CONTENTS

Decorrelating Errors in Quantum Gates by Random Gate Synthesis

Anthony Polloreno
Rigetti Computing, Berkeley, CA

Kevin C. Young*
Sandia National Laboratories, Livermore, CA

(Dated: August 1, 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 correlation, provided the correlation is sufficiently weak and decays sufficiently quickly with distance, 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 temporal noise correlations. We propose to introduce similar methods to effect the spatial decorrelation of errors in quantum circuits, thereby increasing the threshold for fault-tolerant computation in such systems.

I. INTRODUCTION

Steady progress has been made in the theory of quantum error correction, proving higher thresholds for increasingly general models of noise [1][citation needed]. These results show that quantum computation is feasible, however recent NISQ [2] devices have noise that is not only often above thresholds[citation needed], but that also violates fundamental assumptions made by the models used in these results, such as Markovianity [3] and independence of errors [citation needed]. With these assumptions violated, promises about system performance and correctness are difficult to make.

There are many ways in which devices can behave in non-Markovian ways,[4, 5], and even if a system is Markovian, many models make simplifying assumptions about the noise on that system - such as positing that it has a simple structure. For instance Pauli channels are often chosen[citation needed]due to their classical simulability[citation needed]. Some authors have then taken the approach of using these easily-understood models to upper bound the error in a real system by *honestly* representing them [6]. This approach is correct and rigorous, but scales poorly.

Other authors have tried to address these problems [7–9] by using circuit composition to remove correlated noise. Given a process that is generated by a sequence of channels with some unknown, potentially channel depdenent error, coherent errors can be averaged away into incoherent error, at the cost of additional compilation steps, which may in general increase circuit depth and require more sophisticated classical preprocessing. Some of these routines have been demonstrated experimentally [10], and been shown to reduce the coherence of noise.

In this Letter, we take a different approach to address these problems at the gate synthesis level. We propose to inject additional decorrelating randomness into the system through the use of *balanced control solutions* (BCSs). BCSs are families of control solutions that all approximate the same target gate, but with balanced errors for any given instance of the noise Hamiltonian. That is, for a target gate, U_T, we seek a family of control solutions, $c_i(t)$, each implementing an approximation U_i to the target gate, such that the family of unitary approximations is *balanced.* A balanced family is one which satisfies, 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 $DPN[\alpha](\rho)$ is a *generalized* depolarizing noise channel with strength α. (For the rest of this paper, we will refer to them just as depolarizing channels.) Such a

* Corresponding author: kyoung@sandia.gov

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. This means that on average, the unitary approximations implement the target unitary followed by a small depolarizing channel. The task of constructing the BCSs will fall to optimal control.

This techique can be used to turn coherent error into incoherent error therefore correlated error into uncorrelated error. One benefit of this is that non-Markovian noise will be made to be more Markovian. This is particularly useful because many routines exist that can assess the quality of a gateset, however in the presence of non-Markovianity most of them become unreliable.

For example, randomized benchmarking and tomography will report incorrect answers without any syndrome [11], and gateset tomography will report that the gateset failed to be Markovian, but will fail to diagnose in what way it was non-Markovian. Because BCSs change the error to a depolorizing channel, the correlations in the noise can be reduced, and the process can be made much closer to Markovian, and as we will show, this comes at no real cost to the fidelity of the implemented gate. Finally, unlike in [7, 8, 10] we note that this method of coherent error mitigation is cheap to implement - it only requires a random choice of pulse definition at run time, in addition to the extra storage required to store the pulse shapes - no classical preprocessing or runtime frame tracking is required.

II. A SIMPLE EXAMPLE

As a somewhat trivial example, consider a single-qubit rotation-angle error, such as result from stochastic laser amplitude fluctuations. A BCS may consist of an X_π pulse, as well as an $\bar{X}_\pi$ 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 $\bar{X}_\pi$ pulse results in an *under*-rotation error. When it comes time to perform the target gate in a quantum circuit, one member of the BCS 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 solve the minimization problem given by equation 9 analytically. In particular, if we choose the weights in 1 such that $\omega_i = 1$, and we choose to represent our gates in the vectorized superoperator representation, then:

