Improving quantum simulation performance via Trotter error mitigation

IBM Open Science Prize 2021

Gideon Lee,[1] Shantanu Jha,[2] and Shoumik Chowdhury[2]

[1] *Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637, USA*

[2] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

(Dated: April 23, 2022)

Quantum simulation often involves complicated gates that may be hard to implement on current hardware. One solution to this is to break down the Hamiltonian of interest into smaller parts that can be simulated using gates native to hardware via trotterization. However, trotterization leads to fundamental trotterization errors for any finite number of trotter steps even in the noiseless case due to non-commuting components. Current error mitigation techniques can be used to reduce effects of noise but cannot go beyond this limit. In this work, we propose a means, Trotter error mitigation (TEM) to go beyond the finite trotterization limit by using data points from finite trotter steps to extrapolate to the infinitesimally small step case. In order to make the measured data coherent enough for TEM, we also apply a slew of standard error mitigation techniques. TEM is not unique to the Heisenberg model simulation at hand but generalizes to any trotterization problem.

1. Introduction

Quantum simulators are platforms that seek to reproduce the physical behavior and/or dynamical evolution of other quantum systems in a programmable fashion [1, 2]. As one of the first proposed applications of quantum computers [3–5], quantum simulation has received considerable scientific attention [6–13], and remains among the most promising near-term avenues for "quantum advantage" [12]. While many *analog* quantum simulators have been proposed [14–23] and implemented [24–34], these platforms are often specially designed to tackle only specific problems. By contrast, *digital* quantum simulators run on quantum computing machines are, in principle, universal, allowing us to simulate the Hamiltonian dynamics of a much wider class of systems [1]. Many such simulators have been implemented, most notably using trapped-ion and superconducting circuit platforms. However, it remains an outstanding challenge to improve the quality of quantum simulation given the current error rates in near-term devices.

1.1. Simulation of the XXX Heisenberg Model

In this work, we will study the XXX Heisenberg spin model, defined via the Hamiltonian:

$$\hat{H}_{\text{Heis}} = \sum_{\langle ij \rangle}^{N} J \left(\hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j + \hat{Z}_i \hat{Z}_j \right). \quad (1)$$

Here $\hat{X}_i, \hat{Y}_i, \hat{Z}_i$ are the Pauli operators for a spin-1/2 particle, and J is the interaction strength. For the IBM Open Science Prize, we focus on the case of $N = 3$ spins. Simulating the dynamics of the corresponding system represents a challenging task for noisy intermediate-scale quantum (NISQ) hardware; however, given the small Hilbert space size of $2^N = 8$, we can still simulate the problem classically. For the remainder of this work, we will assume $N = 3$.

Under the Hamiltonian above, the unitary dynamics are governed by the propagator $\hat{U}_{\text{heis}} = e^{-it\hat{H}_{\text{heis}}}$. The goal of quantum simulation, then, is to approximate this unitary by native single- and two-qubit gates available on the quantum computer. In this challenge, we use the `ibmq-jakarta` device from IBM. Here '`jakarta`' is a 7-qubit superconducting quantum processor available through the cloud. In order to simulate the Heisenberg model dynamics, we implement the Trotterized evolution of $\hat{U}_{\text{heis}}(t)$ up to $t = \pi$, and run on qubits Q_1, Q_3, Q_5 of the device. Given our knowledge of the actual dynamics from classical simulation (cf. Fig. 1), we can then evaluate the performance of our scheme by performing quantum state tomography.

2. Theory

For convenience, we define

$$\hat{S}^{(ij)} = \hat{X}_i \hat{X}_j + \hat{Y}_i \hat{Y}_j + \hat{Z}_i \hat{Z}_j \quad (2)$$

and $\hat{\xi} = [\hat{S}^{(01)}, \hat{S}^{(12)}]$. The relevant operator in the Heisenberg Hamiltonian is $\hat{S}^{(H)} = \hat{S}^{(01)} + \hat{S}^{(12)}$. Now, setting $J = 1$, the time evolution in the Heisenberg

