



OPEN

## Efficient noise mitigation technique for quantum computing

Ali Shaib<sup>1</sup>, Mohamad Hussein Naim<sup>1</sup>, Mohammed E. Fouda<sup>2✉</sup>, Rouwaida Kanj<sup>1</sup> & Fadi Kurdahi<sup>2</sup>

Quantum computers have enabled solving problems beyond the current machines' capabilities. However, this requires handling noise arising from unwanted interactions in these systems. Several protocols have been proposed to address efficient and accurate quantum noise profiling and mitigation. In this work, we propose a novel protocol that efficiently estimates the average output of a noisy quantum device to be used for quantum noise mitigation. The multi-qubit system average behavior is approximated as a special form of a Pauli Channel where Clifford gates are used to estimate the average output for circuits of different depths. The characterized Pauli channel error rates, and state preparation and measurement errors are then used to construct the outputs for different depths thereby eliminating the need for large simulations and enabling efficient mitigation. We demonstrate the efficiency of the proposed protocol on four IBM Q 5-qubit quantum devices. Our method demonstrates improved accuracy with efficient noise characterization. We report up to 88% and 69% improvement for the proposed approach compared to the unmitigated, and pure measurement error mitigation approaches, respectively.

Quantum Computation is an emerging field that sits at an intersection between computer science, electrical engineering and quantum physics. By now, scientists and engineers alike are optimistic of its computational and technological potential, achieving a so-called computational supremacy<sup>1,2</sup>. Countless of applications for quantum devices have been proposed from building a quantum internet<sup>3,4</sup> to enhancing and feeding into scientific discoveries<sup>5</sup> to revolutionizing simulation<sup>6</sup>, neural network<sup>7</sup> and optimization<sup>8</sup>. Some of these applications are already being tested and even practically implemented<sup>9</sup>. That said, quantum computer implementation technologies are still a work in progress. Results on different qubit technologies, computer models, and interconnects and more have recently been published<sup>10–12</sup>.

Building large-scale quantum computers is still a challenging task due to a plethora of engineering obstacles<sup>13</sup>. One prominent challenge is the intrinsic noise. In fact, implementing scalable and reliable quantum computers requires implementing quantum gates with sufficiently low error rates. There has been substantial progress in characterizing noise in a quantum system<sup>14–16</sup> and in building error correcting schemes that can detect and correct certain types of errors<sup>17–19</sup>.

Numerous protocols have been constructed to characterize the noise in quantum devices. Many of these protocols fail in achieving one of the following desirables: scalability to large-scale quantum computers and efficient characterization of the noise. Quantum Process Tomography<sup>20</sup> is a protocol that can give a complete description of the dynamics of a quantum black box, however, it's not scalable to large-scale quantum systems. Randomized Benchmarking (RB) is another protocol that's typically used to estimate the error rate of some set of quantum gates<sup>21,22</sup>. Although RB is a scalable protocol in principle, it can only measure a single error rate that's used to approximate the average gate infidelity thus providing an incomplete description of noise. Various other protocols based on RB protocol are able to characterize the correlations of noise between the different qubits, however, these protocols lack scalability<sup>21,23,24</sup>. Quantum Error Mitigation<sup>25</sup> (QEM) is a recently emerging field that aims to improve the accuracy of near-term quantum computational tasks<sup>26</sup> many times through data post-processing. It is considered a much more feasible alternative, for the time being, to Quantum Error Correction (QEC) which revolves around encoding quantum states in multi-qubit entangled states<sup>27,28</sup> to achieve universal fault-tolerant quantum computation. QEM protocols include zero-noise Richardson extrapolation<sup>29</sup>, probabilistic error cancellation through sampling circuits and taking their weighted average<sup>30,31</sup>, and exploiting state-dependent bias through invert-and-measure techniques to map the predicted state to the strongest one<sup>32</sup>. Others have worked on using diverse ansatzes/models to predict observables' noise-free values<sup>33</sup>, distribution correction factors<sup>34</sup> or even circuit noise metrics (to behave as an objective function for a quantum compiler<sup>35</sup>).

<sup>1</sup>Electrical and Computer Engineering Department, American University of Beirut, Beirut 1107 202, Lebanon. <sup>2</sup>Center for Embedded and Cyber-Physical Systems, University of California-Irvine, Irvine, CA 92697-2625, USA. ✉email: foudam@uci.edu

To mention a few more, there is quantum subspace expansion, symmetry verification and countless learning-based techniques available recently<sup>36,37</sup>.

Measurement Error Mitigation (MEM) is a QEM protocol that models the noise in a quantum circuit as a measurement noise matrix  $E_{meas}$  applied to the ideal output of the circuit. The columns of  $E_{meas}$  are the probability distributions obtained through preparing and immediately measuring all possible  $2^n$  basis input states<sup>38</sup>.

