This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Quantum Error Mitigation with Artificial Neural Network

CHANGJUN KIM[1, 2], DANIEL K. PARK[1, 2], AND JUNE-KOO KEVIN RHEE.[1, 2]
[1]School of Electrical Engineering, KAIST, Daejeon, 34141, Republic of Korea
[2]ITRC of Quantum Computing for AI, KAIST, Daejeon, 34141, Republic of Korea

Corresponding author: June-Koo Kevin Rhee (e-mail: rhee.jk@kaist.edu).

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2018-0-01402) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) and by the National Research Foundation of Korea (Grant No. 2019R1I1A1A01050161). We acknowledge the use of the IBM Q for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Q team.

ABSTRACT A quantum error mitigation technique based on machine learning is proposed, which learns how to adjust the probabilities estimated by measurement in the computational basis. Neural networks in two different designs are trained with random quantum circuits consisting of a set of one- and two-qubit unitary gates whose measurement statistics in the ideal (noiseless) and real (noisy) cases are known. Once the neural networks are trained, they infer the amount of probability adjustment to be made on the measurement obtained from executing an unseen quantum circuit to reduce the error. The proposed schemes are tested with two-, three-, five-, and seven-qubit quantum circuits of depth up to 20 by computer simulations with realistic error models and experiments using the IBM quantum cloud platform. In all test cases, the proposed mitigation technique reduces the error effectively. Our method can be used to improve the accuracy of noisy intermediate-scale quantum (NISQ) algorithms without relying on extensive error characterization or quantum error correction.

INDEX TERMS Quantum Computing, Quantum Error Mitigation, Machine Learning, Artificial Neural Network

I. INTRODUCTION

The development of full-fledged quantum computers that promise revolutionary opportunities is being challenged due to computational errors that occur when theoretical ideas are implemented on real quantum devices. The theory of quantum error correction (QEC) [1]–[3] and fault-tolerance guarantees scalable quantum computation, but the computational resource overhead is non-negligible [4]–[6]. Moreover, the typical error rate of current quantum devices is near or above the fault-tolerance threshold, hindering the power of quantum error correction. With this background, the idea of quantum error mitigation (QEM) emerged recently. Unlike QEC, QEM does not necessarily aim to fully remove the entropy increased by unwanted interaction with the environment to recover the logical state. Instead, it aims to merely improve the accuracy of estimating the final answer in a given computational task without having to encode logical quantum state in a multi-qubit entangled state. Since QEM does not require extra quantum resources, it is expected to improve the quantum computation to some extent even when the error rate is above the fault-tolerance threshold value. Thus, QEM is an excellent fit for improving the performance of NISQ computing [7]–[11]. Moreover, a scalable error mitigation technique, if exists, can contribute to the fault-tolerant quantum computing beyond the NISQ era by reducing the error at the physical level.

Several QEM methods exists. Examples of the quantum circuit error mitigation technique are based on extrapolation [12]–[14], probabilistic error cancellation [13], [14], quantum subspace expansion [15], [16], and symmetry verification [17], [18]. A machine learning protocol based on training with Clifford circuits was used in Ref. [19] to improve probabilistic error cancellation by replacing the need of reconstructing an error model in the experiment. Data regression utilizing the classical simulability of Clifford circuits is also used in Ref. [20]. Some of these methods are experimentally verified in Refs. [21]–[23]. Also, error mitigation methods specifically designed for readout errors are presented in Refs. [24], [25].

In this paper, we propose a mitigation method that uses

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access

classical machine learning. An artificial neural network (ANN) is trained with shallow-depth quantum circuits whose measurement outcomes are known. The counts of gates applied to individual qubits and the actual outcome probability of all computational basis measurements are given as the input layer to the neural network for training. The ANN is trained with respect to a loss function that quantifies the error with respect to a true output state derived by computer simulation. Here, the error includes both cumulative gate error and qubit measurement error. Then given a new quantum circuit, the ANN infers the error of the probability distribution of the measurement results. This work serves as a stimulating example that shows classical machine learning can improve quantum computation.

The remainder of this paper is organized as follows. Section II describes the quantum error mitigation methods considered in this work. Section III presents the proof-of-principle implementation of QEM methods. The simulation and experimental setups for the test, such as the noise models and the device parameters, are described in section III-A and the results are presented in section III-B. Conclusions are drawn in section IV.

II. QUANTUM ERROR MITIGATION

A. EXACT METHODS FOR SIMPLE NOISE MODELS

Analytical methods can be so powerful to eliminate some types of errors of quantum gates in limited cases. Although the assumptions made for the application of these methods may not be realistic in a general experimental setup, the analytical methods presented here can serve as a reference to which the machine learning methods are compared.

1) Depolarizing Error

A two-qubit density matrix can be expressed with Pauli matrices as follows

$$\rho = \frac{I \otimes I}{2^2} + \sum_{k,l}^{4} \gamma_{k,l} \sigma_k \otimes \sigma_l, \tag{1}$$

where $\sigma_k, \sigma_l \in \{I, X, Y, Z\}$ and $\gamma_{1,1} = 0$. Since Eq. (1) describes an arbitrary two-qubit state, the effect of a noisy channel can be thought of as the modification of the coefficients $\gamma_{k,l}$. $\tilde{\gamma}_{k,l}$ denotes the erroneous coefficients.

