Lawrence Berkeley National Laboratory

Title
Mitigating Depolarizing Noise on Quantum Computers with Noise-Estimation Circuits

Authors
Urbanek, Miroslav
Nachman, Benjamin
Pascuzzi, Vincent R
et al.

Mitigating depolarizing noise on quantum computers with noise-estimation circuits

Miroslav Urbanek,[1, *] Benjamin Nachman,[2] Vincent R. Pascuzzi,[2]
Andre He,[2, †] Christian W. Bauer,[2] and Wibe A. de Jong[1, ‡]

[1]*Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[2]*Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

A significant problem for current quantum computers is noise. While there are many distinct noise channels, the depolarizing noise model often appropriately describes average noise for large circuits involving many qubits and gates. We present a method to mitigate the depolarizing noise by first estimating its rate with a noise-estimation circuit and then correcting the output of the target circuit using the estimated rate. The method is experimentally validated on a simulation of the Heisenberg model. We find that our approach in combination with readout-error correction, randomized compiling, and zero-noise extrapolation produces close to exact results even for circuits containing hundreds of $CNOT$ gates. We also show analytically that zero-noise extrapolation is improved when it is applied to the output of our method.

Noisy intermediate-scale quantum (NISQ) computers [1] are current and near-term quantum computers that are not fault-tolerant. The presence of noise and errors limits their utility. Even quantum algorithms designed for NISQ devices, for example the variational quantum eigensolver [2], are hampered by imperfections of real devices. Error rates are still too large to solve relevant scientific problems on existing quantum computers. Consequently, there has been a lot of effort to reduce noise and mitigate errors present on these devices.

An important class of errors are readout errors. They manifest themselves as readouts of incorrect qubit values during a measurement, e.g., reading one while the qubit is in the zero state and vice versa. Readout errors can be successfully mitigated with readout-error correction. Several method with various degree of sophistication have been developed [3–8].

Another large source of errors are gate errors. They can be classified into coherent and incoherent errors. Coherent errors preserve state purity. They are typically small miscalibrations in control parameters. Coherent errors usually produce similar errors in consecutive executions of a quantum circuit and lead to a systematic bias in the output. Incoherent errors can be understood as either coherent errors with randomly varying control parameters or as processes that entangle the system with its environment. Incoherent errors are easier to handle than coherent errors, because they can often be modeled as depolarizing noise. A method for converting coherent errors into incoherent errors is randomized compiling [9–11]. Its benefits for NISQ computers have been demonstrated experimentally [12–14].

Another practical technique of error mitigation is zero-noise extrapolation [15–25]. A circuit is executed multiple times with various degrees of noise and the measured output is extrapolated to the zero-noise limit. Recently, a method to mitigate the depolarizing noise using randomized measurements has been demonstrated [26]. Other mitigation methods have been developed as well [27–40].

In this letter, we introduce a new mitigation method.

From a given quantum circuit, which we call a target circuit, we construct a circuit with a similar structure that we call an estimation circuit. We execute the estimation circuit to measure the depolarizing noise rate and then use the measured rate to correct the output of the target circuit. We experimentally demonstrate that the combination of readout-error correction, randomized compiling, mitigation with estimation circuits, and zero-noise extrapolation produces results that are very close to the exact results.

We first describe the method, introduce a simple class of estimation circuits, and present our full mitigation protocol. We then show improvements obtained for our test case, which is a simulation of the Heisenberg model. Finally, we develop a theoretical argument showing that zero-noise extrapolation performed after mitigation with estimation circuits increases the extrapolation accuracy.

Methods.—A simple model of incoherent noise is the depolarizing noise model given by [41]

$$\epsilon(\rho) = (1-p)\rho + p\frac{I}{2^n}, \qquad (1)$$

where ϵ denotes the noise channel, ρ is the density matrix, p is the probabilistic error rate that depends on the device and also on the circuit, and n is the number of qubits. Notice that if ρ is initially a pure state, one can reconstruct the initial state from the noisy density matrix. For $p > 0$, the initial pure state is the state with the largest weight in $\epsilon(\rho)$. Alternatively, if one knows p, the initial density matrix ρ can be reconstructed simply by calculating the inverse $\epsilon^{-1}(\rho)$.

