Decorrelating Errors in Quantum Gates by Random Gate Synthesis

Anthony Polloreno*
Rigetti Computing, Berkeley, CA

Kevin C. Young
Sandia National Laboratories, Livermore, CA

(Dated: September 26, 2018)

Thresholds for fault-tolerant quantum computation are often calculated assuming a noise model in which errors are uncorrelated. While convenient for simulation, these error models are often unphysical. Work by Preskill and others has shown that arbitrarily long computations may be performed even in the presence of spatial and temporal correlation, provided the correlation is sufficiently weak and decays sufficiently quickly, but at the cost of a significantly lower threshold. The success of algebraic decorrelation methods, such as dynamical decoupling, demonstrate that quantum control techniques are capable of reducing noise correlations. We propose to introduce similar methods at the gate synthesis level to effect the decorrelation of errors in quantum circuits, thereby increasing the threshold for fault-tolerant computation in such systems. We show numerically and experimentally on a superconducting qubit that these methods can reduce the magnitude of the diamond norm of the error by at least an order of magnitude.

I. INTRODUCTION

Steady progress has been made in the theory of quantum error correction, proving higher thresholds for increasingly general models of noise [1–6]. These results show that quantum computation is feasible in principle, however recent NISQ [7] devices have noise that is not only often above known thresholds, but that also violates fundamental assumptions made by the models used in these results[8–10], such as Markovianity [11] and independence of errors[12]. With these assumptions violated, many properties of system performance and correctness can no longer be guaranteed.

Moreover, even when noise can be accurately modeled as Markovian, many existing works make further simplifying approximations about its structure to make problems tractable. For example, Pauli channels are often used to model systems due to their classical simulability[13], even in the absence of physical motivation[14–21]. One approach, then, would be to use these thresholds to give a loose lower bound for thresholds in other systems[20]. Such an approach is correct and rigorous, but will produce overly-pessimistic bounds that may not be reflective of numerically demonstrated performnace.

Many characterization routines also make assumptions about the form of noise, and if these assumptions are violated the utility and correctness of these routines may be reduced. For example, randomized benchmarking and tomography will report incorrect answers without any syndrome in the case of non-Markovian noise [22]. Slightly more helpful, gate-set tomography will report that a system fails to be Markovian, without saying in what way..

Other authors have approached these problems at varying levels of abstraction, each with their own benefits and short-comings. At the level of unitary synthesis, one can consider randomizing over errors that might arise from gate approximations like the Solovay-Kitaev decomposition [6, 23]. While this affords the ability to prove theorems that guarantee quadratic improvements to performance, they do not offer routines which are accessible and implementable with current coherence times. In addition, present quantum computing work focusses primarily on gate synthesis, rather than unitary synthesis.

At the level of compilation, Frame Randomization[24, 25] can be used to average over noise by twirling error over an appropriately selected unitary 2-design.[26] This has the advantage of being realizable in existing systems[25], but it can require intensive use of classical resources, such as precompilation and generation of many waveform variants, and becomes more difficult when the gates being used are not in the Clifford group, as is the case in various quantum computing architectures.

The general theme of these results is randomized decoupling[27, 28] – by introducing classical randomness into quantum computations, both coherent and non-Markovian noise can be transformed to incoherent, independent, and Markovian noise. In this paper, we explore a different method of solving these problems. We propose to inject additional decorrelating randomness into the system during physical gate synthesis through the use of *balanced optimal control solutions* (BOCSs). The task of generating BOCSs will fall to optimal control.

