Quantum Error Mitigation and Its Progress

Suguru Endo

Abstract

Current quantum hardware is significantly affected by computation errors. Error reduction is thus required to obtain meaningful results from quantum computers. Quantum error mitigation (QEM) methods, which are a class of hardware-friendly error-reduction methods not relying on the encoding of quantum information, are being actively researched. In this article, I review the major QEM methods. I then introduce the recent progress in QEM technologies proposed by our research group. First, I review the worlds' first quantum sensing method incorporating QEM. I then review the generalized quantum subspace expansion method, quite a general unified framework of QEM.

Keywords: quantum computing, quantum error mitigation, post-processing of measurement results

1. Quantum error mitigation

A pressing challenge for quantum computers is suppressing the effects of computational errors due to the loss of quantum coherence. Quantum error mitigation (QEM) is a relatively recent concept proposed for mitigating computation errors while keeping the hardware load to a minimum [1]. QEM is often compared with quantum error correction (QEC). In QEC, multiple physical quantum bits (qubits) are used to represent a single logical qubit. This redundancy is used to detect computational errors, and errors are actively corrected on the basis of this information. However, because the number of qubits in quantum hardware is at most several hundred qubits, QEC reduces the effective number of qubits. QEC thus cannot make the best of the computation power of near-term quantum devices. Therefore, QEM was introduced as a set of methods that can reduce computational errors without reducing the effective number of qubits by avoiding the use of redundancy. Progress related to QEM implementation has been remarkable. In a recent paper, IBM claimed the world's first accomplishment of a practical task with a 127-qubit quantum processor [2]. This breakthrough shows that QEM has an extremely useful role. The exponential-extrapolation error mitigation method I proposed [3] also shows extremely high performance. There are a variety of other QEM methods. In QEM, the correct result of a calculation is generally estimated by post-processing the output from multiple quantum circuits using a classical computer. A conceptual diagram of QEM is shown in **Fig. 1(a)**.

QEM cannot fundamentally suppress errors in the quantum state. However, it can mitigate errors in the expectation values of observables. **Figure 1(b)** illustrates the function of QEM. Because many quantum algorithms that are expected to be implemented in current quantum computers and first-generation fault-tolerant quantum computation use the expectation values of observables, QEM is considered to be highly useful. It should be noted that the cost of QEM is an increase in measurement shots; an exponentially greater number of measurements in accordance with the frequency of computational errors in quantum hardware is required. Intuitively, this is because QEM has the effect of amplifying the expectation values of observables that generally decays exponentially with respect to the number of quantum gates and the gate error rate. The variance of calculation results thus increases exponentially. Mathematical proof related to the exponential increase in the number of measurements have been shown in several papers, including