The computational basis state of a qubit is defined as the eigenstates of Z. Therefore, the measurement in the computational basis uncovers only the information about $\tilde{\gamma}_{1,4}$, $\tilde{\gamma}_{4,1}$, and $\tilde{\gamma}_{4,4}$. Thus, the goal of the error mitigation in this work can be reduced to estimating $\gamma_{1,4}$, $\gamma_{4,1}$, and $\gamma_{4,4}$ from the estimators of $\tilde{\gamma}_{1,4}$, $\tilde{\gamma}_{4,1}$, and $\tilde{\gamma}_{4,4}$. Once $\gamma_{1,4}$, $\gamma_{4,1}$, and $\gamma_{4,4}$ are estimated, the correct probabilities in the noiseless case can be calculated by

$$\mathrm{Tr}(|ab\rangle\langle ab| \rho) = \begin{cases} \frac{1}{4} + \gamma_{1,4} + \gamma_{4,1} + \gamma_{4,4}, & ab = 00 \\ \frac{1}{4} - \gamma_{1,4} + \gamma_{4,1} - \gamma_{4,4}, & ab = 01 \\ \frac{1}{4} + \gamma_{1,4} - \gamma_{4,1} - \gamma_{4,4}, & ab = 10 \\ \frac{1}{4} - \gamma_{1,4} - \gamma_{4,1} + \gamma_{4,4}, & ab = 11 \end{cases}. \tag{2}$$

Given the gate-independent depolarizing error (see Appendix A) with single-qubit gate error rate denoted by ε_1 and two-qubit gate error rate denoted by ε_2, the error-free coefficients γ can be expressed in terms of the noisy coefficients $\tilde{\gamma}$, which can be estimated from the actual measurement probabilities $\mathrm{Tr}(|ab\rangle\langle ab| \rho_\epsilon)$ (c.f. Eq. (2)), as

$$\begin{aligned} \tilde{\gamma}_{1,4} &= (1 - \frac{4}{3}\varepsilon_1)^{G_{12}} (1 - \frac{16}{15}\varepsilon_2)^{G_2} \gamma_{1,4} \\ \tilde{\gamma}_{4,1} &= (1 - \frac{4}{3}\varepsilon_1)^{G_{11}} (1 - \frac{16}{15}\varepsilon_2)^{G_2} \gamma_{4,1} \\ \tilde{\gamma}_{4,4} &= (1 - \frac{4}{3}\varepsilon_1)^{G_{11}} (1 - \frac{4}{3}\varepsilon_1)^{G_{12}} (1 - \frac{16}{15}\varepsilon_2)^{G_2} \gamma_{4,4} \end{aligned}, \tag{3}$$

where G_{1k} is the number of single-qubit gates acting on the kth qubit and G_2 is the number of two-qubit gates.

2) Measurement Error

A simple measurement error mitigation (MEM) scheme is presented in Ref. [26], which we summarize in this section. The measurement error in an N-qubit experiment can be modeled as a linear transformation of the ideal measurement probabilities as

$$\mathbf{P}_{\mathrm{error}} = \mathbf{E}_{\mathrm{meas}} \mathbf{P}_{\mathrm{ideal}}, \tag{4}$$

where $\mathbf{P}_{\mathrm{ideal}}$ and $\mathbf{P}_{\mathrm{error}}$ are column vectors with 2^N elements indicating the probabilities for measuring the basis states in ideal and real cases, respectively, and $\mathbf{E}_{\mathrm{meas}}$ is a $2^N \times 2^N$ matrix that characterizes the measurement error. The jth column of $\mathbf{E}_{\mathrm{meas}}$ is constructed by measuring the input state ideally prepared in the jth computational basis state. In other words, the matrix element at row i and column j indicates the probability to obtain i when measuring the jth computational basis state. With this matrix in hand, the measurement error mitigation is performed by applying $\mathbf{E}_{\mathrm{meas}}^{-1}$ to the probability vector obtained in a given experiment.

B. ERROR MITIGATION VIA MACHINE LEARNING

The primary goal in this work is to mitigate errors that cannot be described solely as a gate-independent depolarizing noise and without relying on any noise characterization methods, such as randomized benchmarking and process tomography [27]–[39]. This section is dedicated to explaining machine learning methods to achieve such a goal.

1) Error Quantity Definition

As a preliminary step, we first define the error quantity that is to be minimized in the circuit of Fig.1. Given an arbitrary input state described by the density matrix ρ_i and the ideal unitary transformation denoted by U, the probability to obtain an outcome $|j\rangle$ can be expressed as

$$P(j|i) := Tr\left(M_j U \rho_i U^\dagger\right), \tag{5}$$

where M_j is a measurement operator. But in practice, the intended unitary transformation cannot be perfectly realized due to unwanted interaction with the environment and imperfections in the control. By denoting the actual (noisy)

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access

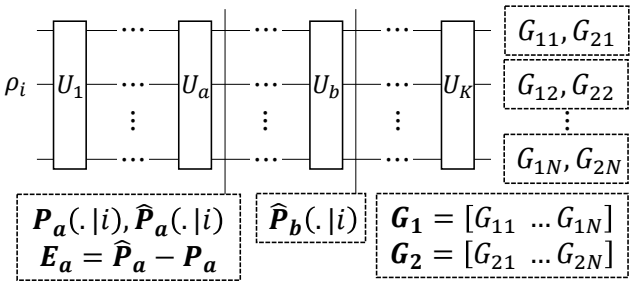


FIGURE 1. A schematic of ANN input extraction from a quantum circuit. ρ_i represents an input state. G_{kl} indicates the number of k-qubit gates applied to the lth qubit. The number of single-qubit gates ($\mathbf{G_1} = [G_{11}, \ldots, G_{1N}]$) and two-qubit gates ($\mathbf{G_2} = [G_{21}, \ldots, G_{2N}]$) applied to each qubit , the error at depth a ($\mathbf{E}_a = \hat{\mathbf{P}}_a - \mathbf{P}_a$), and the noisy measurement outcome at depth b ($\hat{\mathbf{P}}_b$) are used as the input values of a ANN.

quantum process by Λ, the actual probability estimated in a given experiment can be expressed as

$$\hat{P}(j|i) := Tr\left(M_j \Lambda\left(\rho_i\right)\right). \tag{6}$$

Based on the above, we can quantify the error as

$$E(j|i) := \hat{P}(j|i) - P(j|i). \tag{7}$$

Our error mitigation technique seeks to directly adjust $\hat{P}$ so that $E(j|i)$ is minimized. We quantify such correction (i.e. quantity of mitigation) as $C(j|i)$. Then, the final amount of error after applying the mitigation technique as

$$E_f(j|i) := E(j|i) - C(j|i) \tag{8}$$

Finally, we define the root mean squared error (RMSE) as

$$E_{\mathrm{RMS}} := \sqrt{\frac{\sum_{j=1}^{m}\sum_{i=1}^{l} E_f(j|i)^2}{lm}}. \tag{9}$$

Of course, $P(j|i)$ is not known to experimenters, and our goal is to find the distribution of $C(j|i)$ which minimizes E_{RMS}.

