Decorrelating Errors in Quantum Gates by Random Gate Synthesis

Anthony Polloreno
Rigetti Computing, Berkeley, CA

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

(Dated: July 22, 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, but that also violates fundamental assumptions made by the models used in these results, such as Markovianity [3] , independance of errors [citation needed], and even the magnitude of the noise [citation needed]. With these assumptions violated, promises about system performance cannot be made as easily.

As examples of these deviations from the theory, there are many ways in which devices can behave in non-Markovian ways, such as drift in controls, [4] and general non-Markovian noise [5]. Even if a given system behaves in a Markovian way, many models make simplifying assumptions about the noise on that system - such as positing that it has a simple structure. Because of its classical simulability, the channel that is often chosen is the Pauli channel [citation needed]. From there, one can ask what models can honestly be represented by using these simple models [citation needed], but the fundamental problem with this approach is the pessimism that comes along with it. As we get closer to having real, usable systems, we want models that not only give us guarantees, but faithfully represent our system.

Other authors have tried to address these problems [6–8] by using circuit composition to try to remove correlated noise. Given a process that is generated by a sequence of channels with some unknown, potentially channel depdenent error, the realization has been made that coherent error in the process can be averaged away into incoherent error, at the cost of an additional compilation step. Some of these routines have even been demonstrated experiementally [9], and been shown to reduce the off diagonal elements in the Pauli Liouville representation [citation needed]of the channels.

A common feature of these routines is that take advantage of the fact that processes are generally enacted as a sequence of gates. By inserting a potentially expensive classical pre-compilation step, they can exploit sequence structure of a process to average away coherent error. In the case of [7], the analysis takes the point of view of treating Pauli and Clifford gates as being free, which is out of reach of modern quantum computers[citation needed]. In [6] and the implementation in [9], the random frames are implemented in each sequence by precompiling the inserted Pauli frame rotations with the Clifford operations. In general, this cannot be done efficiently for non-Clifford operations, and there for this routine is demanding given the severe circuit depth constraints of modern quantum computers.[citation needed]

To remedy these problems, we take a different approach 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 channel is defined as:

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

* Corresponding author: kyoung@sandia.gov

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 and correlated error into uncorrelated error. The former gives the benefit that the error incurred from incoherent errors compound linearly, while coherent errors compound quadratically. The latter gives the benefit of reducing non-Markovian effects. This is a particularly useful property of BCSs, as diagnosing non-Markovianity in a gateset is challenging, but essential to perform reliable quantum computation. 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 [10], 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 massively reduced, and the process can be made much closer to Markovian, at no real cost to the fidelity of the implemented gate. Unlike in [6, 7, 9] this routine can be implemented natively in control electronics, and only requires that the onboard sequencer has the abiltiy to generate random numbers - no pre or post processing is required.

II. GRAPE

Generating BCSs can be done in a variety of ways, using any quantum optimal control technique [11, 12] to find a family of controls. For simplicity in this paper we use the GRAPE algorithm to generate candidate pulseshapes to approximate the target gate. First described in [13], 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

$$U_i = \exp\{-i\Delta t(H_0 + \sum_{j=1}^{n} u_{ij}H_{ij}\} \tag{3}$$

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 $1 \leq i \leq n$ **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 [13] Khaneja et al. derive a simple update formula that is correct to first order. In particular one can show that:

$$\frac{\partial J(U)}{\partial u_{ij}} = -2Re\{\left\langle U_{j+1}^\dagger ... U_N^\dagger U_T | i\Delta t H_j U_j ... U_1 \right\rangle \left\langle U_j ... U_1 | U_{j+1}^\dagger ... U_N^\dagger U_T \right\rangle\} + \mathcal{O}(\Delta t^2) \tag{4}$$

III. OPTIMAL CONTROL PROBLEM

A. Random Gate Synthesis

To explore the utility of BCSs, we consider the following Hamiltonian:

$$H(t) = \delta_0 H_0 + \sum_{i=1}^{n}(1+\delta_i)c_i(t)H_i \tag{5}$$

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, frequeny, and phase of the laser might drift over time. [citation needed](Blueprint for a microwave trapped ion quantum computer) 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]

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

In this paper we have modified the update step in 4 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 [14] 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.