II. THEORY

BOCSs are families of control solutions, $c_i(t)$, where each member of the family approximates the same target gate with slightly different errors, for any given instance a noise Hamiltonian. That is, for a target gate U_T, we seek

* Email: anthony@rigetti.com

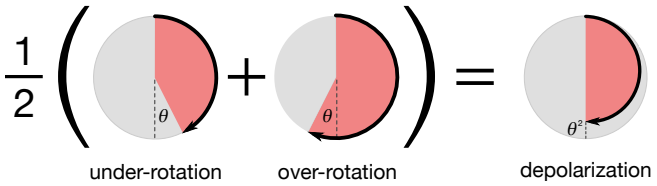

FIG. 1. An example of a balanced control solution. Using optimal control, two implementations of a Z_π gate are designed to have equal and opposite sensitivity to errors (if one implementation over-rotates by angle θ, then the other *under*-rotates by θ). Each time the gate is used, one of these implementations is chosen at random. The resulting quantum channel is equivalent to a perfect implementation of the gate followed by dephasing of $\mathcal{O}\left(\theta^2\right)$.

a family of approximating gates $U_i = \mathcal{T}e^{-i\int c_i(t)H(t,\delta)}$ such that the family of unitary approximations is *balanced*. A family of controls is balanced if there is a weighting ω_i of its members such that produces a *balanced channel*, satisfying for some small α,

$$\frac{1}{N}\sum_{i=1}^{N}\omega_i U_i\rho U_i^\dagger = DPN[\alpha]\left(U_T\rho U_T^\dagger\right) \tag{1}$$

where $\omega_i > 0$ and $DPN[\alpha](\rho)$ is a depolarizing noise channel, with strength α. Such a channel is defined as:

$$DPN[\alpha](\rho) \to (1-\alpha)\rho + \alpha\sum p_i\sigma_i\rho\sigma_i \tag{2}$$

with p_i summing to one.

Thus, following from the mixing lemma proven by Campbell and Hastings in [6] and [23], one property of BOCSs is that if each approximation of the target gate is correct to within ϵ in operator norm and $\alpha \leq \epsilon^2$, then the diamond distance of the BOCS to U_T is also no larger than ϵ^2. Because any sufficiently populated BOCS will have a set of controls that result in error of order $\mathcal{O}(\epsilon^2)$ on their balanced channel, it is then clear that in priniple BOCSs can always quadratically decrease the error in our approximation (in diamond norm). Moreover, in the setting where our system's performance is non-Markovian, e.g. drift, optimizing our BOCSs *robustly* over a range of control parameters can reduce the non-Markovianity of the noise.

III. A SIMPLE EXAMPLE

As a somewhat trivial example, consider a single-qubit rotation-angle error, such as from stochastic laser amplitude fluctuations. A BOCS may consist of an X_π pulse, as well as an $X_{-\pi}$ pulse (i.e., a clockwise and counter clockwise rotation of the qubit). In the case of excess amplitude, the X_π pulse will result in an over-rotation error, while the $X_{-\pi}$ pulse will result in an *under*-rotation error. When it comes time to perform the target gate in a quantum circuit, one member of the BOCS is chosen uniformly at random. This has the effect of decreasing the norm of the noise channel and decorrelating the over-rotation error (Figure 1). In this simple example, we can analytically find a solution to Equation 1. Specifically, by choosing weights $\omega_i = 1$, we see:

$$\begin{aligned}&\frac{1}{2}(X^*_{\pi+\epsilon}\otimes X_{\pi+\epsilon} + X^*_{-(\pi+\epsilon)}\otimes X_{-(\pi+\epsilon)})\\ &= (\sin^2\frac{\pi+\epsilon}{2}I\otimes I + \cos^2\frac{\pi+\epsilon}{2}X\otimes X)X\otimes X\\ &\approx DPN[\epsilon^2]X\otimes X\end{aligned} \tag{3}$$

Therefore, for a rotational error of angle $\epsilon > 0$, we see that X_π and $X_{-\pi}$ form a BOCS, with $\alpha \approx \epsilon^2$.