2) Training Data Generation

The training data for machine learning is generated by using random quantum circuits of depth K whose output state at depth $k = 1, \ldots, K$ is known *a priori* (by pre-calculation) for given input state. The ideal measurement statistics at each depth is compared with the actual measurement statistics obtained under noisy computation during the training. We extend the notations introduced in the previous section with a subscript k as $P_k, \hat{P}_k, E_k, E_{f,k}$ to indicate the quantities at depth k of a given quantum circuit. We also introduce probability vectors $\mathbf{P}_k$ and $\hat{\mathbf{P}}_k$ to denote the measurement probabilities. Of these matrices, the ith column corresponds to having the input basis state $|i\rangle$, and the jth row corresponds to the probability to measure the basis state $|j\rangle$. The error vector $\mathbf{E}_k = \hat{\mathbf{P}}_k - \mathbf{P}_k$ is also introduced to denote the error between the actual measurement outcome and the ideal measurement outcome. In Fig. 1, by computing the circuit at depth a, we can have $\mathbf{P_a}$ and by measuring the circuit at depth a and b, we can have $\hat{\mathbf{P}}_\mathbf{a}$ and $\hat{\mathbf{P}}_\mathbf{b}$ respectively.

The training dataset is generated using a set of random circuits of chosen depth, consisting of elementary single- and two-qubit gates. The single-qubit gates are the phase gate $S = |0\rangle\langle 0| + i\,|1\rangle\langle 1|$, the T gate $T = \sqrt{S}$, and the hadamard gate $H = (X + Z)/\sqrt{2}$. The controlled-Z gate $CZ = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Z$ is used as the two-qubit gate.

3) Machine Learning Structure

We propose and investigate two different designs of machine learning methods, namely an artificial neural network (ANN) and a concatenated artificial neural network (Concatenated ANN). The input of the ANNs consists of three parts: the known error information $\mathbf{E}_a$, the measured probability vector $\hat{\mathbf{P}}_b$ for $0 \leq a < b \leq K$, and the number of one- and two-qubit gates applied to each qubit between depth a and depth b denoted by $\mathbf{G_1}(a, b)$ and $\mathbf{G_2}(a, b)$, respectively. During the training, the input state for a quantum circuit is fixed. Thus, for a system of N qubits, there are $2N + 2^{N+1}$ input nodes: $2N$ for the number of single- and two-qubit gates, $\mathbf{G_1}(a, b)$ and $\mathbf{G_2}(a, b)$, and 2^{N+1} for ith columns of $\mathbf{E}_a$ and $\hat{\mathbf{P}}_b$ given the input state $|i\rangle$. Then the known correction $C_b(j|i)$ is placed in the output node for training. Thus, there are 2^N nodes at the output. The ANN has two hidden layers and all layers are fully connected. The sigmoid activation is used except for the last layer. The output of the ANN is the weighted sum of the values in the last layer. Concatenated ANN has one more hidden layer. The first hidden layer is split into two groups such that the first group is connected only to the input nodes for the number of gates while the second is connected only to the input nodes for $\mathbf{E}_a$ and $\hat{\mathbf{P}}_b$. As an ANN works better with normalized data, we add batch normalization for the number of gates after the first hidden layer. Then these two groups are concatenated (fully connected) to the next layer. Afterward, the concatenated ANN has the same structure as that of the ANN explained above. The neural network structures used in this work are depicted in Fig. 2.

The ANN trained with various input states $|i\rangle$ and pairs of a and b is used to infer the amount of correction for a target quantum circuit of depth $L > K$. The input nodes consist of $\mathbf{E}_l$, $\hat{\mathbf{P}}_L$, and the number of gates applied to each qubit between l and L, and the ANN outputs $C_L(j|i)$. The use of $\mathbf{E}_l$ improves the error mitigation when the error is known for some $l > 0$. In the absence of such information, the error mitigation is performed with only $\mathbf{E}_0$, which indicates the measurement error. The workflow to train the machine is in Fig. 3.

III. SIMULATION AND EXPERIMENT

As a proof of principle, we use simulation with two different noise models and experiment with an IBM quantum computer to validate the performance of QEM.

A. SETUP

In this example, 100 random circuits with $K = 10$ are used for training, and $l = 0$ and $L = 20$ for inference.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access

