Feed-Forward Probabilistic Error Cancellation with Noisy Recovery Gates

Leo Kurosawa
Graduate School of Computer Science and Engineering
University of Aizu
Aizu-Wakamatsu, Fukushima, Japan
m5271018@u-aizu.ac.jp

Yoshiyuki Saito
Graduate School of Computer Science and Engineering
University of Aizu
Aizu-Wakamatsu, Fukushima, Japan
d8241104@u-aizu.ac.jp

Xinwei Lee
School of Computing and Information Systems
Singapore Management University
80 Stamford Rd, Singapore
xwlee@smu.edu.sg

Xinjian Yan
Graduate School of Science and Technology
University of Tsukuba
Tsukuba, Ibaraki, Japan
yanxinjian@cavelab.cs.tsukuba.ac.jp

Ningyi Xie
Graduate School of Science and Technology
University of Tsukuba
Tsukuba, Ibaraki, Japan
nyxie@cavelab.cs.tsukuba.ac.jp

Dongsheng Cai
Faculty of Engineering, Information and Systems
University of Tsukuba
Tsukuba, Ibaraki, Japan
cai@cs.tsukuba.ac.jp

Nobuyoshi Asai
School of Computer Science and Engineering
University of Aizu
Aizu-Wakamatsu, Fukushima, Japan
nasai@u-aizu.ac.jp

Abstract—**Probabilistic Error Cancellation (PEC) aims to improve the accuracy of expectation values for observables. This is accomplished using the probabilistic insertion of recovery gates, which correspond to the inverse of errors. However, the inserted recovery gates also induce errors. Thus, it is difficult to obtain accurate expectation values with PEC since the estimator of PEC has a bias due to noise induced by recovery gates. To address this challenge, we propose an improved version of PEC that considers the noise resulting from gate insertion, called Feed-Forward PEC (FFPEC). FFPEC provides an unbiased estimator of expectation values by cancelling out the noise induced by recovery gates. We demonstrate that FFPEC yields more accurate expectation values compared to conventional PEC method through numerical simulations with bit-flip and depolarizing noises.**

Index Terms—**Quantum Error Mitigation, Probabilistic Error Cancellation, Quantum Algorithm**

I. INTRODUCTION

Fascinated by the possibilities of high-speed computing with quantum computers, many quantum algorithms have been proposed to tackle problems that require a lot of calculation time even in classical computers, such as quantum chemical calculations [1], [2] and combinatorial optimization problems [3], [4]. These algorithms require expectation values of some observables to obtain a solution. However, since a Noisy Intermediate-Scale Quantum (NISQ) computer [5] is not equipped with error correction systems for noises, the accuracy of an expectation value obtained from measurement outcomes of a quantum computer is inevitably degraded by the noises. To obtain a meaningful result from a NISQ computer, quantum error mitigation (QEM) techniques without additional qubits have been proposed to suppress the effects of noises [6]–[19].

Probabilistic Error Cancellation (PEC) is one of the most well-known QEM techniques used to reduce noises caused by quantum gate operations [6], [8], [13], [14], [16]. The key idea of PEC is to cancel out noise Λ induced by a quantum gate operation by adding an operation corresponding to the inverse of the noise. This operation to cancel the noise is represented as a linear combination of quantum gates. To cancel out the noise in this way, PEC requires a full characterization of the noise such as its probability to occur and its inverse. If we have full knowledge of the noise, it is possible to obtain an unbiased estimator for the expectation value of an ideal quantum circuit, through statistical post-processing of measurement outcomes.

Using PEC, an unbiased estimator for an ideal expectation value can be obtained with a variance of $O(\gamma^2)$, where $\gamma \geq 1$

is a factor that increases exponentially with the number of quantum gates in a quantum circuit. To obtain an expectation value with a high accuracy, a large number of shots is essential to reduce the variance. So, this γ is called the *sampling overhead*. From a practical standpoint, it is important to reduce the value of γ. Some methods have been proposed to minimize the sampling overhead γ by decomposing a quantum circuit with a set of gate operations that a quantum computer supports [20], [21].

The characterization of noise and the reduction of sampling overhead described above are necessary to obtain more accurate expectation values with PEC. However, in practical, there is a limit to the accuracy of expectation values that can be obtained even if the noise can be well-characterized because recovery gates (gates inserted to cancel noise) induce noise again. Thus, the PEC estimator has a bias due to noise induced by applying the recovery gate [22]. Since recovery gates are also run on noisy quantum computers, it is difficult to reduce this bias to zero. Exploring methods to reduce the bias of the PEC estimator to zero is necessary to obtain an accurate expectation value, while reluctantly accepting noises caused by recovery gates.

In this paper, we study a method to achieve zero bias in the PEC estimator even with noises induced by recovery gate applied. We propose an error mitigation method, named the Feed-Forward PEC (FFPEC), which addresses noises induced by recovery gate operations in PEC process to reduce the bias to zero. The key idea of FFPEC is to cancel out noises induced by the recovery gates by defining an inverse map of noise including these noises. Although the sampling cost γ of FFPEC is slightly larger than that of the standard PEC, the expectation values for the FFPEC estimator is expected to be closer to an exact expectation value than for the standard PEC.

The present paper is organized as follows. First, we give a background for standard PEC technique and two well-known noise models, bit-flip error and depolarizing error in Sec. II. In Sec. III, we present the Feed-Forward PEC. Then, numerical experiments are conducted to evaluate the performance of FFPEC by comparing it with the standard PEC in Sec. IV. Finally, we conclude about FFPEC in Sec. V.

II. BACKGROUND

We first briefly introduce the Probabilistic Error Cancellation (PEC) [6], [23]. We also introduce two commonly used noise models, bit-flip and depolarizing, and their inverses for PEC.