B. Channel 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 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:

$$\frac{1}{N}\sum_{i=1}^{N} w_i \mathcal{E}^\dagger (U_T \rho U_T^\dagger)\mathcal{E} \tag{7}$$

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

$$\begin{aligned} &\underset{w_0,\ldots,w_N}{\textbf{minimize}} \{\sum_{i\neq j}^{N} |\sigma_i \Lambda(\sigma_j)|^2\} \\ &\textbf{where } \Lambda(\sigma_j) := \sum_{i=1}^{N} w_i \mathcal{E}_i^\dagger \sigma_j \mathcal{E}_i \\ &\textbf{subject to} \sum_{i=1}^{N} w_i = 1 \end{aligned} \tag{8}$$

This can be solved with a constrained minimization algorithm, such as Sequential Least Squares Programming[15].

Previous authors have considered minimizing the diamond distance to the nearest Pauli or Clifford Channel [16], 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 V A and V B.

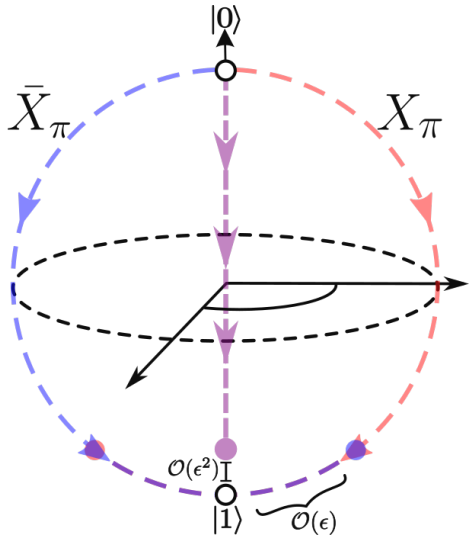

FIG. 1. An example of a BCS. The red path is a π pulse over-rotating clockwise, while the blue path is a π pulse over-rotating counter-clockwise. The purple path is the average of the two.

IV. 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 8 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:

$$\begin{aligned} &\frac{1}{2}(X^*_{\pi+\epsilon} \otimes X_{\pi+\epsilon} + \bar{X}^*_{\pi+\epsilon} \otimes \bar{X}_{\pi+\epsilon}) \\ &= (\sin^2\frac{\pi+\epsilon}{2} I\otimes I + \cos^2\frac{\pi+\epsilon}{2} X\otimes X)X\otimes X \\ &\approx ((1-\epsilon^2) I\otimes I + \epsilon^2 X\otimes X)X\otimes X \end{aligned} \tag{9}$$

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

V. 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 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 $1E-3$.

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:

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

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.

VI. 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 20.5 μs T1, and a 11.00 μs T2, and the pulse shapes being considered were 30ns long top hat pulse sequences, generated via a set-and-hold signal generator using the amplitudes shown in 5.

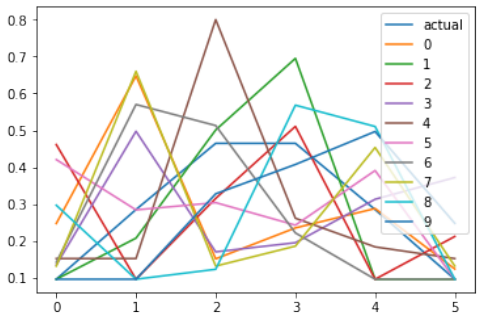

FIG. 5. The pulseshapes used in the routine. The original pulse was generated by iteratively applying $RX(\frac{\pi}{2})$ pulses, fitting the resulting sinusoid, and updating the controls. The variants were generated via optimal control, seeding the GRAPE algorithm with the initial pulse.