Recently, the authors in<sup>39</sup> developed a protocol based on the RB that relies on the concept of a Gibbs Random Field (GRF) to completely and efficiently estimate the error rates of the Pauli Channel and detect correlated errors between the qubits in a quantum computer. Their effort paves the way to enable quantum error correction and/or mitigation schemes. Herein, we refer to their efficient learning protocol as the {EL protocol}. In this paper, we build upon the EL protocol and decompose the average noise of a quantum circuit of specific depth into State Preparation and Measurement (SPAM) error and average gate error. We propose a linear algebraic based protocol and proof to efficiently construct and model the average behavior of noise in a quantum system for any desired circuit depth without having to run a large number of quantum circuits on the quantum computer or simulator. We then rely on this model to mitigate the noisy output of the quantum device. Noise mitigation schemes are often categorized into two, either active involving online processing dependent on each run and operation, or passive in which it is independent of the run and operation. Our proposed technique is considered a passive one since only one noise characterization is made for the quantum device at hand and then used to mitigate errors for any arbitrary circuit of specific depth on that device. For an  $n$ -qubit quantum system, the average behavior of the noise can be well approximated as a special form of a Pauli Channel<sup>40–42</sup>. A Pauli channel  $\varepsilon$  acts on a qubit state  $\rho$  to produce

$$\varepsilon(\rho) = \sum_i p_i P_i \rho P_i \quad (1)$$

where  $p_i$  is an error rate associated with the Pauli operator  $P_i$ . The  $p_i$ 's form a probability distribution ( $\sum_i p_i = 1$ ), and are related to the eigenvalues,  $\lambda$ , of the Pauli Channel defined as

$$\lambda_i = 2^{-n} \text{Tr}(P_i \varepsilon(P_i)) \quad (2)$$

Thus, when a state  $\rho$  is subjected to the noisy channel  $\varepsilon$ ,  $p_i$  describes the probability of a multiqubit Pauli error  $P_i$  affecting the system, while  $\lambda_i$  describes how faithfully a given multispin Pauli operator is transmitted.  $\rho$  and  $\lambda$  are related by Walsh–Hadamard transform,  $W$  where

$$\lambda = W\rho \quad (3)$$

While RB only estimates the average value of all  $\lambda_i$  of the Pauli Channel, the EL protocol estimates the individual  $\lambda_i$ . A complete characterization of the Pauli channel requires learning more than the eigenvalues or error rates associated with single-qubit Pauli operators such as  $\sigma_z^{(1)}$  or  $\sigma_x^{(3)}$ ; it requires learning all of the noise correlations in the system, that is, also learning the eigenvalues and error rates associated with multiqubit Pauli operators such as  $\sigma_z^{(1)} \otimes \mathbf{1}^{(2)} \otimes \sigma_x^{(3)}$  and how they vary compared to the ones obtained under independent local noise. Estimating these correlations is essential for performing optimal QEC and/or QEM. However, these correlations increase exponentially as the number of qubits increases, so having an efficient noise characterization protocol is crucial to direct the error mitigation efforts to capture the critical noise correlations.

Our method relies on the error rates vector  $\rho$  of the Pauli-Channel to decompose the average behavior of noise for circuits of depth  $m$  into two noise components: a SPAM error matrix denoted by the matrix  $N$  and a depth dependent component comprising an average gate error matrix denoted by the matrix  $M$ . We evaluate our model for the average noise by predicting the average probability distribution for circuits of depth  $m$  and computing the distance between this predicted distribution and the empirically obtained one. Finally, we use our proposed decomposition to mitigate noisy outputs of random circuits and compare our mitigation protocol with the MEM protocol<sup>38</sup>. We applied our noise characterization and mitigation protocols on the following IBM Q 5-qubit quantum computers: Manila, Lima, and Belem<sup>43</sup>.

## Results

**Proposed protocol theory.** The ideal output probability distribution of an  $n$ -qubit quantum circuit with depth  $m$  is perturbed by the SPAM and the average gate errors. Our aim is to construct a comprehensive linear algebraic model that takes into account both these errors for an arbitrary depth  $m$ . Matrix algebra can then be employed to mitigate the noise as follows:

$$C_{ideal} = Q_m^{-1} C_{noisy} \quad (4)$$

where  $Q_m$  is the characterized noise matrix for circuits of depth  $m$ ,  $C_{ideal}$  and  $C_{noisy}$  are the ideal and noisy outputs of a given circuit of depth  $m$ , respectively. The straight-forward approach would be to construct  $Q_m$  from empirical simulations in a similar fashion to the  $E_{meas}$  noise matrix that was characterized in the MEM scheme. The columns of  $Q_m$  comprise the empirical average probability distributions for basis input states  $|in\rangle \in \{|0\rangle, |1\rangle, \dots, |2^n - 1\rangle\}$ , denoted by  $\hat{q}(m, |in\rangle)$ , where  $\hat{q}(m, |in\rangle)$  are obtained through sampling a number of depth  $m$  circuits to incorporate the average gate and SPAM errors.

$$Q_m = [\hat{q}(m, |0\rangle) \hat{q}(m, |1\rangle) \dots \hat{q}(m, |2^n - 1\rangle)] \quad (5)$$