A. Probabilistic Error Cancellation

We first introduce the notations required to explain the PEC procedure. We consider an n qubits system. Let $\mathcal{Q}$ denote a unitary trace-preserving completely positive (TPCP) map of an ideal quantum circuit on the n qubits system, ρ_{in} denote the input quantum state to this circuit $\mathcal{Q}$ and ρ_{out} denote the output quantum state after applying the circuit $\mathcal{Q}$ to the input state ρ_{in} with a noisy quantum computer. The circuit $\mathcal{Q}$ can be represented as $\mathcal{Q}(\rho) = Q\rho Q^\dagger$ with a corresponding unitary

matrix Q for any quantum state ρ. In addition, the quantum circuit $\mathcal{Q}$ consists of N_G ideal quantum gates as follows:

$$\mathcal{Q} := \mathcal{U}_{N_G} \circ \mathcal{U}_{N_G-1} \circ \cdots \mathcal{U}_2 \circ \mathcal{U}_1, \tag{1}$$

where $\circ$ represents a composition of TPCP maps, and each $\mathcal{U}_i$ is a unitary TPCP map corresponding to a quantum gate U_i. Suppose that a quantum circuit $\mathcal{Q}$ is executed on a noisy quantum computer and that each noise Λ_i is induced for each gate $\mathcal{U}_i$. So, the output quantum state ρ_{out} is written as:

$$\rho_{\mathrm{out}} = \Lambda_{N_G} \circ \mathcal{U}_{N_G} \circ \Lambda_{N_G-1} \circ \mathcal{U}_{N_G-1} \cdots \Lambda_1 \circ \mathcal{U}_1(\rho_{\mathrm{in}}). \tag{2}$$

We denote an ideal output quantum state as $\rho_{\mathrm{out}}^{\mathrm{ideal}} := \mathcal{Q}(\rho_{\mathrm{in}})$, then an ideal expectation value of an observable M is written as $\langle M \rangle_{\mathrm{ideal}} := \mathrm{Tr}\big[M \rho_{\mathrm{out}}^{\mathrm{ideal}}\big]$.

The objective of PEC is to provide an expectation value with a good accuracy while executing a quantum circuit on a noisy quantum computer. The key idea of PEC is to cancel out noises Λ_i which occurs after U_i executed by adding its inverse Λ_i^{-1}. By cancelling out noises in such a way, PEC can provide an expectation value $\langle M \rangle_{\mathcal{Q}}$ via statistical post-processing. To simplify, we describe the PEC process using a single quantum gate U. After that, the PEC process is explained for a general quantum circuit. We assume that there exists a gate set $\{\mathcal{R}_i\}$ that is implemented on a quantum computer so that the inverse map of noise Λ^{-1} can be represented as a linear combination with elements of this set as a basis:

$$\Lambda^{-1} = \sum_i q_i \mathcal{R}_i, \tag{3}$$

where each q_i is a real number (pseudo-probability) and satisfies $\sum_i q_i = 1$ [6]. Then we have

$$\mathcal{U} = \Lambda^{-1}\Lambda\mathcal{U} = \sum_i q_i \mathcal{R}_i \Lambda\mathcal{U} = \gamma \sum_i \sigma_i \mathrm{sgn}(q_i)\mathcal{R}_i\Lambda\mathcal{U}, \tag{4}$$

with $\gamma = \sum_i |q_i|$, $\sigma_i = |q_i|/\gamma$. Factor γ is the sampling overhead, and each σ_i is the insertion probability of gate $\mathcal{R}_i$, and $\mathrm{sgn}(q_i) = \pm 1$ means a parity corresponding to gate $\mathcal{R}_i$.

Using γ, the expectation value of observable M can be expressed as follows:

$$\langle M \rangle_{\mathcal{U}} = \gamma \sum_i \sigma_i \left\langle \mu_i^{\mathrm{eff}} \right\rangle, \tag{5}$$

where $\mu_i^{\mathrm{eff}} = \mathrm{sgn}(q_i)m_i$ and m_i is the measurement outcome of the state $\mathcal{R}_i\Lambda\mathcal{U}(\rho_{\mathrm{in}})$, which is generated by applying the gate $\mathcal{R}_i$ with probability σ_i after applying gate $\mathcal{U}_i$. Multiplying the outcome μ_i by the parity $\mathrm{sgn}(q_i)$ corresponding to $\mathcal{R}_i$ gives μ_i^{eff}. The expectation value of the random variable $\gamma\mu^{\mathrm{eff}}$ is approximately an unbiased estimator of $\langle M \rangle_{\mathcal{U}}$.

To obtain an expectation value $\langle M \rangle_{\mathcal{Q}}$ of an observable M with a quantum circuit $\mathcal{Q}$, we apply the process stated above for $\mathcal{Q}$. Applying the above process to the quantum circuit $\mathcal{Q} = \prod_{k=1}^{N_G} \mathcal{U}_k$, we have

$$\prod_{k=1}^{N_G} \mathcal{U}_k = \prod_{k=1}^{N_G} \gamma^{(k)} \sum_{i_1,\cdots,i_{N_G}} \prod_{k=1}^{N_G} \mathrm{sgn}(q_{i_k}) \prod_{k=1}^{N_G} \sigma_{i_k} \prod_{k=1}^{N_G} \mathcal{R}_{i_k}\Lambda_k\mathcal{U}_k, \tag{6}$$

with $\mathcal{U}_k = \gamma^{(k)} \sum_{i_k} \mathrm{sgn}(q_{i_k})\sigma_{i_k}\mathcal{R}_{i_k}\Lambda_k\mathcal{U}_k$. We can obtain a random variable μ^{eff} by executing a circuit modified by operation $\mathcal{R}_{i_k}$ added with probability σ_{i_k} and multiplying sign $\prod_{k=1}^{N_G} \mathrm{sgn}(q_{i_k})$ to measurement outcome. The expectation value $\langle \mu^{\mathrm{eff}} \rangle$ is obtained by repeating this process, and finally we obtain the expectation value $\langle M \rangle_{\mathcal{Q}} = \gamma^{\mathrm{tot}} \langle \mu^{\mathrm{eff}} \rangle$, where $\gamma^{\mathrm{tot}} = \prod_{k=1}^{N_G} \gamma^{(k)}$. The estimator $\langle M \rangle_{\mathcal{Q}}$ is unbiased for the exact expectation value $\langle M \rangle_{\mathrm{ideal}}$.