IV. OPTIMAL CONTROL PROBLEMS

A. Random Gate Synthesis

Generating BOCSs can be done in a variety of ways, using any of the many available quantum optimal control techniques [29–31]. For our numerics, we chose to use the GRAPE algorithm to generate candidate pulse-shapes to approximate the target gate. First described in [29], the GRAPE (GRadient Ascent Pulse Engineering) algorithm is a technique for finding piecewise constant control sequences that approximate a desired unitary, U_T. Defining our uncontrolled Hamiltonian as H_0, our control Hamiltonians as $H_{i\neq 0}$, and our *control matrix* c_{ij} as containing control amplitude associated with the i^{th} time step and the j^{th} hamiltonian, we can write our approximate unitary at any timestep as

$$U_i = \exp\{-i\Delta t(H_0 + \sum_{j=1}^{n} c_{ij}H_j\} \tag{4}$$

Then, to measure the simularity of our approxiate unitary U to our target unitary U_T, we can define a cost function $J(U) = Tr\{U_T^\dagger U\}$.

To optimize this cost function we can perform the following standard update loop for some threshold value $\varepsilon > 0$ and step size $\delta > 0$:

Gradient Ascent
while $J(U) < (1-\varepsilon)$ **do**
  $c_{ij} \to c_{ij} + \delta\frac{\partial J(U)}{\partial c_{ij}}$
  for $1 \leq i \leq n$ **do**
    $U_i \to \exp\{-i\Delta t(H_0 + \sum_{i=0}^{n} c_{ij}H_j)\}$
  end for
  $U \to \prod_{i=1}^{n} U_i$
end while

In general these gradients can be computed by propagating partial derivatives of the cost function with respect

to control parameters through each timestep of the via the chain rule. However, in [29] Khaneja et al. derive a simple update formula that is correct to first order. In particular one can show that:

$$\frac{\partial J(U)}{\partial u_{ij}} = -2Re\{\left\langle U_{j+1}^{\dagger}...U_N^{\dagger}U_T|i\Delta t H_j U_j...U_1\right\rangle \left\langle U_j...U_1|U_{j+1}^{\dagger}...U_N^{\dagger}U_T\right\rangle\} + \mathcal{O}(\Delta t^2) \tag{5}$$

In our numerical results in Section V A and V B, we consider the controls to have quasi-static[41](non-Markovian) Gaussian distributed errors, so to generate controls that are robust to this error we modify our gradient to instead be:

$$\frac{\partial \tilde{J}(U)}{\partial u_{ij}} = \int p(\vec{\delta})\frac{\partial J(U(\vec{\delta}))}{\partial u_{ij}}d\vec{\delta} \tag{6}$$

with $p(\vec{\delta})$ Gaussian distributed, as has been done in previous works [32] to ensure that the optimal control results are robust over a wide range of errors. To make this averaging tractable, we approximate this integral using Gaussian quadrature, approximating the cost functions as a low order polynomial[33].

Concretely, we consider a Hamiltonian of the following form:

$$H(t) = \delta_0 H_0 + \sum_{i=1}^{n}(1+\delta_i)c_i(t)H_i \tag{7}$$

for control Hamiltonians H_i, free evolution Hamiltonian H_0 and random variables δ_i, that model some small uncertainty in parameters in the Hamiltonian. Such a model might describe a superconducting qubit quantum processor where control amplitudes for the RF pulses or qubit transition frequencies vary over time, or a trapped ion quantum computer where the intensity, frequency, or phase of the laser might drift[8, 10, 34]. Correlations between different δ_i might arise, for instance, if two of the controls have the same noise source, e.g. RX and RY gates in superconducting qubit architectures might use the same AWG and pulse envelope, and all of the lab equipment may suffer from the same temperature drift of control electronics.

B. BOCS Approximation

After using GRAPE or another optimal control routine to synthesize a collection of controls, we must find the weights ω_i such that the collection of controls form a BOCS as described in Equation 1. To do this, for each control U_i we find the unitary error channel $\mathcal{E}_i$ such that $\mathcal{E}_i U_i = U_T$, where U_T is the target gate. The we see that for any stochastic application of these channels, the resulting map is given by:

$$\frac{1}{N}\sum_{i=1}^{N} w_i \mathcal{E}_i^{\dagger}(U_T \rho U_T^{\dagger})\mathcal{E}_i \tag{8}$$