Building  $\mathbf{Q}_m$ , however, through empirical simulations can be expensive especially when the circuit depth is large. Herein, we propose a method for an efficient estimation of  $\mathbf{Q}_m$  where the individual probability distributions  $\hat{\mathbf{q}}(m, |\mathbf{in}\rangle)$  are estimated as follows:

$$\mathbf{q}'(m, |\mathbf{in}\rangle) = N_{in} \mathbf{M}_{in}^m |\mathbf{in}\rangle \quad (6)$$

where  $N_{in}$  and  $\mathbf{M}_{in}$  are input-specific matrices that represent the SPAM error matrix and average gate error for input  $|\mathbf{in}\rangle$ , respectively. Both  $\mathbf{M}_{in}$  and  $N_{in}$  are extracted empirically using random circuits from a set of small circuit depths  $T$  and then used in mitigating the outputs for circuits with higher depths. We first show the construction of  $N_0$  and  $\mathbf{M}_0$ .

The construction of  $N_0$  and  $\mathbf{M}_0$  proceeds by estimating the error rates vector  $\mathbf{p}$  associated with the Pauli Channel based on the assumption in Eq. 1 for the average behavior of the noisy quantum device at hand using the EL protocol. The protocol proceeds by constructing  $K$  random identity circuits of depth  $m \in T^{23,39}$ . Each circuit is constructed by initializing the qubits to the all-zeros state  $|\mathbf{0}\rangle$  followed by choosing a random sequence  $s \in S_m$ , the set of all length  $m$  sequences of one-qubit Clifford gates applied independently on each qubit, followed by an inverse gate for the chosen sequence to ensure an identity circuit. It then estimates the resulting empirical probability distribution  $\hat{\mathbf{q}}(m, |\mathbf{0}\rangle)$  by averaging over all the empirical probability distributions  $\hat{\mathbf{q}}(m, s, |\mathbf{0}\rangle)$  for the constructed random identity circuits of depth  $m$ , that is,

$$\hat{\mathbf{q}}(m, |\mathbf{0}\rangle) = \frac{1}{K} \sum \hat{\mathbf{q}}(m, s, |\mathbf{0}\rangle) \quad (7)$$

$\hat{\mathbf{q}}(m, |\mathbf{0}\rangle)$  is a vector with  $2^n$  entries each corresponding to the possible observed outcome. A Walsh–Hadamard transform is then applied on each  $\hat{\mathbf{q}}(m, |\mathbf{0}\rangle)$  to obtain

$$\mathbf{\Lambda}(m) = W \hat{\mathbf{q}}(m, |\mathbf{0}\rangle) \quad (8)$$

Each parameter  $\Lambda_i(m)$  in  $\mathbf{\Lambda}(m)$  is fitted to the model

$$\Lambda_i(m) = A_i \lambda_i^m \quad (9)$$

where  $A_i$  is a constant that absorbs SPAM errors and the vector  $\boldsymbol{\lambda}$  of all fitted parameters  $\lambda_i$  is a SPAM-free estimate to the eigenvalues of the Pauli Channel defined in Eq. 2. Notice that we can rewrite Eq. 9 as

$$\mathbf{\Lambda}(m) = \mathbf{A} \boldsymbol{\lambda}^m \quad (10)$$

where  $\mathbf{A}$  is a diagonal matrix where the diagonal entries are  $A_i$  and  $\boldsymbol{\lambda}^m$  is an element-wise exponentiation of a vector. An inverse Walsh–Hadamard Transform is then applied on  $\boldsymbol{\lambda}$  to get the error rate vector  $\mathbf{p}$  of the Pauli Channel as

$$\mathbf{p} = W^{-1} \boldsymbol{\lambda} \quad (11)$$

$\mathbf{p}$  is then projected onto a probability simplex to ensure  $\sum_i p_i = 1$ . Introducing the GRF model by the EL protocol allows the scalability of estimating  $\mathbf{p}$  with the increase in the number of qubits. The GRF model assumes the noise correlations are bounded between a number of neighboring qubits depending on the architecture of the quantum computer at hand. Thus, decreasing the number of noise correlations to be estimated.

The final outcome  $\mathbf{p}$  of the EL protocol represents the SPAM-free probability distribution of the average noise in the quantum computer. Each element  $p_i \in \mathbf{p}$  corresponds to the probability of an error of the form *binary*( $i$ ) on an input state  $|\mathbf{0}\rangle$ . For example, for a 5-qubit quantum computer,  $p_0$  corresponds to the probability of no bit flips on the input state, i.e., error of the form *IIIII*,  $p_1$  to the error of the form *IIIX*,  $p_2$  to the error of the form *IIIXI*, etc...