In this example we generated 10 candidate pulses using the GRAPE algorithm with a high convergence tolerance, with the intention that the optimizer would find solutions with some degree of noticeable coherent error. For each pulseshape, which enacts some unitary U_i, we performed an RPE-like experiment[17] where the σ_y and σ_z expectations values were measured at the end of a sequence of n applications of U_i, where n ranged from 1 to 100, and each circuit was sampled 1000 times. The results are plotted in [REF SUBPLOTS OF 6]. Then, using the same optimization tool chain as in V, we picked out the coeffients p_i such that p_i, U_i are a BCS, and performed the same experiment, but where for each shot of the experiment, we precompiled a different circuit. In particular, in our sampling procedure, circuit $U_{i_0}, U_{i_1}, ..., U_{i_{n-1}}$ occurs with probability $p_{i_0}p_{i_1}...p_{i_{n-1}}$. While in this example high fidelity, low coherent error controls can be produced [REFERENCE SUBPLOT OF 5], this demonstrates that this method can reduce coherent errors drastically.

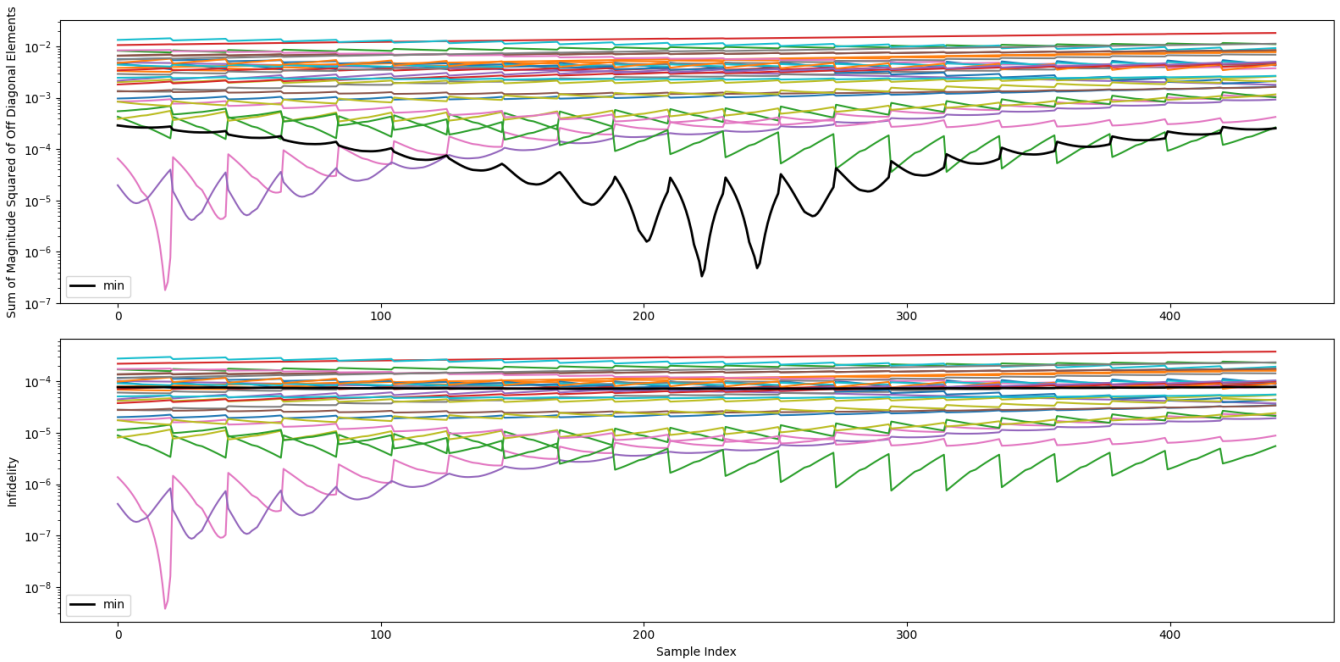

FIG. 2. In the above we have used our method to produce a BCS for a $RX(\frac{\pi}{2})$ pulse, with the Hamiltonian given in 10. A) shows the sum of the squared norms of the off diagonal terms in the Pauli Transfer Matrix of $\mathcal{E}$. B) gives the corresponding infidelity for each member of the BCS, and the optimal choice of weights.

