becomes minimized such that the four qubits Q_{ri} are partitioned into the two groups $[Q_{r0}, Q_{r2}]$ and $[Q_{r1}, Q_{r3}]$, and it implies that all the edges of the square are to be cut, i.e., the maximum number of edges to be cut is four. Correspondingly, as shown in Fig. 17a the ideal minimum value of the cost function is -4.0, and we obtain the maximum cut number four by multiplying minus one. Consequently, our QEM scheme works for the noisy QAOA simulation.

QEM scheme implementation

In this section, we demonstrate our QEM scheme using two IBM Q Experience processors, ibmq_belem and ibm_perth[86]. In Fig. 4 in Sect. V of the Supplementary Information, we show schematics of spatial configurations for qubits in these machines: Fig. 4a is the illustration for ibmq_belem whereas Fig. 4b is that for ibm_perth. As similar to the quantum simulations demonstrated in "Numerical simulations", we examine the efficacy of our QEM scheme for the real quantum devices by varying the depths of the quantum algorithms to be run. We do this for two quantum algorithms, U_{pre1}^{QC} in Fig. 5a and a quantum algorithm for a two-qubit system defined by $U_{imp,2}^{QC} = (U_{CZ}[Q_{ri}; Q_{rj}])^{n^{rep}} \cdot X_i X_j$, where n^{rep} is an integer. For the usage of ibmq_belem, we choose qubits Q_1 and Q_3 as register qubits whereas we use $Q_2 (Q_4)$ as an ancilla bit for mitigating the quantum noise effect on $Q_{r1} (Q_{r3})$. On the other hand, we use Q_1 and Q_3 as register qubits while we use an ancilla bit $Q_2 (Q_5)$ for QEM of the quantum noise effect on $Q_1 (Q_3)$ for the usage of ibm_perth. In the following, we rewrite Q_1 and Q_3 as Q_{r1} and Q_{r3}, respectively, to emphasize that they are used as register bits. Our QEM scheme can be applied provided that noise parameters (strengths) are given a priori. Thus, we start from the acquisition of the quantum noise strengths of these devices (noise characterization). Then, we discuss the results of our QEM method obtained with these machines (implementation).

Noise characterization

Let us first discuss from how to obtain the AD strength $\tau (= \gamma \Delta t)$ or the T_1 time. The relation between the decay rate γ and the T_1 time is $T_1 = \frac{1}{\gamma(2\bar{n}+1)}$, where $\bar{n}$ is the Bose-Einstein distribution function. For superconducting qubit systems, qubit frequencies and temperatures are about 5.0 GHz (see also Tables II, IV, and V in the Supplementary Information) and 10 mK[9,17], respectively. As a result, the Bose-Einstein distribution function $\bar{n}$ is estimated to be $\bar{n} \sim 10^{-11}$ and we approximate $\bar{n}$ to be zero. The value of T_1 can be obtained as follows. First, we prepare a single qubit and apply the X gate so that the qubit is in the $|1\rangle$ state (excited state). Next, we let the qubit relax until certain time t_a^{relax} ($a = 1, 2, \ldots, N_m$ and $0 < t_1 < t_2 < \cdots < t_{N_m}$), measure the output state of the qubit which is either $|0\rangle$ or $|1\rangle$, and repeat this process to obtain the probability weights of the $|0\rangle$ and $|1\rangle$ states. The total exposure time of the qubit is $t_a^{tot,T_1} = \Delta t^X + t_a^{relax}$, where Δt^X denotes a X-gate operation time. This is equivalent to measuring the dynamics of the expectation value $\langle P^1 \rangle(t_a^{tot,T_1})$ (the dynamics of the probability such that $|1\rangle$ is to be measured as the output state), and by doing an exponential fitting on the $\langle P^1 \rangle(t_a^{tot,T_1})$ plots we can extract the value of T_1. We can also extract a T_2 time and here let us consider a T_2^* time which is obtained by the Ramsey measurement explained in the following and hereinafter we just denote T_2^* time as T_2 time. Basically, what we do is first we apply the Hadamard gate, let the qubit relax for t_a^{relax}, apply the second Hadamard gate, measure the output state, and repeat this process: The actual gate operations implemented on these experiments are not the two Hadamard gates owing to the issue of transmon qubit systems and its details is described in Sect. IV in the Supplementary Information. The above procedure is essentially equivalent to measuring the
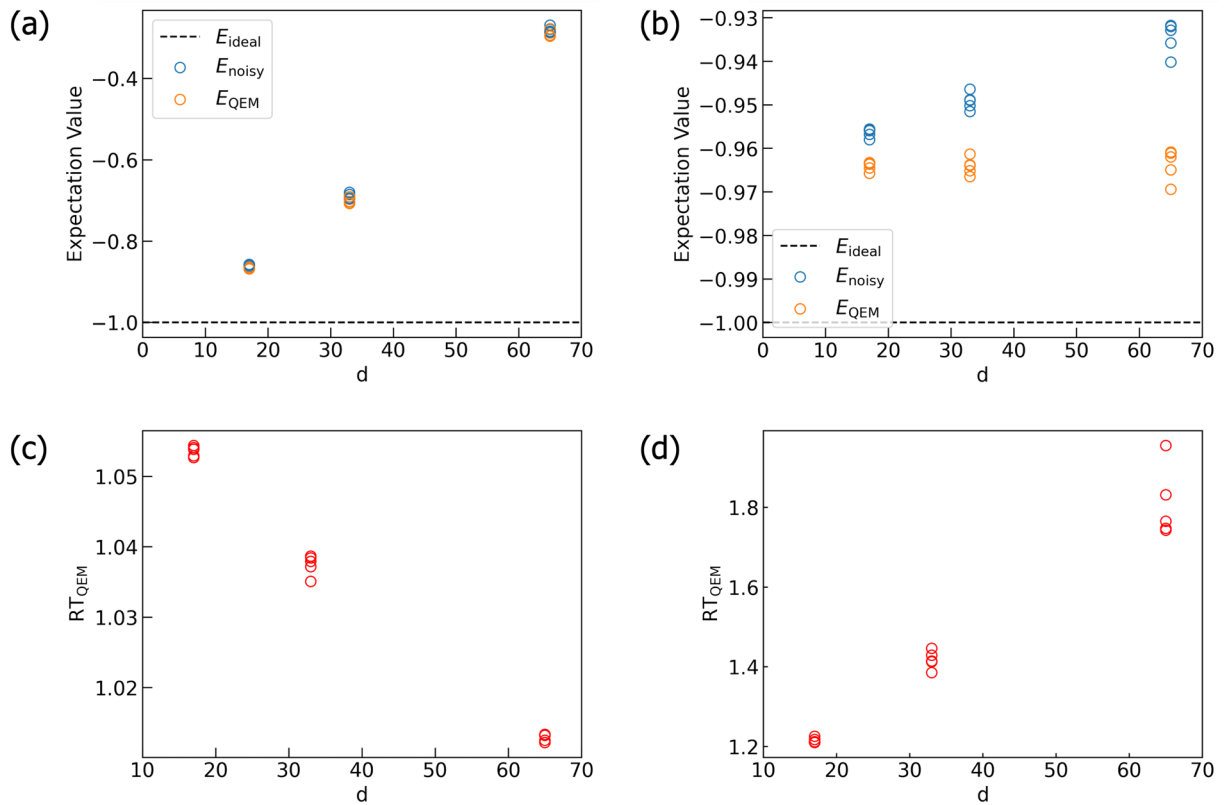
Figure 20. Plots in (**a**) and (**c**) are quantum computation results obtained by ibm_perth and (**b**) and (**d**) are quantum simulation results. The quantum algorithm which has been run is U^{QC}_{pre1}. Both (**a**) and (**b**) are the results of expectation values of Z_1 and (**c**) and (**d**) are the ones of RT_{QEM}. All the quantum circuits have been executed under $N_{QC} = 2^{14}, N_{samp} = 5$. The horizontal axes represent $d = n^{rep} + 1$, where $n^{rep} = 16, 32, 64$. The quantum computation has been conducted at 07:17, 09/26/2023.

dynamics of $\langle P^1 \rangle (t_a^{tot,T_2})$, where $t_a^{tot,T_2} = 2\Delta t^H + t_a^{relax}$ with Δt^H denoting Hadamard-gate operation time. From the $\langle P^1 \rangle (t_a^{tot,T_2})$ curve we obtain a T_2 time. In the Supplementary Information IV, we explain how to estimate the T_1 and T_2 times with presenting the experimental data plots of the expectation values: Figs. 2 and 3 show the data plots of the T_1 and T_2 times, respectively, and Table I lists the values of them for ibmq_belem.

In Sect. V in the Supplementary Information, we present the physical properties (single-qubit and two-qubit gate times, qubit frequencies) of ibmq_belem and ibm_perth: Tables II and III (Tables IV–VII) list the physical properties of ibmq_belem (ibm_perth). Here let us briefly explain the gate properties. The native gates of these machines are the following, identity gate, virtual Z gate, $\sqrt{X}$ gate, X gate, CNOT, and reset operation (reset into $|0\rangle$). The Hadamard gate is decomposed into a series of the native gates as $H = R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right)$.