If we consider the Pauli-Liouville representation[35] of this error channel, the diagonal terms are the *stochastic* terms that arise from classical uncertainty, while the off-diagonal terms may more generally arise from *coherent* errors[25]. Thus to approximate a depolarizing channel we define our optimal control problem to be the following, which minimizes the off-diagonal terms:

$$\begin{aligned} &\underset{w_0,...,w_N}{\textbf{minimize}}\{\sum_{\substack{i,j \\ i\neq j}}^{N}|\sigma_i \Lambda(\sigma_j)|^2\} \\ &\textbf{where } \Lambda(\sigma_j) := \sum_{i=1}^{N} w_i \mathcal{E}_i^{\dagger}\sigma_j \mathcal{E}_i \\ &\textbf{subject to} \sum_{i=1}^{N} w_i = 1 \end{aligned} \tag{9}$$

In the numerics section below, this was solved with a constrained minimization algorithm (Sequential Least Squares Programming[36]). However, it can be solved more efficiently as a least squares problem. In particular, consider the $M^2 - M$ off diagonal elements $\mathcal{E}_i$. Vectorize their squares, and concatenate the resulting vectors, for each $\mathcal{E}_i$. Then we just need to find ω that produces the smallest $\vec{b}$ in the following underconstrainted least squares problem:

$$A\omega = \vec{b}$$

where:

$$\mathrm{A} = \begin{pmatrix} \mathcal{E}_{0_{10}} & \mathcal{E}_{1_{10}} & \cdots & \mathcal{E}_{N-1_{10}} \\ \mathcal{E}_{0_{20}} & \mathcal{E}_{1_{20}} & \cdots & \mathcal{E}_{N-1_{20}} \\ \vdots & \vdots & \ddots & \\ \mathcal{E}_{0_{M-1M}} & \mathcal{E}_{1_{M-1M}} & & \mathcal{E}_{N-1_{M-1M}} \end{pmatrix}$$

In particular, we want the solution ω such that $||\vec{b}||_2^2$ is minimized. This is readily given by the solution to a constrained least squares minimization problem as $\frac{1}{N}(\tilde{A^{-1}}_{AVG})^{-1}\tilde{A^{-1}}_{i_{AVG}}$.[citation needed]http://eeweb.poly.edu/iselesni/lecture_notes/least_squares/least_squares_SP.pdf

Previous authors have considered minimizing the diamond distance to the nearest Pauli or Clifford Channel [37], and while this gives a good theoretical framework, it requires the more computationally challenging task of optimizing over the diamond norm, and does not constrain the resulting channel to be decomposable into a given family of controls. Our routine, on the other hand, optimizes over an easy to compute sum, and produces a channel defined explicitly in terms of given family of controls.

V. NUMERICAL RESULTS

In the following two subsections, we present numerical results of our routine, first for a one qubit example, and then for a two-qubit example. The code and BOCSs generated for both examples are available online at [38].

A. 1Q Gates

For the one-qubit case, we generate BOCSs for $RX(\frac{\pi}{2})$ and $RY(\frac{\pi}{2})$, that together with $RZ(\theta)$ rotations are universal for one-qubit computation. (In particular, choosing just one of these two gates would be sufficient.) Our control Hamiltonian is given as:

$$H = \epsilon\sigma_z + (1+\delta)(c_x(t)\sigma_x + c_y(t)\sigma_y) \tag{10}$$

where $\epsilon, \delta \sim \mathcal{N}(0, .001)$ We assume that the errors on σ_x and σ_y are perfecly correlated, as mentioned in Section IV. In our simulation we chose an total evolution time of $T = \pi$, a number of steps $N = 100$, and a threshold infidelity of $1E-3$ for generating the control solutions with GRAPE.