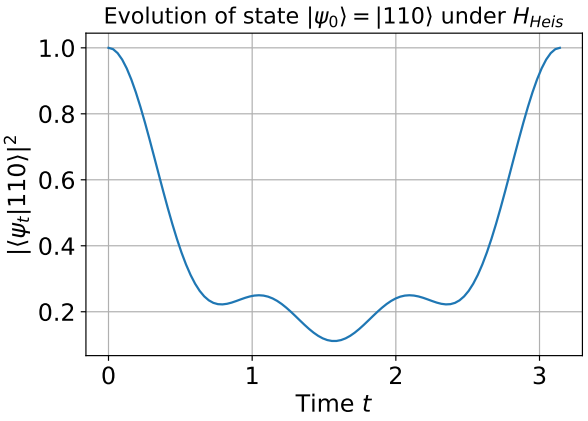

FIG. 1. Time-evolution of the initial state $|\psi_0\rangle = |110\rangle$ under the action of $\hat{H}_{\text{heis}}$. The state $|\psi_t\rangle$ satisfies the Schrödinger equation $i\hbar\partial_t |\psi_t\rangle = \hat{H}_{\text{heis}} |\psi_t\rangle$. Note that at $t = \pi$, the state returns to itself, $|\psi_\pi\rangle = |\psi_0\rangle$.

model is governed by the unitary

$$\hat{U}_{\text{heis}}(t) = \exp\left[-it\hat{S}^{(H)}\right]. \tag{3}$$

Performing a first-order Lie-Trotter decomposition [35–37] gives:

$$\hat{U}_{\text{trot}}(t, N) = \left[e^{-i(t/N)\hat{S}^{(01)}} e^{-i(t/N)\hat{S}^{(12)}}\right]^N \tag{4}$$

Next, using the Baker-Campbell-Hausdorff (BCH) and Zassenhaus formulae, we may rewrite the individual trotter gate as,

$$\begin{aligned} &\exp\left[-i\frac{t}{N}S^{(01)}\right]\exp\left[-i\frac{t}{N}S^{(12)}\right] \\ &= \exp\left[-i\frac{t}{N}S^{(H)}\right]\exp\left[\frac{1}{2}\frac{t^2}{N^2}\xi\right] + O\left(\frac{t^3}{N^3}\right) \\ &\simeq \exp\left[-i\frac{t}{N}S^{(H)} + \frac{1}{2}\frac{t^2}{N^2}\xi\right] + O\left(\frac{t^3}{N^3}\right) \end{aligned} \tag{5}$$

Taking the Nth power,

$$\hat{U}_{trot}(t, N) \simeq \exp\left(-it\hat{S}^{(H)} + \frac{t^2}{2N}\hat{\xi}\right) + O\left(\frac{t^2}{N^2}\right) \tag{6}$$

For the purposes of error mitigation, the primary object of interest is the deviation of the expectation value of a given Hermitian observable $\mathcal{O}$ from its ideal value. Here, we consider the case of errors solely due to trotterization. For a given initial state ρ_0, evolution time t, and Trotter steps N, the Trotter error in $\langle\mathcal{O}\rangle$ may be quantified as:

$$\Delta\mathcal{O}(t, N) \equiv \text{Tr}\left[\mathcal{O}\left(\hat{U}_{\text{heis}}\rho_0\hat{U}^\dagger_{\text{heis}} - \hat{U}_{\text{trot}}\rho_0\hat{U}^\dagger_{\text{trot}}\right)\right] \tag{7}$$

In principle, one can calculate the exact form of $\Delta\mathcal{O}$ given a specific observable $\mathcal{O}$ and state ρ_0. This is possible using known tools such as the matrix polynomial form of the Zassenhaus formula [38]. For practical reasons, however, we simply note that the form of Eq. 6 suggests that $\Delta\mathcal{O}$ is controlled and bounded as $N \to \infty$. This corresponds to the limit in which the Trotterized unitary $\hat{U}_{\text{trot}}(t, N)$ approaches the 'ideal' unitary $\hat{U}_{\text{heis}}(t)$. Empirically, we find that exponential fits of the form $y(x) = y_0 + Be^{mx}$ work best over a large range of N that we expect to apply to NISQ devices.

2.1. Error Mitigation via Trotter Error Mitigation

To exploit the convergence of Trotter errors for error mitigation, we propose the following procedure:

1. Run circuits for a range of trotterization steps from N_{min} to N_{max}.
2. On each circuit take the desired measurements to estimate $\langle\mathcal{O}\rangle$.
3. Fit a curve of the form $y(x) = y_0 + Be^{mx}$ where x are the sizes of the trotterization steps and y is the measured expectation value.
4. y_0 furnishes our estimate for the ideal value of value of $\langle\mathcal{O}\rangle$ in the absence of Trotter errors.