By using the data in Sects. IV and V in the Supplementary Information we are now able to implement our QEM scheme. Before showing the results, let us comment five things about our experiment. Firstly, the values of T_1 and T_2 times differ from qubit to qubit on a real device due to imperfection (inhomogeneities of T_1 and T_2). To take these inhomogeneities into account and extract T_1 and T_2 times, we need to perform the above two procedures independently on each qubit. Secondly, a gate time Δt is actually gate dependent as shown in Tables II–VII in the Supplementary Information. By taking the inhomogeneities of both the T_1 and T_2 times and the gate times into account, instead of τ, which does not depend on both the qubits and the quantum gates, we perform our QEM scheme by using a perturbative parameter $\tau_{jk} = \gamma_j \Delta t_k$, where $j(= 0, \ldots, N_q - 1)$ is the index for qubit numbering whereas $k(= 1, \ldots, d)$ is that for labeling the gates. In the following, we denote the T_1 and T_2 times of the qubit Q_j by $T_{1,j}$ and $T_{2,j}$, respectively. Thirdly, T_1 and T_2 times fluctuate temporally on a real device. To run our QEM scheme, it is necessary to record the data of T_1 and T_2 times day-by-day and use these values. Thus, we indicate the time and the date in the coordinated universal time (UCT) when we list the data of T_1 and T_2 times. Fourthly, the CZ gate operation is decomposed into the form $U_{CZ}[Q_{r1}; Q_{r3}] = \left(R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right) \right)_3 \cdot U_{CX}[Q_{r1}; Q_{r3}] \cdot \left(R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right) \right)_3$, and according to the transpilation of IBM Q Experience programming $(U_{CZ}[Q_{r1}; Q_{r3}])^2$ has been processed as $(U_{CZ}[Q_{r1}; Q_{r3}])^2 = \left(R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right) \right) \cdot U_{CX}[Q_{r1}; Q_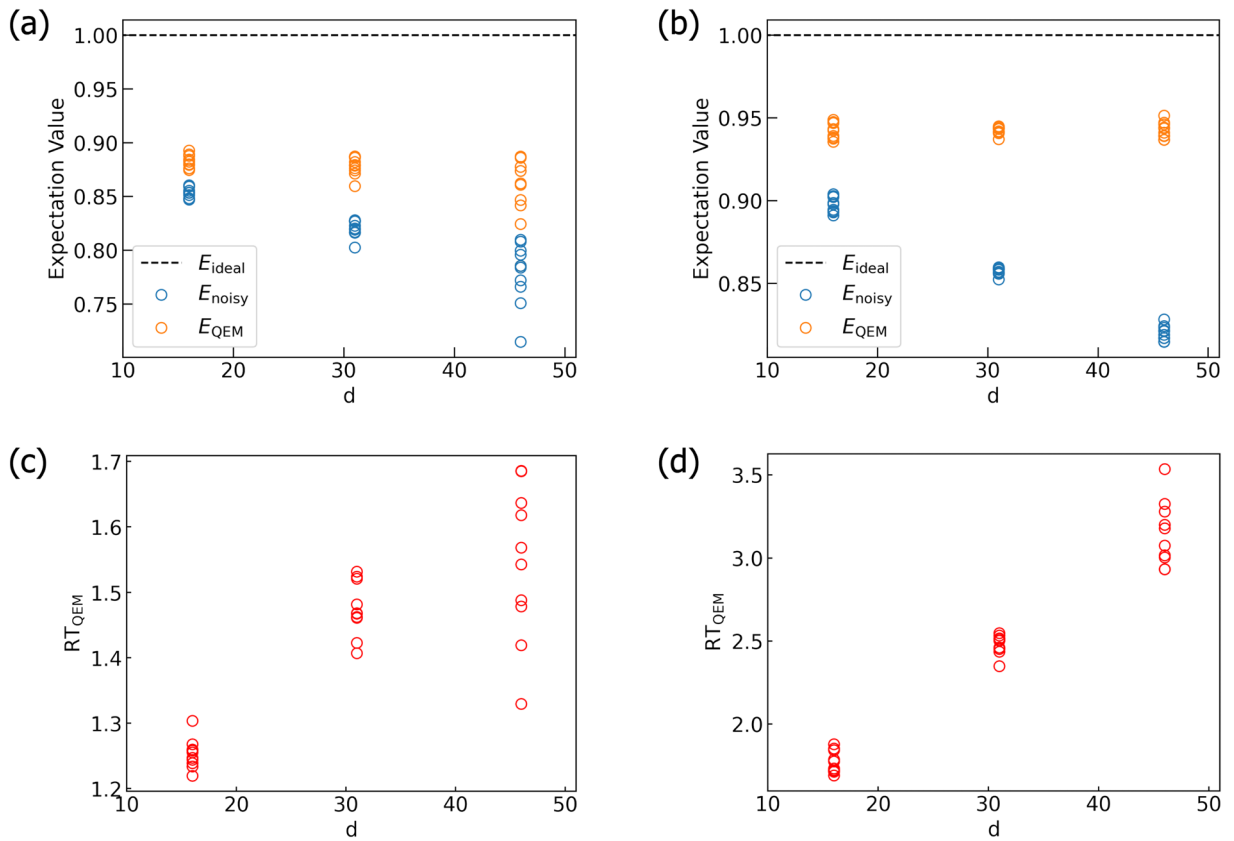{r3}] \cdot \mathbf{1} \cdot \mathbf{1} \cdot U_{CX}[Q_{r1}; Q_{r3}] \cdot \left(R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right) \right)$, and $U^{QC}_{imp,2}$ has been executed as $U^{QC}_{imp,2} = \left(R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right) \right)_3 \cdot U_{CX}[Q_{r1}; Q_{r3}] \cdot (\mathbf{1} \cdot \mathbf{1} \cdot U_{CX}[Q_{r1};$

Figure 21. Quantum computation and simulation results for $U_{\text{imp},2}^{\text{QC}}$. (**a**) ((**b**)) and (**c**) ((**d**)) are the quantum computation (simulation) results of expectation values $\langle Z_1 Z_3 \rangle$ and ratios RT_{QEM}, respectively. All the horizontal axes represent $d = 3n^{\text{rep}} + 1$ with $n^{\text{rep}} = 5, 10, 15$. The quantum computations have been conducted via ibm_perth at 06:55, 09/21/2023 under $N_{\text{QC}} = 2^{14}, N_{\text{samp}} = 10$.

$Q_{r3}])^{n^{\text{rep}}-1} \cdot \left(R_z\left(\frac{\pi}{2}\right) \cdot \sqrt{X} \cdot R_z\left(\frac{\pi}{2}\right) \right)_3 \cdot X_1 X_3$[86]. Note that **1** denotes the identity operator. Fifthly, the time when the quantum computation has been conducted (see the captions of Figs. 20, 21, and 22) indicates the time when the original circuit and the quantum-noise-effect-circuit group have been executed. Finally, the measured (experimental) value of τ, which we denote by τ_{exp}, can be differ from the true value of τ due to, for instance, imperfection of experimental apparatuses. Although that is the case, the conduction of our QEM method is said to be a success provided that the condition $\text{RT}_{\text{QEM}} > 1$ is satisfied.

Implementation

As mentioned previously, in real quantum devices both T_1 and T_2 times and gate operation times are inhomogeneous, and moreover, both AD and PD effects exist. By taking these elements into account, we perform our QEM scheme by improving the formulas given by Eqs. (3) and (8) as

$$\langle \hat{O} \rangle_{\rho_{d\cdots 1}}^{\text{QEM}} \equiv \langle \hat{O} \rangle_{\rho_{d\cdots 1}^{real}} - \sum_{j=0}^{N_q-1} \sum_{k=1}^{d} \langle \hat{O} \rangle_{\left(\Delta_{1,jk}^{\text{AD\&PD}} \rho_{d\cdots 1} \right)^{real}},$$

$$\Delta_{1,jk}^{\text{AD\&PD}} \rho_{d\cdots 1} = \left(\prod_{l=k+1}^{d} U_l \right) \cdot \left[-\frac{\Delta t_k}{2T_{2,j}} \left(\rho_{k\cdots 1} - Z_j \rho_{k\cdots 1} Z_j \right) + \frac{\Delta t_k}{T_{1,j}} \left(\tilde{\sigma}_j \rho_{k\cdots 1} \tilde{\sigma}_j^+ - P_j^1 \rho_{k\cdots 1} P_j^1 \right) \right] \cdot \left(\prod_{l=k+1}^{d} U_l \right)^{\dagger}.$$

(20)

We note that the quantum circuits such that the additional operations $\{Z, \tilde{\sigma}^-, P^1\}$ are inserted after the virtual Z gates are not needed for (or do not contribute to) the calculation of Eq. (20) because the operation time of the virtual Z gate is zero.

First, let us discuss from the results for $U_{\text{pre1}}^{\text{QC}}$ which are shown in Fig. 20. Here all the horizontal axes represent $d = n^{\text{rep}} + 1$, where the repetition number n^{rep} is taken to be 16, 32, and 64. Fig. 20a presents a quantum computation result of expectation values $\langle Z_1 \rangle$ obtained by ibm_perth whereas the plots in Fig. 20b is a quantum simulation result of $\langle Z_1 \rangle$[85]. On the other hand, Fig. 20c displays a quantum computation result of RT_{QEM} with ibm_perth and Fig. 20d plots a quantum simulation result of RT_{QEM}[85]. In Fig. 20a, we observe that both the
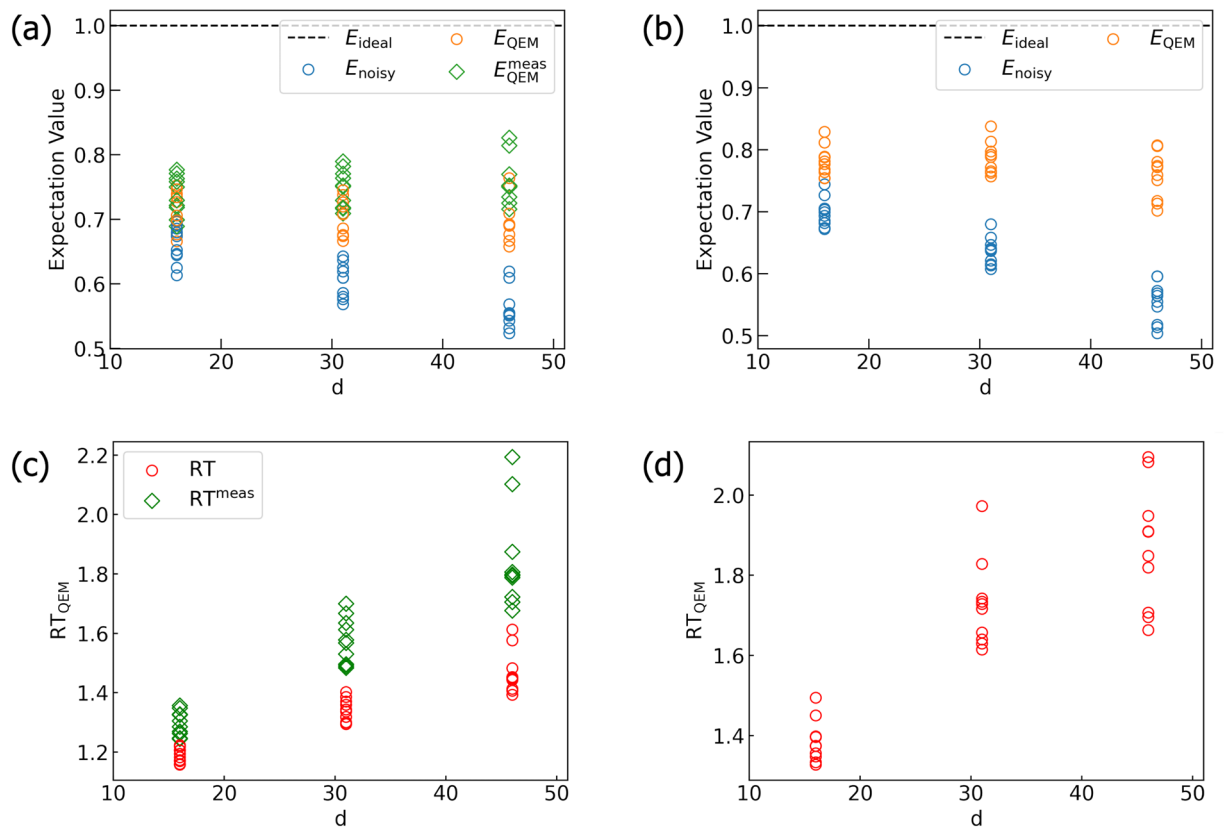
Figure 22. Quantum computation (($\mathbf{a}$) and ($\mathbf{c}$)) and simulation results (($\mathbf{b}$) and ($\mathbf{d}$)) for $U_{\text{imp},2}^{\text{QC}}$. ($\mathbf{a}$) and ($\mathbf{b}$) are the plots of $\langle Z_1 Z_3 \rangle$ and ($\mathbf{c}$) and ($\mathbf{d}$) are those of RT_{QEM}. All the horizontal axes represent $d = 3n^{\text{rep}} + 1$ with $n^{\text{rep}} = 5, 10, 15$. The quantum computations have been conducted via ibmq_belem at 11:43, 09/11/2023 under $N_{\text{QC}} = 2^{10}, N_{\text{samp}} = 10$.

expectation values with and without QEM get farther from the ideal expectation value (black dashed line) rapidly as we increase n^{rep}, and correspondingly, the monotonic decreasing of RT_{QEM} is exhibited in Fig. 20c, which are similar to the characteristics in Fig. 6a,c. Compared with these results, however, RT_{QEM} decreases gradually since not only the noisy expectation values but also the expectation values with QEM increase rapidly. In contrast, in Fig. 20b, we see that while the expectation values without QEM or noisy expectation values (blue open circles) increase gradually as d (or n^{rep}) gets larger the expectation values with QEM (orange open circles) take almost the same value. In Fig. 20d, the monotonic increasing of RT_{QEM} is exhibited. In the experiment of the implementation of $U_{\text{pre1}}^{\text{QC}}$, both the expectation values and RT_{QEM} obtained by the quantum computation and those by the quantum simulation show qualitatively different behaviors. On the other hand, all the ratios RT_{QEM} are greater than one. As a result, our QEM scheme has worked, however, not effectively.