Observables are given by Hermitean operators acting on the system Hilbert space. They can be decomposed into sums of strings of identity and Pauli matrices,

$$O = \sum_i c_i \prod_{j=1}^{n} \sigma^{i,j}, \qquad (2)$$

where c_i are real coefficients and $\sigma^{i,j} \in \{I, \sigma_x, \sigma_y, \sigma_z\}$ are identity or Pauli matrices acting on qubit j. The trace

of $S = \prod_{j=1}^{n} \sigma^{i,j}$ is either $\mathrm{tr}(S) = 2^n$ if S is a product of identity matrices or $\mathrm{tr}(S) = 0$ otherwise. The expectation value of an observable O for a state represented by a density matrix ρ is $\langle O \rangle = \mathrm{tr}(\rho O)$. The expectation value of O for a noisy density matrix (1) is therefore given by

$$\overline{\langle O \rangle} = \mathrm{tr}[\epsilon(\rho)O] = (1-p)\langle O \rangle + \frac{p}{2^n}\,\mathrm{tr}(O), \qquad (3)$$

where we denote the noisy expectation value by an overline. Notice that $\overline{\langle S \rangle} = 1$ for strings S consisting of identity matrices only and $\overline{\langle S \rangle} = (1-p)\langle S \rangle$ otherwise.

We can therefore decompose any observable O as $O = cI + O'$, where c is a constant, I is the identity operator, and $\mathrm{tr}(O') = 0$. Its expectation value is $\langle O \rangle = c + \langle O' \rangle$. If we assume that the system decoherence is well described by the depolarizing noise model and if we know p, we can correct a noisy expectation value by calculating

$$\langle O \rangle = \frac{\overline{\langle O \rangle} - cp}{1-p}, \qquad (4)$$

where $\langle O \rangle$ is the corrected expectation value. We assume $c = 0$ without loss of generality in the following, because c is just a constant shift of the expectation value known in advance.

To correct the expectation value of any observable under the depolarizing noise model, we have to estimate the value of p. We do it by executing a circuit that is similar to our target circuit but has a known output. We assume that the target circuit consists of single-qubit and $CNOT$ gates only and that $CNOT$ gates are the leading source of gate errors. Our approach to construct an estimation circuit is to remove all single-qubit gates from the target circuit and to keep only the $CNOT$ gates in it. Since the initial state on a quantum computer is the zero state, ideal $CNOT$ gates do not transform the initial state at all. The final state is again a zero state on an ideal quantum computer. We can therefore estimate $1 - p$ by measuring the probability of obtaining the zero state with the estimation circuit. The main assumption is the estimation and the target circuit are affected by a similar p because they have similar structures.

It is not always necessary to remove all single-qubit gates. The estimation circuit can be any circuit that has a known output sensitive to noise and that has a similar structure as the target circuit. It may be beneficial to preserve some single-qubit gates to keep it similar to the target circuit or add extra gates to reduce systematic errors. We therefore add a layer of random rotations as the first circuit layer and its inverse as the last circuit layer to increase the robustness of the estimation.

An alternative recent approach uses near-Clifford circuits, which one can simulate classically, to perform the mitigation [42]. The main difficulty of this approach is that the output of random near-Clifford circuits is similar to an output obtained with a completely mixed density
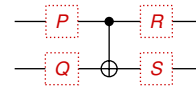

FIG. 1. Randomized compiling. Each $CNOT$ gate is dressed with the P, Q, R, and S gates.

TABLE I. Gate choices in randomized compiling. Each P, Q, R, and S assignment produces a dressed gate equal to a $CNOT$ gate. An assignment is chosen independently and randomly for each $CNOT$ gate in the circuit.