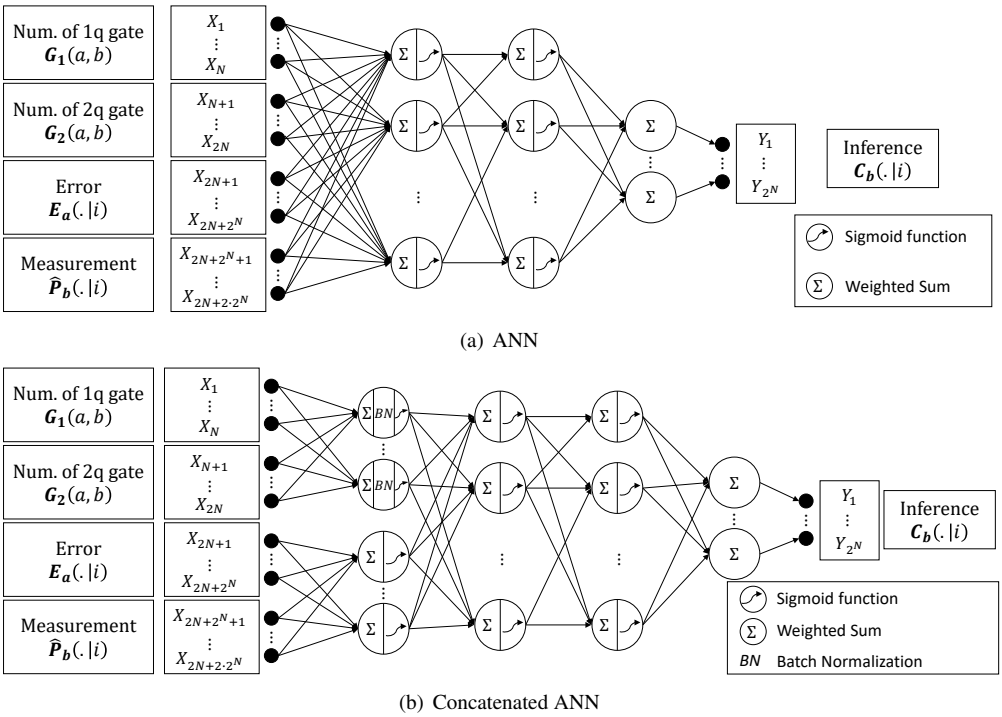


FIGURE 2. The neural network architectures used in this work. Two different ANN models are used, (a) ordinary ANN and (b) concatenated ANN. The ordinary ANN has two hidden layers. All layers are fully connected using weight summation and sigmoid is used for activation function. In the concatenated ANN, the first two layers are concatenated after applying batch normalization to the number of gates. After that, all hidden layers are fully connected using weight summation and activated by sigmoid. The ANN infers the error of a quantum circuit at depth b with the given information at depth $a < b$. $\mathbf{G_1}(a,b)$ and $\mathbf{G_2}(a,b)$ are the vector of the number of single- and two-qubit gates applied to each qubit between depth a to b, respectively. Input nodes are the ith column vectors of $\mathbf{E}_a$ and $\hat{\mathbf{P}}$, which are for a particular input state $|i\rangle$. Then ANN outputs the ith column vector of the correction matrix $\mathbf{C}_b$. The input quantity in the circuit is explained in Fig. 1.

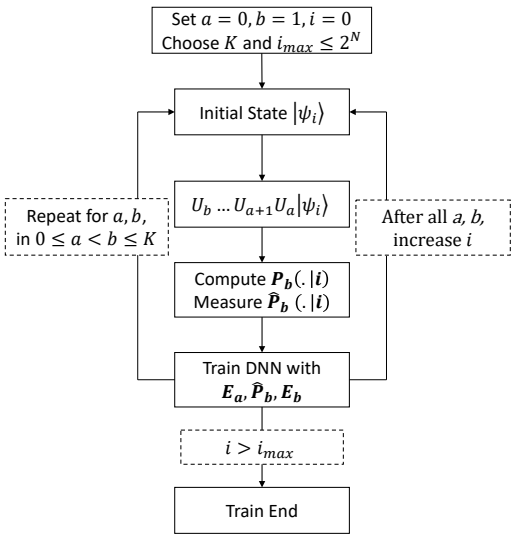
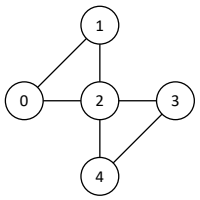


FIGURE 3. The workflow for training a ANN for one sample circuit. Initial states are chosen from the computational basis set. The workflow repeats until all circuits in the training set are used.

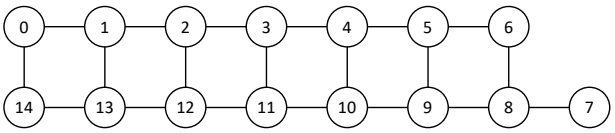
For each circuit, we use four states randomly chosen from the computational basis states are used as input. Each probability vector is estimated by repeating the same quantum circuit followed by the projective measurement 8192 times. Simulations were carried out with quantum circuits with the number of qubits ranging from two to seven. First, we perform the simulation with the depolarizing noise defined in Appendix. A. This noise model allows for the exact error cancellation via analytical solution as discussed in Sec. II-A. Thus, we use such oversimplification to test the performance of our method to the theoretical upper bound. For testing under more realistic noise, we use the noise model and parameters provided by the IBM Q cloud service [40]. In this case, the noise model consists of the depolarizing noise, longitudinal and transversal relaxations characterized by T_1 and T_2, and the measurement error. Hereinafter, we refer to this noise model as *IBM Q noise* model. The noise model is described in detail in Ref. [41]. For simulations with the IBM Q noise model, the error mitigation performance was tested with and without the measurement error mitigation technique described in Sec. II-A2. For simulations with up to five qubits, we took the data of the ibmq_5_yorktown quantum processor [42], which consists of five superconducting qubits, from June 25th, 2020. The single-qubit gate error rate was 6.25×10^{-4} and 6.70×10^{-4} for qubit 0 and qubit 1 respectively. Two-qubit gate error rate was 1.65×10^{-2}. For the seven-qubit simulation, we use the data of a quantum processor with 15 superconducting qubits named ibmq_16_melbourne [43]. For this device, the single-qubit error rate ranges from 4.49×10^{-4} to 4.18×10^{-3}, and the two-qubit error rate ranges from 1.09×10^{-2} to

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access



(a) Ibmq_5_yorktown



(b) Ibmq_16_melbourne

FIGURE 4. Physical qubit layouts of the IBM superconducting-qubit devices used in this work. (a) A five-qubit device called ibmq_5_yorktown is used for the two-qubit experiment. The noise parameters for this device are used for simulations with up to five qubits. The qubits used for the experiment are labeled as 0 and 1 in the figure. (b) A 15-qubit device called ibmq_16_melbourne. The noise parameters for this device are used for the seven-qubit simulation. Qubits 0, 1, 2, 3, 11, 12 and 13 are used for the simulation.

1.66×10^{-1}. The physical qubit layouts of the IBM quantum devices are shown in Fig. 4.

The experimental error mitigation was tested for two-qubit quantum circuits by using ibmq_5_yorktown.

