Error mitigation via stabilizer measurement emulation

A. Greene,1 M. Kjaergaard,1 M. E. Schwartz,2 G. O. Samach,1,2 A. Bengtsson,1,3 C. McNally,1 M. O’Keeffe,2 D. K. Kim,2 M. Marvian,1,4 A. Melville,2 B. M. Niedzielski,2 A. Vepsäläinen,1 R. Winik,1 J. Yoder,2 D. Rosenberg,2 S. Lloyd,1,4 T. P. Orlando,1 I. Marvian,5 S. Gustavsson,1 and W. D. Oliver1,2,6,7

1Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
2MIT Lincoln Laboratory, Lexington, MA 02421, USA
3Microtechnology and Nanoscience, Chalmers University of Technology, Göteborg, SE-412 96, Sweden
4Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
5Departments of Physics & Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA
6Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
7Department of Electrical Engineering & Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

(Dated: February 12, 2021)

INTRODUCTION

Over the past few years, impressive strides have been made in the field of quantum computing. Quantum advantage has been reported [1] and there is now an ecosystem of cloud-based quantum processors and companies interested in using them [2–5]. However, high error rates continue to limit circuit depth such that solving real-world problems with today’s quantum computers remains a challenge [6, 7]. Errors in quantum devices can be grouped into two categories: incoherent (stochastic) and coherent (deterministic). Although coherent errors do not reduce the purity of a qubit state, they are surprisingly pernicious. For small errors, infidelity grows linearly with incoherent errors and quadratically with coherent errors [8]. As quantum processors grow in complexity and computational power, control signal cross-talk and coherent errors from spectator qubits become increasingly problematic for gate performance and increasingly challenging to characterize [9]. Unlike incoherent errors, they compound each other in unpredictable ways [10] that lead to substantially worse performance than predicted by randomized benchmarking techniques.

Feedback-based quantum error correction (QEC) is designed to suppress the high error rates in quantum hardware, and attaining QEC is a long-term goal of the quantum computation community [11–15]. QEC works by making error detection measurements that are carefully designed to preserve quantum information, and then performing restorative gates on the system conditioned on the error pattern observed in order to return the processor to an error-free state. However, performing feedback is an enormous technical challenge. Compared to other operations on today’s superconducting qubit quantum processors, readout has the highest error rate by an order of magnitude [1, 16, 17]. Readout is also the longest operation on these processors by up to two orders of magnitude [17, 18], so teams working to get a feedback operation below the threshold for fault-tolerant computation must wrestle with a substantial incoherent error as well as a relatively low assignment fidelity [19]. It thus behooves us to get as much mileage as we can out of error reduction techniques which do not utilize feedback (i.e., feed-forward error mitigation), especially in the era of Noisy Intermediate-Scale Quantum (NISQ) technology [20].

Dynamical decoupling (DD) is the primary feed-forward control technique in use today [21–25]. It takes advantage of temporal symmetries in the structure of the noise in order to reverse errors, and is highly effective against low frequency incoherent noise. The canonical example of such a technique is the Hahn echo, pioneered in the context of magnetic resonance imaging [26]. However, most sources of coherent error (such as imperfectly calibrated gates and cross-talk) are discrete events with no inherent temporal symmetry to exploit.

Randomized compiling [27, 28] is a promising new feed-forward control technique that specifically addresses coherent errors from complex gates by inserting stochastically chosen high-fidelity gates to make those errors inco-
herent. A twirling gate $T_k$, drawn uniformly at random from the group generated by Pauli gates and the phase gate, is inserted before each error-prone gate $G_k$. To preserve the original computation, $G_k$ is followed by a correction gate $T_k^{-1}G_kT_k = G_k$. This renders coherent errors incoherent, which reduces error rates and lowers the gate fidelity threshold required for fault-tolerant quantum computation.