Next, let us explain the results of the implementation which are shown in Fig. 21. Figure 21a,c display quantum computation results via ibm_perth and Fig. 21b,d present quantum simulation results[85]: Fig. 21a,b are the results of expectation values $\langle Z_1 Z_3 \rangle$ while Fig. 21c,d are those of RT_{QEM}. All the horizontal axes describe $d = 3n^{\text{rep}} + 1$ with $n^{\text{rep}} = 5, 10, 15$: the factor three represents the implementation of one CNOT gate $U_{CX}[Q_{r1}; Q_{r3}]$ and two identity operations acting on both Q_{r1} and Q_{r3}. Such identity operations are coming from the operation $\left(R_z \left(\frac{\pi}{2} \right) \cdot \sqrt{X} \cdot R_z \left(\frac{\pi}{2} \right) \right)^2 = \mathbf{1} \cdot \mathbf{1}$: see also the fourth comment in "Noise characterization". Note that the number of the implementation of the virtual Z gate is not included in the depths of the quantum algorithms. Let us explain from the results in Fig. 21a,b. We observe that both of these results show qualitatively the same behavior, i.e., while the noisy expectation values (blue open circles) gradually get farther from the ideal expectation value (black dashed line) as n^{rep} increases, the expectation values with QEM (orange open circles) approximately take the same value. On the other hand, in Fig. 21c,d the ratios RT_{QEM} exhibit the monotonic increasing, and moreover the ratio RT_{QEM} is greater than one for every d: RT_{QEM} shown in Fig. 21c are larger than those in Fig. 20c for every d. In addition to the quantum computation with ibm_perth, we have also conducted a quantum computation using ibmq_belem and a quantum simulation for $U_{\text{imp},2}^{\text{QC}}$[85]: Fig. 22a,c are the quantum computation results and Fig. 22b,d are the quantum simulation results. As similar to the labeling in Fig. 21, Fig. 22a,b (22c,d) are the results of $\langle Z_1 Z_3 \rangle$ (RT_{QEM}), and all the vertical axes represent $d = 3n^{\text{rep}} + 1$ with $n^{\text{rep}} = 5, 10, 15$. Overall, the characteristics of these results are similar to those for ibm_perth. Here we plot two types of expectation values with QEM which are indicated by the orange open circles and the green squares and

Method	N_a	N_{AQC}
Ours	k	$\mathcal{O}\big((N_q d)^k\big)$
PEC	0	$\mathcal{O}\big(\exp(2b\epsilon_{\mathrm{err}}N_q d)\big)$
ZNE	0	0
ESD/VD	$(n-1)N_q + \mathrm{lor}(n-1)N_q$	0

Table 1. Comparison between our QEM method in kth-order perturbation theory and the other three methods. Here b is a positive constant and ϵ_{err} is a gate error rate. The number of ancilla bits N_a is equal to $(n-1)N_q$ for VD without ancilla bits.

two types of RT$_{\mathrm{QEM}}$ by the red open circles and the green squares. The quantities plotted by the orange and red open circles have been calculated by using the data of the T_1 and T_2 times obtained at 05:08 whereas the others obtained by the data taken at 12:33: see Table I in the Supplementary Information. The time when the latter data has been taken (12:33) is closer to the time when the quantum computation has been conducted (11:43), and thus we consider that the latter RT$_{\mathrm{QEM}}$ are greater than the former ones: see also the third comment in "Noise characterization". As a result, we observe RT$_{\mathrm{QEM}} > 1$ for every quantum computation in our experiment, and we consider that our QEM works for the IBM Q Experience processors.

To summarize our experiment of the implementation of our QEM scheme, the quantum computations for $U_{\mathrm{pre1}}^{\mathrm{QC}}$ and $U_{\mathrm{imp,2}}^{\mathrm{QC}}$ show the different behaviors although the same machine has been used: the former case exhibits the different behavior from the simulation result while the latter case shows qualitatively the similar behavior. Moreover, the ratios RT$_{\mathrm{QEM}}$ for $U_{\mathrm{imp,2}}^{\mathrm{QC}}$ are larger than those for $U_{\mathrm{pre1}}^{\mathrm{QC}}$, which are in contrast to the results in Figs. 6 and 7: On the whole, RT$_{\mathrm{QEM}}$ for $U_{\mathrm{pre2}}^{\mathrm{QC}}$ are basically smaller than those for $U_{\mathrm{pre1}}^{\mathrm{QC}}$. One way to interpret the characteristics in Fig. 21c and 22c, the increasing behavior of RT$_{\mathrm{QEM}}$ with respect to n^{rep}, is as follows. When the noise strength is too small the noisy expectation values become sufficiently close to the ideal ones and in such circumstances the conduction of our QEM scheme can give rise to negative effects since computers can treat up to certain digits. Indeed, such a characteristic has been observed in the quantum simulation results in Fig. 9c,d and 17b in the range $0 < \vartheta_\tau \leq 0.1$: for a superconducting qubit system $T_1 \approx 100\,\mu$s and $\Delta t \approx 100$ ns and ϑ_τ is estimated to be $\vartheta_\tau \approx 0.063$. Thus, in order to utilize our QEM method effectively we need to use it under circumstances with moderate quantum noise strengths or for running moderately long quantum algorithms under weak quantum noise effects: such a way of interpretation, however, cannot be adapted to understand the characteristics in Fig. 20c. Consequently, the interpretation of the discrepancy between the quantum simulation result and the quantum computational result for $U_{\mathrm{pre1}}^{\mathrm{QC}}$ and the discrepancy between the characteristic of RT$_{\mathrm{QEM}}$ for $U_{\mathrm{imp,2}}^{\mathrm{QC}}$ and that for $U_{\mathrm{pre1}}^{\mathrm{QC}}$ remain unresolved for this experiment.

We note that our experiment has been conducted under restricted conditions such as the time for which we could have used the IBM Q Experience processors and the machines which have been available. Provided that we have no such restrictions, let us give several comments on how to improve our results. The first way is to increase the value of N_{samp}. As indicated in Figs. 10 and 18, by increasing the value of N_{samp} we expect that the expectation values with QEM approach to unique values and we become clearer to see whether our QEM scheme is working or not. The second way is to mitigate other types of errors. By combining our QEM scheme with QEM methods for other errors such as state preparation and measurement errors or errors according to other quantum noise channels such as crosstalk[96], we anticipate that the value of RT$_{\mathrm{QEM}}$ becomes larger.

In addition to the above discussion, let us consider the effectiveness of the implementation of our QEM scheme on different quantum hardware and here we choose ion trap qubit systems. Ion trap qubit systems are engineered, for instance, as linear chains and a two-qubit gate operation can be exploited such that it can be performed on any pair of qubits[97]: In contrast, in order to implement a two-qubit gate on two separated qubits, say Q_a and Q_b, on the superconducting quantum devices which have been used in this experiment we need to insert SWAP operations acting on qubits which locate between Q_a and Q_b[86]. Therefore, all quantum algorithms as well as quantum-noise-effect circuit groups are able to be implemented as indicated by the quantum circuits for them. In other words, we can harness our QEM scheme on ion trap qubit devices without reformulating the quantum circuits. Next, let us discuss quantum noise in ion trap qubit systems. The quantum noise occur in, for instance, hyperfine-state type ion-trap qubit systems are considered to be the phase damping and T_2 times are about 10 s while single-qubit gate times are around 10 microseconds and two-qubit gate times are about $100\,\mu$s[7,12]. By setting $T_2 = (2\gamma_{\mathrm{p}})^{-1}$, we have $\tau_{\mathrm{p}} = \gamma_{\mathrm{p}}\Delta t \approx 5 \times 10^{-6}$, where we have set $\Delta t = 100\,\mu$s. τ_{p} is sufficiently small and therefore, we expect that our QEM scheme also works for ion-trap NISQ devices. From these ingredients, we expect that the implementation our QEM scheme works more effectively and is more suitable for ion trap qubit systems compared with the superconducting qubit systems.

Comparison with other methods

We make comparisons between our method and other methods. Although many types of QEM methods have been proposed up to now[26,27,50–76], here we focus on the following three methods, probabilistic error cancellation (PEC)[26,27,50,53,57,60,65,68–71,73,74,76], zero noise extrapolation (ZNE)[26,27,50–53,72,74,76], and error suppression by dearangements (ESD)[67,75,76] and virtual distillation (VD)[66,76]. In Table 1, we summarize and present the comparison between our method and the others in terms of the number of ancilla bits N_a and the number of additional
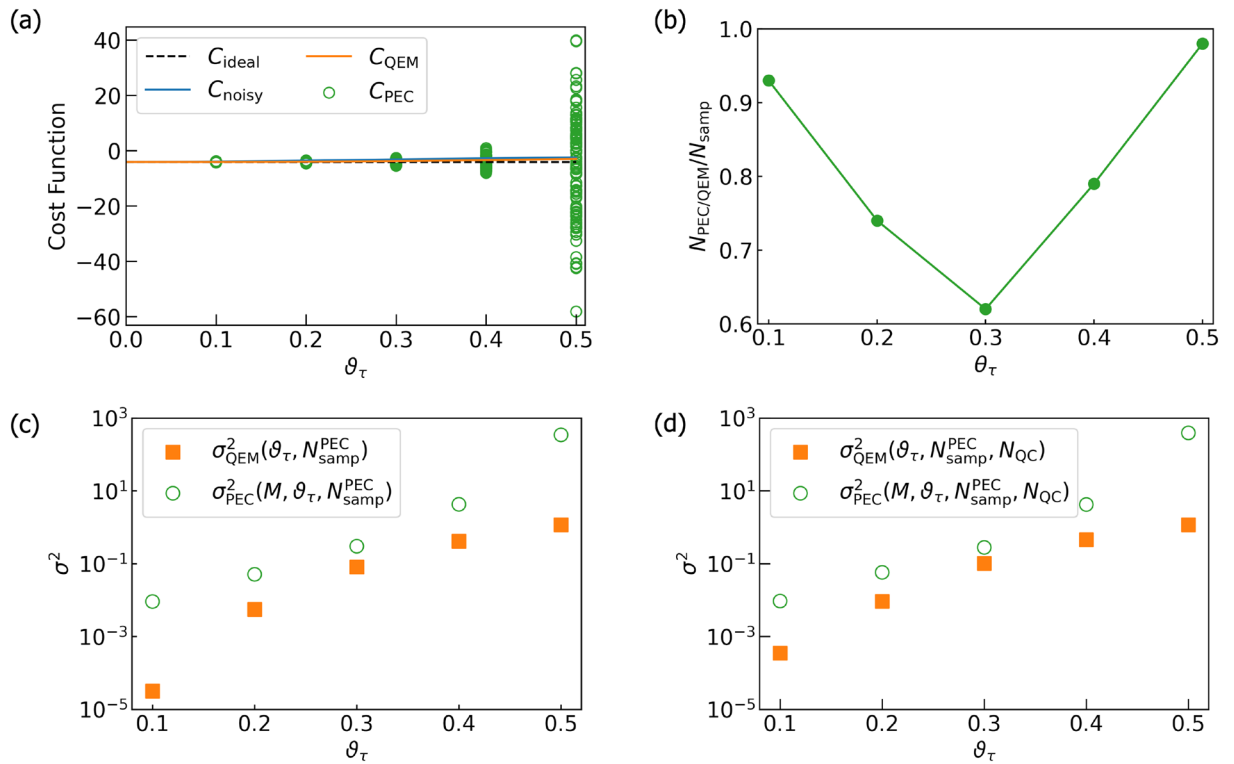
24