B. RESULT

Table 1 presents simulation and experimental results for two-qubit error mitigation applied to 100 random test circuits that are not in the training set. The values in the table are the root mean squared error defined in Eq. 9 averaged over all test circuits, and the parenthesis contains the ratio of the E_{RMS} after error mitigation to before. The smaller the number in the table is the better the error mitigation performance. In all cases, the applied error mitigation technique always reduces the RMSE. As expected, the analytical method performs the best when only the depolarizing noise is present. The machine learning based methods also work reasonably well when compared to the analytical method; the RMSE of all three approaches is in the same order of magnitude. For more complex but relevant error models (i.e. IBM Q noise model), the machine learning based methods prevail. For the real quantum device, the concatenated ANN performs the best.

The simulation results for the seven-qubit system error mitigation are shown in Table 2. The machine learning reduces the error, although the amount of reduction is less than that for the two-qubit system. Indeed, our machine learning based error mitigation method always reduces the error for all cases that we tested as shown in Fig. 5. We also plot the ratio of the E_{RMS} after error mitigation to before in Fig. 6. The figure shows that the Concatenated ANN achieves slightly better results than the ANN. The numerical values for the simulation of three- and five-qubit systems are provided in Appendix B. Another interesting observation is that the RMSE obtained from the simulation with the measurement error mitigation explained in Sec. II-A2 does not provide a noticeable enhancement to the simulation result without it. This implies that our machine learning based error mitigation takes care of the measurement error to some extent.

TABLE 1. E_{RMS} in two-qubit error mitigation obtained from simulations under the depolarizing and IBM Q noise model and experiments with the IBM quantum device. The value in the parenthesis shows the ratio of the E_{RMS} after error mitigation to before.

Error Model	Unmitigated	ANN	CANN	Analytical
Depolarizing	2.5e-4	8.0e-5 (0.32)	7.0e-5 (0.28)	4.1e-5 (0.11)
IBM Q Noise w/o MEM	4.2e-4	1.6e-4 (0.37)	1.5e-4 (0.35)	3.3e-4 (0.70)
IBM Q Noise with MEM	2.1e-4	1.6e-4 (0.75)	1.3e-4 (0.63)	1.9e-4 (0.89)
Experiment	1.8e-3	1.1e-3 (0.63)	9.7e-4 (0.54)	1.7e-3 (0.97)

TABLE 2. E_{RMS} in seven-qubit error mitigation obtained from simulations under the depolarizing and IBM Q noise model. The value in the parenthesis shows the ratio of the E_{RMS} after error mitigation to before.

Error Model	Unmitgated	ANN	CANN
Depolarizing	3.7e-5	3.0e-5 (0.80)	3.1e-5 (0.83)
IBM Q Noise w/o MEM	8.2e-4	6.4e-4 (0.77)	6.4e-4 (0.77)
IBM Q Noise with MEM	8.3e-4	6.5e-4 (0.77)	6.5e-4 (0.78)

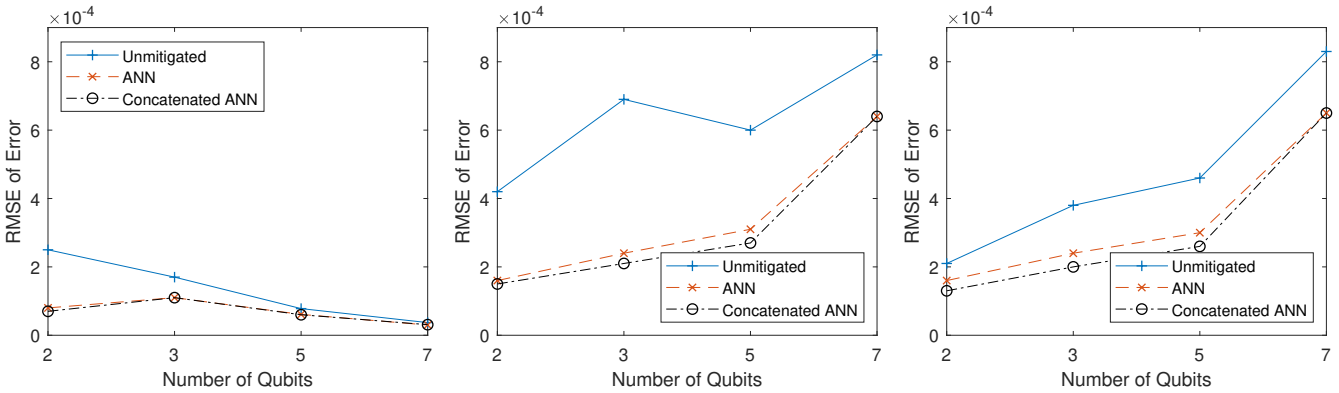
IV. CONCLUSION

We developed a quantum error mitigation scheme based on classical machine learning methods. The artificial neural network and the concatenated artificial neural network are used for machine learning. The neural networks are trained with quantum circuits for which the error behavior can be trained by numerical simulations and hardware experiments, in order to infer the amount of error in the result measurement of quantum output states. The performance of our method was tested using simulations with realistic noise models and experiments accomplished by the IBM quantum cloud service for quantum circuits consisting of two, three, five, and seven qubits. In all cases, our method was able to reduce the error. In particular, in the two-qubit experiment, our method reduced the error by a factor of 1/2.