In this work, we introduce a related feed-forward control technique for addressing coherent errors, which uses stochastic gates but requires no classical pre-computation. Previous work on stabilizer codes has shown that errors can be reduced simply by measuring the stabilizers, even without feedback [29-33]. We take this a step further and investigate the effects of ‘simulating’ the quantum channel of measurement via the stochastic application of single-qubit gates, which is faster than real measurement by more than an order of magnitude. We find that this technique leads to an improvement in circuit performance. This “quantum measurement emulation” (QME) technique extends naturally from the stabilizer code formalism and is particularly well-suited for discrete coherent errors that are challenging for dynamical decoupling to address.

**OPERATING PRINCIPLE OF ‘EMULATING’ A QUANTUM MEASUREMENT CHANNEL**

Consider the measurement of a state $\rho$ along the $z$-axis in an experiment without feedback, described by the Kraus operators $P_0 = |0\rangle\langle 0|$ and $P_1^z = |1\rangle\langle 1|$: 

$$M_z(\rho) : \rho \mapsto \sum_{i=0,1} P_i^z \rho P_i^{z\dagger}$$

$$M_z \left( \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \right) = \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix} \tag{1}$$

The same quantum channel (i.e. mapping) is implemented by 

$$D_z(\rho) : \rho \mapsto \frac{1}{2} I \rho I + \frac{1}{2} Z \rho Z$$

$$D_z \left( \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \right) = \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix} \tag{2}$$

which is the dephasing operator [34]. Equation (2) provides an operational interpretation of the act of measurement: the state is dephased in the basis of the measurement. Equation (2) can also be interpreted as a probabilistic application of either the identity operation (I) or a Pauli $Z$ operation.

Taken together this observation implies that when averaging over an ensemble of identically prepared qubits, applying a $Z$ gate with probability 0.5 (and otherwise doing nothing) has the same effect on the qubit’s density matrix $\rho$ as performing a measurement in the $z$ basis. We call this stochastic operation *quantum measurement emulation*:

$$\boxed{QME_Z} = \begin{cases} 1 & \text{with } p = \frac{1}{2} \\ Z & \text{with } p = \frac{1}{2} \end{cases} \tag{3}$$

For an intuitive picture, consider a vector on the Bloch sphere. Fig 1a visualizes the effect of $z$-axis measurement as a dephasing channel. The measurement projects the quantum state onto the $z$-axis and reduces the length of the Bloch vector. b Stochastic gate application can create the same quantum channel as $z$-axis measurement. c Various qubit initial states reconstructed with state tomography (grey squares), plotted after applying $z$-axis measurement. d State tomography data of qubit $\rho$, plotted after applying $I$ (grey squares), Pauli $Z$ gate (blue triangles) and $QME_Z$ (green circles) for varying initial states. We see that $QME_Z$ implements the same channel as $z$-axis measurement.
ments for a variety of input states. As expected from Equation (2), the probabilistic application of $Z$ and $1$ gates effectively projects the qubit state onto the $z$-axis, with some small deviations due to measurement errors. To highlight the effective equivalence between this and actual $z$-basis measurement, we prepare the same initial states and plot $\langle Z \rangle$ (Fig 1c).

The equivalence we have described between $M_z(\rho)$ and $D_z(\rho)$ can be generalized to measurement along any axis, including those in a multi-qubit basis. The observable for an $n$-qubit measurement along a multi-qubit axis $\nu$ corresponds to a unitary $S_\nu \in SU(2^n)$. Measurement along $\nu$ is performed by replacing the projectors in Equation (1) with the projectors for the $\pm 1$ eigenstates of $S_\nu$. Consequently, the generalized dephasing channel is given by:

$$D_\nu(\rho) : \rho \mapsto \frac{1}{2} |\psi\rangle \langle \psi | \rho |\psi\rangle \langle \psi | + \frac{1}{2} S_\nu \rho S_\nu^\dagger$$

(4)

This implies that a quantum measurement emulation of any Pauli operator can be performed with the stochastic application of single-qubit gates. As an example, we show below how the quantum channel for $\langle ZZ \rangle$ measurement can be emulated using single-qubit $Z$ gates.

$$QME_{ZZ} = \begin{cases} 
-|1\rangle \langle 1| & \text{with } p = \frac{1}{2} \\
-|1\rangle \langle 1| & \text{with } p = \frac{1}{2}
\end{cases}$$

(5)

An important example of this is when $S_\nu$ is the stabilizer for a logical qubit. Stabilizer codes are designed such that stabilizer measurements do not corrupt the quantum information stored in the logical qubit, and the measurement results can be used as a parity check for error detection or error correction. In this case, performing $QME_\nu$ emulates a stabilizer measurement which can be used to mitigate coherent errors which rotate the state outside of the codespace.