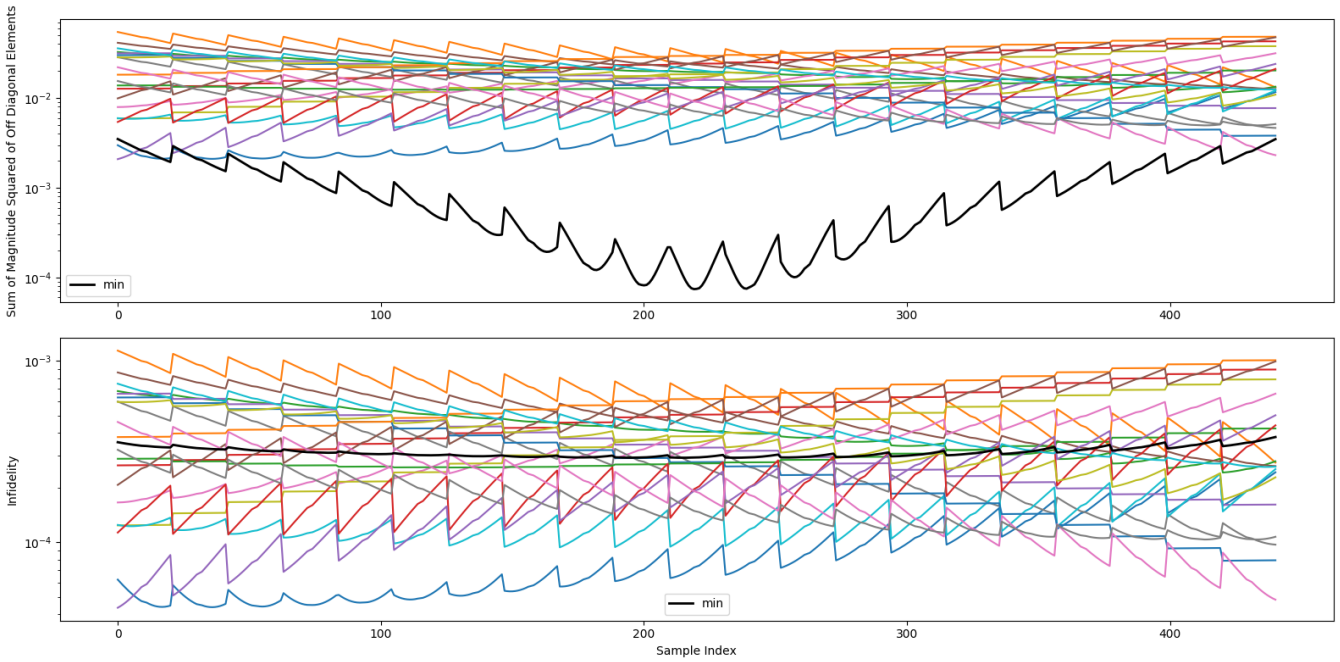

FIG. 3. In the above we have used our method to produce a BCS for a $RY(\frac{\pi}{2})$ pulse, with the Hamiltonian given in 10. A) shows the sum of the squared norms of the off diagonal terms in the Pauli Transfer Matrix of $\mathcal{E}$. B) gives the corresponding infidelity for each member of the BCS, and the optimal choice of weights.