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

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

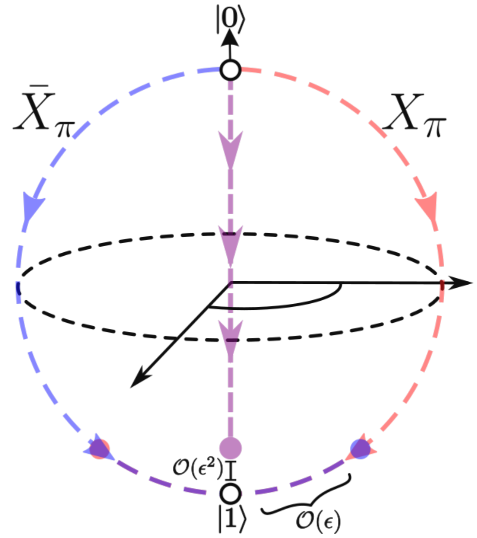


FIG. 1: An example of a BCS. The red path is a π pulse over-rotating clockwise, while the blue path is a π pulse over-rotating counter-clockwise. The purple path is the average of the two.

III. OPTIMAL CONTROL PROBLEMS

A. Random Gate Synthesis

Generating BCSs can be done in a variety of ways, using any quantum optimal control technique [12, 13] to find a family of controls. For simplicity in this paper we use the GRAPE algorithm to generate candidate pulse-shapes to approximate the target gate. First described in [14], 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* u_{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

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

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

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 $J(U_n) < (1-\varepsilon)$ **do**
  $u_{ij} \to u_{ij} + \delta \frac{\partial J(U)}{\partial u_{ij}}$
  for $1 \le i \le n$ **do**
    $U_i \to \exp\{-i\Delta t(H_0 + \sum_{i=0}^{n} u_{ij} H_i)\}$
  end for
  $U \to \prod_1^n U_i$
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 [14] Khaneja et al. derive a simple update formula that is correct to first order. In particular one can show that:

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

Given such a Hamiltonian, a threshold approximation fidelity $\mathcal{F}$, and a target gate U_T, we can use the technique described in **??** to generate candidate pulse sequences that enact unitaries U_i that approximate U_T, such that $Tr(U_T^{\dagger}U_i) \ge \mathcal{F}$.

In this paper we have modified the update step in 5 to instead use approximately the following gradient:

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

with $p(\vec{\delta})$ Gaussian distributed. This technique has been used in previous works such as [15] to ensure that the optimal control results are robust over a wide range of errors, and we approximate this integral in this paper by using Gaussian quadrature.[citation needed]Doing this ensures that the family of controls produced by our routine, U_i, perform moderately well over a range of control errors that might occur.

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 vary over time, or a trapped ion quantum computer where the intensity, frequency, and phase of the laser might drift over time.[16] Correlations between different δ_i might arise if two of the controls have the same noise source. Examples of shared noise sources include X and Y gates in superconducting qubit architectures might use the same AWG and pulse envelope, and diurnal temperature drift of control electronics. [citation needed]

B. BCS Approximation

After using GRAPE or another optimal control routine to synthesize a collection of controls, we must find the weights w_i such that the collection of controls form a BCS as described in (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. If we consider the Pauli-Liouville representation[17] 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* operations. In particular, we see that we can write a convex sum over these channels as:

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

Now, to approximate a depolarizing channel we define our optimal control problem to be the following, which minimizes the off-diagonal terms:

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

This can be solved with a constrained minimization algorithm, such as Sequential Least Squares Programming[18].

Previous authors have considered minimizing the diamond distance to the nearest Pauli or Clifford Channel [19], and while this gives a good theoretical framework, it requires optimizing over the diamond norm, and in particular does not have the restriction that the optimal channel be decomposable into a given family of controls. Our routine, on the other hand, optimizes over an easy to compute sum. In the next section we give a simple example, followed by numerical results for one qubit and two qubits gates in sections IV A and IV B.