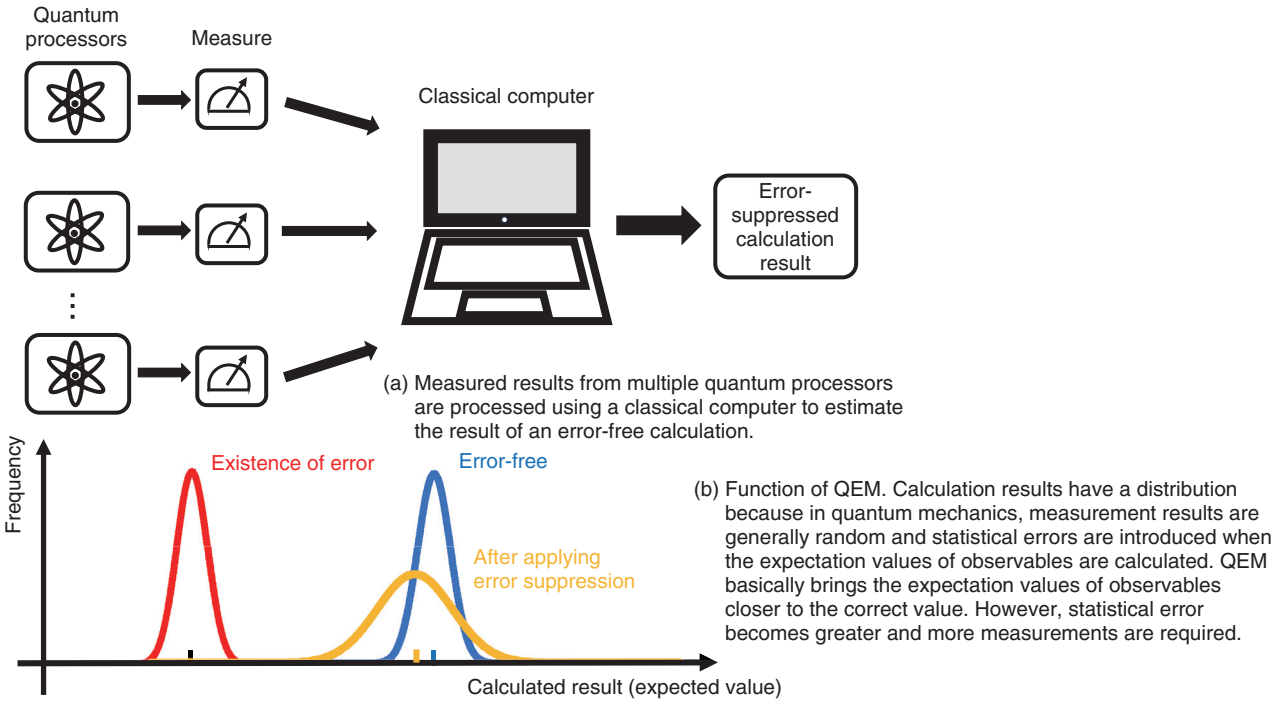


Fig. 1. Conceptual diagram of QEM.

that by our research group [4], using a quantum information-theoretic approach. In the following section, I discuss the major QEM methods: extrapolation [3, 5], quasi-probability (also called probabilistic error cancellation [3, 5]), virtual distillation [6], and subspace expansion [7].

I then describe our research groups' latest achievements, a quantum sensing method incorporating QEM [8] and quite a general unified framework of QEM called the generalized quantum subspace expansion method [9]. For readers who wish to simply have an overview of QEM, it is sufficient to understand the extrapolation methods. For those who wish to learn more, I encourage you to read the other sections. Referring to the review paper I wrote [1] when needed will provide an in-depth understanding of QEM.

1.1 Extrapolation methods

As the name suggests, extrapolation methods estimate the ideal error-free calculation result by extrapolating multiple measurement results [3, 5]. They are simple yet powerful methods used in many experiments. An overview of extrapolation methods is shown in **Fig. 2**. The horizontal axis shows the error rate and the vertical axis shows the result (the expectation value of an observable). Of course, we cannot freely reduce the error rate, but it is relatively easy to increase calculation errors. For example, it is possible to increase the frequency of errors by slowly carrying out gate operations or by carrying out extra gate operations. By extrapolating the original calculation result and calculation results associated with increased error rates, the ideal error-free calculation result is then estimated. When extrapolation methods were first proposed, they used Richardson extrapolation with linear and polynomial functions [5]. Observing that calculation results generally decay exponentially with the frequency of calculation errors, I proposed extrapolation using an exponential function [3]. The exponential extrapolation showed extremely good performance in an actual experiment [2]. However, extrapolation methods cannot guarantee computation accuracy and can be said to be relatively heuristic.

The number of measurements, a cost factor in QEM, can be easily understood with extrapolation methods. Considering linear extrapolation as an example, for calculation error rate ε_0, we express the experimentally obtained average value of the observable as $\langle O(\varepsilon_0)\rangle$ and that with twice the error rate as $\langle O(2\varepsilon_0)\rangle$. From extrapolation, the error-mitigated result can be written as $O_{\mathrm{est}} = 2\langle O(\varepsilon_0)\rangle - \langle O(2\varepsilon_0)\rangle$.