Figure 23. Numerical comparisons between our QEM method and PEC. We take $\vartheta_\tau = 0.1 \times a$ with $a = 1, 2, \dots, 5$. The results in (**a**)–(**c**) are the ones for $N_{QC} \to \infty$. (**a**) Plots of the ideal cost function C_{ideal} (black dotted line), the cost function without QEM C_{noisy} (blue curve), the cost function with QEM C_{QEM} (orange curve), and the cost function with PEC C_{PEC} which are plotted by 100 ($N_{samp}^{PEC} = 100$) green circles for each ϑ_τ. We take $M = 181$. (**b**) The ratio $N_{PEC/QEM}/N_{samp}$ which satisfy $\mathrm{RT}_{PEC/QEM} > 1$. (**c**) Plots of the variances $\sigma_{QEM}^2(\vartheta_\tau, N_{samp}^{PEC})$ and $\sigma_{PEC}^2(M, \vartheta_\tau, N_{samp}^{PEC})$ in Eq. (23). We take $M = 181$ and $N_{samp}^{PEC} = 100$. (**d**) Plots of the variances $\sigma_{QEM}^2(\vartheta_\tau, N_{samp}^{PEC}, N_{QC})$ and $\sigma_{PEC}^2(M, \vartheta_\tau, N_{samp}^{PEC}, N_{QC})$ by taking $N_{QC} = 2^{10}$.

circuits N_{AQC}. We choose PEC from these three methods and numerically compare with our method and the reason we do this is the following. Both methods are theoretically similar such that they evaluate quantum noise effects on quantum states by quantum computational operations and perform QEM which are represented as sums of expectation values yielded by ensembles of quantum circuits including original circuits and quantum circuits containing additional operations.

Comparison with PEC

First, we make a comparison between our method and PEC[26,27,50,53,57,60,65,68–71,73,74,76]. These two methods have a theoretical similarity such that they use additional quantum computational operations to mitigate quantum noise effects, however, their treatments are technically different. In PEC, quantum noise effects are mitigated by constructing inverse processes of quantum noises comprised of the sixteen basis operations and quasiprobabilities called recovery operations. In other words, both original circuits and additional quantum circuits are probabilistically generated owing to quasiprobabilities. Suppose that the recovery operation for the quantum noise under consideration, which acts on a single-qubit state, is composed of N_{quasp} nonzero quasiprobabilities. The elementary operations of the recovery operation (the terms composing the recovery operation) is inserted after each operation of U_l and the maximum number of circuits for performing PEC is $N_{quasp}^{N_q d}$. Once all these circuits are run obeying the quasibrobabilities the quantum noise effects are completely mitigated (non-perturbavtive method with respect to the noise strength). We mathematically describe the recovery operation as follows. By writing the non-zero quasiprobabilities and the associated basis operations as η_α and $\mathcal{B}_\alpha$ ($\alpha = 1, \dots, N_{quasp}$), respectively, the recovery operation executed after the operation of U_l is described in terms of these quantities as $\mathcal{E}_{QN,l}^{-1} = \bigotimes_{j_l=1}^{N_{q,l}^{PEC}} \left(\sum_{\alpha_{j_l}=1}^{N_{quasp}} \eta_{\alpha_{j_l}} \mathcal{B}_{\alpha_{j_l}} \right)$, where $N_{q,l}^{PEC}$ is the number of qubits on which the recovery operations in the lth layer act, $\mathcal{B}_{\alpha_{j_l}}$ is the basis operation acting on the jth qubit Q_j, and $\eta_{\alpha_{j_l}}$ is the associated quasiprobability of $\mathcal{B}_{\alpha_{j_l}}$: note that $\eta_{\alpha_{j_l}}$ include both positive and negative values. Furthermore, let us say that we compute an expectation value $\langle O \rangle$ with a repetition number N_{QC} and accuracy ϵ and write an associated quantum mechanical variance by $\Delta_O^2[N_{QC}, \epsilon]$. When we perform PEC under the same repetition number N_{QC} and the accuracy ϵ the variance of the expectation value with PEC being performed, which we denote by $\Delta_{O,PEC}^2[N_{QC}, \epsilon]$, becomes larger than the original variance $\Delta_O^2[N_{QC}, \epsilon]$ as $\Delta_{O,PEC}^2[N_{QC}, \epsilon] = \mathcal{C}_{PEC}^2 \Delta_O^2[N_{QC}, \epsilon]$, where

$$\mathcal{C}_{\mathrm{PEC}}^2 = \left(\prod_{l=1}^{d} \sum_{\alpha=1}^{N_{\mathrm{quas}}^{N_{q,l}^{\mathrm{PEC}}}} |\eta_\alpha| \right)^2 = \exp(2b\epsilon_{\mathrm{err}}d)$$ with b a positive constant and ϵ_{err} is a gate error rate which is

assumed to be gate independent (or a typical gate error rate)[27,50,53,60]. Let us say that we generate M quantum circuits (we call M as PEC sampling number) obeying the quasiprobability and an elementary circuit of the M quantum circuits can be either the original circuit or one of the additional circuits. By writing the set of quantum circuits for PEC which is composed of $N_{\mathrm{quas}}^{N_q d}$ quantum circutis by $\mathcal{S}^{\mathrm{PEC}}$ and the ith generated circuit by C_i^{PEC} ($\in \mathcal{S}^{\mathrm{PEC}}$), the exact error cancellation owing to PEC is described by $\lim_{M\to\infty} \langle O \rangle_{\mathrm{PEC},M} = \lim_{M\to\infty} \sum_{i=1}^{M} \langle O \rangle_{C_i^{\mathrm{PEC}}} = \langle O \rangle_{\mathrm{ideal}}$, where $\langle O \rangle_{C_i^{\mathrm{PEC}}}$ is the expectation value of O obtained by the quantum circuit C_i^{PEC}, which is generated with the probability $\frac{1}{M}$[50]. How many times a quantum circuit, say $\mathcal{C}_a$, appears in the M sampling depends on its quasiprobability. The computational resource, however, is finite in practice and in order to implement PEC in a real circumstance we take the sampling number M such that $\left| \langle O \rangle_{\mathrm{PEC},M} - \langle O \rangle_{\mathrm{ideal}} \right| \leq \epsilon$. The sampling number M scales in ϵ as $\mathcal{O}(\epsilon^{-2})$[50,53,60,73,74,76]. On the other hand, our QEM method is conducted by the estimation of the kth-order quantum noise effect $\Delta_k^{\mathrm{AD}} \rho_{d\cdots1}$ and we subtract it from the noisy expectation value. In contrast to the quasiprobabilities, the coefficients which compose $\Delta_k^{\mathrm{AD}} \rho_{d\cdots1}$ such as ± 1 and $\pm \frac{1}{4}$ are not probabilistic but deterministic (see Eq. (9) and the argument in Sect. I in the Supplementary Information), and furthermore, to evaluate $\Delta_k^{\mathrm{AD}} \rho_{d\cdots1}$ we deterministically prepare and execute the quantum-noise-effect circuit group whose size (the number of elementary circuits composing it) is $\mathcal{O}((dN_q)^k)$. The quality of our QEM method gets better as we increase k and this corresponds to the increasement of M in PEC. Let us make numerical comparisons between these two methods for QAOA discussed in "QAOA" and show them in Fig. 23. We perform the quantum simulations by taking the AD strength $\vartheta_\tau = 0.1 \times a$ with $a = 1, 2, \ldots, 5$. We take $k = 1$ (first-order perturbation theory) for our QEM method and for PEC we take $M = 3dN_q + 1$, which is the number of quantum circuits we need to perform our first-order QEM scheme, and in this case $d = 15, N_q = 4$ and $M = 181$. The reason why we take $M = 3dN_q + 1$ is the following. Let us say that we conduct each QEM method with the same amount of quantum computational resource (under the same condition) and this can be regarded as the number of quantum circuits to be used for QEM. This is because, as mentioned previously, both methods are described as sums of expectation values generated by ensembles of quantum circuits consist of original circuits and quantum circuits including additional operations. By taking account this, next let us introduce an error defined by $\delta_{\mathrm{QEM}} = \left| \langle \hat{O} \rangle_{\rho_{d\cdots1}} - \langle \hat{O} \rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}} \right|$, which represents the absolute of the difference between the ideal expectation value and the expectation value with our QEM method being performed.

On the other hand, we introduce an error $\delta_{\mathrm{PEC}} = \left| \langle \hat{O} \rangle_{\rho_{d\cdots1}} - \langle O \rangle_{\mathrm{PEC},M} \right|$ with setting $M = 3dN_q + 1$. The numerical comparison between our method and PEC can be redescribed such that we perform each method using the common quantum computational resource which is the $3dN_q + 1$ quantum circuits and examine whether δ_{QEM} is smaller than δ_{PEC} or not: the method having smaller error can be represented as the better QEM method. Hereinafter, let us rewrite the expectation value with QEM and that with PEC as $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\mathrm{QEM}}} \to \langle \hat{O} \rangle_{\mathrm{QEM}}(\vartheta_\tau)$ and $\langle O \rangle_{\mathrm{PEC},M} \to \langle O \rangle_{\mathrm{PEC}}(\vartheta_\tau, M)$, respectively, to express the noise-strength dependencies. Furthermore, we rewrite the ideal expectation value and the noisy expectation value as $\langle \hat{O} \rangle_{\rho_{d\cdots1}} \to \langle \hat{O} \rangle_{\mathrm{ideal}}$ and $\langle \hat{O} \rangle_{\rho_{d\cdots1}^{\mathrm{real}}} \to \langle \hat{O} \rangle_{\mathrm{noisy}}(\vartheta_\tau)$, respectively. Note that the operator O is chosen to be H_C in Eq. (17). The recovery operation of AD acting on single-qubit states after the operation of U_l is described by[73]

$$\mathcal{E}_{\mathrm{AD},l}^{-1} = \bigotimes_{j_l=1}^{5} \frac{1 + \sqrt{1 - \epsilon^{\mathrm{AD}}(\vartheta_\tau)}}{2(1 - \epsilon^{\mathrm{AD}}(\vartheta_\tau))} \mathbf{1}_{j_l} + \frac{1 - \sqrt{1 - \epsilon^{\mathrm{AD}}(\vartheta_\tau)}}{2(1 - \epsilon^{\mathrm{AD}}(\vartheta_\tau))} \boldsymbol{\mathcal{Z}}_{j_l} \tag{21}$$
$$- \frac{\epsilon^{\mathrm{AD}}(\vartheta_\tau)}{1 - \epsilon^{\mathrm{AD}}(\vartheta_\tau)} (\boldsymbol{\mathcal{P}}_{|0\rangle})_{j_l},$$