B. 2Q Gates

For a two-qubit example, we again consider the single qubit gates $RX(\frac{\pi}{2})$ and $RY(\frac{\pi}{2})$ on both qubits, along with single qubit $RZ(\theta)$ rotations. The entangling operation we chose is $ZZ(\frac{\pi}{2})$, which together with the single-qubit operations are universal for quantum computation. Our control Hamiltonian is given as:

$$\begin{aligned} H = &\sum_{j=1}^{2}(\epsilon_j\sigma_z^j + (1+\delta_j)(c_x^j x(t)\sigma_x^j + c_y^j(t)\sigma_y^j)) \\ &+ (1+\delta)c(t)\exp\left(-i\frac{\sigma_z^1 \otimes \sigma_z^2}{4}\right) \end{aligned} \tag{11}$$

We again consider $\epsilon_j, \delta_j, \delta \sim .001$. In this simulation we again used a threshold infidelity of $1E-3$, but we increased the total evolution time to $T = 4\pi$, and increased the number of steps to $N = 400$ so that the size of each time step was the same as in the one qubit example, however the total evolution time was greater to allow GRAPE more opportunities to find non-trivial pulseshapes.

VI. EXPERIMENTAL RESULTS

Here we present experimental results from implementing our routine on a fixed-frequency superconducting transmon qubit. In particular we used qubit 8 on the Rigetti 19Q-Acorn chip, whose characterization can be found in [39]. To implement a BOCS on this qubit, four incorrectly calibrated approximately Gaussian pulses were produced by scaling the pulseshape for a calibrated $50\mu s$ $RX(\frac{\pi}{2})$ pulse by 106.4%, 103.9%, 93.7% and 91.2%.

Using Equation 9, we then generated the optimal weighting ω_i. To benchmark the quality of the new balanced channel, we then performed six randomized benchmarking experiments: one for each over- and under-calibrated pulse, one for the calibrated pulse, and one for the balanced channel. We used $N = 1000$ shots per experiment and $K = 10$ sequences per sequence length, for sequence lengths up to $L = 64$[40]. In each case, our Clifford operations were decomposed into RX($\frac{\pi}{2}$)) and RY($\frac{\pi}{2}$)) pulses. In our implementation, these gates are implemented using the same pulse envelope definitions and control electronics, phase shifted by $\frac{\pi}{2}$ radians, and are therefore subject to identical miscalibration errors. The results are shown in Figure 2 for sequence lengths $L = 64$. In this experiment, fitting the randomized benchmarking data with scipy's curve_fit using the Trust Region Reflective method reports one-qubit gate fidelities of 99.4% for the calibrated pulse, 98.8% for Pulse1, 99.4% for Pulse2, 98.9% for Pulse3, 98.6% for Pulse4, and 99.0% for the Randomized pulse. As one can see from looking at Figure 2, despite all controls performing comparably to the Calibrated control, only the Randomized control has similarly tight error bars. In the cases of each miscalibrated pulse, the error bars extend significantly, in often asymmetric ways, making usage of chi-square statistics, and errors on the fit less straightforward.

This is consistent with the result shown in [41] that for particular non-Markovian error models, noise will manifest as gamma distributed points for each sequence length. On the other hand, Markovian noise, such as depolarizing noise, will result in Gaussian distributed fidelity estimates for each randomized benchmarking sequence length. We see that the coherently miscalibrated controls have long tails, consistent with gamma distributed random variables, while the calibrated and randomized implementations both have much shorter tails, consistent with Gaussian distributed random variables.

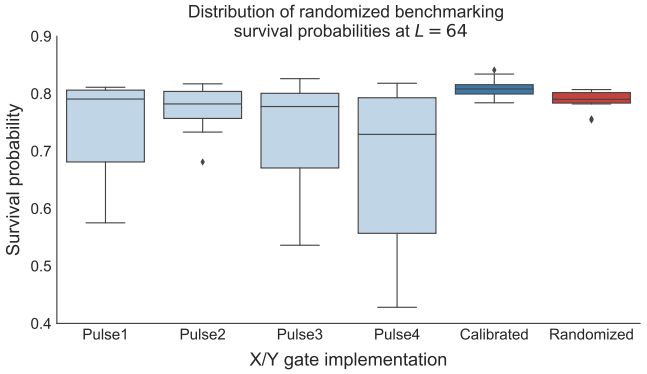

FIG. 2. Randomized benchmarking experiments ran using different pulse definitions. The four plots on the left are from the incorrectly calibrated pulse, while the top right is the calibrated pulse, and the bottom right is the BCS.