P	Q	R	S	P	Q	R	S	P	Q	R	S	P	Q	R	S
I	I	I	I	Y	I	Y	X	X	I	X	X	Z	I	Z	I
I	X	I	X	Y	X	Y	I	X	X	X	I	Z	X	Z	X
I	Y	Z	Y	Y	Y	X	Z	X	Y	Y	Z	Z	Y	I	Y
I	Z	Z	Z	Y	Z	X	Y	X	Z	Y	Y	Z	Z	I	Z

matrix. One therefore has to select a particular subset of circuits that produce biased outputs. The authors used machine learning to find appropriate near-Clifford circuits with this property. Our method does not require any such selection. We simply remove single-qubit gates to obtain a biased circuit that can be simulated trivially.

We implemented our method in combination with readout-error correction, randomized compiling, and zero-noise extrapolation. Readout-error correction is performed using the unfolding method [8, 43].

Coherent errors are dominant gate errors. They are not covered well by the depolarizing noise model. Randomized compiling [9] can convert coherent errors into incoherent errors. In particular, we consider single-qubit gates being the easy gates and $CNOT$ gates being the hard gates. We perform randomized compiling by inserting a layer of randomizing single-qubit gates before and after each layer of $CNOT$ gates as shown in Fig. 1. The randomizing gates are the identity and the Pauli gates. Each $CNOT$ gate is preceded and succeeded by a pair of gates so that the overall action of the four single-qubit gates and a $CNOT$ gate is exactly equal to a $CNOT$ gate. All possible gate choices are listed in Table I. The layer of single-qubit gates after a layer of $CNOT$ gates can be composed with a layer of single-qubit gates before the next layer of $CNOT$ gates. The circuit structure therefore consists of layers of $CNOT$ gates interspersed with layers of single qubit gates. We use randomized compiling for both the estimation and the target circuit.

In the original formulation of zero-noise extrapolation [16, 19], the authors varied gate duration assuming that slower gates have larger errors. They ran experiments for several duration values and extrapolated the measured results to the zero duration. A pulse-level control is required to implement this method. The technique has been extended to systems with gate-level control [17]. The main assumption is that $CNOT$ gates are the domi-

nant source of errors. The authors replaced each *CNOT* gate with a sequence of three or five *CNOT* gates, which are equivalent to a single *CNOT* gate, executed their circuits, and extrapolated to the zero-gate limit. This idea has been further extended to replace only a subset of *CNOT* gates with sequences of *CNOT* gates [21]. Both methods assume a certain dependence of errors on the number of *CNOT* gates. In this work, we execute three versions of each circuit, where each *CNOT* gate is replaced by one, three, or five consecutive *CNOT* gates, and perform quadratic extrapolation to the limit corresponding to zero *CNOT* gates.

Experiment.—Our test case is time evolution of the Heisenberg model. We consider a quench of a one-dimensional XX chain of noninteracting spin-1/2 particles [44]. Its Hamiltonian is given by

$$H = -J \sum_{j=1}^{n-1} \left(\sigma_x^j \sigma_x^{j+1} + \sigma_y^j \sigma_y^{j+1} \right), \qquad (5)$$

where $J > 0$ is a coupling constant, and σ_x^i and σ_y^i are Pauli matrices acting on qubit j. The system is initially prepared in a domain-wall configuration $|\psi_0\rangle = |\dots 111000 \dots\rangle$ with qubits in the first and second half of the chain in the one and zero state, respectively. We consider $\hbar = 1$ for simplicity in the following.

The propagator $\exp(-iHt)$ is approximated by its second-order Trotter–Suzuki decomposition [45, 46] to enable its implementation on a quantum computer. The approximated propagator for one time step is given by

$$e^{-iH\delta t} \approx e^{-iF\delta t/2} e^{-iG\delta t} e^{-iF\delta t/2}, \qquad (6)$$

where F and G contain terms in H that act only on even and odd bonds, respectively, and δt is a time step. Since all terms in both F and G commute with each other, we can decompose the exponentials in Eq. (6) into products of two-qubit exponentials of the form $\exp[iJ(\sigma_x^j \sigma_x^{j+1} + \sigma_y^j \sigma_y^{j+1})\delta t/d]$, where $d \in \{1, 2\}$. Each such exponential can be implemented by a circuit consisting of two *CNOT* gates and a number of single-qubit gates. One time step is therefore implemented by three layers of two-qubit circuits acting on even, odd, and even bonds. Each two-qubit circuit is subsequently decomposed into two *CNOT* and multiple single-qubit gates. We measure the time evolution of the local magnetization of the last spin in the chain, $M_n(t) = \langle \psi(t) | \sigma_z^n | \psi(t) \rangle$.