where $\mathbf{1}_{j_l}$ is the identity operator acting on Q_j, $\boldsymbol{\mathcal{Z}}_{j_l}[\rho] = Z_j \rho Z_j$, $(\boldsymbol{\mathcal{P}}_{|0\rangle})_{j_l}$ is the reset operator acting on Q_j, and $\epsilon^{\mathrm{AD}}(\vartheta_\tau)$ is the error rate given by $\epsilon^{\mathrm{AD}}(\vartheta_\tau) = \sin^2\left(\frac{\vartheta_\tau}{2}\right) = 1 - e^{-\tau}$. Let us explain from the result in Fig. 23a. Here we plot four types of quantities, the ideal cost function $\langle \hat{O} \rangle_{\mathrm{ideal}}$ (black dotted line), the noisy (no QEM) cost function $\langle \hat{O} \rangle_{\mathrm{noisy}}(\vartheta_\tau)$ (blue curve), the cost function with QEM $\langle \hat{O} \rangle_{\mathrm{QEM}}(\vartheta_\tau)$ (orange curve), and the cost function with PEC $\langle \hat{O} \rangle_{\mathrm{PEC}}(M, \vartheta_\tau, l)$ with $l = 1, \ldots, N_{\mathrm{samp}}^{\mathrm{PEC}}$, where $N_{\mathrm{samp}}^{\mathrm{PEC}}$ is the repetition number of the PEC simulation under the same condition in terms of M and ϑ_τ and $\langle \hat{O} \rangle_{\mathrm{PEC}}(M, \vartheta_\tau, l)$ is the cost function acquired in the lth round of the PEC simulation. We note that $N_{\mathrm{samp}}^{\mathrm{PEC}}$ is different from N_{samp} introduced in "Numerical simulations". In contrast to our QEM scheme, PEC is a probabilistic theory and when we perform a PEC simulation for $N_{\mathrm{samp}}^{\mathrm{PEC}}$ times and obtain $N_{\mathrm{samp}}^{\mathrm{PEC}}$ data of the cost function $\langle \hat{O} \rangle_{\mathrm{PEC}}(M, \vartheta_\tau, l)$, in general $\langle \hat{O} \rangle_{\mathrm{PEC}}(M, \vartheta_\tau, l_1) \neq \langle \hat{O} \rangle_{\mathrm{PEC}}(M, \vartheta_\tau, l_2)$ for $l_1 \neq l_2$. This is why we introduce the repetition number $N_{\mathrm{samp}}^{\mathrm{PEC}}$ for the PEC simulation. On the other hand, $\langle \hat{O} \rangle_{\mathrm{QEM}}(\vartheta_\tau, l_1) = \langle \hat{O} \rangle_{\mathrm{QEM}}(\vartheta_\tau, l_2) = \langle \hat{O} \rangle_{\mathrm{QEM}}(\vartheta_\tau)$ for $l_1 \neq l_2$ and the repetition number $N_{\mathrm{samp}}^{\mathrm{PEC}}$ is unnecessary for

our QEM scheme. We compute $\langle\hat{O}\rangle_{\text{ideal}}, \langle\hat{O}\rangle_{\text{noisy}}(\vartheta_\tau), \langle\hat{O}\rangle_{\text{QEM}}(\vartheta_\tau)$ with our original code while we compute $\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, l)$ with the software package Mitiq[74] and Cirq[98], and the PEC simulations are done under the condition $M = 181$ and $N_{\text{samp}}^{\text{PEC}} = 100$ and the data points of $\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, l)$ are plotted with green circles for each ϑ_τ. To quantify which of the two cost functions, $\langle\hat{O}\rangle_{\text{QEM}}(\vartheta_\tau)$ and $\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, l)$, is closer to $\langle\hat{O}\rangle_{\text{ideal}}$, we introduce a measure defined by

$$\text{RT}_{\text{PEC/QEM}}(M, \vartheta_\tau, l) = \frac{|\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, l) - \langle\hat{O}\rangle_{\text{ideal}}|}{|\langle\hat{O}\rangle_{\text{QEM}}(\vartheta_\tau) - \langle\hat{O}\rangle_{\text{ideal}}|}. \tag{22}$$

Like RT_{QEM} in Eq. (13), the ratio $\text{RT}_{\text{PEC/QEM}}(M, \vartheta_\tau, l)$ becoming greater than one implies that our QEM scheme is working better than PEC. By denoting the number of $\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, l)$ data points satisfying $\text{RT}_{\text{PEC/QEM}}(M, \vartheta_\tau, l) > 1$ as $N_{\text{PEC/QEM}}$, in Fig. 23b we show the ratio $N_{\text{PEC/QEM}}/N_{\text{samp}}^{\text{PEC}}$. We see that the ratio $\text{RT}_{\text{PEC/QEM}}(M, \vartheta_\tau, l)$ is greater than 50% for every ϑ_τ. We also obtain the ratio $N_{\text{PEC/QEM}}/N_{\text{samp}} = 0.99$ for the realistic noise strength $\vartheta_\tau = 0.045$ ($T_1 \sim 100\,\mu\text{sec}$ and $\Delta t \sim 100\,\text{nsec}$) in the current superconducting devices.

In addition to the ratio $\text{RT}_{\text{QEM}}(M, \vartheta_\tau, l)$ in Eq. (22), we compute mean squared errors defined by

$$\sigma_{\text{PEC}}^2(M, \vartheta_\tau, N_{\text{samp}}^{\text{PEC}}) = \sum_{l=1}^{N_{\text{samp}}^{\text{PEC}}} \frac{1}{N_{\text{samp}}^{\text{PEC}}} \left(\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, l) - \langle\hat{O}\rangle_{\text{ideal}}\right)^2,$$

$$\sigma_{\text{QEM}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}}) = \sum_{l=1}^{N_{\text{samp}}^{\text{PEC}}} \frac{1}{N_{\text{samp}}^{\text{PEC}}} \left(\langle\hat{O}\rangle_{\text{QEM}}(\vartheta_\tau, l) - \langle\hat{O}\rangle_{\text{ideal}}\right)^2. \tag{23}$$

Note that $\sigma_{\text{QEM}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}}) = \sigma_{\text{QEM}}^2(\vartheta_\tau)$ (no $N_{\text{samp}}^{\text{PEC}}$ dependency). By using the variances in Eq. (23) we can re-describe the comparison between the quality of our QEM and that of PEC as the comparison of the magnitudes of the two variances, i.e., the method exhibiting smaller variance has the better QEM quality. We show this numerical result in Fig. 23c and we obtain $\sigma_{\text{QEM}}^2(\vartheta_\tau) < \sigma_{\text{PEC}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}})$ for every ϑ_τ. The results in Fig. 23a–c are the ones for $N_{\text{QC}} \to \infty$. In Fig. 23d, we compute the variances $\sigma_{\text{QEM}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}})$ and $\sigma_{\text{PEC}}^2(M, \vartheta_\tau, N_{\text{samp}}^{\text{PEC}})$ for $N_{\text{QC}} = 2^{10}$ (each circuit is executed by taking $N_{\text{QC}} = 2^{10}$) and denote them as $\sigma_{\text{QEM}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}})$ and $\sigma_{\text{PEC}}^2(M, \vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}})$, respectively. Note that we have used $N_{\text{samp}}^{\text{PEC}}$ data of expectation values $\langle\hat{O}\rangle_{\text{PEC}}(M, \vartheta_\tau, N_{\text{QC}}, l)$ to compute $\sigma_{\text{PEC}}^2(M, \vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}})$: for computing both $\sigma_{\text{PEC}}^2(M, \vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}})$ and $\sigma_{\text{QEM}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}})$ we set $N_{\text{samp}} = 1$. We do not observe any significant difference in the variance of PEC between the finite and infinite N_{QC} cases, as can be seen by comparing Fig. 23c,d. However, due to the finite $N_{\text{QC}} = 2^{10}$, we observe a slight increase in the variance of our QEM in the range $0.1 \le \vartheta_\tau \le 0.3$. Nonetheless, our QEM still exhibits $\sigma_{\text{QEM}}^2(\vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}}) < \sigma_{\text{PEC}}^2(M, \vartheta_\tau, N_{\text{samp}}^{\text{PEC}}, N_{\text{QC}})$ for every ϑ_τ.

As a result, the quality of our QEM method outperforms that of PEC under such conditions. Let us examine for realistic cases when M is sufficiently larger than 181 and we consider that there are two types of effects. The first one is that, owing to the concept of PEC, $\langle O\rangle_{\text{PEC}}(\vartheta_\tau, M)$ gets closer to $\langle O\rangle_{\text{ideal}}$ owing to an increasement of M, which is a positive effect. The second one, which is a negative effect, is that the total amount of errors associated with the additional operations gets bigger by increasing M which makes $\langle O\rangle_{\text{PEC}}(\vartheta_\tau, M)$ to become farther from $\langle O\rangle_{\text{ideal}}$. At some point, say $M = M_c$, we consider that the second (negative) effect gets larger than the first (positive) effect because the second effect is not to be mitigated and gets bigger. As a result, $\langle O\rangle_{\text{PEC}}(\vartheta_\tau, M)$ becomes farther from $\langle O\rangle_{\text{ideal}}$ for $M > M_c$. It has been shown in Ref.[60] that the errors associated with the additional operations can be mitigated by combining with ZNE. The necessity of ZNE, however, implies that we need an additional computational resource to mitigate such errors. In contrast, our QEM method is conducted self-consistently such that the quantum noise effects on both the original circuit and the quantum-noise-effect circuit groups are mitigated. In other words, we do not need additional resources to mitigate the quantum noise effects on the quantum-noise-effect circuit groups. We consider that such a self-consistency is the advantage of our method compared with PEC. Furthermore, the size of the quantum-noise-effect circuit group (the computational cost) is practically polynomial in $N_q d$ while the number of quantum circuits for performing PEC is $N_{\text{quasp}}^{N_q d}$ (exponential in $N_q d$), and thus the practical computational cost of our QEM method is lower than that of PEC. In total, we consider that our QEM method is superior to PEC even for sufficiently large M.

Comparison with ZNE

Next, let us make a comparison between ZNE[26,27,50–53,72,74,76] and our method. The similarity between ZNE and our method (as well as PEC) is that both of these methods require additional quantum circuits. In ZNE, first a quantum mechanical expectation of an physical observable, say O_{phys}, is measured which includes a quantum computational error given by a noise strength (or an error rate) γ_0. Then, we create extra quantum circuits which ideally yield the same expectation value of O_{phys} as the original quantum circuit does but with noise strengths larger than γ_0. Such boostings of the noise strengths can be done, for instance, by insertions of identity gate operations[27,51,72]. Let us denote the original circuit by C_{org}, which is subject to the quantum noise with the strength γ_0, while we denote the extra circuits subject to noises with strengths γ_j by C_j with $j = 1, 2, \ldots, N_{\text{ZNE}}$, i.e., for ZNE $N_{\text{AQC}} = N_{\text{ZNE}}$. Here we labeled the extra circuits C_j so that $\gamma_0 < \gamma_1 < \cdots \gamma_{N_{\text{ZNE}}}$. The highest order of the quantum noise effects which can be mitigated is N_{ZNE} and this is one of the powerful advantage of ZNE. The difference between our method