B. Bit-flip and Depolarizing noises

We introduce two noise models, bit-flip and depolarizing [24], which are used in our experiments. To explain these noises, we denote $\mathcal{I}, \mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ as trace-preserving completely positive (TPCP) maps corresponding to the pauli matrices I, X, Y, and Z, respectively. Similarly, $\mathcal{II}, \mathcal{IX}, \ldots, \mathcal{ZZ}$ also represent TPCP maps corresponding to couples of pauli matrices $II, IX, \ldots, ZZ$. In addition, single qubit state and two qubits state denote as ρ and ρ', respectively. The error rate is p for both noises.

1) Bit-flip for single qubit state: Let $\mathcal{B}_1$ denote a TPCP map of the bit-flip noise on a single qubit state ρ as follows:

$$\mathcal{B}_1(\rho) = (1-p)\mathcal{I}(\rho) + p\mathcal{X}(\rho). \quad (7)$$

The inverse map of $\mathcal{B}_1$ is written as follows [25]:

$$\mathcal{B}_1^{-1}(\rho) = (1-q)\mathcal{I}(\rho) + q\mathcal{X}(\rho), \quad (8)$$

where

$$q = \frac{-p}{1-2p}. \quad (9)$$

The sampling cost is $\gamma = 1/(1-2p)$, which gives that the insertion probability of pauli X gate is p.

2) Bit-flip for two qubits state: Let $\mathcal{B}_2$ denote a TPCP map of the bit-flip noise on two qubits state ρ' as follows:

$$\mathcal{B}_2(\rho') = (1-p)^2\mathcal{II}(\rho') + (1-p)p\mathcal{IX}(\rho') \\ + p(1-p)\mathcal{XI}(\rho') + p^2\mathcal{XX}(\rho'). \quad (10)$$

The inverse map of $\mathcal{B}_2$ is written as follows:

$$\mathcal{B}_2^{-1}(\rho') = (1-q)^2\mathcal{II}(\rho') + (1-q)q\mathcal{IX}(\rho') \\ + q(1-q)\mathcal{XI}(\rho') + q^2\mathcal{XX}(\rho'), \quad (11)$$

where

$$q = \frac{-p}{1-2p}, \quad (12)$$

which is obtained by solving equation $\mathcal{B}_2^{-1}(\mathcal{B}_2(\rho')) = \rho'$. Note that there are two solutions for q since this equation is a quadratic for q, thus the solution q [Eq. (12)] that minimizes the sampling cost must be chosen [6]. The sampling cost is $\gamma = 1/(1-2p)^2$, which gives that the insertion probabilities of pauli gates IX, XI, and XX are $(1-p)p$, $p(1-p)$, and p^2, respectively.

3) Depolarizing for single qubit state: Let $\mathcal{D}_1$ denote a TPCP map of the depolarizing noise on a single qubit state ρ as follows:

$$\mathcal{D}_1(\rho) = \left(1 - \frac{3p}{4}\right)\mathcal{I}(\rho) + \frac{p}{4}(\mathcal{X}(\rho) + \mathcal{Y}(\rho) + \mathcal{Z}(\rho)). \quad (13)$$

The inverse map of $\mathcal{D}_1$ is written as follows [25]:

$$\mathcal{D}_1^{-1}(\rho) = \left(1 - \frac{3q}{4}\right)\mathcal{I}(\rho) + \frac{q}{4}(\mathcal{X}(\rho) + \mathcal{Y}(\rho) + \mathcal{Z}(\rho)), \quad (14)$$

where

$$q = \frac{-p}{1-p}. \quad (15)$$

The sampling cost is $\gamma = (1+p/2)/(1-p)$, which gives that the insertion probabilities of pauli gates X, Y, and Z are all the same value $p/(4+2p)$.

4) Depolarizing for two qubits state: Let $\mathcal{D}_2$ denote a TPCP map of the depolarizing noise on two qubits state ρ' as follows:

$$\mathcal{D}_2(\rho') = \left(1 - \frac{15p}{16}\right)\mathcal{II}(\rho') + \frac{p}{16}(\mathcal{IX}(\rho') + \cdots + \mathcal{ZZ}(\rho')). \quad (16)$$

The inverse map of $\mathcal{D}_2$ is written as follows [25]:

$$\mathcal{D}_2^{-1}(\rho') = \left(1 - \frac{15q}{16}\right)\mathcal{II}(\rho') \\ + \frac{q}{16}(\mathcal{IX}(\rho') + \mathcal{IY}(\rho') + \cdots + \mathcal{ZZ}(\rho')), \quad (17)$$

where

$$q = \frac{-p}{1-p}. \quad (18)$$

The sampling cost is $\gamma = (1+7p/8)/(1-p)$, which gives that the insertion probabilities of pauli gates $IX, IY, \ldots$, and ZZ are all the same value $p/(16+14p)$.

III. FEED-FORWARD PEC

We introduce Feed-Forward Probabilistic Error Cancellation (FFPEC) to obtain a more accurate expectation value by defining a new inverse map of noise taking a noise induced by recovery gate insertion. The calculation process of FFPEC is the same as PEC calculation process with the new inverse map instead of a conventional inverse map of noise. In the following, we give the idea of FFPEC how to determine a new inverse map of noise with two well-known noise models, bit-flip and depolarizing. Moreover, we show insertion probabilities of recovery gates and sampling costs for bit-flip and depolarizing for FFPEC. Here, we assume that the noise caused by the gate operation and the probability p of its occurrence are given and small $p \ll 0.5$. Since recovery gate added is also gate operation, the probability of noise induced by the recovery gate is required to be the same as p.