The proposed error mitigation is designed to reduce the error of a target experiment conducted under the same noise model with which the neural networks are trained. Thus our method can effectively mitigate time-dependent noise with frequencies lower than the frequency of which the training is carried out. To optimize the error mitigation performance in a systematic way, the training should be done as a subroutine of the device calibration. Also, the ANN should be trained more frequently than the known frequency of the most dominant time-dependent noise. The development of a noise-independent error mitigation method will be of great

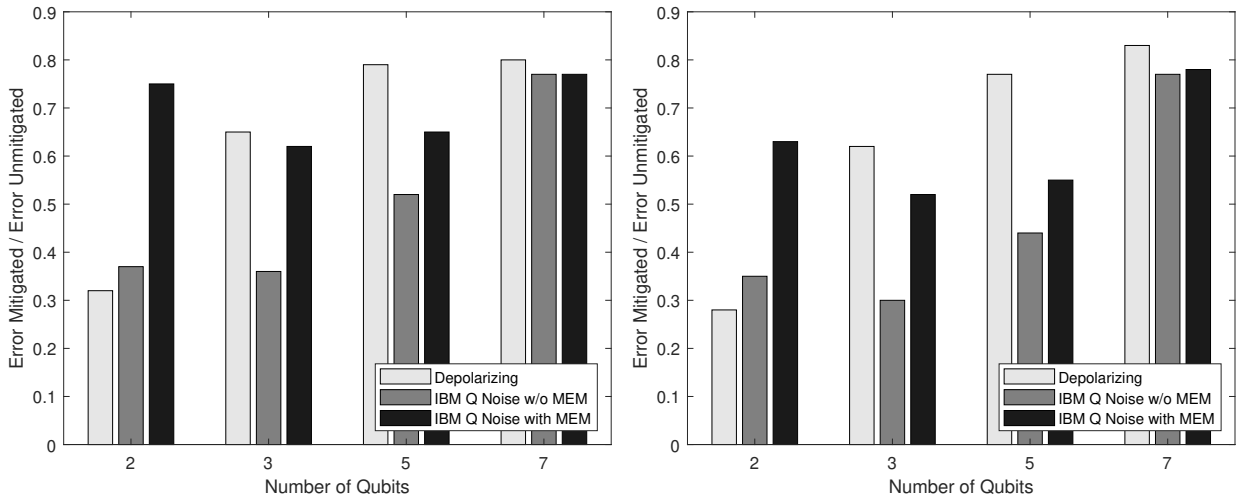
This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access



(a) E_{RMS} of depolarizing error (b) E_{RMS} of IMB Q Noise without MEM (c) E_{RMS} of IMB Q Noise with MEM

FIGURE 5. Simulation results for the error mitigation with ANN, Concatenated ANN, and no error mitigation under the (a) depolarizing noise model, (b) IBM Q noise model without measurement error mitigation, and (c) IBM Q noise model with measurement error mitigation. The vertical axis represents the E_{RMS} error averaged over 100 test circuits, and the horizontal axis shows the number of qubits.



(a) ANN error ratio (b) Concatenated ANN error ratio

FIGURE 6. The ratio of E_{RMS} after error mitigation to before for (a) ANN and (b) Concatenated ANN as a function of the number of qubits.

impact, which we plan to investigate in future work.

Another crucial future work is the development of a scalable ANN architecture in which the number of nodes increases polynomially with the number of quantum circuit elements. If the process of training can be done in polynomial time, it will be possible to apply the error mitigation method to a large number of qubits beyond the NISQ era. Although our method is limited in practice to small quantum systems, it sets an important milestone in the advancement of quantum computing, opening up tremendous opportunities for the application of machine learning to improve quantum control.

APPENDIX A DEPOLARIZING ERROR

The density matrix of one qubit can be expressed as

$$\rho = \frac{I}{2} + \sum_i \gamma_i \sigma_i, \tag{10}$$

where $\sigma_i \in \{X, Y, Z\}$. Under the depolarizing error with an error rate of ε_1, the density matrix changes to

$$\begin{aligned}
\rho_\varepsilon &= (1-\varepsilon_1)\rho + \frac{\varepsilon_1}{2^2-1}\Big(\sum_i \sigma_i \rho \sigma_i\Big) \\
&= (1-\varepsilon_1)\Big(\frac{I}{2} + \sum_i \gamma_i \sigma_i\Big) \\
&\quad + \frac{\varepsilon_1}{2^2-1}\sum_i \sigma_i \Big(\frac{I}{2} + \sum_j \gamma_j \sigma_j\Big)\sigma_i \\
&= (1-\varepsilon_1)\Big(\frac{I}{2} + \sum_i \gamma_i \sigma_i\Big) - \frac{\varepsilon_1}{2^2-1}\Big(\frac{3}{2} I + \sum_i \gamma_i \sigma_i\Big) \\
&= \frac{I}{2} - \Big(1 - \frac{4}{3}\varepsilon_1\Big)\sum_i \gamma_i \sigma_i,
\end{aligned} \tag{11}$$

where $\sigma_i \in \{X, Y, Z\}$.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access

We can express a two qubit density matrix as

$$\rho = \frac{I \otimes I}{2^2} + \sum_{i,j}^{4} \gamma_{i,j} \sigma_i \otimes \sigma_j, \tag{12}$$

where $\sigma_i, \sigma_j \in \{I, X, Y, Z\}$ and $\gamma_{1,1} = 0$. Under the two-qubit depolarizing error with an error rate of ε_2, the density matrix transforms as

$$\begin{aligned}
\rho_\varepsilon &= (1-\varepsilon_2)(\frac{I \otimes I}{2^2} + \sum_{i,j}^{4} \gamma_{i,j} \sigma_i \otimes \sigma_j) \\
&\quad + \frac{\varepsilon_2}{4^2-1}(\sum_{i,j}^{4}(\sigma_i \otimes \sigma_j) \\
&\quad (\frac{I \otimes I}{2^2} + \sum_{k,l}^{4} \gamma_{k,l} \sigma_k \otimes \sigma_l)(\sigma_i \otimes \sigma_j)) \\
&= (1-\varepsilon_2)(\frac{I \otimes I}{2^2} + \sum_{i,j}^{4} \gamma_{i,j} \sigma_i \otimes \sigma_j) \\
&\quad + \frac{\varepsilon_2}{4^2-1}(\frac{15}{2^2} I \otimes I - \sum_{i,j}^{4}(\gamma_{i,j} \sigma_i \otimes \sigma_j)) \\
&= \frac{I \otimes I}{4} + (1 - \frac{16}{15}\varepsilon_2) \sum_{i,j}^{4}(\gamma_{i,j} \sigma_i \otimes \sigma_j)
\end{aligned} \tag{13}$$