and ZNE, on the other hand, is that in our method (as well as in PEC and ESD and VD) QEM tasks are performed by quantum computational operations (gate operations and measurements) whereas in ZNE quantum computers are used for calculating expectation values while the QEM tasks are done by classical computers. The advantages of ZNE are that it does not require ancilla bits and QEM can be performed without knowledge of quantum noises. There is a drawback, however, that its quality becomes poor when γ_0 is too big. Furthermore, in experiments we need to obtain the values $\gamma_0, \gamma_1, \ldots, \gamma_{\text{ZNE}}$ with high precision. On the other hand, our method requires both the ancilla bits and knowledge of quantum noises but can be applied to quantum noises for arbitrary noise strengths although higher computational cost is demanded for doing higher-order perturbation and both the values of decay rates and gate operation times are needed with high precision. Next let us compare the two methods with respect to the depth of a quantum algorithm and a coherence time with which we identify as a T_1 time. Here we assume that all single-qubit and two-qubit gate times are identical and denote the single-qubit time and the two-qubit gate time by t_1^{g} and t_2^{g}, respectively. Moreover, we assume that all T_1 times are identical with respect to qubits. Let us first consider from the conduction of Richardson-extrapolation-based ZNE and for simplicity we make an approximation $t_1^{\text{g}} \ll t_2^{\text{g}}$ so that the depth of a quantum algorithm d can be identified with the depth of (the number of the layers of) two-qubit gates to be implemented. Let us say that we enhance the noise strength γ_i by the factor $c_i = r^i$ with taking $r_i = 1 + 0.1 \times i$ with $i = 1, 2, \ldots$ Such a situation can be created by inserting an identity operation for the time $0.1 \times i \times t_2^{\text{g}}$ after the operation of each U_k. In this way, we can effectively make the operation time of U_k to be $r_i t_2^{\text{g}}$ or we can effectively make the decay rate to be $r_i \gamma_i$ while the operation time of U_k to be t_2^{g}. The operation time to execute the quantum algorithm $\prod_{k=1}^{d} U_k = U_d \cdot U_{d-1} \cdots U_2 \cdot U_1$ plus the additional identity operations is $r_i d t_2^{\text{g}}$ and it must satisfy the condition $r_i d t_2^{\text{g}} < T_1$. For superconducting qubits $t_2^{\text{g}} \approx 100$ nsec and $T_1 \approx 100\,\mu\text{s}$[10,11,86] and under such conditions we obtain $r_i d < 1000$. For $d = 800$, we have $r_1 d = 880, r_2 d = 960, r_3 d > 1000$, which means that we are able to perform QEM up to second order: For $d = 900$ we have $r_1 d = 990, r_2 d > 1000$, which means that we are able to perform QEM up to linear order, and for $d > 910$ we become unable perform QEM. For larger c_i the upper limit of d such that QEM is valid gets smaller. From these considerations, we see that we can perform high-order QEM for small d while for large d we can only perform low-order QEM and for sufficiently large d we cannot apply QEM. Such a characteristics comes from the d dependence of the total additional identity operation time. Furthermore, the exponential extrapolation also becomes invalid for large d since it is only effective for small error rate ϵ_{err} such that $\epsilon_{\text{err}} d = \mathcal{O}(1)$[27]. Next, let us consider the conduction of our QEM. The time for performing our kth-order QEM method is at most $dt_2^{\text{g}} + (2t_2^{\text{g}} + t_1^{\text{g}})k \simeq dt_2^{\text{g}} + 2.1kt_2^{\text{g}}$ (we assume that all additional operations are P^1), where we have set $t_1^{\text{g}} \simeq 0.1t_2^{\text{g}}$[86]. Thus, the operational time for implementing the k additional operations, which is $2.1kt_2^{\text{g}}$, does not depend on d. For $d = 800, 900,$ and 910 the highest orders of quantum noise effects which we can mitigate are $95, 47,$ and 42, respectively. Consequently, we consider that for small d satisfying $r_{\text{ZNE}} dt_2^{\text{g}} < T_1$ the quality of ZNE is better while for large d our method has a better quality and such a tendency does not change even T_1 times are extended and the gate times t_2^{g} get shorter. At the end, we mention that ZNE is not applicable to time-dependent noise[27,50,51] whereas our method is by reformulating τ as a function of time.

Comparison with ESD and VD

Finally, we make a comparison between our method and ESD[67,75,76] and VD[66,76]. Since these two are similar approaches we bring them together and abbreviate it as ESD/VD. Like PEC and our method, ESD/VD is composed of gate operations and/or quantum measurement on an ancilla bit. ESD/VD has two advantages, it can be applied to various types of quantum noise and additional quantum circuits are unneeded. The procedure of ESD/VD is comprised of three parts, a preparation of n copies of an original circuit which generate $(\rho_{d\cdots1}^{\text{real}})^{\otimes n}$, a performance of derangement operation, which is a generalization of SWAP operation (in Ref.[66] it is called cyclic shift operator), and an operation of controlled-O operator with O denoting the physical operator of which we take an expectation value[67,75]: In Ref.[66], the scheme which does not use an ancilla-assisted measurement has been presented and in this case $N_a = (n-1)N_q$ (see Table 1). Let us express $\rho_{d\cdots1}^{\text{real}}$ in a spectral decomposition form denoted by $\rho_{d\cdots1}^{\text{real}} = \lambda_{\text{dom}}|\psi_{\text{dom}}\rangle\langle\psi_{\text{dom}}| + \sum_{a=1}^{2^{N_q}-1} \lambda_a|\psi_a\rangle\langle\psi_a|$, where the eigenvalues satisfy the descending order $\lambda_{\text{dom}} > \lambda_1 \cdots > \lambda_{2^{N_q}-1}$. Due to such a process, we obtain the expectation value $\langle\psi_{\text{dom}}|O|\psi_{\text{dom}}\rangle$. In ESD/VD, there are two problems which need to be handled with, the mismatch between the ideal state $\rho_{d\cdots1}$ and the dominant eigenvector state $\rho^{\text{dom}} = |\psi_{\text{dom}}\rangle\langle\psi_{\text{dom}}|$ which is called coherent mismatch or noise floor, and the mitigation of the quantum noise effects associated with the operations of the derangement operator and the controlled-O operator. It was argued in Ref.[67] that the quantum computational error coming from the coherent mismatch and that due to the quantum noise effects associated with the operations constructing the quantum circuit for ESD/VD, which we call ESD/VD circuit, can be mitigated by combining with ZNE. As argued in the previous discussion for ZNE, we consider, however, that such a hybrid scheme works provided that the composite circuit of the original circuit and the ESD/VD circuit can be executed within a coherence time of a qubit. Furthermore, we need large numbers of qubits to perform ESD/VD and this limits sizes of overhead. On the other hand, our method needs the additional quantum circuits (quantum-noise-effect circuit group) but both the errors associated with the operations of U_k and those associated with the insertions of the additional operations implemented on the quantum-noise-effect circuit group are self-consistently mitigated and saves an additional computational resource for mitigating the errors associated with the additional operations since our scheme does not need to be combined with other methods for doing this. Moreover, the number of ancilla bits which we need to perform k-th order QEM is k, which does not enlarge the overhead that much. Such properties, in principle, enable us to perform QEM for both short-term (NISQ algorithms) and long-term algorithms.

Conclusion and outlook

In this paper, we have established our QEM scheme for reducing the quantum noise (decoherence) effects on the single-qubit states which occur during the gate operations. We have formulated it as the perturbation theory with respect to the noise strength (in the case of AD effect it is τ), which are evaluated by the gate time and decay rate (T_1 time and/or T_2 time), and is represented by the ensemble of quantum circuits, namely the quantum-noise-effect circuit groups. The numbers of quantum circuits composing the quantum-noise-effect circuit groups are polynomial with respect to the product of the depth of the quantum algorithm under consideration and the number of register bits, which can be considered that the conduction of our QEM scheme is not so high-cost computational performance. To demonstrate the validity of our QEM scheme, we have performed the noisy quantum simulations of the qubits under the AD effects for four types of quantum algorithms based on the linear-order perturbation theory. It is to be noted that before we conduct our QEM scheme, we need to be careful with if the expectation values on which we are aiming to perform QEM do not include the linear-order term in τ or if they are equivalent to the ideal simulation results like the computation of the expectation value of the identity operator. As long as these two are not the cases then our linear-order QEM scheme works as we have demonstrated in "Numerical simulations", and it is valid in a broad region of τ, which implies its effectiveness and powerfulness. Our QEM scheme can be generalized to error mitigation for other types of quantum noises including the generalized amplitude damping, the phase damping, the composition of these two, and the stochastic Pauli noises like the depolarizing channel. Furthermore, it can be extended to cases of error mitigation for higher-order quantum noise effects and once this is established we expect that we become able to perform quantum computations in high accuracy even for long-depth quantum algorithms. In "QEM scheme implementation", we have discussed the experimental results of the implementation of our QEM scheme. Consequently, we have observed $RT_{QEM} > 1$ for every quantum computation and our QEM scheme has worked for the IBM Q Experience processors. In this work, we have focused on the decoherence effects acting on the single-qubit states and established the QEM scheme for them. In real quantum devices, however, there exist many types of errors and complex quantum noise channels. We expect that by conducting an elaborate noise (device) characterization we become able to improve the quality of QEM. For instance, by combining our QEM scheme with other error-mitigation techniques such as those for state preparation (initialization) and measurement and imperfections of gate operations or those for other quantum noise channels like crosstalk we anticipate to realize QEM with higher quality.

Our QEM scheme is solely conducted by gate operations and measurements on ancilla bits and can be applied to any type of quantum algorithm. Such a characteristic enables us to programmably perform high-accurate quantum computing solely by the quantum-computational operations (software manipulations). Furthermore, our QEM scheme can be performed with any type of quantum hardware such as solid-state systems and atomic-molecule and optical systems and with quantum devices of any generation, and both the computational cost whose order is polynomial in $N_q d$ and its accuracy can be coherently controlled. These three characteristics are the big advantages of our scheme.

One of the important outlooks of this work is quantum computations by large-scale quantum devices or future (next-generation) quantum devices using various types of quantum hardware such as superconducting circuits and ion-trap qubit systems. In such a case, we consider that we also need to take into account quantum noises acting on many-body quantum states such as collective quantum noises[87–89,99–102] and correlated noises[59,103–108]. Another important problem is the establishment of QEM schemes for time-dependent quantum noises including non-Markovian quantum noises[109,110].

Our QEM scheme can be extended to mitigation of these complex quantum noise effects provided that they are formulated as groups of quantum circuits. When such formalisms are being constructed, we expect that we become able to realize QEM scheme which mitigates various types of quantum noise effects. We expect that this leads to conduction of quantum computing for big-size problems with high-quality results being obtained. We believe that this paves the way to realize high-quality quantum computing for application to problems in many branches of science and engineering including material science, quantum chemistry, combinatorial optimization problems, and machine learning using large-scale quantum computers.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Feynman, R. P. Simulating physics with computers. *Int. J. Theoretical Phys.* **21**, 467–488 (1982).
2. Deutsch, D. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proc. R. Soc. Lond. A. Math. Phys. Sci.* **400**, 97 (1985).
3. Lloyd, S. Universal quantum simulators. *Science* **273**, 1073 (1996).
4. DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschritte der Physik Progress Phys.* **48**, 771 (2000).
5. M. A. Nielsen & I. Chuang. Quantum computation and quantum information (2002).
6. Georgescu, I. M., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.* **86**, 153. https://doi.org/10.1103/RevModPhys.86.153 (2014).
7. Linke, N. M. *et al.* Experimental comparison of two quantum computing architectures. *Proc. Natl. Acad. Sci.* **114**, 3305 (2017).
8. Wendin, G. Quantum information processing with superconducting circuits: A review. *Rep. Progress Phys.* **80**, 106001 (2017).
9. Krantz, P. *et al.* A quantum engineer's guide to superconducting qubits. *Appl. Phys. Rev.* **6**, 021318 (2019).
10. Kjaergaard, M. *et al.* Superconducting qubits: Current state of play. *Annu. Rev. Condensed Matter Phys.* **11**, 369 (2020).
11. Huang, H.-L., Wu, D., Fan, D. & Zhu, X. Superconducting quantum computing: A review. *Sci. China Inform. Sci.* **63**, 1 (2020).