**USING QME TO MITIGATE COHERENT ROTATION ERRORS**

After a coherent error has occurred, applying the appropriate QME will project the qubit onto the measurement axis, effectively converting the coherent error into an incoherent error. Fig. 2a shows a qubit initially in the $|0\rangle$ state (white) which is subjected to a coherent rotation error $R_x(\theta)$ (grey). When QME is performed along the $z$-axis (denoted $QME_Z$), the qubit state is projected onto the $z$-axis. After QME, the coherent error which rotated the qubit state on the surface of the Bloch sphere is now manifest as an incoherent error which shortens the length of the vector along the measurement axis. After applying $QME_Z$ the trace distance between the initial and measured states is only second order in the size of the original perturbation. Here, trace distance $T(\rho, \sigma) = \frac{1}{2} \| \rho - \sigma \|_1$ is a more sensitive metric than fidelity — trace distance measures the maximum distinguishability between states, whereas fidelity measures the distinguishability along the axis of the qubit state.

To demonstrate QME as a state stabilization technique, we prepare the qubit in the $|0\rangle$ state and induce coherent errors of increasing strength (Fig 2b). We plot $1 - T(\rho, \rho_{\text{initial}})$, the trace distance to the initially prepared state with (green circles) and without (grey squares) QME (Fig. 2a and 2d), overlaid with theoretical values from simulations performed using the measured initial state and no free parameters. Qubits are prepared in $|0\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)$ in a simple code stabilized by $ZZ$, subjected to a small $R_x(\theta) \otimes R_y(\theta)$ coherent error and plot the trace distance to the initially prepared state with (green circles) and without (grey squares) $QME_{ZZ}$ for increasing error strength $\theta$. In both of these examples, QME provides a first-order insensitivity to $\theta$. (In these plots, $1 - T(\rho, \rho_{\text{initial}})$ has been normalized to 1 in order to account for state preparation and measurement (SPAM) errors.)

**FIG. 2. QME for Error Mitigation**

a) Diagram depicting QME as an error mitigation technique. After the initial state is prepared (white circle), a coherent error rotates the state away (grey square). Applying $QME_Z$ projects that error into an incoherent error, pushing the qubit state back to the $z$-axis (green circle). Notice that after performing QME, the trace distance to the initial state is greatly reduced. b) Qubit is prepared in the $|0\rangle$ state and subjected to coherent errors of increasing strength. We plot the trace distance with (grey squares) and without (green circles) QME, overlaid with theoretical values from simulations performed using the measured initial state and no free parameters. Qubits are prepared in $|0\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)$ in a simple code stabilized by $ZZ$, subjected to a small $R_x(\theta) \otimes R_y(\theta)$ coherent error and plot the trace distance to the initially prepared state with (green circles) and without (grey squares) $QME_{ZZ}$ for increasing error strength $\theta$. In both of these examples, QME provides a first-order insensitivity to $\theta$. (In these plots, $1 - T(\rho, \rho_{\text{initial}})$ has been normalized to 1 in order to account for state preparation and measurement (SPAM) errors.)
values from simulations performed using the measured initial state and no free parameters. For small values of the error strength $\theta$, the corrected trace distance curve is nearly flat, indicating a first-order insensitivity to coherent errors.

To use QME to reduce the effects of deleterious coherent errors, one must know the axis along which to ‘measure’. In the single-qubit case, this corresponds to knowing the axis of the ideal qubit state — in which case, the qubit can only store classical information. Stabilizer codes, on the other hand, are carefully designed so that parity measurements may be performed without disrupting the quantum information. When QME is applied within the framework of a stabilizer code, coherent errors that would have moved information outside of the codespace are projected into incoherent errors. This both reduces the trace distance and prevents the build-up of coherent errors.