The one-qubit depolarizing error acting on the first qubit changes the two-qubit density matrix as

$$\begin{aligned}
\rho_\varepsilon &= (1-\varepsilon_1)\rho + \frac{\varepsilon_1}{2^2-1} \sum_{i=2}^{4} (\sigma_i \otimes I)\rho(\sigma_i \otimes I) \\
&= (1-\varepsilon_1)(\frac{I \otimes I}{2^2} + \sum_{i,j}^{4} \gamma_{i,j} \sigma_i \otimes \sigma_j) \\
&\quad + \frac{\varepsilon_1}{2^2-1} \sum_{i=2}^{4} (\sigma_i \otimes I) \\
&\quad (\frac{I \otimes I}{2^2} + \sum_{k,l}^{4} \gamma_{k,l} \sigma_k \otimes \sigma_l)(\sigma_i \otimes I)
\end{aligned} \tag{14}$$

$$\begin{aligned}
&= (1-\varepsilon_1)(\frac{I \otimes I}{2^2} + \sum_{i,j}^{4} \gamma_{i,j} \sigma_i \otimes \sigma_j) \\
&\quad + \frac{\varepsilon_1}{2^2-1}(\frac{3}{2^2} I \otimes I - \sum_{i,j}^{4} \gamma_{i,j} \sigma_i \otimes \sigma_j)) \\
&= \frac{I \otimes I}{4} + (1 - \frac{4}{3}\varepsilon_1) \sum_{i,j}^{4}(\gamma_{i,j} \sigma_i \otimes \sigma_j).
\end{aligned} \tag{15}$$

When the single-qubit and two-qubit depolarizing errors are combined, the two-qubit density matrix can be expressed with the number of single-qubit gates N_1 and the number of two-qubit gates N_2 as

$$\rho_\varepsilon = \frac{I \otimes I}{4} + (1 - \frac{4}{3}\varepsilon_1)^{N_1}(1 - \frac{16}{15}\varepsilon_2)^{N_2} \sum_{i,j}^{4}(\gamma_{i,j} \sigma_i \otimes \sigma_j). \tag{16}$$

APPENDIX B MITIGATION RESULTS

TABLE 1. Error mitigation result for the five-qubit case.

Error Model	Unmitigated	ANN	CANN
Depolarizing	7.8e-5	6.1e-5 (0.79)	6.0e-5 (0.77)
IBM Q Noise w/o MEM	6.0e-4	3.1e-4 (0.52)	2.7e-4 (0.44)
IBM Q Noise with MEM	4.6e-4	3.0e-4 (0.65)	2.6e-4 (0.55)

TABLE 2. Error mitigation result for the three-qubit case.

Error Model	Unmitigated	ANN	CANN
Depolarizing	1.7e-4	1.1e-4 (0.65)	1.1e-4 (0.62)
IBM Q Noise w/o MEM	6.9e-4	2.4e-4 (0.36)	2.1e-4 (0.30)
IBM Q Noise with MEM	3.8e-4	2.4e-4 (0.62)	2.0e-4 (0.52)

REFERENCES

[1] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. Physical review A, 52(4):R2493, 1995.

[2] Andrew M Steane. Error correcting codes in quantum theory. Physical Review Letters, 77(5):793, 1996.

[3] A Robert Calderbank and Peter W Shor. Good quantum error-correcting codes exist. Physical Review A, 54(2):1098, 1996.

[4] Austin G Fowler, Matteo Mariantoni, John M Martinis, and Andrew N Cleland. Surface codes: Towards practical large-scale quantum computation. Physical Review A, 86(3):032324, 2012.

[5] Barbara M Terhal. Quantum error correction for quantum memories. Reviews of Modern Physics, 87(2):307, 2015.

[6] David JC MacKay, Graeme Mitchison, and Paul L McFadden. Sparse-graph codes for quantum error correction. IEEE Transactions on Information Theory, 50(10):2315–2330, 2004.

[7] John Preskill. Quantum computing in the nisq era and beyond. Quantum, 2:79, 2018.

[8] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. Nature communications, 5:4213, 2014.

[9] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature, 549(7671):242–246, 2017.

[10] James I Colless, Vinay V Ramasesh, Dar Dahlen, Machiel S Blok, ME Kimchi-Schwartz, JR McClean, J Carter, WA De Jong, and I Siddiqi. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. Physical Review X, 8(1):011021, 2018.

[11] Nikolaj Moll, Panagiotis Barkoutsos, Lev S Bishop, Jerry M Chow, Andrew Cross, Daniel J Egger, Stefan Filipp, Andreas Fuhrer, Jay M Gambetta, Marc Ganzhorn, et al. Quantum optimization using variational algorithms on near-term quantum devices. Quantum Science and Technology, 3(3):030503, 2018.

[12] Ying Li and Simon C Benjamin. Efficient variational quantum simulator incorporating active error minimization. Physical Review X, 7(2):021050, 2017.

[13] Suguru Endo, Simon C Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications. Physical Review X, 8(3):031027, 2018.

[14] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. Error mitigation for short-depth quantum circuits. Physical review letters, 119(18):180509, 2017.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031607, IEEE Access

[15] Jarrod R McClean, Mollie E Kimchi-Schwartz, Jonathan Carter, and Wibe A De Jong. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. Physical Review A, 95(4):042308, 2017.

[16] Jarrod R McClean, Zhang Jiang, Nicholas C Rubin, Ryan Babbush, and Hartmut Neven. Decoding quantum errors with subspace expansions. Nature Communications, 11(1):1–9, 2020.

[17] Xavi Bonet-Monroig, Ramiro Sagastizabal, M Singh, and TE O'Brien. Low-cost error mitigation by symmetry verification. Physical Review A, 98(6):062339, 2018.

[18] Sam McArdle, Xiao Yuan, and Simon Benjamin. Error-mitigated digital quantum simulation. Physical review letters, 122(18):180501, 2019.

[19] Armands Strikis, Dayue Qin, Yanzhu Chen, Simon C. Benjamin, and Ying Li. Learning-based quantum error mitigation, 2020.