12. Bruzewicz, C. D., Chiaverini, J., McConnell, R. & Sage, J. M. Trapped-ion quantum computing: Progress and challenges. *Appl. Phys. Rev.* **6**, 021314 (2019).
13. Kaushal, V. *et al.* Shuttling-based trapped-ion quantum information processing. *AVS Quant. Sci.* **2**, 014101 (2020).
14. Nakamura, Y., Pashkin, Y. A. & Tsai, J. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature* **398**, 786 (1999).
15. Makhlin, Y., Schön, G. & Shnirman, A. Quantum-state engineering with Josephson-junction devices. *Rev. Mod. Phys.* **73**, 357. https://doi.org/10.1103/RevModPhys.73.357 (2001).
16. Xiang, Z.-L., Ashhab, S., You, J. Q. & Nori, F. Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems. *Rev. Mod. Phys.* **85**, 623. https://doi.org/10.1103/RevModPhys.85.623 (2013).
17. Kwon, S., Tomonaga, A., LakshmiBhai, G., Devitt, S. J. & Tsai, J.-S. Gate-based superconducting quantum computing. *J. Appl. Phys.* **129**, 041102 (2021).
18. Cirac, J. I. & Zoller, P. Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091. https://doi.org/10.1103/PhysRevLett.74.4091 (1995).
19. Sørensen, A. & Mølmer, K. Quantum computation with ions in thermal motion. *Phys. Rev. Lett.* **82**, 1971. https://doi.org/10.1103/PhysRevLett.82.1971 (1999).
20. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281. https://doi.org/10.1103/RevModPhys.75.281 (2003).
21. Wineland, D. J. Quantum information processing and quantum control with trapped atomic ions. *Physica Scripta* **2009**, 014007 (2009).
22. Häffner, H., Roos, C. F. & Blatt, R. Quantum computing with trapped ions. *Phys. Rep.* **469**, 155 (2008).
23. Peruzzo, A. *et al.* A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 1 (2014).
24. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 023023 (2016).
25. Kandala, A. *et al.* Hardware-efficient variational quantum Eigensolver for small molecules and quantum magnets. *Nature* **549**, 242 (2017).
26. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. *Rev. Mod. Phys.* **92**, 015003. https://doi.org/10.1103/RevModPhys.92.015003 (2020).
27. Endo, S., Cai, Z., Benjamin, S. C. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Jpn.* **90**, 032001 (2021).
28. E. Farhi, J. Goldstone, & S. Gutmann. A quantum approximate optimization algorithm. arXiv preprint arXiv:1411.4028 (2014).
29. G. E. Crooks. Performance of the quantum approximate optimization algorithm on the maximum cut problem. arXiv preprint arXiv:1811.08419 (2018)
30. Wang, Z., Hadfield, S., Jiang, Z. & Rieffel, E. G. Quantum approximate optimization algorithm for MaxCut: A fermionic view. *Phys. Rev. A* **97**, 022304 (2018).
31. R. Shaydulin & Y. Alexeev. Evaluating quantum approximate optimization algorithm: A case study. in *2019 tenth international green and sustainable computing conference (IGSC) (IEEE, 2019)*, 1–6.
32. Zhou, L., Wang, S.-T., Choi, S., Pichler, H. & Lukin, M. D. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Phys. Rev. X* **10**, 021067 (2020).
33. J. Abhijith, A. Adedoyin, J. Ambrosiano, P. Anisimov, A. Bärtschi, W. Casper, G. Chennupati, C. Coffrin, H. Djidjev, D. Gunter, *et al*. Quantum algorithm implementations for beginners. arXiv e-prints , arXiv (2018).
34. Schuld, M. & Petruccione, F. *Supervised Learning with Quantum Computers* Vol. 17 (Springer, 2018).
35. Mitarai, K., Negoro, M., Kitagawa, M. & Fujii, K. Quantum circuit learning. *Phys. Rev. A* **98**, 032309. https://doi.org/10.1103/PhysRevA.98.032309 (2018).
36. Benedetti, M., Lloyd, E., Sack, S. & Fiorentini, M. Erratum: Parameterized quantum circuits as machine learning models (2019 Quant. Sci. Tech. 4 043001). *Quant. Sci. Technol.* **5**, 019601 (2019).
37. Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505 (2019).
38. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
39. Palma, G. & Suominen, K. AK Ekert Proc. Roy. Soc. *London A* **452**, 567 (1996).
40. Resch, S. & Karpuzcu, U. R. Benchmarking quantum computers and the impact of quantum noise. *ACM Comput. Surveys (CSUR)* **54**, 1 (2021).
41. Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493. https://doi.org/10.1103/PhysRevA.52.R2493 (1995).
42. Devitt, S. J., Munro, W. J. & Nemoto, K. Quantum error correction for beginners. *Rep. Progress Phys.* **76**, 076001 (2013).
43. Lidar, D. A. & Brun, T. A. *Quantum Error Correction* (Cambridge University Press, 2013).
44. Roffe, J. Quantum error correction: An introductory guide. *Contemp. Phys.* **60**, 226 (2019).
45. Viola, L. & Lloyd, S. Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A* **58**, 2733 (1998).
46. Viola, L., Knill, E. & Lloyd, S. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.* **82**, 2417 (1999).
47. Khodjasteh, K. & Lidar, D. A. Fault-tolerant quantum dynamical decoupling. *Phys. Rev. Lett.* **95**, 180501 (2005).
48. Masuyama, Y. *et al.* Extending coherence time of macro-scale diamond magnetometer by dynamical decoupling with coplanar waveguide resonator. *Rev. Sci. Instrum.* **89**, 125007 (2018).
49. Wise, D. F., Morton, J. J. & Dhomkar, S. Using deep learning to understand and mitigate the qubit noise environment. *PRX Quant.* **2**, 010316. https://doi.org/10.1103/PRXQuantum.2.010316 (2021).
50. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
51. Kandala, A. *et al.* Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491 (2019).
52. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).
53. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).
54. Bonet-Monroig, X., Sagastizabal, R., Singh, M. & O'Brien, T. E. Low-cost error mitigation by symmetry verification. *Phys. Rev. A* **98**, 062339. https://doi.org/10.1103/PhysRevA.98.062339 (2018).
55. McArdle, S., Yuan, X. & Benjamin, S. Error-mitigated digital quantum simulation. *Phys. Rev. Lett.* **122**, 180501 (2019).
56. Jattana, M. S., Jin, F., De Raedt, H. & Michielsen, K. General error mitigation for quantum circuits. *Quant. Inform. Process.* **19**, 1 (2020).
57. Xiong, Y., Chandra, D., Ng, S. X. & Hanzo, L. Sampling overhead analysis of quantum error mitigation: Uncoded vs. coded systems. *IEEE Access* **8**, 228967 (2020).
58. A. Zlokapa & A. Gheorghiu. A deep learning model for noise prediction on near-term quantum devices, arXiv preprint arXiv:2005.10811 (2020).
59. Bravyi, S., Sheldon, S., Kandala, A., Mckay, D. C. & Gambetta, J. M. Mitigating measurement errors in multiqubit experiments. *Phys. Rev. A* **103**, 042605. https://doi.org/10.1103/PhysRevA.103.042605 (2021).
60. Sun, J. *et al.* Mitigating realistic noise in practical noisy intermediate-scale quantum devices. *Phys. Rev. Appl.* **15**, 034026. https://doi.org/10.1103/PhysRevApplied.15.034026 (2021).