In order to proceed with the proof for our proposed decomposition of Eq. (6) for input state  $|\mathbf{0}\rangle$ , we first state the following lemma (the detailed proof of the lemma can be found Section I in the supplementary):

**Lemma 1** *Let  $\boldsymbol{\lambda}$  and  $\mathbf{p}$  be the respective eigenvalues and error rates of a Pauli Channel with  $n$  qubits, then  $\boldsymbol{\lambda}^m = W \mathbf{M}^m |\mathbf{0}\rangle$  where  $\mathbf{M}$  is a  $2^n \times 2^n$  matrix such that  $M_{ij} = p_{i \oplus j}$  ( $i \oplus j$  is the bitwise exclusive-OR operator).*

Using Lemma 1 and Eqs. 8 and 10,  $\hat{\mathbf{q}}(m, |\mathbf{0}\rangle)$  can be estimated as

$$\mathbf{q}'(m, |\mathbf{0}\rangle) = W^{-1} \mathbf{A} W \mathbf{M}^m |\mathbf{0}\rangle \quad (12)$$

The transition matrix  $\mathbf{M} = \mathbf{M}_0$  represents the average error per gate while the  $N = W^{-1} \mathbf{A} W = N_0$  matrix represents the SPAM errors for an input state  $|\mathbf{0}\rangle$ . Notice that the average noise for depth  $m$  circuits on an input state  $|\mathbf{0}\rangle$  behaves as a sequence of  $m$  average noise gates  $\mathbf{M}_0$  followed by SPAM errors  $N_0$ .

The construction of  $N_{in}$  and  $\mathbf{M}_{in}$  for input state  $|\mathbf{in}\rangle$  proceeds similar to the procedure of constructing  $N_0$  and  $\mathbf{M}_0$ , however, a permutation of  $\hat{\mathbf{q}}(m, |\mathbf{in}\rangle)$  is required before applying a Walsh–Hadamard transform to ensure that each element  $p_i(|\mathbf{in}\rangle)$  in the input-specific error rate vector  $\mathbf{p}(|\mathbf{in}\rangle)$  corresponds to the probability of an error of the form *binary*( $i$ ) on an input state  $|\mathbf{in}\rangle$ . This permutation is done by applying an input-specific permutation matrix  $\pi_{in}$  on  $\hat{\mathbf{q}}(m, |\mathbf{in}\rangle) \forall m$  where  $\pi_{in_{ij}} = 1$  if  $i \oplus j = in$  and 0 otherwise. It must be noted that  $\mathbf{Q}_m$  can be nearly singular especially for large  $m$  as can be directly inferred from Eq. 12 ( $|\mathbf{M}^m| \sim 0$ ). In this case, the results are random hence the desire to use some enhancement technique to address this issue as much as feasible. For

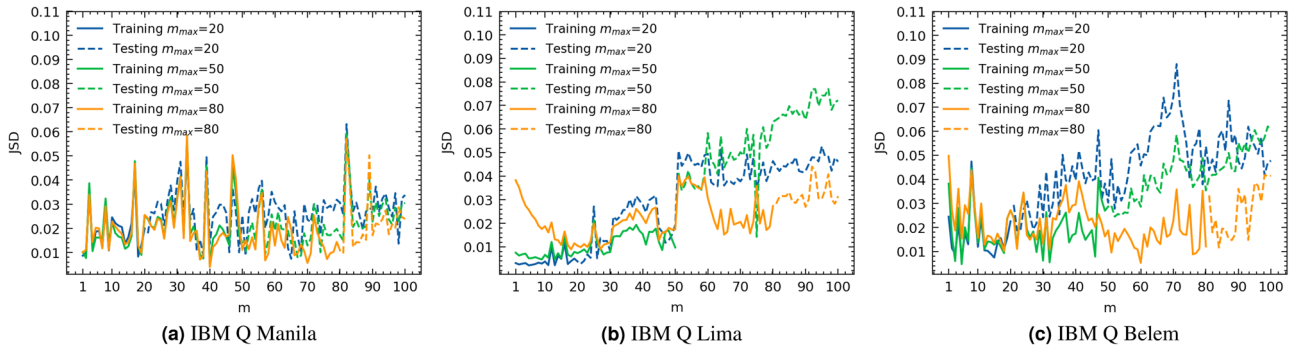
example, one way to do so is by utilizing unfolding techniques<sup>44</sup>. In this work, we did not enhance the results as we only experimented with relatively smaller values of  $m$ .

**Experiments.** In this section, we evaluate the accuracy of the model in Eq. 12 in predicting the average probability output,  $\hat{q}(m, |0\rangle)$ , for identity circuits of higher depths by estimating  $A_0$  and  $p(|0\rangle)$  using only simulations of lower depths identity circuits. Denote by  $q'(m, |0\rangle)$  the predicted average probability distribution obtained using Eq. 12. We select a *training set of depths*  $T = \{1, 2, \dots, m_{max}\}$  to estimate  $A_0$  and  $p$  using the EL protocol followed by the construction of the average gate error matrix  $M_0$  and SPAM error matrix  $N_0$  where  $M_{0ij} = p_{i \oplus j}(|0\rangle)$  and  $N_0 = W^{-1}A_0W$ . A new *testing set of depths*  $T' = \{m_{max} + 1, m_{max} + 2, \dots, 100\}$  is then selected where we compute the *Jensen–Shannon Divergence (JSD)* between  $\hat{q}(m', |0\rangle)$  and  $q'(m', |0\rangle) \forall m' \in T'$ . The *JSD* measures the similarity between the two probability distributions<sup>45</sup>. The lower the *JSD*, the closer the two distributions are. More information about the *JSD* can be found in Section II in the supplementary. Figure 1 presents the computed *JSD* for different quantum computers while varying  $m_{max}$ . Figure 2 presents the average and standard deviation for the test *JSD* values for the different quantum computers. The average test *JSD* varies between 0.024 and 0.056 for the different  $m_{max}$  values with lower average *JSD* values noted for high  $m$  for  $m_{max} = 80$  as indicated in Fig. 2b.