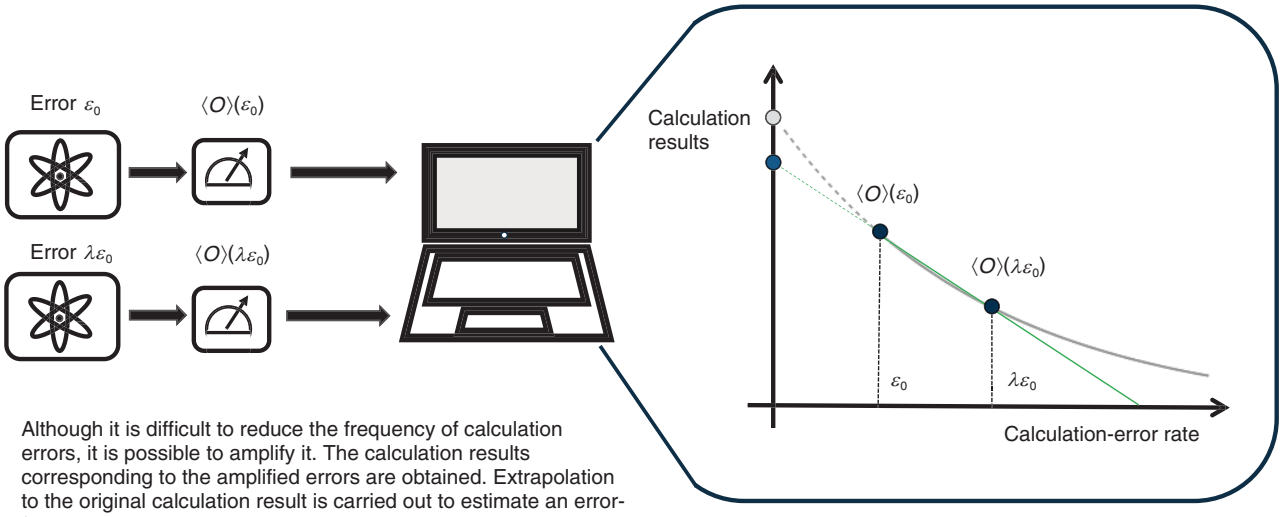


Fig. 2. Overview of extrapolation methods.

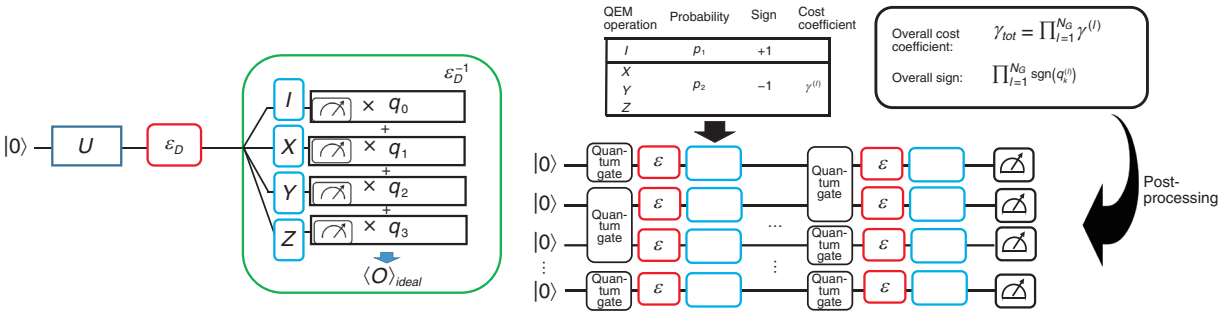
When calculating variance, if no correlation is assumed between $\langle O(\varepsilon_0)\rangle$ and $\langle O(2\varepsilon_0)\rangle$, we obtain $\mathrm{Var}[O_{est}] = 4\mathrm{Var}[\langle O(\varepsilon_0)\rangle] + \mathrm{Var}[\langle O(2\varepsilon_0)\rangle]$. This shows that the variance is amplified after applying QEM and that more measurements are needed to obtain the correct calculation result.

1.2 Quasi-probability method

The quasi-probability method counteracts the effect of gate noise by effectively constructing the inverse of the noise based on the noise model obtained through noise characterization techniques, such as process tomography or gate set tomography [3, 5]. We denote the quantum process corresponding to the noise as $\mathcal{E}$ (it may be easier to think of it as the quantum mechanical version of a transition matrix) and its inverse map as $\mathcal{E}^{-1}$. While $\mathcal{E}^{-1}$ can be mathematically constructed, it is not generally a "physical process" and cannot be directly operated on a quantum computer. By constructing $\mathcal{E}^{-1}$ by a set of operations $\{B_k\}_k$ for QEM that we can execute with fewer calculation errors, we can decompose $\mathcal{E}^{-1}$ as $\mathcal{E}^{-1} = \sum_k q_k B_k$. Usually, we assume that $\{B_k\}_k$ are single qubit operations.