We implemented this model on the IBM Q Paris device using six qubits Q23, Q24, Q25, Q22, Q19, and Q20 with 8192 shots for each circuit. The circuit is shown in Fig. 2. It contains 14 *CNOT* gates per time step. The longest circuit for 15 time steps contains 210 *CNOT* gates.

We created the estimation circuit from the target circuit by removing all its single-qubit gates. A layer of random single-qubit gates and a layer of their inverses were added as the first and the last circuit layer, respectively. These gates have no effect in a noiseless calculation and help to reduce systematic errors in real gate implementations. Versions with one, three, and five *CNOT* gates per each *CNOT* gate were created. We then constructed and executed 448 randomized instances of each circuit. Readout errors were corrected using the unfolding technique as the first step in data processing. All expectation values were averaged over the randomized instances.

We estimated $1 - p$ by measuring the $\langle \sigma_z^6 \rangle$ expectation value with estimation circuits. Ideally, $\langle \sigma_z^6 \rangle = 1$, so the depolarizing noise rate is given by $1 - p = \langle \sigma_z^6 \rangle$. The mitigation was performed using Eq. (4) with $c = 0$. We then applied zero-noise extrapolation. Data points obtained with circuits with $r = 1$, 3, and 5 *CNOT* gates were extrapolated to $r = 0$ using a quadratic fit. The final results are shown in Fig. 3. The maximal error obtained with the target and with the mitigated circuit is 0.39 and 0.11, respectively, in both cases at $t = 2.4$.

So far, we have considered only a global depolarizing noise channel that affects all qubits in a circuit. The improvement obtained using zero-noise extrapolation on top of the mitigation with estimation circuits can be better understood by considering a local depolarizing model. A local depolarizing noise channel affects only qubits a given quantum gate is applied to. Since error rates of single-qubit gates are typically significantly lower than noise rates of entangling gates, it suffices to consider a noise channel that affects only *CNOT* gates. The first-order approximation of the noisy expectation value obtained with the target circuit in this model is given by [21, 47]

$$\overline{\langle O_t(r) \rangle} = (1-\varepsilon)^{rN_C} \sum_{i=0}^{N_C} \left[(1-\varepsilon)^{-r} - 1 \right]^i \langle O_t \rangle_i$$
$$\approx \langle O \rangle \left[1 + r\varepsilon \left(\frac{\langle O_t \rangle_1}{\langle O_t \rangle_0} - N_C \right) \right], \qquad (7)$$

where r is the noise factor, ε is the depolarizing noise rate of a single *CNOT* gate, N_C is the number of *CNOT* gates, and $\langle O_t \rangle_i$ are sums of expectation values of all circuits where i *CNOT* gates have been replaced by noise channels that output maximally-mixed states on respective pairs of qubits. Here, $r = 1$, 3, and 5 for circuits with one, three, and five *CNOT* gates in place of each individual *CNOT* gate. The form of Eq. (7) follows from Eq. (1) where the global depolarizing noise with rate p is replaced by a local depolarizing noise with rate ε. A local noise channel associated with a *CNOT* gate affects only the Hilbert subspace corresponding to the gate qubits. Maximally-mixed term $I/2^n$ in Eq. (1) is therefore replaced by a term that is maximally-mixed only in this subspace. Eq. (7) is then obtained by considering noise channels for all *CNOT* gates and collecting individual terms into $\langle O_t \rangle_i$. A detailed discussion can be found in Ref. [21].