To demonstrate the power of stabilizer measurement emulation, we use a simple code consisting of Bell states and stabilized by $ZZ$, with $|\tilde{0}\rangle = \frac{1}{\sqrt{2}} |00\rangle + |11\rangle$ and $|\tilde{1}\rangle = \frac{1}{\sqrt{2}} |00\rangle - |11\rangle$. To mitigate coherent errors for this code, we will apply $\text{QME}_{ZZ}$ as outlined in Eq. 5; namely, we apply identity gates or simultaneous $Z$ gates on both qubits with 50% probability. Note that extending QME to multiple qubits does not consume more time or require more complicated gates: with $p = 0.5$ we perform an identity operation, and with $p = 0.5$ we apply a $Z$ gate to each qubit. As with the state stabilization experiment, we apply a coherent error of varying strength (now an $R_x(\theta) \otimes R_y(\theta)$ error), either do nothing or apply QME, and measure the resulting trace distance to the initial state with a state tomography measurement (Fig. 2b, 2c). Here too, applying QME leads to a distinctly smaller error. Though this plot demonstrates improved performance specifically for an $R_x(\theta) \otimes R_y(\theta)$ error, that particular error was chosen arbitrarily; any error that rotates the state outside of the $|\tilde{0}\rangle, |\tilde{1}\rangle$ manifold will behave similarly.

**Using QME to Mitigate Coherent Errors from 2QB Gates**

Two-qubit gates are the dominant source of coherent errors in all current generation quantum processors, so we explore the effects of QME in a circuit comprised of a sequence of $CZ$ gates (Fig. 3a). For this experiment we use a different code, stabilized by $XX$, with $|\tilde{1}\rangle = \frac{1}{\sqrt{2}} |00\rangle + |11\rangle$ and $|\tilde{0}\rangle = \frac{1}{\sqrt{2}} |01\rangle + |10\rangle$. ($CZ$ is not a logical operation in this code, but $CZ^2 = 1 \otimes 1$ is a logical operation in every code.) To set a performance baseline, we prepare the $|\tilde{0}\rangle$ state, apply $N$ $CZ$ pairs and record the trace distance (state fidelity) to the initial state as shown in Fig 3b(c). We then introduce error mitigation in the form of an $\text{QME}_{XX}$ inserted between every pair of $CZ$ gates.

Executing the circuit without QME results in the steep plunge and revival that is the hallmark of the accumulation of coherent errors, seen in both the trace distance and the fidelity. The device used in this experiment has $CZ$ gates with an average fidelity of 0.997 (measured via interleaved Clifford randomized benchmarking). Nevertheless, after 10 steps (20 gates) the output fidelity is nearly 0.5 — a clear demonstration of the deleterious accumulation of coherent errors. Adding QME leads to a smooth monotonic decay, which indicates that the coherent error has been made incoherent. Though adding QME increases the circuit depth, it results in a substantial improvement in circuit performance.

**Outlook**

We have demonstrated a new feed-forward control technique for mitigating coherent errors in quantum information processing which is tailored to discrete coherent errors. By using stochastically-applied single-qubit gates to ‘emulate’ quantum measurement along the appropriate axis, coherent errors can be made incoherent. Unlike randomized compiling, QME does not require the
computation of correction gates needed in randomized compiling, with the trade-off that it only protects against errors that rotate the qubit out of the logical codespace. QME also offers protection against coherent errors that might have occurred during twirling gates. We show how QME can be used in the context of stabilizer codes to improve circuit performance in terms of both trace distance and fidelity. This still holds in arbitrarily generated circuits where simple dynamical decoupling schemes do not offer an advantage. QME is a promising addition to the collection of quantum computing control techniques.