A. Bit-Flip for single qubit state

We present how to determine our inverse map of bit-flip on a single qubit state for FFPEC. In conventional PEC, the inverse map of bit-flip $\mathcal{B}_1^{-1}$ [Eq. (8)] is used to cancel out the bit-flip noise. To cancel out the noise in PEC, pauli X gate is inserted after a single qubit gate operation with the insertion probability $\sigma = p$. FFPEC employs a modified inverse map by taking the bit-flip noise on pauli X gate inserted as a recovery gate as shown in Fig. 1. So, we define a noisy version of $\mathcal{B}_1^{-1}$ as follows:

$$\tilde{\mathcal{B}}_1^{-1}(\rho) = (1 - q)\mathcal{I}(\rho) + q\tilde{\mathcal{X}}(\rho) \tag{19}$$
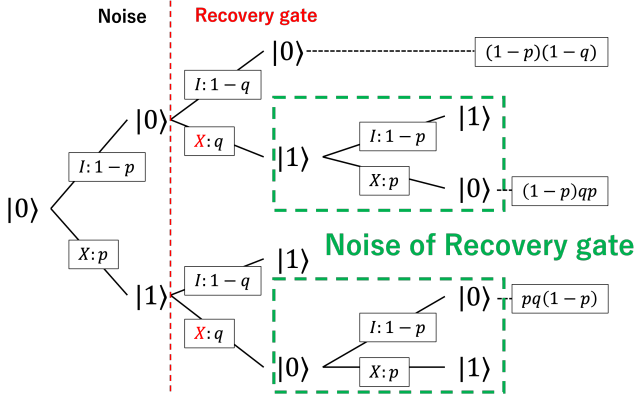$$= [(1 - q) + qp]\mathcal{I}(\rho) + q(1 - q)\mathcal{X}(\rho), \tag{20}$$

where q is a real number and $\tilde{\mathcal{X}}$ is a noisy version of $\mathcal{X}$ for bit-flip written as follows:

$$\tilde{\mathcal{X}}(\rho) = \mathcal{B}_1(\mathcal{X}(\rho)) = (1 - p)\mathcal{X}(\rho) + p\mathcal{I}(\rho). \tag{21}$$

For FFPEC, the parameter q is essential and we can obtain the q by solving equation $\tilde{\mathcal{B}}_1^{-1}(\mathcal{B}_1(\rho)) = \rho$, then we have

$$q = \frac{-p}{1 - 3p + 2p^2} = \frac{-p}{(1 - 2p)(1 - p)}. \tag{22}$$

The sampling cost is $\gamma = (1 - p + 2p^2)/[(1 - 2p)(1 - p)]$ and the insertion probability of pauli X gate is $p/(1 - p + 2p^2)$.

Fig. 1: Tree diagram with a noisy recovery gate in a single qubit bit-flip. FFPEC finds q taking into account noise due to the recovery gate.

B. Bit-Flip for two qubits state

A modified version of inverse of bit-flip for two qubits state is also defined by the same consideration as for single qubit state, stated the above section. Let $\tilde{\mathcal{B}}_2^{-1}$ denote a noisy version of $\mathcal{B}_2^{-1}$ as follows:

$$\tilde{\mathcal{B}}_2^{-1}(\rho') = (1 - q)^2 \mathcal{I}\mathcal{I}(\rho') + (1 - q)q\mathcal{I}\tilde{\mathcal{X}}(\rho') + q(1 - q)\tilde{\mathcal{X}}\mathcal{I}(\rho') + q^2\tilde{\mathcal{X}}\tilde{\mathcal{X}}(\rho'), \tag{23}$$

where q is a real number, and $\mathcal{I}\tilde{\mathcal{X}}, \tilde{\mathcal{X}}\mathcal{I}$, and $\tilde{\mathcal{X}}\tilde{\mathcal{X}}$ are noisy version of $\mathcal{I}\mathcal{X}, \mathcal{X}\mathcal{I}$, and $\mathcal{X}\mathcal{X}$ as follows:

$$\mathcal{I}\tilde{\mathcal{X}}(\rho') = (1 - p)\mathcal{I}\mathcal{X}(\rho') + p\mathcal{I}\mathcal{I}(\rho'), \tag{24}$$
$$\tilde{\mathcal{X}}\mathcal{I}(\rho') = (1 - p)\mathcal{X}\mathcal{I}(\rho') + p\mathcal{I}\mathcal{I}(\rho'), \tag{25}$$
$$\tilde{\mathcal{X}}\tilde{\mathcal{X}}(\rho') = (1 - p)^2 \mathcal{X}\mathcal{X}(\rho') + (1 - p)p\mathcal{X}\mathcal{I}(\rho') + p(1 - p)\mathcal{I}\mathcal{X}(\rho') + p^2\mathcal{I}\mathcal{I}(\rho'). \tag{26}$$

Substituting these three maps into Eq. (23), we have the inverse $\tilde{\mathcal{B}}_2^{-1}$:

$$\tilde{\mathcal{B}}_2^{-1}(\rho') = [(1 - q)^2 + (1 - q)pq + q(1 - q)p + p^2q^2]\mathcal{I}\mathcal{I}(\rho') + [(1 - q)q(1 - p) + q^2p(1 - p)]\mathcal{I}\mathcal{X}(\rho') + [q(1 - q)(1 - p) + q^2(1 - p)p]\mathcal{X}\mathcal{I}(\rho') + [q^2(1 - p)^2]\mathcal{X}\mathcal{X}(\rho'). \tag{27}$$

We can obtain the q by solving equation $\tilde{\mathcal{B}}_2^{-1}(\mathcal{B}_2(\rho')) = \rho'$, then we have

$$q = \frac{-p}{(1 - 2p)(1 - p)}. \tag{28}$$

Note that there are two solutions for q since the equation is a quadratic for q, thus the solution q [Eq. (28)] that minimizes the sampling cost must be chosen. The sampling cost is $\gamma = (1 - p + 2p^2)^2/[(1 - 2p)(1 - p)]^2$. The insertion probabilities of pauli gates IX, XI, and XX are $(1 - 2p + 2p^2)p/(1 - p + 2p^2)^2$, $p(1 - 2p + 2p^2)/(1 - p + 2p^2)^2$, and $p^2/(1 - p + 2p^2)^2$, respectively. In practical, the identity I included in two recovery gates IX and XI is not inserted. Only X gate is inserted.