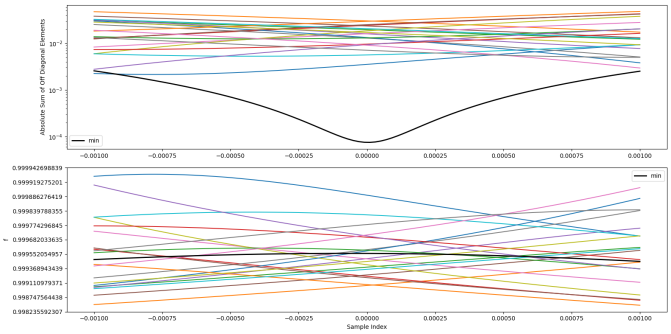
IV. NUMERICAL RESULTS

A. 1Q Gates

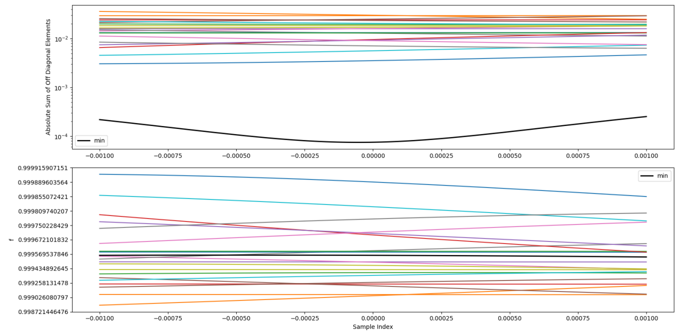
In this section, we present numerical results on generating one-qubit gates that together with $RZ(\theta)$ rotations are universal for one-qubit computation. Our control Hamiltonian is given as:

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

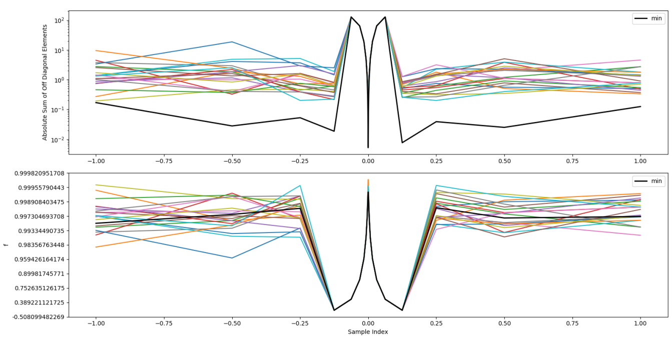
where $\epsilon \sim \delta \sim \mathcal{N}(0, .001)$ We assume that the errors on σ_x and σ_y are perfecly correlated, as mentioned in



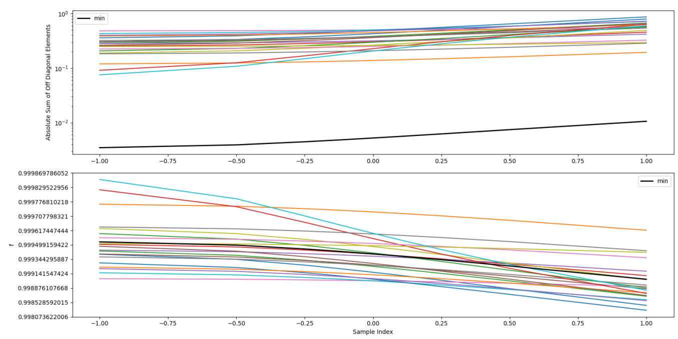
(a) This is one example of a 1D slice varying over one control.



(b) This is another example of a 1D slice varying over one control.



(a) This is one example of a 1D slice varying over one control. There are five of these in total, but only one is interesting.



(b) This is one example of a 1D slice varying over one control. There are five of these in total, but only one is interesting.

Section III. In our simulation we chose an total evolution time of $T = \pi$, and a number of steps $N = 100$, with a threshold infidelity of $1E-3$.

The results can be seen in ??, where we have plotted each control as a function of one detuning value, fixing the others to zero. In the top half of each plot, we show the variation in the sum of the absolute valua of the off diagonal elements of the Pauli-Liouville Representation, while in the lower plot we see the variation in fidelity. While the controls were optimized by considering Gaussian noise with $\sigma = .001$ for each control, we have plotted the performance of the controls over a wider range to show more structure. We see over an order of magnitude improvement in the sum of the absolute values of the off-diagonal elements, while the fidelity remains above the specified target.

B. 2Q Gates

In this section, we present numerical results on a generating one and two-qubit gates that together with $RZ(\theta)$ rotations are universal for two-qubit computation. Our control Hamiltonian is given as:

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

We again assume that the standard deviations are .001 on all parameters δ_j and c_i^j and that errors on the single qubit σ_x and σ_y rotations are perfectly correlated. In this simulation we again had 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.