For example, when we consider the depolarizing noise of error probability p to be $\mathcal{E}_D(\rho) = \left(1 - \frac{3}{4}p\right)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z)$, its inverse map is $\mathcal{E}_D^{-1}(\rho) = \left(1 + \frac{3p}{4(1-p)}\right)\rho - \frac{p}{4(1-p)}(X\rho X + Y\rho Y + Z\rho Z)$. Here, $q_0 = \left(1 + \frac{3p}{4(1-p)}\right)$, $q_1 = q_2 = q_3 = \frac{p}{4(1-p)}$, $B_0(\rho) = \rho$, $B_1(\rho) = X\rho X$, $B_2(\rho) = Y\rho Y$, $B_3(\rho) = Z\rho Z$. Because $\sum_k q_k = 1$, and $\mathcal{E}^{-1}$ is generally not a physical process resulting in q_k being a quasi-probability that can be negative, this method is called the quasi-probability method. A negative probability cannot be directly implemented, but the expected value "same as the case of sampling using negative probability" can be effectively calculated by post-processing of the measurement results. Consider a simple 1-qubit system as an example, which is conceptually shown in **Fig. 3(a)**. The ideal quantum state is $\rho_{ideal} = U|0\rangle\langle 0|U^\dagger$. However, because of depolarizing noise $\mathcal{E}_D$, the actual quantum state is $\rho_{noisy} = \mathcal{E}_D(\rho_{ideal})$. Expressing the observable to be measured as O, because the noiseless expectation value is $\langle O_{ideal}\rangle = q_0\mathrm{Tr}[\rho_{noisy}O] + q_1\mathrm{Tr}[X\rho_{noisy}XO] + q_2\mathrm{Tr}[Y\rho_{noisy}YO] + q_3\mathrm{Tr}[Z\rho_{noisy}ZO]$, the expectation value of the observable can be measured by adding together the measurement outcomes of quantum states ρ_{noisy}, $X\rho_{noisy}X$, $Y\rho_{noisy}Y$ and $Z\rho_{noisy}Z$ with the appropriate weight of quasi-probability. Even if a quasi-probability with a negative value exists, a non-physical inverse map can be constructed by multiplying a negative sign to measurement outcomes and performing post-processing.

It is important to implement the inverse map in a quantum circuit with multiple qubits for practical purposes. Consider when a quasi-probability method is applied to the noise $\mathcal{E}_l$ (l = 1,2,…N_G, where N_G is the number of gates) of multiple quantum gates. The conceptual diagram is shown in **Fig. 3(b)**. We



(a) Quasi-probability method for 1-qubit system. The average of the observables measured for the quantum state, with Pauli operator added, is multiplied by the quasi-probability and added to form an inverse map of the overall calculation error.

(b) QEM is executed randomly after the quantum gate containing the computational error. This figure shows an example of depolarizing noise, where the QEM operation becomes a Pauli operation. The measurement result is multiplied by the cost coefficient and sign, and the average of the outcomes gives the error-mitigated result.

Fig. 3. Overview of the quasi-probability method.

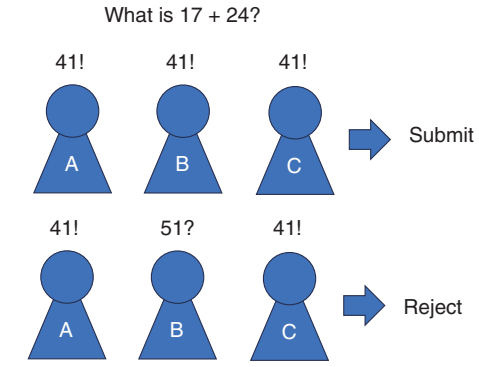
construct an inverse map for each error: $\mathcal{E}_l^{-1} = \sum_k q_k^{(l)} B_k = \gamma^{(l)} \sum_k p_k^{(l)} \mathrm{sgn}\left(q_k^{(l)}\right) B_k$ ($\sum_k p_k^{(l)} = 1$, $p_k^{(l)} > 0$, $\gamma^{(l)} = \sum_k |q_k^{(l)}| > 1$, $\mathrm{sgn}\left(q_k^{(l)}\right) = q_k^{(l)}/|q_k^{(l)}|$, where $\gamma^{(l)}$ is the cost coefficient. After each quantum gate (or before, depending on the formulation), operation B_k is generated with probability $p_k^{(l)}$, and the products of sign $\prod_{l=1}^{N_G} \mathrm{sgn}\left(q_k^{(l)}\right)$ and cost coefficient $\gamma_{tot} = \prod_{l=1}^{N_G} \gamma^{(l)}$ are multiplied to the measurement result. By repeating this, the average of the result gives the error-mitigated result. Because the variance of the calculated result is approximately amplified by γ_{tot}^2 compared with the case without error mitigation, an exponentially large number of measurements according to the number of gates is required.