Acknowledgments: AG acknowledges funding from the 2019 Google US/Canada PhD Fellowship in Quantum Computing. MK acknowledges support from the Carlsberg Foundation during part of this work. IM acknowledges funding from NSF grant FET-1910859. This research was funded in part by the U.S. Army Research Office Grant W911NF-18-1-0411 and the Assistant Secretary of Defense for Research & Engineering under Air Force Contract No. FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

Competing interests: The authors declare no competing interests.

[1] F. Arute, et al., *Nature* **574**, 505–510 (2019).
[2] M. Henderson, J. Gallina, M. Brett, *arXiv:2004.03079* (2020).
[3] J. M. Chow, et al., *Nature Communications* **5** (2014).
[4] R. S. Smith, M. J. Curtis, W. J. Zeng, *arXiv:1608.03355* (2016).
[5] A. R. R. Carvalho, H. Ball, M. J. Biercuk, M. R. Hush, F. Thomsen, *arXiv:2010.08057* (2020).
[6] A. Martin, et al., *arXiv:1904.05803* (2019).
[7] Y. Cao, et al., *Chemical Reviews* **119**, 10856–10915 (2019).
[8] A. Carignan-Dugas, J. J. Wallman, J. Emerson, *New Journal of Physics* **21**, 053016 (2019).
[9] S. Kriinner, et al., *arXiv:2005.05914* (2020).
[10] Y. R. Sanders, J. J. Wallman, B. C. Sanders, *New Journal of Physics* **18**, 012002 (2015).
[11] D. Lidar, T. Brun, *Quantum Error Correction* (2013).
[12] J. P. B. Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, B. J. Brown (2020).
[13] C. Chamberland, G. Zhu, T. J. Yoder, J. B. Hertzberg, A. W. Cross, *Physical Review X* **10** (2020).
[14] L. Egan, et al., *Fault-tolerant operation of a quantum error-correction code* (2020).
[15] C. Chamberland, et al., *Building a fault-tolerant quantum computer using concatenated cat codes* (2020).
[16] P. Jurcevic, et al., *arXiv:2008.08571* (2020).
[17] J. Stehlik, et al., *Tunable coupling architecture for fixed-frequency transmons* (2021).
[18] D. Sank, et al., *Physical Review Letters* **117** (2016).
[19] D. Ristè, C. C. Bultink, K. W. Lehnert, L. DiCarlo, *Physical Review Letters* **109** (2012).
[20] J. Preskill, *Quantum* **2**, 79 (2018).
[21] L. Viola, S. Lloyd, *Phys. Rev. A* **58**, 2733 (1998).
[22] K. Khodjasteh, D. A. Lidar, *Physical Review Letters* **95** (2005).
[23] P. Zanardi, *Physics Letters A* **258**, 77–82 (1999).
[24] L. Viola, E. Knill, S. Lloyd, *Physical Review Letters* **82**, 2417–2421 (1999).
[25] G. S. Uhrig, *Physical Review Letters* **98** (2007).
[26] E. L. Hahn, *Phys. Rev. 80*, 580 (1950).
[27] J. J. Wallman, J. Emerson, *Phys. Rev. A* **94** (2016).
[28] A. Hashim, et al., *Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor* (2020).
[29] B. Pokharel, N. Anand, B. Fortman, D. A. Lidar, *Physical Review Letters* **121** (2018).
[30] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, H. Neven, *Nature Communications* **11** (2020).
[31] M. Müller, S. Gherardini, N. Dalla Pozza, F. Caruso, *Physics Letters A* **384**, 126244 (2020).
[32] D. Layden, E. Martín-Martínez, A. Kempf, *Physical Review A* **91** (2015).
[33] P. Facchi, D. A. Lidar, S. Pascazio, *Physical Review A* **69** (2004).
[34] M. N. ad I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2012).
[35] M. K. et al., *arXiv:2001.08838* (2020).
Supplementary material for “Error mitigation via stabilizer measurement emulation”

A. Greene,1 M. Kjaergaard,1 M. E. Schwartz,2 G. O. Samach,1,2 A. Bengtsson,1,3 C. McNally,1 M. O’Keeffe,2 D. K. Kim,2 M. Marvian,1,4 A. Melville,2 B. M. Niedzielski,2 A. Vepsäläinen,1 R. Winik,1 J. Yoder,2 D. Rosenberg,2 S. Lloyd,1,4 T. P. Orlando,1 I. Marvian,5 S. Gustavsson,1 and W. D. Oliver1,2,6,7

1Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
2MIT Lincoln Laboratory, Lexington, MA 02421, USA
3Microtechnology and Nanoscience, Chalmers University of Technology, Göteborg, SE-412 96, Sweden
4Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
5Departments of Physics & Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA
6Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
7Department of Electrical Engineering & Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

DEVICE

Supplementary Figure 1.