C. Depolarizing for single qubit state

We explain how to determine an inverse map of depolarizing on a single qubit state for FFPEC. Let $\tilde{\mathcal{D}}_1^{-1}$ denote a noisy version of $\mathcal{D}_1^{-1}$ as follows:

$$\tilde{\mathcal{D}}_1^{-1}(\rho) = \left(1 - \frac{3q}{4}\right)\mathcal{I}(\rho) + \frac{q}{4}(\tilde{\mathcal{X}}(\rho) + \tilde{\mathcal{Y}}(\rho) + \tilde{\mathcal{Z}}(\rho)), \tag{29}$$

where $\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}$, and $\tilde{\mathcal{Z}}$ are noisy version of $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ for depolarizing, respectively. These are represented as follows:

$$\tilde{\mathcal{X}}(\rho) = \left(1 - \frac{3p}{4}\right)\mathcal{X}(\rho) + \frac{p}{4}(\mathcal{I}(\rho) + \mathcal{Z}(\rho) + \mathcal{Y}(\rho)), \tag{30}$$
$$\tilde{\mathcal{Y}}(\rho) = \left(1 - \frac{3p}{4}\right)\mathcal{Y}(\rho) + \frac{p}{4}(\mathcal{Z}(\rho) + \mathcal{I}(\rho) + \mathcal{X}(\rho)), \tag{31}$$
$$\tilde{\mathcal{Z}}(\rho) = \left(1 - \frac{3p}{4}\right)\mathcal{Z}(\rho) + \frac{p}{4}(\mathcal{Y}(\rho) + \mathcal{X}(\rho) + \mathcal{I}(\rho)). \tag{32}$$

To obtain the parameter q for FFPEC, equation $\tilde{\mathcal{D}}_1^{-1}(\mathcal{D}_1(\rho)) = \rho$ is solved. Then we have

$$q = \frac{-4p}{4 - 5p + p^2} = \frac{-4p}{(1 - p)(4 - p)}. \tag{33}$$

The sampling cost is $\gamma = (4 + p + p^2)/[(1 - p)(4 - p)]$ and the insertion probabilities of pauli gates X, Y, and Z are all the same value $4p/(4 + p + p^2)$.

D. Depolarizing for two qubis state

A modified inverse of depolarizing for two qubits is also determined by a similar way as for single qubit state explained the previous section. Let $\tilde{\mathcal{D}}_2^{-1}$ denote a noisy version of $\mathcal{D}_2^{-1}$ as follows:

$$
\tilde{\mathcal{D}}_2^{-1}(\rho') = \left(1 - \frac{15q}{16}\right)\mathcal{II}(\rho')
$$
$$
+ \frac{q}{16}(\widetilde{\mathcal{IX}}(\rho') + \widetilde{\mathcal{IY}}(\rho') + \cdots + \widetilde{\mathcal{ZZ}}(\rho')), \tag{34}
$$

where $\widetilde{\mathcal{IX}}, \widetilde{\mathcal{IY}}, \ldots, \widetilde{\mathcal{ZZ}}$ are noisy versions of Pauli maps corresponding to $\mathcal{IX}, \mathcal{IY}, \ldots, \mathcal{ZZ}$. Each $\widetilde{\mathcal{IX}}, \widetilde{\mathcal{IY}}, \ldots, \widetilde{\mathcal{ZZ}}$ is represented as follows:

$$
\widetilde{\mathcal{IX}}(\rho') = \left(1 - \frac{15p}{16}\right)\mathcal{IX}(\rho') + \frac{p}{16}\sum_{\mathcal{P}\in\{\mathcal{I},\mathcal{X},\mathcal{Y},\mathcal{Z}\}^{\otimes 2}\setminus\{\mathcal{IX}\}}\mathcal{P}(\rho'), \tag{35}
$$

$$
\widetilde{\mathcal{IY}}(\rho') = \left(1 - \frac{15p}{16}\right)\mathcal{IY}(\rho') + \frac{p}{16}\sum_{\mathcal{P}\in\{\mathcal{I},\mathcal{X},\mathcal{Y},\mathcal{Z}\}^{\otimes 2}\setminus\{\mathcal{IY}\}}\mathcal{P}(\rho'), \tag{36}
$$

$$
\vdots
$$

$$
\widetilde{\mathcal{ZZ}}(\rho') = \left(1 - \frac{15p}{16}\right)\mathcal{ZZ}(\rho') + \frac{p}{16}\sum_{\mathcal{P}\in\{\mathcal{I},\mathcal{X},\mathcal{Y},\mathcal{Z}\}^{\otimes 2}\setminus\{\mathcal{ZZ}\}}\mathcal{P}(\rho'). \tag{37}
$$

To obtain the parameter q, equation $\tilde{\mathcal{D}}_2^{-1}(\mathcal{D}_2(\rho')) = \rho'$ is solved. Then we have

$$
q = \frac{-16p}{16 - 17p + p^2} = \frac{-16p}{(1-p)(16-p)}. \tag{38}
$$

The sampling cost is $\gamma = (16+13p+p^2)/[(1-p)(16-p)]$ and the insertion probabilities of pauli gates $IX, \ldots,$ and ZZ are all the same value $p/(16+13p+p^2)$. The identity I included in $IX, IY, \ldots,$ and ZI gates is inserted as a recovery gate.