VII. CONCLUSION AND FUTURE WORK

We have shown numerically that using a balanced optimal control solution (BOCS) can reduce non-Markovian coherent error on a quantum channel by at least an or-

der of magnitude in diamond norm, over a wide range of quasi-static values of noise. In addition, we have demonstrated that these approximate controls can be generated through optimal control (GRAPE), and that the minimization problem is tractable.

Future directions for this work include demonstrating the routine experimentally on a two-qubit gate, moving the random gate selection from a precompilation step to runtime logic onboard the FPGA, investigating other optimization routines such as CRAB [30] and GOAT[31], and using more sophisticated benchmarking routines such as GST[9] to quantitatively investigate the performance of our method.

Another interesting area of research would be using model-free approaches. The numerical work in the paper assumes access to a model of the system, however an experimentalist may not have a model readily available to describe the system, e.g. in the presence of unknown on-chip crosstalk, or an uncalibrated transfer function of the system. Moreover, even if a model is available, it might be computationally inconvenient to simulate, i.e. for more than a few qubits, or for non-adiabatic gates, where the timestep is required to be much smaller than the gate-time.

In these situations, one approach would be to use *in-situ* optimal control techniques [42–44] to generate candidate controls, and then use an optimizer like Nealder-Mead to perform the minimization. While performing optimization this way would still be slow (requiring full process tomography, in general, to reconstruct the process matrix of a candidate balanced channel), one could intead select pulse sequences that are sensitive to known parameters of interest in the Hamiltonian and just run the minimization with respect to those terms. (e.g. if we suspect there is unwanted $Z \otimes Z$ crosstalk between qubits.)

VIII. ACKNOWLEDGEMENTS

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

[1] D. Aharonov, A. Kitaev, and J. Preskill, Physical Review Letters **96** (2006), 10.1103/physrevlett.96.050504.
[2] N. P. Breuckmann, K. Duivenvoorden, D. Michels, and B. M. Terhal, (2016), arXiv:1609.00510.
[3] M. E. Beverland, (2016), 10.7907/z96m34sc.
[4] A. Kubica, M. E. Beverland, F. Brandão, J. Preskill, and K. M. Svore, Physical Review Letters **120** (2018), 10.1103/physrevlett.120.180501.
[5] C. Wang, J. Harrington, and J. Preskill, Annals of Physics **303**, 31 (2003).
[6] E. Campbell, Physical Review A **95** (2017), 10.1103/physreva.95.042306.
[7] J. Preskill, "Quantum computing in the nisq era and beyond," (2018), arXiv:1801.00862.
[8] J. Kelly, P. O'Malley, M. Neeley, H. Neven, and J. M. Martinis, "Physical qubit calibration on a directed acyclic graph," (2018), arXiv:1803.03226.
[9] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, Nature Communications **8** (2017), 10.1038/ncomms14485.
[10] P. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, E. Lucero, J. Mutus, O. Naaman, C. Neill, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. White, S. Boixo, R. Babbush, V. Smelyanskiy, H. Neven, and J. Martinis, Physical Review Letters **121** (2018), 10.1103/physrevlett.121.090502.
[11] A. Y. Kitaev, Russian Mathematical Surveys **52**, 1191 (1997).
[12] E. Knill, R. Laflamme, and W. H. Zurek, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **454**, 365 (1998).
[13] D. Gottesman, (1998), arXiv:quant-ph/9807006.
[14] P. Aliferis and A. W. Cross, Physical Review Letters **98** (2007), 10.1103/physrevlett.98.220502.
[15] E. Knill, Nature **434**, 39 (2005).
[16] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Physical Review A **83** (2011), 10.1103/physreva.83.020302.
[17] G. Duclos-Cianci and D. Poulin, Physical Review Letters **104** (2010), 10.1103/physrevlett.104.050504.
[18] J. R. Wootton and D. Loss, Physical Review Letters **109** (2012), 10.1103/phaliferis2008accuracyysrevlett.109.160503.
[19] H. Bombin, R. S. Andrist, M. Ohzeki, H. G. Katzgraber, and M. A. Martin-Delgado, Physical Review X **2** (2012), 10.1103/physrevx.2.021004.
[20] D. Puzzuoli, C. Granade, H. Haas, B. Criger, E. Magesan, and D. G. Cory, Physical Review A **89** (2014), 10.1103/physreva.89.022306.
[21] P. Aliferis, D. Gottesman, and J. Preskill, Quantum Information and Computation **8**, 0181 (2008).
[22] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, Physical Review A **87** (2013), 10.1103/physreva.87.062119.
[23] M. B. Hastings, "Turning gate synthesis errors into incoherent errors," (2016), arXiv:1612.01011.
[24] J. J. Wallman and J. Emerson, Physical Review A **94** (2016), 10.1103/physreva.94.052325.
[25] M. Ware, G. Ribeill, D. Riste, C. A. Ryan, B. Johnson, and M. P. da Silva, "Experimental demonstration of pauli-frame randomization on a superconducting qubit," (2018), arXiv:1803.01818.
[26] A. Roy and A. J. Scott, Designs, codes and cryptography **53**, 13 (2009).