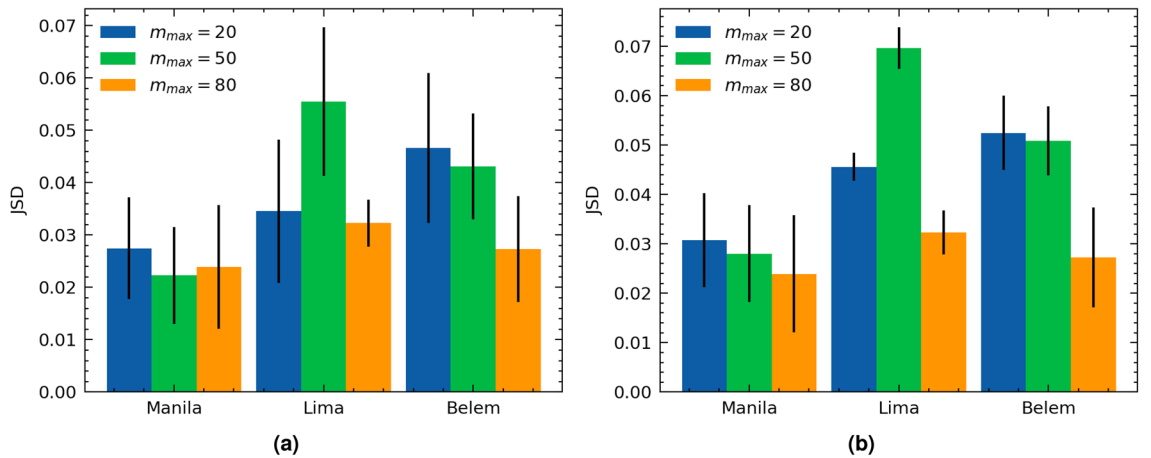
We rely on  $q'(m, |in\rangle)$  to construct and evaluate the mitigation power of  $Q_m$  for different depths. We first select a *training set of depths*  $T = \{1, 20, 40, 60, 80, 100\}$  to estimate  $A_{in}$  and  $p(|in\rangle)$  for each input state  $|in\rangle$  using the EL protocol followed by the construction of  $M_{in}$  using  $p(|in\rangle)$  and  $N_{in} = W^{-1}A_{in}W$ . We then estimate  $\hat{q}(m, |in\rangle)$  as  $q'(m, |in\rangle)$  for all inputs using Eq. 6 in order to construct  $Q_m$  using Eq. 5. We then choose a new *testing set of depths*  $T' = \{10, 30, 50, 70, 90\}$  so that  $Q_m$  is used in mitigating the outputs for circuits of depth  $m \in T'$  where for a given identity circuit of depth  $m$  with input  $|in\rangle$  and sequence  $s$  of gates, the mitigated output  $\hat{q}(m, s, |in\rangle)_{mit}$  is obtained as

$$\hat{q}(m, s, |in\rangle)_{mit} = Q_m^{-1} \hat{q}(m, s, |in\rangle) \tag{13}$$

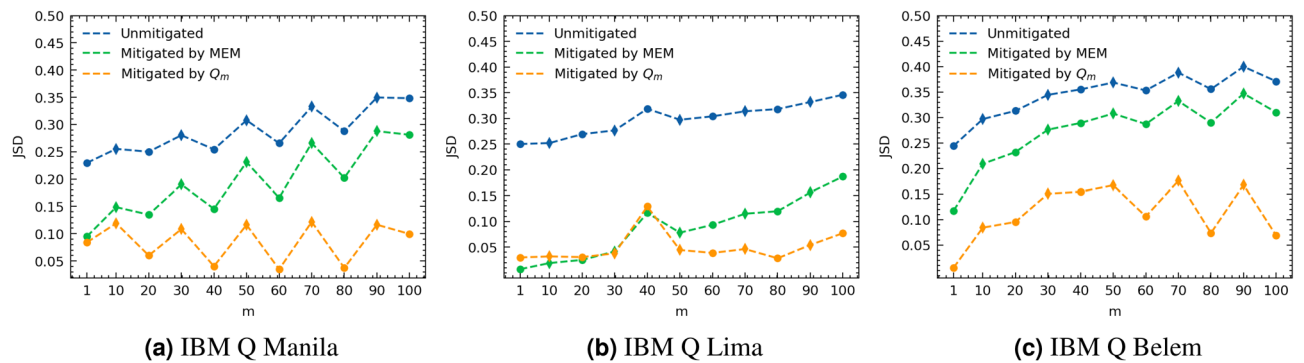
$\hat{q}(m, s, |in\rangle)_{mit}$  is projected onto a probability simplex to ensure a probability distribution. The *JSD* between  $\hat{q}(m, s, |in\rangle)_{mit}$  and the ideal output  $|in\rangle$  is computed and then averaged over all input states and all random circuits of depth  $m$ . We also compare our proposed mitigation protocol using  $Q_m$  with the MEM scheme (Fig. 3).