Although a suitable error-characterization method has not been proposed when the quasi-probability was first proposed, I found that gate-set tomography is an efficient error-characterization method for this method [3]. I also discovered a set of operations $\{B_k\}_k$ for QEM that allows for removal of arbitrary computational errors [3]. I have also shown that the quasi-probability method can be applied not only to gate models but also temporally continuous noise models such as those described by the Lindblad master equation $\frac{d}{dt}\rho = -i[H, \rho] + \sum_k (2 L_k \rho L_k^\dagger - L_k^\dagger L_k \rho - \rho L_k^\dagger L_k)$, extending QEM to analog quantum systems [10].

1.3 Virtual distillation method

Virtual distillation executes QEM by preparing multiple copies of a noisy quantum state ρ_{noisy}, executing entanglement measurements between them and post-processing the results using a classical computer. This enables us to simulate the error-suppressed quantum state as if we distill a noiseless quantum [6]. An example of the "classical" counterpart of this method is as follows: we ask several students to solve the same problem, e.g., elementary school students who often get erroneous results for simple arithmetic problems. Only when all the calculation results are the same is the answer submitted; otherwise, the results are discarded (**Fig. 4**). The more students involved in calculating the result, the higher the percentage of the correct answer. However, the probability of success (i.e., the probability of all students calculating the right answer) decreases exponentially with the number of students.

With virtual distillation, we can calculate the expected value of the physical quantity corresponding to the distilled quantum state $\rho_{vd} = \frac{\rho_{noisy}^n}{\mathrm{Tr}[\rho_{noisy}^n]}$, where n is the number of copies of the noisy quantum state. When we have the spectral decomposition $\rho_{noisy} = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ ($p_0 \geq p_1 \geq \ldots$), it is expected that the eigenstate corresponding to the largest eigenvalue is a good approximation of the ideal quantum state when the noise is small. Now, as n increases, ρ_{vd} asymptotically approaches $|\psi_0\rangle$. The contribution of $|\psi_k\rangle$ ($k = 1, 2, \ldots$) is suppressed exponentially with respect to n. However, the number of required measurements increases exponentially with n. The advantage



An intuitive classical explanation of the virtual distillation method. Calculation accuracy is increased by having several participants solve the same problem. The result is rejected if even one of them produces a different result.

Fig. 4. Overview of virtual distillation method.

of this method is that it can mitigate errors with high accuracy if the errors are stochastic, even without information about the error model. However, coherent errors caused by rotation errors of quantum gates and the insufficient expression capability due to the lack of depth of ansatz quantum circuits in variational quantum eigensolver cannot be mitigated with this method, no matter the increase in the number of copies.