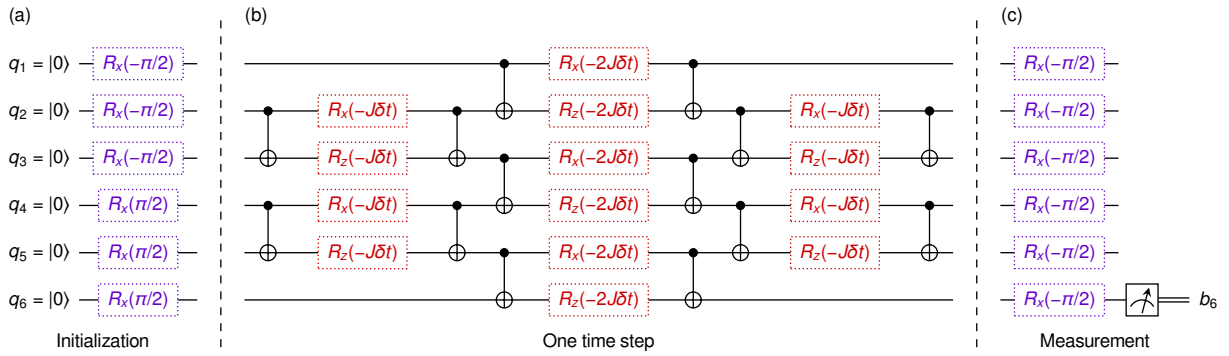The expectation value for the mitigated circuit is given

FIG. 2. Quantum circuit for the simulation of the XX chain. a) Preparation of the initial domain-wall state and basis transformation to a convenient basis. The dotted gates were replaced by random rotations in the estimation circuit. b) One step of the time evolution. Multiple steps are obtained by repeating this subcircuit. The dotted gates were removed in the estimation circuit. c) Basis transformation and measurement of the last qubit. The dotted gates were replaced by the inverses of the random rotations from the initialization step in the estimation circuit.
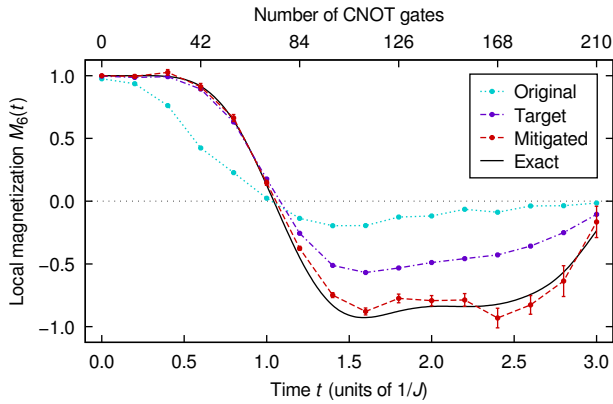
FIG. 3. Comparison of the original and mitigated results for the time evolution of the local magnetization in the XX chain. The original results were obtained using the original circuits without any mitigation. There are 14 $CNOT$ gates per time step and the longest original circuit contains 210 $CNOT$ gates. Target results use readout-error correction, randomized compiling, and zero-noise extrapolation. Mitigated results use readout-error correction, randomized compiling, mitigation with estimation circuits, and zero-noise extrapolation. Data for extrapolation were obtained with circuits where each $CNOT$ gate was replaced by one, three, or five $CNOT$ gates. Each circuit was executed with 448 random instances. Error bars represent the standard error of processed data. Exact solution takes the Trotter–Suzuki decomposition into account.

by

$$\overline{\langle O_\mathrm{m}(r)\rangle} = \overline{\langle O_\mathrm{t}(r)\rangle} \frac{\langle O_\mathrm{e}\rangle}{\langle O_\mathrm{e}(r)\rangle}$$
$$\approx \langle O\rangle \left[1 + r\varepsilon\left(\frac{\langle O_\mathrm{t}\rangle_1}{\langle O_\mathrm{t}\rangle_0} - \frac{\langle O_\mathrm{e}\rangle_1}{\langle O_\mathrm{e}\rangle_0}\right)\right], \quad (8)$$

where $\langle O_\mathrm{m}\rangle$ and $\langle O_\mathrm{e}\rangle$ denote the expectation values measured with the mitigated and the estimation circuit, respectively. Notice that $\langle O_\mathrm{t}\rangle_0 = \langle O\rangle$ and $\langle O_\mathrm{e}\rangle_0 = \langle O_\mathrm{e}\rangle$.