IV. NUMERICAL EXPERIMENTS

We show some numerical experiments to investigate the accuracy of expectation values obtained by FFPEC. To run noisy simulations, we use Qiskit [26].

A. Problem Settings

Bit-flip and depolarizing noise models are used. The error rate for both noises is $p = 0.01$. Table I summarizes the insertion probabilities of recovery gates for PEC and FFPEC at $p = 0.01$. In the case of bit-flip noise, the insertion probabilities of recovery gates for FFPEC are relatively about 1% higher than that for PEC for both single qubit and two qubits gates. In the case of depolarizing noise, on the other hand, the insertion probabilities of recovery gates for FFPEC are slightly higher than those of PEC. It is about 0.25% for single qubit gate and by about 0.06% for two qubits gate. To investigate these effects of the insertion probabilities on the accuracy of expectation values, the four experiments

listed below are addressed for each noise, using three circuits consisting of basic quantum gates as shown in Fig. 2:

(a) investigation of the estimated accuracy for the single qubit gate with circuit 2(a),

(b) investigation of the estimated accuracy for the two qubits gate with circuit 2(b),

(c) investigation of the estimated accuracy of expectation values for a quantum circuit with mixed single qubit and two qubits gates with circuit 2(c), and

(d) investigating the relationship between the number of quantum gates and the accuracy with circuit 2(c).

In the experiments, initial state $|0\rangle^{\otimes 6}$ and observable $ZZZZZZ$ are used. The exact expectation value of the observable in this case is either 1 or -1. For each listed experiment, the expectation values are obtained for different number of measurements $2000, 4000, 6000, 8000,$ and 10000. For each number of measurements, 4000 samplings are taken to evaluate the accuracies, averages, and standard deviations. The accuracies of PEC and FFPEC estimators are also evaluated by the absolute error between this average and the exact solution.

TABLE I: Total insertion probability of recovery gates for one noisy gate with $p = 0.01$.

Noise	Type	PEC	FFPEC	Relative difference
Bit-Flip	single qubit gate	0.010000	0.010099	0.009897
	two qubits gate	0.019896	0.020096	0.009847
Depolarizing	single qubit gate	0.007463	0.007481	0.002469
	two qubits gate	0.009294	0.009299	0.000614

TABLE II: Sampling overheads with $p = 0.01$.

| Noise	Type	PEC	FFPEC
Bit-Flip	γ for single qubit gate	1.020408	1.020614
	γ for two qubits gate	1.041233	1.041654
	γ^{tot} of (a)	3.360744	3.401724
	γ^{tot} of (b)	3.360744	3.401724
	γ^{tot} of (c)	3.360744	3.401724
Depolarizing	γ for single qubit gate	1.015152	1.015189
	γ for two qubits gate	1.018939	1.018951
	γ^{tot} of (a)	2.465199	2.470738
	γ^{tot} of (b)	1.755701	1.756314
	γ^{tot} of (c)	2.080421	2.083121

B. Numerical results and Discussions

The sampling costs of each circuit for FFPEC in bit-flip are 0.04 larger than that for PEC, summarized in Table II. In depolarizing, the sampling costs for FFPEC are slightly larger than that for PEC but are matched to 1 decimal place for the both circuits 2(a) and 2(c), and to 2 decimal places for the circuit 2(b). Thus, the standard deviation of expectation values obtained from the same number of measurements is almost the same as shown in the left panels of Fig. 3 for the result of bit-flip, and Fig. 4 for the result of depolarizing, respectively.
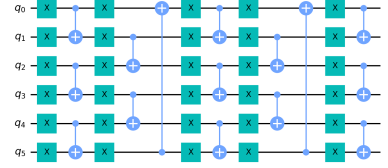
We focus on the averages of expectation values. In the bit-flip results (right side of Fig. 3), the average absolute
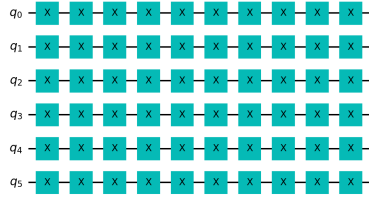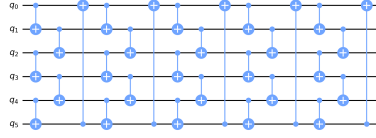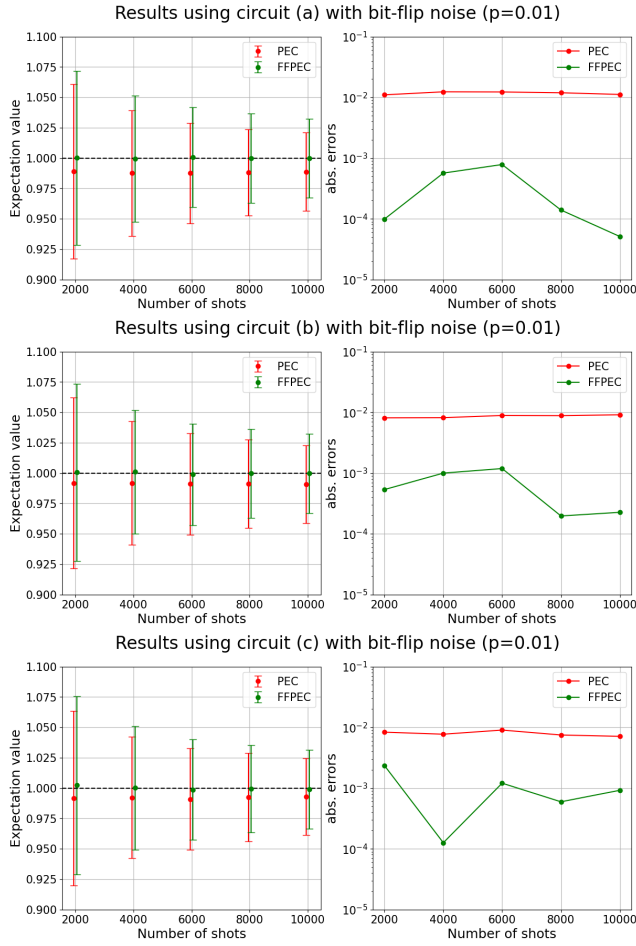
Fig. 2: Three types of 6 qubits quantum circuits used in experiments. The circuit 2(a) consists of 60 X gates. The circuit 2(b) consists of 30 CNOT gates. The circuit 2(c) is a combination of circuit 2(a) and circuit 2(b). The circuit depth is 10 for all three circuits.