Throughout this work, we will focus on the case where the observables of interest $\mathcal{O}_i$ are 3-qubit Pauli strings. We then exploit the proposed 'Trotter Error Mitigation' (TEM) technique to extrapolate the Trotter errors to zero for the measured observables.

2.1.1. Noiseless Case

To illustrate the convergence with step size, we pick an arbitrary time, $T_{\text{test}} = 0.315\pi \times (1/J)$, and observable, $\mathcal{O}_{\text{test}} = \hat{X}_1\hat{X}_3\hat{Z}_5$. We pick an arbitrary time *that does not coincide with any special points in the period* of the model's evolution so as to demonstrate that this fitting works even in absence of special values in the expectations such as $0, \pm 0.5, \pm 1$. We evolve the initial state $|\psi_0\rangle = |110\rangle$ under the Heisenberg Hamiltonian, as well as under a range of

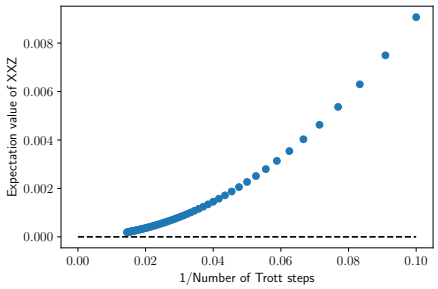

FIG. 2. Blue dots plot the expectation values of a selected observable $\hat{X}\hat{X}\hat{Z}$ for initial state $|110\rangle$ evolving under $\hat{U}_{trot}(t, N)$ for $t = 0.315\pi$ and $N = 5$ to $N = 49$. Vertical axis is the expectation value, whereas horizontal axis is the time step $0.315\pi/N$. The dashed line indicates the expectation value under the ideal evolution $\hat{U}_{\text{heis}}(t)$. We observe that the Trotterized expectation values approach the ideal value as time step goes to zero, as expected.

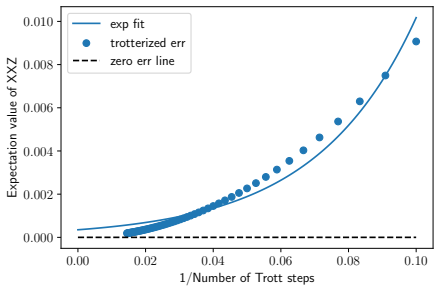

FIG. 3. Sample fit of exponential curve $y = y_0 + Be^{mx}$ against trotterized expectation values. For clarity, the entire y axis has been shifted by the ideal expectation value. We observe y_0 furnishes a lower error than any single trotterized expectation value.

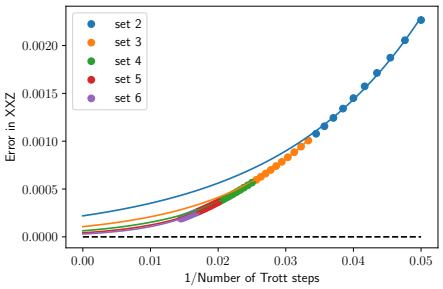

FIG. 4. Fits of an exponential curve against trotterized expectation values. Scatter plot refers to trotterized values and line plots to the exponential fits. Observe that all the fits have lower error than the the trotterized errors themselves.

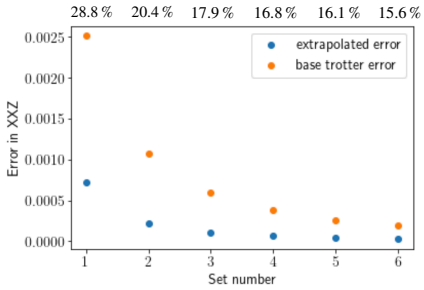

FIG. 5. Comparisons of the best trotterization error in set $\mathcal{N}_i$ against the extrapolated error using the same set. Horizontal axis refer to the set numbering i as described in the text. We observe that the higher the values of N are the better the extrapolation, and the better the fractional improvement as well. Numbers above the plot refer to the fraction of final error over initial error. We observe that the fractional improvement grows with set number, and represents roughly an order of magnitude improvement.

Trotter steps $N = 10$ to $N = 69$ and calculate the expectation value $\langle \mathcal{O}_{\text{test}} \rangle$ on each of the final states. These values are plotted in Fig. 2. The fit against an exponential curve is plotted in Fig. 3, with all expectation values shifted by the ideal expectation value such that the plots approach is 0 (i.e. plotting errors).