**Figure 1.**  $JSD(\hat{q}(m, |0\rangle), q'(m, |0\rangle))$  for training sets of depths  $T$  and testing sets of depths  $T'$  with variable maximum training depth  $m_{max} \in \{20, 50, 80\}$  on different IBM Q 5-qubit quantum computers.



**Figure 2.** The average and standard deviation of  $JSD(\hat{q}(m, |0\rangle), q'(m, |0\rangle))$ ; (a) over all depths  $m \in [m_{max} + 1, 100]$  and (b) over depths  $m \in [80, 100]$  while varying the maximum training depth  $m_{max}$  on different IBM Q 5-qubit quantum computers.



**Figure 3.** Average  $JSD$  between the ideal output  $|\mathbf{in}\rangle$  and each of the unmitigated output  $\hat{q}(m, s, |\mathbf{in}\rangle)$ , mitigated output by the MEM protocol, and mitigated output by our proposed noise model for each depth  $m$  on IBM Q 5-qubit quantum computers.

We report upto 88% improvement in the  $JSD$  value for the proposed approach compared to the unmitigated approach, and upto 69% improvement compared to MEM approach. Note that for the results presented here, we rely on the average SPAM free error rate,  $\mathbf{p}_{avg} = \frac{1}{2^m} \sum_{\mathbf{in}=0}^{2^m-1} \mathbf{p}(|\mathbf{in}\rangle)$  to construct  $\mathbf{M}_{in} = \mathbf{M}_{avg}$  for all inputs. We compare the results using  $\mathbf{p}_{avg}$  and  $\mathbf{p}(|\mathbf{in}\rangle)$  in the supplementary Section V.  $N_{in}$  remains input specific. Further elaborations on the results are presented in supplementary Section VI.

**Complexity.** So far in the estimation of  $\mathbf{M}_{in}$  and  $\mathbf{N}_{in}$  for each input state  $|\mathbf{in}\rangle$  using the EL protocol,  $K$  random circuits are generated for each depth  $1 \leq m \leq m_{max}$  where the EL protocol requires  $O(2^{2n})$  for the Walsh–Hadamard transform which can be reduced into  $O(n^2)$  using fast Walsh–Hadamard transform. Thus, the overall complexity of the construction of  $\mathbf{M}_{in}$  and  $\mathbf{N}_{in}$  for all input states is  $O(m_{max}Kn^22^n)$ . Furthermore, the GRF model factors the error rates vector into a product of  $\sim O(n)$  factors, depending on the architecture of the quantum computer, where each factor depends on a subset of adjacent qubits of cardinality  $N \ll n$  (typically  $N = 4$ ). Thus, the complexity is reduced further into  $O(nm_{max}KN^22^n)$ . The construction of  $\mathbf{Q}_m$  would be based on Eq. 6 for each input state  $|\mathbf{in}\rangle$  where  $\mathbf{M}_{in}^m$  can be computed efficiently using the Singular Value Decomposition (SVD) of  $\mathbf{M}_{in}$ , thus the construction of  $\mathbf{Q}_m$  is  $O(2^{3n})$ . For the MEM scheme, the construction of  $\mathbf{E}_{meas}$  requires only generating  $K$  circuits with no gates for each input state  $|\mathbf{in}\rangle$ , thus the complexity is  $O(K2^n)$ . For mitigation, both protocols are based on matrix inversion, thus the complexity for mitigation is  $O(2^{3n})$ .

## Discussion

The proposed mitigation protocol builds upon the SPAM-free noise characterization protocols for low circuit depths to generate a SPAM-error matrix  $\mathbf{N}_{in}$  and an average gate error matrix  $\mathbf{M}_{in}$  for each input state  $|\mathbf{in}\rangle$ . It then constructs a noise mitigation matrix  $\mathbf{Q}_m$  for arbitrary circuit depths  $m$  where the columns of  $\mathbf{Q}_m$  are the estimated average probability distributions  $\mathbf{q}'(m, |\mathbf{in}\rangle) = \mathbf{N}_{in}\mathbf{M}_{in}^m$ . The mitigated output  $\hat{q}(m, s)_{mit}$  of a given circuit of depth  $m$  with sequence  $s$  of gates is obtained by applying  $\mathbf{Q}_m^{-1}$  on the empirical circuit output  $\hat{q}(m, s)$ .

We evaluated the accuracy of our model in estimating the average probability distributions for high depth circuits and evaluated our mitigation protocol on the IBM Q 5-qubit quantum devices: Belem, Lima, and Manila. For the model accuracy evaluations, for the different  $m_{max}$  values, we reported on average a test  $JSD(\hat{q}(m, |\mathbf{0}\rangle), \mathbf{q}'(m, |\mathbf{0}\rangle))$  value around 0.022–0.028 for Manila, 0.03–0.055 for Lima, and 0.028–0.048 for Belem. For  $m_{max} = 20$ , the test  $JSD$  values varied between 0.005 and 0.05 for Lima computer, 0.01 and 0.06 for Manila computer, and 0.02 and 0.09 for Belem computer. We note that for  $m_{max} = 20$  the test spans  $m \in [21-100]$ . For higher depths  $m \in [80-100]$ , on average  $m_{max} = 80$  resulted in better model error than  $m_{max} = 50$  and  $m_{max} = 20$ . Results for IBM Q Athens are presented in the supplementary.