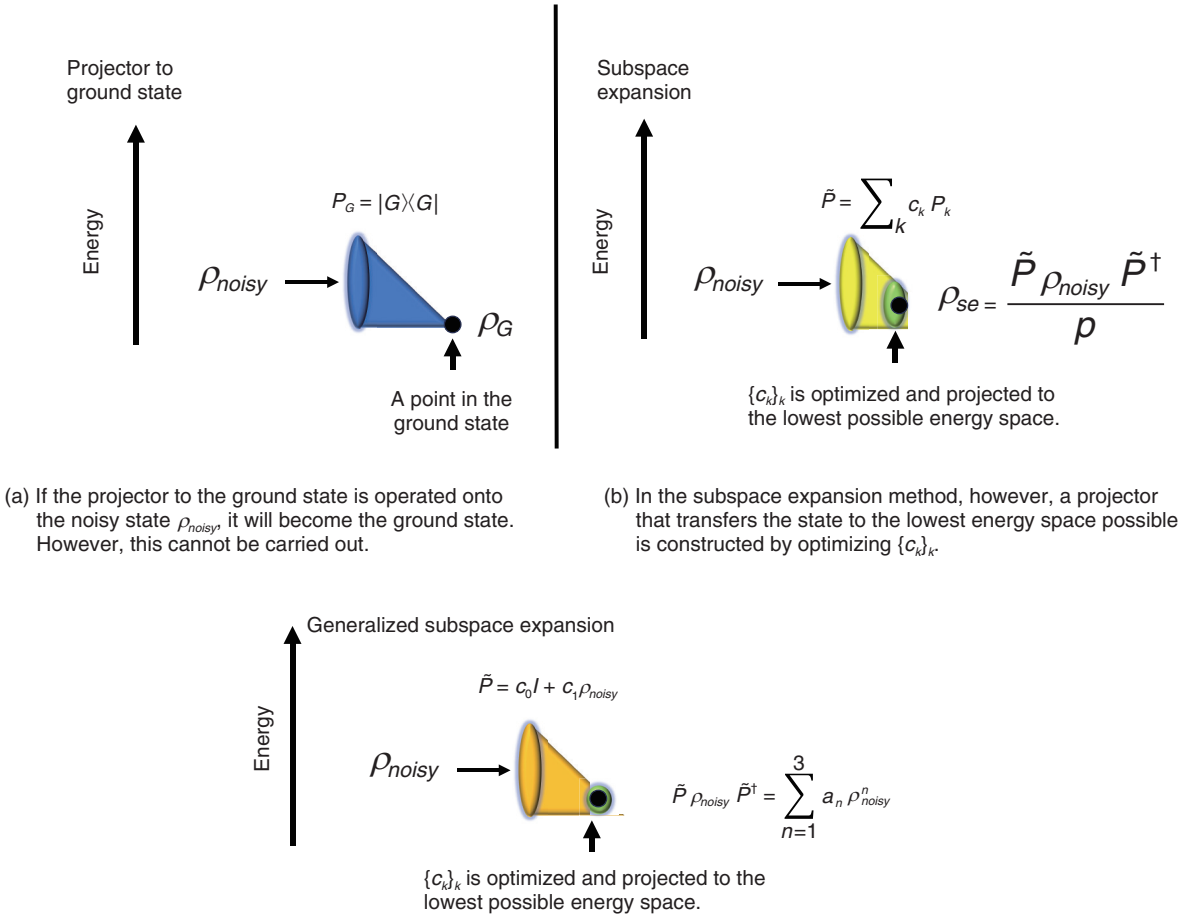
1.4 Subspace expansion method

The subspace expansion method constructs a projector (strictly speaking, this operator does not satisfy the mathematical properties of a projector but is called one here for convenience) [7]. Consider a case in which the actual quantum state immediately before measurement differs from the ideal quantum state because of noise. For example, the variational quantum eigenvalue solver is a method for determining the ground state $\rho_G = |G\rangle\langle G|$ of molecules, etc; however, the actual quantum state may be another quantum state ρ_{noisy} because of errors. If we can construct the projector onto the ground state $P_G = |G\rangle\langle G|$, an error-free quantum state can be obtained as $\frac{P_G \rho_{noisy} P_G}{p_G} = \rho_G$ with p_G being the projection probability (**Fig. 5(a)**). However, because $|G\rangle$ is an extremely large quantum state in reality, the expression of P_G cannot be obtained in the first place, and the projection cannot be executed accurately. We thus seek to construct a projection operator (which strictly speaking, does not satisfy the mathematical properties of a projection operator but called one here for convenience) that can project the noisy state onto a space with the lowest possible energy. Using Pauli operators P_k to express such a projection operator as $\tilde{P} = \sum_k c_k P_k$ (where c_k is a complex number), we optimize $\{c_k\}_k$ using a classical computer so that the energy of the projected quantum state $\rho_{se} = \frac{\tilde{P} \rho_{noisy} \tilde{P}^\dagger}{p}$ (where p is the projection probability) can be minimized (**Fig. 5(b)**). What P_k to choose is arbitrary. Methods for constructing P_k from, for example, excitation operators of a molecule's spin-orbitals, have been proposed [7]. This method can suppress coherent errors to a certain extent but is known to be unsuitable for suppressing stochastic errors such as bit flips.