Now, we partition the trotterized expectation values into sets of 10, $\mathcal{N}_1 = \{10, 11, ..., 19\}, ..., \mathcal{N}_6 = \{60, 61, ..., 69\}$. Our goals are two-fold. First, we want to show that fitting an exponential curve does lead to a better estimate of the expectation value, and second, we want to show that the estimate increases as N_{min}, N_{max} increase. For each of these sets, we fit our $y = y_0 + Be^{mx}$, and pick the y-intercept as our estimator $\widehat{\langle \mathcal{O}^{(i)}_{\text{test}} \rangle}$. In Fig. 4 we plot these sets against their respective fits, and in Fig. 5 we show how the trotterized error (lowest error in the set) compares to the extrapolated error.

2.1.2. Noisy case

From the noiseless case, it seems like we should just pick the largest possible values of N and extrapolate from those. However, this method is pointless in the noiseless case, because we can simply let $N \to \infty$ in which case we obtain exact simulation results.

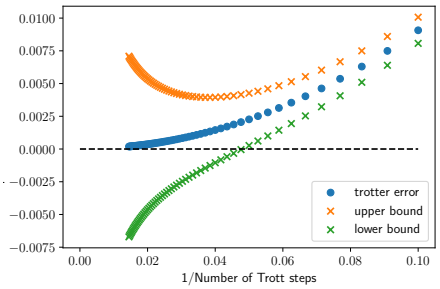

FIG. 6. Upper and lower bounds for the effect of small errors on the Trotter error for an expectation value.

Conversely on a noisy device, the value of N is limited by the overhead of each trotterization step. Generically, we expect the trotter error to decrease up to some critical step number N_c after which increasing the number of steps increases the trotterization error as well. Picking the maximum N in the extrapolation set is then simply a matter of picking some $N \lesssim N_c$.

To pick a minimum N for the extrapolation set, one must keep in mind two issues. First, when N is too small, the measured values have very correlation with the ideal values. The reason for this is that the error has some kind of $1/N$ dependence, which is roughly the size of the unitary itself when N is small. The larger the minimum N is, the better we avoid this problem. Second, we should keep a respectable number of data points in our extrapolation set. From various numerical experiments, it looks like 5 to 10 data points is a good number, and we hope to conduct further work uncover more analytical reasoning behind these heuristic suggestions with more general conditions on what kind of scalings hold for what kind of models.

To illustrate the effects of the noise overhead on this method, suppose we have some small probability p of an error, say some random unitary, occurring with each trotter gate. Generically, we can upper-bound the effect of this noise on our measured expectation value as an $\pm Np$ perturbation to the measured expectation value. Take for instance $p = 10^{-4}$. This error rate is smaller than the error we anticipate on real devices, but lends itself to clearer plots that illustrate the principle. The upper and lower bounds of the resulting error in the $\hat{X}\hat{X}\hat{Z}$ example look like Fig. 6.

From Fig. 6, we see that for step size > 0.06, we maintain good convergence towards the zero error value. In our work, we identify this regime by running circuits on `jakarta`, and only attempt to fit this region.

Finally, we emphasize that the theoretical results and simulations in this section can be applied to arbitrary observables $\mathcal{O}$ and times t. Thus, we expect the specifics of these plots to have little correlation with the actual problem posed by IBM. As we will see, these methods work well for the actual problem statement as well, which speaks to the generalizability of our method.

3. Methods

In order to simulate the Heisenberg Hamiltonian dynamics on the `ibmq-jakarta` device, we employ a number of error mitigation and calibration strategies in tandem. Our mitigation strategies are fully general, and, crucially, do not rely on the specific symmetries of the problem at hand, e.g. on the form of $\hat{H}_{\text{Heis}}$ or the initial state $|110\rangle$.

3.1. Measurement Error Calibration

The first technique that we implement for improving quantum simulation performance is measurement error calibration [39, 40]. This mitigation strategy aims to correct for readout/measurement errors during the final stage of a quantum computation.