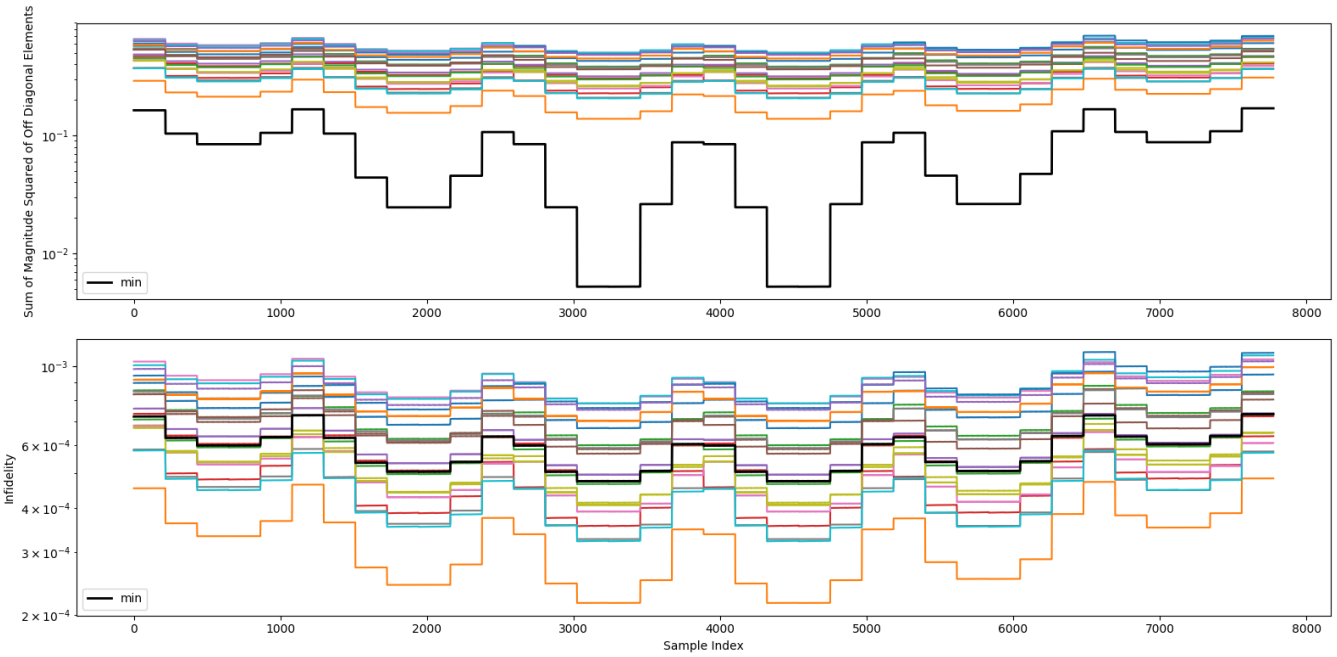

FIG. 4. In the above we have used our method to produce a BCS for a ZZ($\frac{\pi}{2}$) entangling gate with controls given by 11. A) shows the sum of the squared norms of the off diagonal terms in the Pauli Transfer Matrix of $\mathcal{E}$. B) gives the corresponding infidelity for each member of the BCS, and the optimal choice of weights.

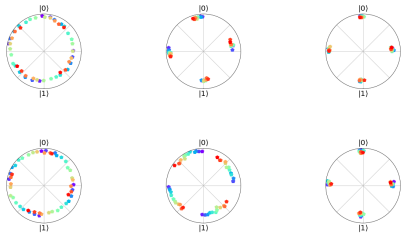

FIG. 6. An experimental implementation of our routine. We should include an array of images, maybe similar to this one, and then comment that in the image representing the BCS (above) that the noise is dominantly incoherent. The colors, ranging from purple to red, correspond to the depth into the sequence, where the darkest purple point is at the top of the rebit at $|0\rangle$. As the color becomes more red, the gate is applied more times, and consequently in the BCS example, moves the qubit's state vector closer to the maximally mixed state at the origin. The angle was computed by taking the arctangent of the σ_y and σ_z expectations, while the ther radius was computed by adding the two expectations in quadrature.

VII. 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 [11] and GOAT[12], 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 [18].

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

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

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

[9] 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.

[10] 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.

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

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

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

[14] 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.

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

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

[17] S. Kimmel, G. H. Low, and T. J. Yoder, Physical Review A **92** (2015), 10.1103/physreva.92.062315.

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