The N_C term in Eq. (7) is replaced by $\langle O_\mathrm{e}\rangle_1/\langle O_\mathrm{e}\rangle_0$ in Eq. (8). The error in the mitigated expectation value is small when $\langle O_\mathrm{t}\rangle_1/\langle O_\mathrm{t}\rangle_0$ is close to $\langle O_\mathrm{e}\rangle_1/\langle O_\mathrm{e}\rangle_0$. The more the target and the estimation circuits are similar, the smaller the difference between these two terms is. Even though $\langle O_\mathrm{e}\rangle_1/\langle O_\mathrm{e}\rangle_0$ depends on a given estimation circuit in a nontrivial way, in most cases $|\langle O_\mathrm{t}\rangle_1/\langle O_\mathrm{t}\rangle_0 - \langle O_\mathrm{e}\rangle_1/\langle O_\mathrm{e}\rangle_0|$ is smaller than $|\langle O_\mathrm{t}\rangle_1/\langle O_\mathrm{t}\rangle_0 - N_C|$, since the target and the estimation circuit have similar structures. The same argument can be made for higher-order terms, so expressions that depend on N_C^i in target expectation values are replaced by $\langle O_\mathrm{e}\rangle_i/\langle O_\mathrm{e}\rangle_0$ in mitigated expectation values. Mitigated expectation values are therefore a better input to zero-noise extrapolation than target expectation values. We constructed estimation circuits by removing all single-qubit gates. One can eventually construct better estimation circuits by removing a subset of single-qubit gates to create near-Clifford circuits that can still be efficiently simulated on classical computers [42].

Conclusion.—We presented a method to mitigate errors and noise on quantum computers that are characterized by the depolarizing noise model. The method prescribes a construction of an estimation circuit to estimate the noise rate that is then used to correct the output of a given circuit. A crucial part of this approach is the randomized compiling that ensures that gate errors can be modeled as incoherent depolarizing noise. We demonstrated that the method works well, especially in combination with readout-error correction and zero-noise extrapolation, on a set of test circuits containing hundreds of $CNOT$ gates. The method is scalable to any number of qubits and gates given that enough randomized samples are collected to achieve low uncertainty [47].

———————

* Corresponding author: urbanek@lbl.gov
† Present address: IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA
‡ Corresponding author: wadejong@lbl.gov