As in **??**, the results of these numerics can be visualized in **??**. Notably, even in the two qubit case where optimization over many controls becomes more difficult, we see a decrease in the magnitude of the off-diagonal terms by over an order of magnitude.

V. EXPERIMENTAL RESULTS

To demonstrate the realizability of this routine, we implemented it on a superconducting qubit, to calibrate an

$RX(\frac{\pi}{2})$ pulse. The qubit being used had a FILL IN μs T1, and a FILL IN μs T2.

In this example we intentionally over and under-calibrated four pulseshapes. For each of these pulseshapes, in addition to their BCS and the calibrated pulse, we then ran a randomized benchmarking experiment.[citation needed] The results are plotted in 4. In each case, our Clifford sequences were decomposed into RX($\frac{\pi}{2}$)) and RY($\frac{\pi}{2}$)) pulses. In the superconducting qubit architecture used, these gates are implemented using the same pulse envelope definition. It was shown in [20] that for a particular error model, non-Markovian, or *quasi-static*, noise can result in more general Gamma Distributed noise, while Markovian noise, such as depolarizing noise, should result in Gaussian distributed fidelity estimates for each randomized benchmarking sequence depth. Our results are consistent with this, as demonstrated by the long tails in the four leftmost plots that vanish in the bottom right plot.

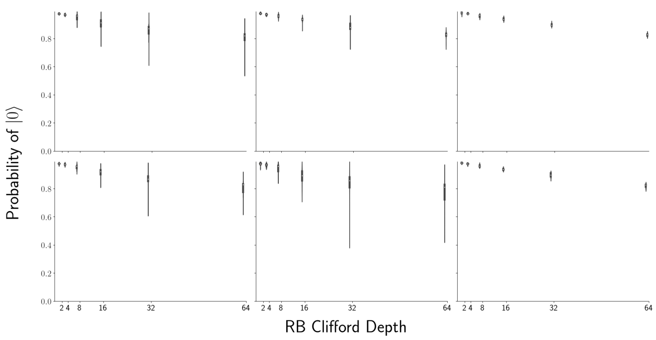


FIG. 4: Randomized benchmarking experiments ran using different pulse definitions. The four plots on the left are from the incorrectly calibrated pulse, while the top right is the calibrated pulse, and the bottom right is the BCS.

VI. CONCLUSION

We have shown numerically and experimentally that drawing from a collection of implementations of a gate with the correct probabilties can reduce coherent error by an order of magnitude, while increasing incoherent error by a smaller amount, at virtually no cost to gate fidelity. We have demonstrated that these approximate controls can be generated through optimal control, and that the minimization problem is tractable. In addition, we have shown that it is possible to perform the routine on existing quantum hardware. Future directions for this work include moving the random gate selection from a precompilation step onto runtime FPGA logic, investigating other optimzation routines such as CRAB [12] and GOAT[13], and using more precise benchmarking routines such as GST[5] to more quantitatively investigate the performance of these routines. The code used in this paper is available at [21].

VII. ACKNOWLEDGEMENTS

[1] D. Aharonov, A. Kitaev, and J. Preskill, Physical Review Letters **96** (2006), 10.1103/physrevlett.96.050504.

[2] J. Preskill, "Quantum computing in the nisq era and beyond," (2018), arXiv:1801.00862.

[3] A. Y. Kitaev, Russian Mathematical Surveys **52**, 1191 (1997).

[4] 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.

[5] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, Nature Communications **8** (2017), 10.1038/ncomms14485.

[6] D. Puzzuoli, C. Granade, H. Haas, B. Criger, E. Magesan, and D. G. Cory, Physical Review A **89** (2014), 10.1103/physreva.89.022306.

[7] J. J. Wallman and J. Emerson, Physical Review A **94** (2016), 10.1103/physreva.94.052325.

[8] E. Campbell, Physical Review A **95** (2017), 10.1103/physreva.95.042306.

[9] B. Heim, K. M. Svore, and M. B. Hastings, "Optimal circuit-level decoding for surface codes," (2016), arXiv:1609.06373.

[10] 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.

[11] 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.

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

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

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

[15] 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.

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

[17] 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.

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

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

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

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