Finally, we report upto 88%  $JSD$  improvement for the proposed approach compared to the unmitigated approach with significant mitigation improvement compared to MEM at mid to higher depths. Specifically, for  $m = 90$ , we reported 58%, 66% and 85%  $JSD$  improvement for the proposed approach compared to the unmitigated on Belem, Manila, and Lima respectively. This is compared 12%, 17% and 51% respectively for the MEM. On average for all the test depths across the different machines, we report 68.4%  $JSD$  improvement for the proposed versus 38.2% for the MEM improvement compared to the unmitigated approach.

In addition, we compared the proposed mitigation technique against zero-noise Richardson extrapolation (ZNRE) method<sup>29</sup>. Our results shows superior performance over ZNRE. the full details of the comparison and the results can be found in the supplementary file in Section VII.

## Methods

In evaluating the accuracy of the model, we run  $K = 1000$  random identity circuits with each submission requesting 1024 shots for each depth  $m \in \{1, 2, \dots, 100\}$ . In evaluating the mitigation power of  $\mathbf{Q}_m$ , we run  $K = 1000$  random identity circuits for depths  $m \in \{1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$  for each basis input state  $|\mathbf{in}\rangle$  with each circuit requesting 1024 shots. The constructed circuits contain single qubit Clifford gates only. We run the circuits on the following IBM Q 5-qubit quantum computers: Manila, Lima, and Belem. Theoretical

derivations and numerical details essential to the study are presented in the Results section. More details can be found in the Supplementary Information. For the configurations and noise profiles of the IBM quantum machines, please go to IBM Quantum Experience at <http://www.research.ibm.com/quantum>.

## Data availability

The data that supports the findings of this study are available from the corresponding author upon reasonable request.