61. Otten, M. & Gray, S. K. Accounting for errors in quantum algorithms via individual error reduction. *NPJ Quant. Inform.* **5**, 1. https://doi.org/10.1038/s41534-019-0125-3 (2019).
62. Otten, M. & Gray, S. K. Recovering noise-free quantum observables. *Phys. Rev. A* **99**, 012338. https://doi.org/10.1103/PhysRevA.99.012338 (2019).
63. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308. https://doi.org/10.1103/PhysRevA.95.042308 (2017).
64. Czarnik, P., Arrasmith, A., Coles, P. J. & Cincio, L. Error mitigation with Clifford quantum-circuit data. *Quantum* **5**, 592. https://doi.org/10.22331/q-2021-11-26-592 (2021).
65. Strikis, A., Qin, D., Chen, Y., Benjamin, S. C. & Li, Y. Learning-based quantum error mitigation. *PRX Quant.* **2**, 040330. https://doi.org/10.1103/PRXQuantum.2.040330 (2021).
66. Huggins, W. J. *et al.* Virtual distillation for quantum error mitigation. *Phys. Rev. X* **11**, 041036. https://doi.org/10.1103/PhysRevX.11.041036 (2021).
67. Koczor, B. Exponential error suppression for near-term quantum devices. *Phys. Rev. X* **11**, 031057 (2021).
68. Piveteau, C., Sutter, D., Bravyi, S., Gambetta, J. M. & Temme, K. Error mitigation for universal gates on encoded qubits. *Phys. Rev. Lett.* **127**, 200505 (2021).
69. Lostaglio, M. & Ciani, A. Error mitigation and quantum-assisted simulation in the error corrected regime. *Phys. Rev. Lett.* **127**, 200506 (2021).
70. Suzuki, Y., Endo, S., Fujii, K. & Tokunaga, Y. Quantum error mitigation as a universal error reduction technique: Applications from the NISQ to the fault-tolerant quantum computing eras. *PRX Quant.* **3**, 010345. https://doi.org/10.1103/PRXQuantum.3.010345 (2022).
71. Piveteau, C., Sutter, D. & Woerner, S. Quasiprobability decompositions with reduced sampling overhead. *NPJ Quant. Inform.* **8**, 1 (2022).
72. Pascuzzi, V. R., He, A., Bauer, C. W., de Jong, W. A. & Nachman, B. Computationally efficient zero-noise extrapolation for quantum-gate-error mitigation. *Phys. Rev. A* **105**, 042406. https://doi.org/10.1103/PhysRevA.105.042406 (2022).
73. Takagi, R. Optimal resource cost for error mitigation. *Phys. Rev. Res.* **3**, 033178. https://doi.org/10.1103/PhysRevResearch.3.033178 (2021).
74. LaRose, R. *et al.* Mitiq: A software package for error mitigation on noisy quantum computers. *Quantum* **6**, 774. https://doi.org/10.22331/q-2022-08-11-774 (2022).
75. Koczor, B. The dominant eigenvector of a noisy quantum state. *New J. Phys.* **23**, 123047. https://doi.org/10.1088/1367-2630/ac37ae (2021).
76. Cai, Z. *et al.* Quantum error mitigation. *Rev. Mod. Phys.*. **95**, 045005. https://doi.org/10.1103/RevModPhys.95.045005 (2023).
77. Wang, H., Ashhab, S. & Nori, F. Quantum algorithm for simulating the dynamics of an open quantum system. *Phys. Rev. A* **83**, 062317. https://doi.org/10.1103/PhysRevA.83.062317 (2011).
78. Hu, Z., Xia, R. & Kais, S. A quantum algorithm for evolving open quantum dynamics on quantum computing devices. *Sci. Rep.* **10**, 1 (2020).
79. Tornow, S., Gehrke, W. & Helmbrecht, U. Non-equilibrium dynamics of a dissipative two-site Hubbard model simulated on the IBM quantum computer. *J. Phys. A: Math. Theor.* **55**, 245302. https://doi.org/10.1088/1751-8121/ac6bd0 (2022).
80. García-Pérez, G., Rossi, M. A. & Maniscalco, S. IBM Q Experience as a versatile experimental testbed for simulating open quantum systems. *NPJ Quant. Inform.* **6**, 1 (2020).
81. Del Re, L., Rost, B., Kemper, A. F. & Freericks, J. K. Driven-dissipative quantum mechanics on a lattice: Simulating a fermionic reservoir on a quantum computer. *Phys. Rev. B* **102**, 125112. https://doi.org/10.1103/PhysRevB.102.125112 (2020).
82. Koppenhöfer, M., Bruder, C. & Roulet, A. Quantum synchronization on the IBM Q system. *Phys. Rev. Res.* **2**, 023026 (2020).
83. de Jong, W. A. *et al.* Quantum simulation of open quantum systems in heavy-ion collisions. *Phys. Rev. D* **104**, L051501 (2021).
84. Hama, Y. Quantum circuits for collective amplitude damping in two-qubit systems. https://doi.org/10.48550/ARXIV.2012.02410 (2020).
85. H. Abraham, AduOffei, R. Agarwal, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, E. Arbel, Arijit02, A. Asfaw, A. Avkhadiev, C. Azaustre, AzizNgoueya, A. Banerjee, A. Bansal, P. Barkoutsos, A. Barnawal, G. Barron, G. S. Barron, L. Bello, Y. Ben-Haim, D. Bevenius, A. Bhobe, L. S. Bishop, C. Blank, S. Bolos, S. Bosch, Brandon, S. Bravyi, Bryce-Fuller, D. Bucher, A. Burov, F. Cabrera, P. Calpin, L. Capelluto, J. Carballo, G. Carrascal, A. Chen, C.-F. Chen, E. Chen, J. C. Chen, R. Chen, J. M. Chow, S. Churchill, C. Claus, C. Clauss, R. Cocking, F. Correa, A. J. Cross, A. W. Cross, S. Cross, J. Cruz-Benito, C. Culver, A. D. Córcoles-Gonzales, S. D. E. Dandachi, M. Daniels, M. Dartiailh, DavideFrr, A. R. Davila, A. Dekusar, D. Ding, J. Doi, E. Drechsler, Drew, E. Dumitrescu, K. Dumon, I. Duran, K. EL-Safty, E. Eastman, G. Eberle, P. Eendebak, D. Egger, M. Everitt, P. M. Fernández, A. H. Ferrera, R. Fouilland, FranckChevallier, A. Frisch, A. Fuhrer, B. Fuller, M. GEORGE, J. Gacon, B. G. Gago, C. Gambella, J. M. Gambetta, A. Gammanpila, L. Garcia, T. Garg, S. Garion, A. Gilliam, A. Giridharan, J. Gomez-Mosquera, Gonzalo, S. de la Puente González, J. Gorzinski, I. Gould, D. Greenberg, D. Grinko, W. Guan, J. A. Gunnels, M. Haglund, I. Haide, I. Hamamura, O. C. Hamido, F. Harkins, V. Havlicek, J. Hellmers, Ł. Herok, S. Hillmich, H. Horii, C. Howington, S. Hu, W. Hu, J. Huang, R. Huisman, H. Imai, T. Imamichi, K. Ishizaki, R. Iten, T. Itoko, JamesSeaward, A. Javadi, A. Javadi-Abhari, W. Javed, Jessica, M. Jivrajani, K. Johns, S. Johnstun, Jonathan-Shoemaker, V. K, T. Kachmann, A. Kale, N. Kanazawa, Kang-Bae, A. Karazeev, P. Kassebaum, J. Kelso, S. King, Knabberjoe, Y. Kobayashi, A. Kovyrshin, R. Krishnakumar, V. Krishnan, K. Krsulich, P. Kumkar, G. Kus, R. LaRose, E. Lacal, R. Lambert, J. Lapeyre, J. Latone, S. Lawrence, C. Lee, G. Li, D. Liu, P. Liu, Y. Maeng, K. Majmudar, A. Malyshev, J. Manela, J. Marecek, and M. Marques, D. Maslov, D. Mathews, A. Matsuo, D. T. McClure, C. McGarry, D. McKay, D. McPherson, S. Meesala, T. Metcalfe, M. Mevissen, A. Meyer, A. Mezzacapo, R. Midha, Z. Minev, A. Mitchell, N. Moll, J. Montanez, G. Monteiro, M. D. Mooring, R. Morales, N. Moran, M. Motta, MrF, P. Murali, J. Müggenburg, D. Nadlinger, K. Nakanishi, G. Nannicini, P. Nation, E. Navarro, Y. Naveh, S. W. Neagle, P. Neuweiler, J. Nicander, P. Niroula, H. Norlen, NuoWenLei, L. J. O'Riordan, O. Ogunbayo, P. Ollitrault, R. Otaolea, S. Oud, D. Padilha, H. Paik, S. Pal, Y. Pang, V. R. Pascuzzi, S. Perriello, A. Phan, F. Piro, M. Pistoia, C. Piveteau, P. Pocreau, A. Pozas-Kerstjens, M. Prokop, V. Prutyanov, D. Puzzuoli, J. Pérez, Quintiii, R. I. Rahman, A. Raja, N. Ramagiri, A. Rao, R. Raymond, R. M.-C. Redondo, M. Reuter, J. Rice, M. Riedemann, M. L. Rocca, D. M. Rodríguez, RohithKarur, M. Rossmannek, M. Ryu, T. SAPV, SamFerracin, M. Sandberg, H. Sandesara, R. Sapra, H. Sargsyan, A. Sarkar, N. Sathaye, B. Schmitt, C. Schnabel, Z. Schoenfeld, T. L. Scholten, E. Schoute, J. Schwarm, I. F. Sertage, K. Setia, N. Shammah, Y. Shi, A. Silva, A. Simonetto, N. Singstock, Y. Siraichi, I. Sitdikov, S. Sivarajah, M. B. Sletfjerding, J. A. Smolin, M. Soeken, I. O. Sokolov, I. Sokolov, SooluThomas, Starfish, D. Steenken, M. Stypulkoski, S. Sun, K. J. Sung, H. Takahashi, T. Takawale, I. Tavernelli, C. Taylor, P. Taylour, S. Thomas, M. Tillet, M. Tod, M. Tomasik, E. de la Torre, K. Trabing, M. Treinish, TrishaPe, D. Tulsi, W. Turner, Y. Vaknin, C. R. Valcarce, F. Varchon, A. C. Vazquez, V. Villar, D. Vogt-Lee, C. Vuillot, J. Weaver, J. Weidenfeller, R. Wieczorek, J. A. Wildstrom, E. Winston, J. J. Woehr, S. Woerner, R. Woo, C. J. Wood, R. Wood, S. Wood, S. Wood, J. Wootton, D. Yeralin, D. Yonge-Mallo, R. Young, J. Yu, C. Zachow, L. Zdanski, H. Zhang, C. Zoufal. Qiskit: An open-source framework for quantum computing (2019). https://doi.org/10.5281/zenodo.2562110.
86. IBM Quantum Experience [online]. https://quantum-computing.ibm.com/ (2023)
87. Carmichael, H. J. *Statistical Methods in Quantum Optics 1: Master Equations and Fokker-Planck Equations* Vol. 1 (Springer Science & Business Media, 1999).
88. Agarwal, G. S. *Quantum Optics* (Cambridge University Press, 2012).
89. Breuer, H.-P. *et al. The Theory of Open Quantum Systems* (Oxford University Press on Demand, 2002).

90. K. Kraus, A. Böhm, J. D. Dollard, & W. Wootters. States, effects, and operations: fundamental notions of quantum theory. in *Lectures in Mathematical Physics at the University of Texas at Austin, Lecture Notes in Physics* **190** (1983).
91. L. K. Grover. in *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing* (1996).
92. Grover, L. K. Quantum mechanics helps in searching for a needle in a Haystack. *Phys. Rev. Lett.* **79**, 325. https://doi.org/10.1103/PhysRevLett.79.325 (1997).
93. Montanaro, A. Quantum algorithms: An overview. *NPJ Quant. Inform.* **2**, 1 (2016).
94. S. Jordan. Quantum algorithm zoo. Retrieved June **27**, 2013 (2011).
95. Y. Minato, K. Higa, & R. Nagai. IBM Quantum de manabu ryoshi konpyuuta (in Japanese) (Shuwa System, 2021) https://books.google.co.jp/books?id=mOdKzgEACAAJ.
96. Sarovar, M. *et al.* Detecting crosstalk errors in quantum information processors. *Quantum* **4**, 321. https://doi.org/10.22331/q-2020-09-11-321 (2020).
97. Wright, K. *et al.* Benchmarking an 11-qubit quantum computer. *Nat. Commun.* **10**, 5464 (2019).
98. C. Developers, Cirq (2022), See full list of authors on Github: https://github.com/quantumlib/Cirq/graphs/contributorshttps://doi.org/10.5281/zenodo.6599601.
99. Dicke, R. H. Coherence in spontaneous radiation processes. *Phys. Rev.* **93**, 99 (1954).
100. Gross, M. & Haroche, S. Superradiance: An essay on the theory of collective spontaneous emission. *Phys. Rep.* **93**, 301 (1982).
101. Duan, L.-M. & Guo, G.-C. Optimal quantum codes for preventing collective amplitude damping. *Phys. Rev. A* **58**, 3491. https://doi.org/10.1103/PhysRevA.58.3491 (1998).
102. Fortunato, E. M., Viola, L., Hodges, J., Teklemariam, G. & Cory, D. G. Implementation of universal control on a decoherence-free qubit. *New J. Phys.* **4**, 5 (2002).
103. V. N. Premakumar & R. Joynt. Error Mitigation in Quantum Computers subject to Spatially Correlated Noise, arXiv preprint arXiv:1812.07076 (2018).
104. Macchiavello, C. & Palma, G. M. Entanglement-enhanced information transmission over a quantum channel with correlated noise. *Phys. Rev. A* **65**, 050301. https://doi.org/10.1103/PhysRevA.65.050301 (2002).
105. Yeo, Y. & Skeen, A. Time-correlated quantum amplitude-damping channel. *Phys. Rev. A* **67**, 064301. https://doi.org/10.1103/PhysRevA.67.064301 (2003).
106. Terhal, B. M. & Burkard, G. Fault-tolerant quantum computation for local non-Markovian noise. *Phys. Rev. A* **71**, 012336. https://doi.org/10.1103/PhysRevA.71.012336 (2005).
107. Novais, E. & Baranger, H. U. Decoherence by correlated noise and quantum error correction. *Phys. Rev. Lett.* **97**, 040501. https://doi.org/10.1103/PhysRevLett.97.040501 (2006).
108. Aharonov, D., Kitaev, A. & Preskill, J. Fault-tolerant quantum computation with long-range correlated noise. *Phys. Rev. Lett.* **96**, 050504. https://doi.org/10.1103/PhysRevLett.96.050504 (2006).
109. Ban, M., Kitajima, S. & Shibata, F. Decoherence of quantum information in the non-Markovian qubit channel. *J. Phys. A Math. General* **38**, 7161 (2005).
110. Yu, T. & Eberly, J. Entanglement evolution in a non-Markovian environment. *Opt. Commun.* **283**, 676 (2010).

Acknowledgements

Author contributions

Y.H. and H.N. accomplished the theoretical analysis. Y.H. and H.N. developed the codes, performed the numerical simulations, and verify the results. All authors contributed to the manuscript preparation and presentation of results.

Competing Interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-52485-7.

Correspondence and requests for materials should be addressed to Y.H.

Reprints and permissions information is available at www.nature.com/reprints.

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