[20] Piotr Czarnik, Andrew Arrasmith, Patrick J. Coles, and Lukasz Cincio. Error mitigation with clifford quantum-circuit data, 2020.

[21] Abhinav Kandala, Kristan Temme, Antonio D Córcoles, Antonio Mezzacapo, Jerry M Chow, and Jay M Gambetta. Error mitigation extends the computational reach of a noisy quantum processor. Nature, 567(7749):491–495, 2019.

[22] Chao Song, Jing Cui, H Wang, J Hao, H Feng, and Ying Li. Quantum computation with universal error mitigation on a superconducting quantum processor. Science advances, 5(9):eaaw5686, 2019.

[23] Shuaining Zhang, Yao Lu, Kuan Zhang, Wentao Chen, Ying Li, Jing-Ning Zhang, and Kihwan Kim. Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system. Nature communications, 11(1):1–8, 2020.

[24] Filip B Maciejewski, Zoltán Zimborás, and Michał Oszmaniec. Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography. Quantum, 4:257, 2020.

[25] Swamit S Tannu and Moinuddin K Qureshi. Mitigating measurement errors in quantum computers by exploiting state-dependent bias. In Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture, pages 279–290, 2019.

[26] Abraham Asfaw and et al. Learn quantum computation using qiskit, 2020.

[27] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: The two-bit quantum gate. Phys. Rev. Lett., 78:390–393, Jan 1997.

[28] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. Phys. Rev. A, 77:012307, Jan 2008.

[29] Easwar Magesan, J. M. Gambetta, and Joseph Emerson. Scalable and robust randomized benchmarking of quantum processes. Phys. Rev. Lett., 106:180504, May 2011.

[30] Easwar Magesan, Jay M. Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. Phys. Rev. A, 85:042311, Apr 2012.

[31] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. Efficient measurement of quantum gate error by interleaved randomized benchmarking. Phys. Rev. Lett., 109:080505, Aug 2012.

[32] Seth T. Merkel, Jay M. Gambetta, John A. Smolin, Stefano Poletto, Antonio D. Córcoles, Blake R. Johnson, Colm A. Ryan, and Matthias Steffen. Self-consistent quantum process tomography. Phys. Rev. A, 87:062119, Jun 2013.

[33] Arnaud Carignan-Dugas, Joel J. Wallman, and Joseph Emerson. Characterizing universal gate sets via dihedral benchmarking. Phys. Rev. A, 92:060302, Dec 2015.

[34] Guanru Feng, Joel J. Wallman, Brandon Buonacorsi, Franklin H. Cho, Daniel K. Park, Tao Xin, Dawei Lu, Jonathan Baugh, and Raymond Laflamme. Estimating the coherence of noise in quantum control of a solid-state qubit. Phys. Rev. Lett., 117:260501, Dec 2016.

[35] Andrew W Cross, Easwar Magesan, Lev S Bishop, John A Smolin, and Jay M Gambetta. Scalable randomised benchmarking of non-clifford gates. npj Quantum Information, 2(1):16012, 2016.

[36] Joel J. Wallman. Randomized benchmarking with gate-dependent noise. Quantum, 2:47, January 2018.

[37] Timothy J. Proctor, Arnaud Carignan-Dugas, Kenneth Rudinger, Erik Nielsen, Robin Blume-Kohout, and Kevin Young. Direct randomized benchmarking for multiqubit devices. Phys. Rev. Lett., 123:030503, Jul 2019.

[38] Kyle Willick, Daniel K. Park, and Jonathan Baugh. Efficient continuous-wave noise spectroscopy beyond weak coupling. Phys. Rev. A, 98:013414, Jul 2018.

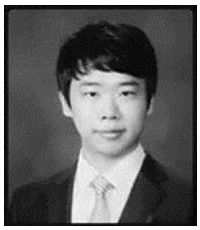
[39] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. Nature Reviews Physics, 2(7):382–390, 2020.

[40] IBM Q Experience, URL: https://quantum-computing.ibm.com/.

[41] Carsten Blank, Daniel K Park, June-Koo Kevin Rhee, and Francesco Petruccione. Quantum classifier with tailored quantum kernel. npj Quantum Information, 6(1):1–7, 2020.

[42] 5 qubit backend: IBM Q team. IBM Q 5 Yorktown backend specification V2.1.0, 2020.

[43] 15 qubit backend: IBM Q team. IBM Q 16 Melbourne backend specification V2.3.0, 2020.



CHANGJUN KIM received the B.S. and M.S. degrees in electrical engineering from KAIST, South Korea, in 2013 and 2015, respectively. He is currently a candidate for a Ph.D. course in electrical engineering, KAIST, South Korea. His research includes quantum machine learning, quantum error mitigation, and quantum artificial intelligence.

DANIEL K. PARK received B.Sc. in Mathematical Physics with Dean's Honours in 2010 from University of Waterloo, Canada, and a Ph.D. degree in Physics-Quantum Information in 2015 from the same university. He was a post-doctoral researcher at Department of Physics in KAIST, South Korea, from 2015 to 2018, and became a research assistant professor in 2018 at School of Electrical Engineering, KAIST. His current research covers development of quantum algorithms for computational and data science.

JUNE-KOO KEVIN RHEE received the B.E. and M.Sc. degrees from Seoul National University, Seoul, South Korea, in 1988 and 1990, respectively, and the Ph.D. degree from the University of Michigan, Ann Arbor, MI, USA, in 1995, all in electrical engineering. He is currently a Professor of electrical engineering with the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea. Prior to joining KAIST, he held research and teaching positions with Princeton University (1995–1996), NEC Research Institute (1996–1998), Corning Incorporated (1998–2002), Samsung Advanced Institute of Technology (2003–2005), and the Information and Communications University (2005–2009). He holds more than 60 patents. He has authored or coauthored more than 180 technical journal and conference papers. His research interests include the areas of optical and wireless networking, quantum communication, and quantum computing. Prof. Rhee was the recipient of the Ministry Certificate of Commendation from the Korean Ministry of Science, ICT, and Planning

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/