[27] L. Viola and E. Knill, Physical Review Letters **94** (2005), 10.1103/physrevlett.94.060502.

[28] L. Viola, in *Proceedings of the 44th IEEE Conference on Decision and Control* (IEEE).

[29] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Journal of Magnetic Resonance **172**, 296 (2005).

[30] T. Caneva, T. Calarco, and S. Montangero, Physical Review A **84** (2011), 10.1103/physreva.84.022326.

[31] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, Physical Review Letters **120** (2018), 10.1103/physrevlett.120.150401.

[32] M. H. Goerz, E. J. Halperin, J. M. Aytac, C. P. Koch, and K. B. Whaley, Physical Review A **90** (2014), 10.1103/physreva.90.032329.

[33] M. Abramowiz and I. Stegun, National Bureau of Standards, US Government Printing Office, Washington, DC (1972).

[34] B. Lekitsch, S. Weidt, A. G. Fowler, K. Mølmer, S. J. Devitt, C. Wunderlich, and W. K. Hensinger, Science Advances **3**, e1601540 (2017).

[35] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki, Physical Review X **4** (2014), 10.1103/physrevx.4.011050.

[36] S. Wright and J. Nocedal, Springer Science **35**, 7 (1999).

[37] E. Magesan, D. Puzzuoli, C. E. Granade, and D. G. Cory, Physical Review A **87** (2013), 10.1103/physreva.87.012324.

[38] A. Polloreno, *Code and data used in this paper.*

[39] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, P. Karalekas, C. B. Osborn, A. Papageorge, E. C. Peterson, G. Prawiroatmodjo, N. Rubin, C. A. Ryan, D. Scarabelli, M. Scheer, E. A. Sete, P. Sivarajah, R. S. Smith, A. Staley, N. Tezak, W. J. Zeng, A. Hudson, B. R. Johnson, M. Reagor, M. P. da Silva, and C. Rigetti, "Unsupervised machine learning on a hybrid quantum computer," (2017), arXiv:1712.05771.

[40] E. Magesan, J. M. Gambetta, and J. Emerson, Physical Review Letters **106** (2011), 10.1103/physrevlett.106.180504.

[41] H. Ball, T. M. Stace, S. T. Flammia, and M. J. Biercuk, Physical Review A **93** (2016), 10.1103/physreva.93.022303.

[42] R.-B. Wu, B. Chu, D. H. Owens, and H. Rabitz, Physical Review A **97** (2018), 10.1103/physreva.97.042122.

[43] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. White, A. Cleland, and J. M. Martinis, Physical Review Letters **112** (2014), 10.1103/physrevlett.112.240504.

[44] C. Ferrie and O. Moussa, Physical Review A **91** (2015), 10.1103/physreva.91.052306.