2. NTT's latest achievements

2.1 Application of QEM to quantum sensing

NTT has developed the world's first framework for quantum sensing incorporating QEM. Quantum sensing is an active area of research in the field of quantum information that uses quantum states to efficiently probe fields such as magnetic fields one wishes to measure. This is done by interacting the quantum states with the field followed by the readout. The process is repeated and the results are accumulated to estimate the value of the magnetic field. What is important about quantum sensing is that when quantum entangled states are used as probes, quantum advantageous scaling can be achieved depending on the number of qubits N. However, if the noise fluctuates at each time of measurement, systematic errors occur in the accumulated value and estimated



(a) If the projector to the ground state is operated onto the noisy state ρ_{noisy}, it will become the ground state. However, this cannot be carried out.

(b) In the subspace expansion method, however, a projector that transfers the state to the lowest energy space possible is constructed by optimizing $\{c_k\}_k$.

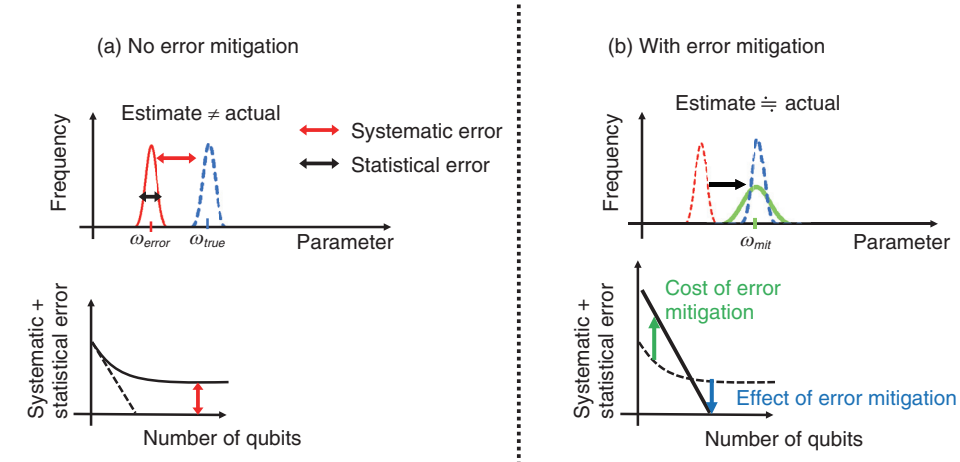
(c) $\tilde{P}$ is extended by not only Pauli operators and their products but also by general operators that include states to construct a projector that transfers the state to the lowest energy space. Thus, extremely general error-mitigated quantum states that completely include the one achieved with virtual distillation and the generalized subspace method can be realized.

Fig. 5. Overview of subspace expansion methods.

magnetic-field value, and quantum advantages cannot be achieved (**Fig. 6(a)**). Our research group has shown that even when noise fluctuates each time the quantum device is executed, virtual distillation can act as a "filter" that removes such noise and accurately mitigates systematic errors [8]. We have also shown that quantum advantageous scaling can be restored (**Fig. 6(b)**).

2.2 Generalized subspace expansion method

Our research group proposed the generalized subspace expansion method, which is quite a general QEM method, which includes subspace expansion and virtual distillation as special cases [9]. I stated above that in subspace expansion, the projection operator $\tilde{P} = \sum_k c_k P_k$ can be optimized so that energy is minimized. The essence of the generalized subspace expansion method is extending P_k to extremely general operators. More specifically, quantum states (and more complex operators that include them) are used as P_k. For example, taking $P_0 = I$, $P_1 = \rho_{noisy}$, the projected state is $\tilde{P}\ \rho_{noisy}\ \tilde{P}^\dagger = |c_0|^2\ \rho_{noisy} + (c_0 c_1^* + c_0^* c_1)\rho_{noisy}^2 + |c_1|^2 \rho_{noisy}^3$, and the expected value of the observables corresponding to an error-mitigated quantum state expanded by a series of powers of a noisy quantum state can be obtained. We call this method the power subspace method (**Fig. 5(c)**). Our research group also proposed the fault-subspace method that uses the essence of extrapolation methods [5] in the construction of projectors. Unification of



When systematic errors exist, increasing the number of qubits in quantum sensing does not achieve quantum advantageous scaling of estimation of parameters of interest.