Fig. 3: Accuracy of expectation values with bit-flip: Results with circuit 2(a) at the top, circuit 2(b) at the middle, and circuit 2(c) at the bottom. The figure on the left shows the average and standard deviation with 4000 samples for each shots. The black dashed line represents the exact expectation value of 1. The figure on the right shows the absolute error between the average and the exact expectation value.
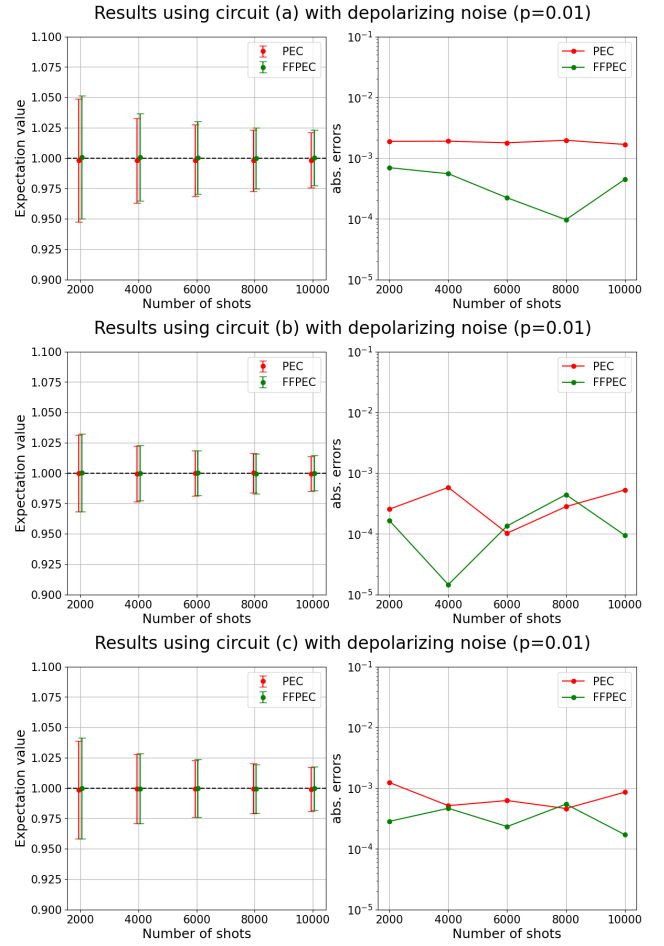
Fig. 4: Accuracy of expectation values with depolarizing: Results with circuit 2(a) at the top, circuit 2(b) at the middle, and circuit 2(c) at the bottom. The figure on the left shows the average and standard deviation with 4000 samples for each shots. The black dashed line represents the exact expectation value of 1. The figure on the right shows the absolute error between the average and the exact expectation value.
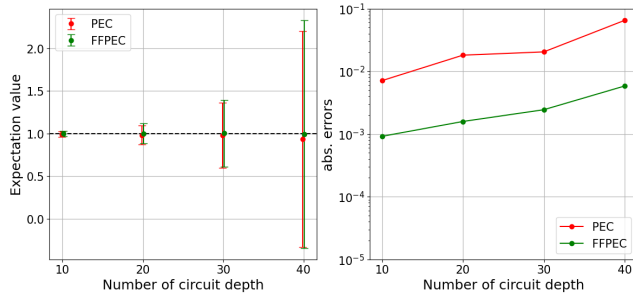
Fig. 5: Relationships between the number of gates in a circuit (depth) and the accuracy with bit-flip: the accuracy of the expectation value using a circuit which is produced by concatenating the circuit 2(c) twice, three times, and four times. Since the depth of circuit 2(c) is 10, the circuit with depth 20, 30, and 40 is a circuit connecting twice, three, and four times circuits 2(c). The expectation value for the depth 20 circuit is -1, thus an absolute value is taken for comparison for this case. The left figure shows the average and standard deviation of 4000 samples, and the right figure shows the absolute error between the average and the exact expectation value. Here, 10000 measurements are taken per sample.

error (red line) of the PEC estimator is about 10^{-2} for all experimental circuits, and the error remains similar even when the number of measurements increases. On the other hand, the error in the average of the FFPEC estimators (green line) is approximately 1 digit better than that of PEC in all of experimental circuits. These results indicate that slight difference of insertion probabilities lead to improve accuracies. FFPEC is capable of providing a more accurate expectation value with the same number of measurements as that of PEC in bit-flip noise.

From the results of depolarizing (right upper side of Fig. 4) with the circuit 2(a), FFPEC provides more accurate expectation values than that of PEC. On the other hand, the accuracy of the expectation values with FFPEC is similar to that of PEC when using circuits containing CNOT gates (2(b) and 2(c)). Since the difference of the insertion probabilities between FFPEC and PEC is 0.25%, 4000 samples would be not enough to detect this difference. Using a larger number of samples, it is possible to reveal this difference. Based on these results, it is expected that even with circuits containing CNOT gates under depolarizing noise, FFPEC can provide expectation values with accuracy same to or better than PEC.