[1] J. Preskill, Quantum **2**, 79 (2018).
[2] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).
[3] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).
[4] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Quantum **4**, 257 (2020).
[5] S. Bravyi, S. Sheldon, A. Kandala, D. C. McKay, and J. M. Gambetta, Phys. Rev. A **103**, 042605 (2021).
[6] L. Funcke, T. Hartung, K. Jansen, S. Kühn, P. Stornati, and X. Wang, Measurement error mitigation in quantum computers through classical bit-flip correction, arXiv (2020), arXiv:2007.03663 [quant-ph].
[7] J. W. O. Garmon, R. C. Pooser, and E. F. Dumitrescu, Phys. Rev. A **101**, 042308 (2020).
[8] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer, npj Quantum Inf. **6**, 84 (2020).
[9] J. J. Wallman and J. Emerson, Phys. Rev. A **94**, 052325 (2016).
[10] Z. Cai and S. C. Benjamin, Sci. Rep. **9**, 11281 (2019).
[11] Z. Cai, X. Xu, and S. C. Benjamin, npj Quantum Inf. **6**, 17 (2020).
[12] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor (2021), arXiv:2010.00215 [quant-ph].
[13] J.-L. Ville, A. Morvan, A. Hashim, R. K. Naik, B. Mitchell, J.-M. Kreikebaum, K. P. O'Brien, J. J. Wallman, I. Hincks, J. Emerson, E. Smith, E. Younis, C. Iancu, D. I. Santiago, and I. Siddiqi, Leveraging randomized compiling for the QITE algorithm, arXiv (2021), arXiv:2104.08785 [quant-ph].
[14] M. Ware, G. Ribeill, D. Ristè, C. A. Ryan, B. Johnson, and M. P. da Silva, Phys. Rev. A **103**, 042604 (2021).
[15] Y. Li and S. C. Benjamin, Phys. Rev. X **7**, 021050 (2017).
[16] K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180509 (2017).
[17] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Phys. Rev. Lett. **120**, 210501 (2018).
[18] S. Endo, S. C. Benjamin, and Y. Li, Phys. Rev. X **8**, 031027 (2018).
[19] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature **567**, 491 (2019).
[20] S. McArdle, X. Yuan, and S. Benjamin, Phys. Rev. Lett. **122**, 180501 (2019).
[21] A. He, B. Nachman, W. A. de Jong, and C. W. Bauer, Phys. Rev. A **102**, 012426 (2020).
[22] J. Sun, X. Yuan, T. Tsunoda, V. Vedral, S. C. Benjamin, and S. Endo, Phys. Rev. Applied **15**, 034026 (2021).
[23] F. G. Fuchs, V. Falch, and C. Johnsen, Eur. Phys. J. Plus **135**, 353 (2020).
[24] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, Learning-based quantum error mitigation, arXiv (2020), arXiv:2005.07601 [quant-ph].
[25] S. Tomkins and R. de Sousa, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (2020) pp. 413–418.
[26] J. Vovrosh, K. E. Khosla, S. Greenaway, C. Self, M. Kim, and J. Knolle, Efficient mitigation of depolarizing errors in quantum simulations, arXiv (2021), arXiv:2101.01690 [quant-ph].
[27] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Phys. Rev. A **98**, 062339 (2018).
[28] I. Schwenk, J.-M. Reiner, S. Zanker, L. Tian, J. Leppäkangas, and M. Marthaler, Phys. Rev. A **97**, 042310 (2018).
[29] J.-M. Reiner, S. Zanker, I. Schwenk, J. Leppäkangas, F. Wilhelm-Mauch, G. Schön, and M. Marthaler, Quantum Sci. Technol. **3**, 045008 (2018).
[30] S. Endo, Q. Zhao, Y. Li, S. Benjamin, and X. Yuan, Phys. Rev. A **99**, 012334 (2019).
[31] M. Otten and S. K. Gray, Phys. Rev. A **99**, 012338 (2019).
[32] M. Otten and S. K. Gray, npj Quantum Inf. **5**, 11 (2019).
[33] M. Otten, C. L. Cortes, and S. K. Gray, Noise-resilient quantum dynamics using symmetry-preserving ansatzes, arXiv (2019), arXiv:1910.06284 [quant-ph].
[34] D. C. Murphy and K. R. Brown, Phys. Rev. A **99**, 032318 (2019).
[35] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo, Phys. Rev. A **100**, 010302(R) (2019).
[36] S. E. Smart and D. A. Mazziotti, Phys. Rev. A **100**, 022517 (2019).
[37] C. Song, J. Cui, H. Wang, J. Hao, H. Feng, and Y. Li, Sci. Adv. **5**, eaaw5686 (2019).
[38] S. Zhang, Y. Lu, K. Zhang, W. Chen, Y. Li, and K. Zhang, Jing-Ningand Kim, Nat. Commun. **11**, 587 (2020).
[39] A. Roggero and A. Baroni, Phys. Rev. A **101**, 022328 (2020).
[40] A. Montanaro and S. Stanisic, Error mitigation by training with fermionic linear optics, arXiv (2021), arXiv:2102.02120 [quant-ph].
[41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).
[42] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Error mitigation with Clifford quantum-circuit data (2021), arXiv:2005.10189 [quant-ph].
[43] M. Urbanek, B. Nachman, and W. A. de Jong, Phys. Rev. A **102**, 022427 (2020).
[44] A. Smith, M. S. Kim, F. Pollmann, and J. Knolle, npj Quantum Inf. **5**, 106 (2019).
[45] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
[46] M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).
[47] See Supplemental Material for a detailed discussion.