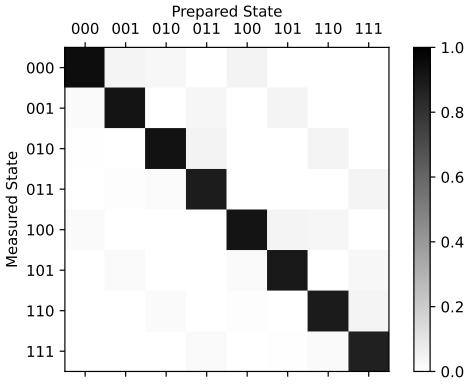

FIG. 7. Calibration matrix for `ibmq-jakarta` device. We prepare each of the 8 basis states $|b\rangle$ for $b \in \{0,1\}^3$ and measure in the computational basis. The resulting matrix gives an estimate for the readout fidelity of the device.

3.2. Zero Noise Extrapolation via Unitary Folding

The next mitigation technique we employ is digital zero-noise extrapolation (ZNE). The general concept of zero-noise extrapolation has gained popularity in recent years, and, among many notable demonstrations, has been used to improve the performance of a Variational Quantum Eigensolver (VQE) algorithm run on a superconducting processor [41]. One approach for ZNE, as taken by Ref. [41], is to use pulse-level control to stretch the time over which the control gates act (while reducing their magnitude so as to execute the same logical circuit). In our work here, however, we pursued a different approach. Specifically, we use the *unitary folding* technique to realize digital ZNE [42]. In our implementation, each Trotter gate $\hat{U}_{\text{trot}}(\pi, N)$ can be written as:

$$\hat{U}_{\text{trot}}(t, N) = \left[\hat{U}_{\text{step}}\left(\frac{t}{N}\right)\right]^N \tag{8}$$

We perform ZNE by *folding* $\hat{U}_{\text{trot}}$ via:

$$\hat{U}_{\text{trot}}(t, N) \rightarrow \left[\hat{U}_{\text{step}}\right]^N \left(\hat{U}^\dagger_{\text{step}}\hat{U}_{\text{step}}\right)^\beta \tag{9}$$

We define the *scaling* factor

$$\lambda = 1 + \frac{2\beta}{N} \tag{10}$$

as the amount by which the circuit depth scales when folding each $\hat{U}_{\text{step}}$ by an amount β. Since $\hat{U}^\dagger_{\text{step}}\hat{U}_{\text{step}} = 1$, we expect the folding to produce no effect on the logical circuit; however, in the presence of realistic hardware noise, this serve to increase the overall noise in the circuit.

By measuring expectation values $\langle\mathcal{O}(\lambda)\rangle$ as a function of the 'noise parameter' λ, we can extrapolate down to the zero noise limit of $\lambda = 0$.

3.3. Optimal Circuit Decomposition

Consider $\hat{U}_{\text{Heis}}(t, N)$. If we perform the Trotter decomposition, we get

$$\begin{aligned}\hat{U}_{\text{trot}}(t, N) &= \left[e^{-i(t/N)\hat{S}^{(01)}} e^{-i(t/N)\hat{S}^{(12)}}\right]^N \\ &\approx \left[e^{-it/N(X_0X_1+Y_0Y_1+Z_0Z_1)} e^{-it/N(X_1X_2+Y_1Y_2+Z_1Z_2)}\right]^N\end{aligned} \tag{11}$$

If we further decompose the two unitaries above into XX, YY, and ZZ gates, then we see that $\hat{U}_{\text{trot}}(t, N)$ can be implemented using six 2-qubit gates (i.e. six CNOTs). However, since $\hat{U}_{\text{trot}}(t, N)$ acts on three qubits, we know that, fundamentally, it should be possible to decompose $\hat{U}_{\text{trot}}(t, N)$ using 3 two-qubit gates [2]. In our numerical study, we use a 3 CNOT decomposition of $\hat{U}_{\text{trot}}(t, N)$ to reduce our overall circuit depth.

3.4. Trotter Error Mitigation

Please see our results in the attached Jupyter notebook for a discussion of our novel technique. Using our error mitigation strategies in tandem, we are able to achieve a final state tomography fidelity of 98.7%.