Received: 9 September 2021; Accepted: 24 February 2023

Published online: 08 March 2023

## References

- Harrow, A. W. & Montanaro, A. Quantum computational supremacy. *Nature* **549**, 203–209. <https://doi.org/10.1038/nature23458> (2017).
- Aaronson, S. & Chen, L. Complexity-theoretic foundations of quantum supremacy experiments, <https://doi.org/10.48550/ARXIV.1612.05903> (2016).
- Wehner, S., Elkouss, D. & Hanson, R. Quantum internet: A vision for the road ahead. *Science* **362**, eaam9288. <https://doi.org/10.1126/science.aam9288> (2018).
- Pirandola, B. Physics: Unite to build a quantum internet. *Nature* **532**, 169–171 (2016).
- Alexeev, Y. *et al.* Quantum computer systems for scientific discovery. *PRX Quantum* **2**, 017001. <https://doi.org/10.1103/prxquantum.2.017001> (2021).
- Georgescu, I., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.* **86**, 153–185. <https://doi.org/10.1103/revmodphys.86.153> (2014).
- Gyongyosi, L. & Imre, S. Training optimization for gate-model quantum neural networks. *Sci. Rep.* **9**, 1–19 (2019).
- King, A. D. *et al.* (2022) Coherent quantum annealing in a programmable 2000-qubit ising chain. *Nat. Phys.*, 18:1324–1328. <https://doi.org/10.48550/ARXIV.2202.05847>
- Gyongyosi, L. & Imre, S. A survey on quantum computing technology. *Comput. Sci. Rev.* **31**, 51–71. <https://doi.org/10.1016/j.cosrev.2018.11.002> (2019).
- Awschalom, D. *et al.* Development of quantum interconnects (QulCs) for next-generation information technologies. *PRX Quantum* **2**, 017002. <https://doi.org/10.1103/prxquantum.2.017002> (2021).
- Foxen, B. *et al.* Demonstrating a continuous set of two-qubit gates for near-term quantum algorithms. *Phys. Rev. Lett.* **125**, 120504. <https://doi.org/10.1103/physrevlett.125.120504> (2020).
- Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
- Almudever, C. G. *et al.* The engineering challenges in quantum computing. In *Proceedings of the Conference on Design, Automation & Test in Europe, DATE '17*, 836–845 (European Design and Automation Association, Leuven, BEL, 2017).
- Gupta, R. S., Edmunds, C. L., Milne, A. R., Hempel, C. & Biercuk, M. J. Adaptive characterization of spatially inhomogeneous fields and errors in qubit registers. *NPJ Quantum Inf.* **6**, 53 (2020).
- Hamilton, K. E. *et al.* Scalable quantum processor noise characterization, 2006.01805 (2020).
- Zaborniak, T. & de Sousa, R. In situ noise characterization of the d-wave quantum annealer. In *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 409–412, <https://doi.org/10.1109/QCE49297.2020.00057> (2020).
- Gottesman, D. An introduction to quantum error correction and fault-tolerant quantum computation, 0904.2557 (2009).
- Steane, A. M. A tutorial on quantum error correction. *Quantum Comput. Algorithms Chaos* 1–32 (2006).
- Gong, M. *et al.* Experimental exploration of five-qubit quantum error correcting code with superconducting qubits. *Natl. Sci. Rev.* **9**, nwab011 (2021).
- Chuang, I. L. & Nielsen, M. A. Prescription for experimental determination of the dynamics of a quantum black box. *J. Modern Opt.* **44**, 2455–2467 (1997).
- Emerson, J., Alicki, R. & Życzkowski, K. Scalable noise estimation with random unitary operators. *J. Opt. B Quantum Semiclassical Opt.* **7**, S347–S352 (2005).
- Knill, E. *et al.* Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
- Gambetta, J. M. *et al.* Characterization of addressability by simultaneous randomized benchmarking. *Phys. Rev. Lett.* **109**, 240504 (2012).
- McKay, D. C., Sheldon, S., Smolin, J. A., Chow, J. M. & Gambetta, J. M. Three-qubit randomized benchmarking. *Phys. Rev. Lett.* **122**, 200502. <https://doi.org/10.1103/PhysRevLett.122.200502> (2019).
- Suzuki, Y., Endo, S., Fujii, K. & Tokunaga, Y. Quantum error mitigation for fault-tolerant quantum computing, 2010.03887 (2021).
- Preskill, J. Quantum computing in the nisq era and beyond. *Quantum*, **2**, 79. <https://doi.org/10.22331/q-2018-08-06-79> (2018).
- Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493–R2496. <https://doi.org/10.1103/PhysRevA.52.R2493> (1995).
- Steane, A. M. Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77**, 793–797. <https://doi.org/10.1103/PhysRevLett.77.793> (1996).
- Kandala, A. *et al.* Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495. <https://doi.org/10.1038/s41586-019-1040-7> (2019).
- Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
- Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027. <https://doi.org/10.1103/PhysRevX.8.031027> (2018).
- Tannu, S. S. & Qureshi, M. K. Mitigating measurement errors in quantum computers by exploiting state-dependent bias. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture, MICRO '52*, 279–290, <https://doi.org/10.1145/3352460.3358265> (Association for Computing Machinery, New York, NY, USA, 2019).
- Czarnik, P., Arrasmith, A., Coles, P. J. & Cincio, L. Error mitigation with clifford quantum-circuit data. *Quantum* **5**, 592 (2021).
- Kim, C., Park, D. & Rhee, J.-K. Quantum error mitigation with artificial neural network. *IEEE Access* **8**, 188853–188860. <https://doi.org/10.1109/ACCESS.2020.3031607> (2020).
- Zlokapa, A. & Gheorghiu, A. A deep learning model for noise prediction on near-term quantum devices 2005.10811 (2020).
- Endo, S., Cai, Z., Benjamin, S. C. & Yuan, X. Hybrid quantum-classical algorithms and quantum error mitigation. *J. Phys. Soc. Japan* **90**, 032001. <https://doi.org/10.7566/jpsj.90.032001> (2021).
- Strikis, A., Qin, D., Chen, Y., Benjamin, S. C. & Li, Y. Learning-based quantum error mitigation. *PRX Quantum* **2**, 040330. <https://doi.org/10.1103/prxquantum.2.040330> (2021).
- Asfaw, A. *et al.* Learn quantum computation using qiskit (2020).
- Harper, R., Flammia, S. T. & Wallman, J. J. Efficient learning of quantum noise. *Nat. Phys.* **16**, 1184–1188 (2020).
- Knill, E. Quantum computing with realistically noisy devices. *Nature* **434**, 39–44. <https://doi.org/10.1038/nature03350> (2005).

41. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94**, 052325. <https://doi.org/10.1103/physreva.94.052325> (2016).
42. Ware, M. *et al.* Experimental pauli-frame randomization on a superconducting qubit. *Phys. Rev. A* **103**, 042604. <https://doi.org/10.1103/physreva.103.042604> (2021).
43. IBM Quantum Experience. <http://www.research.ibm.com/quantum>.
44. Nachman, B., Urbanek, M., de Jong, W. A. & Bauer, C. W. Unfolding quantum computer readout noise. <https://doi.org/10.48550/ARXIV.1910.01969> (2019).
45. Menéndez, M., Pardo, J., Pardo, L. & Pardo, M. The jensen-shannon divergence. *J. Franklin Inst.* **334**, 307–318. [https://doi.org/10.1016/S0016-0032\(96\)00063-4](https://doi.org/10.1016/S0016-0032(96)00063-4) (1997).

### Author contributions

A.S., M.F. and R.K. conceived the idea and designed the study. A.S. and M.N. collected the data and run the simulations. A.S. and R.K. prepared the manuscript with input and critical feedback from all authors. All authors analyzed and discussed the results. All authors critically reviewed and developed the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-30510-5>.

**Correspondence** and requests for materials should be addressed to M.E.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023