Figure 5 shows the relationship between the number of circuit depths and the accuracy of expectation values under bit-flip noise. The circuit by connecting two (three or four) pieces of the circuit 2(c) is used.

As the number of gates increases, the sampling cost also increases, so if the number of measurements is fixed, the standard deviation increases as shown in Fig. 5 (left side). Furthermore, since increasing the number of depths inevitably leads to increasing the number of induced noises, the accuracy of expectation values become worse, as shown in Fig. 5 (right side). FFPEC also exhibits a decrease in accuracy with an increase in the number of depths, but it provides better

accuracy than PEC.

Similar to PEC, FFPEC also has a limitation in terms of achievable accuracy due to the noise caused by recovery gates. However, we expect that FFPEC can provide the same as or more accurate expectation values than PEC with the same number of measurements even if the circuit depth is larger.

V. CONCLUSION

We explore an improved version of PEC to achieve zero bias for the estimator of expectation values even with recovery gates inducing noises that is a realistic situation. We propose Feed-Forward Probabilistic Error Cancellation (FFPEC) method, which can provide a zero bias estimator even in such situation. The key idea of FFPEC is to define a new inverse map of noise to cancel out noises induced by recovery gates. To demonstrate the effect of FFPEC, new versions of inverse maps for bit-flip and depolarizing noises are provided and numerical experiments are conducted. Numerical results show that FFPEC can provide more accurate expectation values than PEC with all experimental circuits in bit-flip noise. In depolarizing noise, FFPEC can provide more accurate expectation values with circuit consisting of only single qubit gates. In the case of using circuits including CNOT gates, on the other hand, FFPEC can give expectation values with the same accuracy as PEC.

REFERENCES

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," *Nature Communications*, vol. 5, no. 1, p. 4213, Jul. 2014.

[2] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev, "VQE method: A short survey and recent developments," *Materials Theory*, vol. 6, no. 1, p. 2, Dec. 2022.

[3] E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm," *arXiv preprint arXiv:1411.4028*, Nov. 2014.

[4] A. Bärtschi and S. Eidenbenz, "Grover Mixers for QAOA: Shifting Complexity from Mixer Design to State Preparation," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Oct. 2020, pp. 72–82.

[5] J. Preskill, "Quantum Computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018.

[6] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," *Physical Review Letters*, vol. 119, no. 18, p. 180509, Nov. 2017.

[7] E. van den Berg, Z. K. Minev, and K. Temme, "Model-free readout-error mitigation for quantum expectation values," *Physical Review A*, vol. 105, no. 3, p. 032620, Mar. 2022.

[8] E. van den Berg, Z. K. Minev, A. Kandala, and K. Temme, "Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors," *Nature Physics*, vol. 19, no. 8, pp. 1116–1121, Aug. 2023.

[9] S. Endo, S. C. Benjamin, and Y. Li, "Practical Quantum Error Mitigation for Near-Future Applications," *Physical Review X*, vol. 8, no. 3, p. 031027, Jul. 2018.

[10] R. Hicks, B. Kobrin, C. W. Bauer, and B. Nachman, "Active readout-error mitigation," *Physical Review A*, vol. 105, no. 1, p. 012419, Jan. 2022.

[11] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, "Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography," *Quantum*, vol. 4, p. 257, Apr. 2020.

[12] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, "Mitigating measurement errors in multi-qubit experiments," *Physical Review A*, vol. 103, no. 4, p. 042605, Apr. 2021.

[13] Y. Guo and S. Yang, "Quantum Error Mitigation via Matrix Product Operators," *PRX Quantum*, vol. 3, no. 4, p. 040313, Oct. 2022.

[14] R. S. Gupta, E. van den Berg, M. Takita, D. Riste, K. Temme, and A. Kandala, "Probabilistic error cancellation for dynamic quantum circuits," *arXiv:2310.07825*, Dec. 2023.

[15] C. Kim, K. D. Park, and J.-K. Rhee, "Quantum Error Mitigation With Artificial Neural Network," *IEEE Access*, vol. 8, pp. 188 853–188 860, 2020.

[16] B. Koczor, J. Morton, and S. Benjamin, "Probabilistic Interpolation of Quantum Rotation Angles," *Physical Review Letters*, vol. 132, no. 13, p. 130602, Mar. 2024.

[17] R. Majumdar and C. J. Wood, "Error mitigated quantum circuit cutting," *arXiv:2211.13431*, Nov. 2022.

[18] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, "Scalable mitigation of measurement errors on quantum computers," *PRX Quantum*, vol. 2, no. 4, p. 040326, Nov. 2021.

[19] H.-C. Nguyen, "Information theoretic approach to readout error mitigation for quantum computers," *Physical Review A*, vol. 108, no. 5, p. 052419, Nov. 2023.

[20] C. Piveteau, D. Sutter, and S. Woerner, "Quasiprobability decompositions with reduced sampling overhead," *npj Quantum Information*, vol. 8, no. 1, pp. 1–9, Feb. 2022.

[21] R. Takagi, "Optimal resource cost for error mitigation," *Physical Review Research*, vol. 3, no. 3, p. 033178, Aug. 2021.

[22] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, "Quantum Error Mitigation," *Reviews of Modern Physics*, vol. 95, no. 4, p. 045005, Dec. 2023.

[23] Y. Suzuki, S. Endo, K. Fujii, and Y. Tokunaga, "Quantum Error Mitigation as a Universal Error Reduction Technique: Applications from the NISQ to the Fault-Tolerant Quantum Computing Eras," *PRX Quantum*, vol. 3, no. 1, p. 010345, Mar. 2022.

[24] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. Cambridge ; New York: Cambridge University Press, 2010.

[25] Y. Guo and S. Yang, "Noise effects on purity and quantum entanglement in terms of physical implementability," *npj Quantum Information*, vol. 9, no. 1, p. 11, Feb. 2023.

[26] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," 2023.