When QEM is carried out, statistical errors increase due to error mitigation. However, systematic errors can be suppressed, so quantum scaling can be recovered.

Fig. 6. Effect of QEM (virtual distillation) on quantum sensing.

the power subspace method and fault-subspace method is also possible. The generalized subspace expansion method inherits the advantages of both subspace expansion and virtual distillation methods and can mitigate both coherent errors and stochastic errors with high accuracy. Therefore, far more accurate QEM is made possible compared with subspace expansion or virtual distillation alone.

References

[1] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, "Hybrid Quantum-classical Algorithms and Quantum Error Mitigation," J. Phys. Soc. Jpn. 90, No. 3, Article ID: 032001, 2021. https://doi.org/10.7566/JPSJ.90.032001

[2] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, "Evidence for the Utility of Quantum Computing before Fault Tolerance," Nature, Vol. 618, pp. 500–505, June 2023. https://doi.org/10.1038/s41586-023-06096-3

[3] S. Endo, S. C. Benjamin, and Y. Li, "Practical Quantum Error Mitigation for Near-future Applications," Phys. Rev. X, Vol. 8, No. 3, 031027, 2018. https://doi.org/10.1103/PhysRevX.8.031027

[4] R. Takagi, S. Endo, S. Minagawa, and M. Gu, "Fundamental Limits of Quantum Error Mitigation," npj Quantum Information, Vol. 8, No. 1, Article number: 114, 2022. https://doi.org/10.1038/s41534-022-00618-z

[5] K. Temme, S. Bravyi, and J. M. Gambetta, "Error Mitigation for Short-depth Quantum Circuits," Phys. Rev. Lett., Vol. 119, No. 18, 180509, 2017. https://doi.org/10.1103/PhysRevLett.119.180509

[6] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, "Virtual Distillation for Quantum Error Mitigation," Phys. Rev. X, Vol. 11, No. 4, 041036, 2021. https://doi.org/10.1103/PhysRevX.11.041036

[7] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, "Hybrid Quantum-classical Hierarchy for Mitigation of Decoherence and Determination of Excited States," Phys. Rev. A, Vol. 95, No. 4, 042308, 2017. https://doi.org/10.1103/PhysRevA.95.042308

[8] K. Yamamoto, S. Endo, H. Hakoshima, Y. Matsuzaki, and Y. Tokunaga, "Error-mitigated Quantum Metrology via Virtual Purification," Phys. Rev. Lett., Vol. 129, No. 25, 250503, 2022. https://doi.org/10.1103/PhysRevLett.129.250503

[9] N. Yoshioka, H. Hakoshima, Y. Matsuzaki, Y. Tokunaga, Y. Suzuki, and S. Endo, "Generalized Quantum Subspace Expansion," Phys. Rev. Lett., Vol. 129, No. 2, 020502, 2022. https://doi.org/10.1103/PhysRevLett.129.020502

[10] J. Sun, X. Yuan, T. Tsunoda, V. Vedral, S. C. Benjamin, and S. Endo, "Mitigating Realistic Noise in Practical Noisy Intermediate-scale Quantum Devices," Phys. Rev. Appl., Vol. 15, No. 3, 034026, 2021. https://doi.org/10.1103/PhysRevApplied.15.034026

Suguru Endo
Associate Distinguished Researcher, Innovative Computing Architecture Laboratory, NTT Computer and Data Science Laboratories.

He received a B.S. and M.S. in quantum information science from Keio University, Kanagawa, in 2014 and 2016 and a Ph.D. in near-term quantum computing from University of Oxford, UK, in 2019. He has been working as a researcher at NTT since 2020. His research interests are hybrid quantum-classical quantum algorithms, quantum error mitigation, early fault-tolerant quantum computing, and bosonic code quantum computing. He was awarded the MIT Technology Review Innovators Under 35 Japan in 2021 and Project Management Institute Future 50 in 2022.