[1] S. Lloyd, Science **273**, 1073 (1996).
[2] F. Tacchino, A. Chiesa, S. Carretta, and D. Gerace, Advanced Quantum Technologies **3**, 1900052 (2020).
[3] Y. I. Manin, Sov. Radio , 13 (1980).
[4] P. Benioff, Journal of statistical physics **22**, 563 (1980).
[5] R. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).
[6] I. Buluta and F. Nori, Science **326**, 108 (2009).
[7] R. Blatt and C. F. Roos, Nature Physics **8**, 277 (2012).
[8] I. Bloch, J. Dalibard, and S. Nascimbene, Nature Physics **8**, 267 (2012).
[9] A. Aspuru-Guzik and P. Walther, Nature physics **8**, 285 (2012).
[10] A. A. Houck, H. E. Türeci, and J. Koch, Nature Physics **8**, 292 (2012).
[11] I. M. Georgescu, S. Ashhab, and F. Nori, Reviews of Modern Physics **86**, 153 (2014).
[12] J. Preskill, Quantum **2**, 79 (2018).
[13] L. Lamata, A. Parra-Rodriguez, M. Sanz, and E. Solano, Advances in Physics: X **3**, 1457981 (2018).
[14] M. P. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Physical Review B **40**, 546 (1989).
[15] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller, Physical Review Letters **81**, 3108 (1998).
[16] M. J. Hartmann, F. G. Brandao, and M. B. Plenio, Nature Physics **2**, 849 (2006).
[17] A. D. Greentree, C. Tahan, J. H. Cole, and L. C. Hollenberg, Nature Physics **2**, 856 (2006).
[18] D. G. Angelakis, M. F. Santos, and S. Bose, Physical Review A **76**, 031805 (2007).
[19] I. Carusotto, S. Fagnocchi, A. Recati, R. Balbinot, and A. Fabbri, New Journal of Physics **10**, 103001 (2008).
[20] L. Tian, Physical review letters **105**, 167001 (2010).
[21] O. Viehmann, J. von Delft, and F. Marquardt, Physical review letters **110**, 030601 (2013).
[22] J.-M. Reiner, M. Marthaler, J. Braumüller, M. Weides, and G. Schön, Physical Review A **94**, 032338 (2016).
[23] P. M. Poggi, N. K. Lysne, K. W. Kuper, I. H. Deutsch, and P. S. Jessen, PRX Quantum **1**, 020308 (2020).
[24] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, nature **415**, 39 (2002).
[25] A. Friedenauer, H. Schmitz, J. T. Glueckert, D. Porras, and T. Schätz, Nature Physics **4**, 757 (2008).
[26] K. Kim, M.-S. Chang, S. Korenblit, R. Islam, E. E. Edwards, J. K. Freericks, G.-D. Lin, L.-M. Duan, and C. Monroe, Nature **465**, 590 (2010).
[27] R. Islam, E. Edwards, K. Kim, S. Korenblit, C. Noh, H. Carmichael, G.-D. Lin, L.-M. Duan, C.-C. Joseph Wang, J. Freericks, et al., Nature communications **2**, 1 (2011).
[28] H. Labuhn, D. Barredo, S. Ravets, S. De Léséleuc, T. Macrì, T. Lahaye, and A. Browaeys, Nature **534**, 667 (2016).
[29] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, et al., Nature **551**, 579 (2017).
[30] P. Roushan, C. Neill, J. Tangpanitanon, V. M. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, et al., Science **358**, 1175 (2017).
[31] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Nature **551**, 601 (2017).
[32] X. Zhang, K. Zhang, Y. Shen, S. Zhang, J.-N. Zhang, M.-H. Yung, J. Casanova, J. S. Pedernales, L. Lamata, E. Solano, et al., Nature communications **9**, 1 (2018).
[33] R. J. MacDonell, C. E. Dickerson, C. J. Birch, A. Kumar, C. L. Edmunds, M. J. Biercuk, C. Hempel, and I. Kassal, Chemical Science **12**, 9794 (2021).
[34] C. J. van Diepen, T.-K. Hsiao, U. Mukhopadhyay, C. Reichl, W. Wegscheider, and L. Vandersypen, Physical Review X **11**, 041025 (2021).
[35] H. F. Trotter, Proceedings of the American Mathematical Society **10**, 545 (1959).
[36] M. Suzuki, Journal of Mathematical Physics **26**, 601 (1985).
[37] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Physical Review X **11**, 011020 (2021).
[38] T. Kimura, Progress of Theoretical and Experimental Physics (2017), arXiv:1702.04681 [math-ph].
[39] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, Physical Review A **103**, 042605 (2021).
[40] "Measurement Error Mitigation," Qiskit Textbook.
[41] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature **567**, 491 (2019).
[42] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng, 2020 IEEE International Conference on Quantum Computing and Engineering (QCE) , 306 (2020).