The device used for these measurements is the same used for [34]. As shown in the SEM image in Supplementary Figure 1, this chip has three asymmetric ‘xmon’-style transmon qubits with individual microwave and flux control lines. All experiments in this paper were performed with the two leftmost qubits; their properties are reported in Table I. Interleaved randomized benchmarking data for single- (two-) qubit gates are shown in Supplementary Figure 2 (3).

| Parameter                         | Qubit 1                      | Qubit 2                      |
|-----------------------------------|------------------------------|------------------------------|
| Idling frequency, $\omega_i/2\pi$ | 4.748 GHz                    | 4.225 GHz                    |
| Anharmonicity, $\eta/2\pi$        | $-175$ MHz                   | $-190$ MHz                   |
| Coupling strength, $g/2\pi$       | 10.6 MHz                     |                              |
| Readout resonator frequency, $f_i/2\pi$ | 7.251 GHz                   | 7.285 GHz                    |
| Junction asymmetry                | 1:5                          | 1:10                         |
| Relaxation time at idling point, $T_1$ | 23 $\mu$s                   | 39 $\mu$s                   |
| Coherence time at idling point, $T_{2R}$ | 13 $\mu$s                   | 25 $\mu$s                   |
| Effective relaxation time undergoing CZ trajectory, $\tilde{T}_1$ | $\approx 17$ $\mu$s (same as idling) |                              |
| Effective coherence time undergoing CZ trajectory, $\tilde{T}_{2R}$ | $\approx 5$ $\mu$s (same as idling) |                              |
| Single-qubit gate fidelity, $t_{\text{1qb}}$ | $\geq 0.999$                | $\geq 0.999$                |
| Single-qubit gate time, $t_{\text{1qb}}$ | 30 ns                       | 30 ns                       |
| Two-qubit gate fidelity, $t_{\text{CZ}}$ | 0.997                       |                              |
| Two-qubit gate time, $t_{\text{CZ}}$ | 60 ns                       |                              |

1 See M. K. et al. arXiv:2001.08838 (2020) for details

TABLE I. Parameters of the two qubits used in this work.
In the main text, Equation 1 and Equation 2 describe how the quantum channel for measuring the expectation value of an arbitrary unitary operator $S_v$ can be performed by stochastically applying $S_v$. Figure 1 demonstrates this equivalence for a measurement along the $z$-axis. Here, we show similar data for emulated measurement along the $x$-axis (Supplementary Figure 4a) and along the $x+y$-axis (Supplementary Figure 4b).

**ERROR MITIGATION**

In Figure 2 of the main text, we demonstrate QME as an error mitigation technique using an intentionally applied $\sigma_x \otimes \sigma_x$ over-rotation error of varying strength. In Figure we show that QME leads to first-order insensitivity to a variety of transversal errors.
All theory curves used in this paper were generated by simulating the described circuit in QuTiP, using the initial $\rho$ as measured by state tomography. In Fig 3 and Fig 4, the nonidealties in the CZ gates were captured by representing each CZ gate as the composition of the ideal operation with a CPHASE($\phi$) over-rotation, single-qubit Z over-rotations $R_Z(\theta_1) \otimes R_Z(\theta_2)$ and a leakage channel parameterized by leakage rate $\lambda$. The parameters $\phi, \theta_1, \theta_2$ and $\lambda$ were fit to the state tomography data taken at each step of the circuit. The simulations for Fig. 3 and Fig. 4 also included decoherence, in the form of amplitude-damping and dephasing channels added after each gate in the circuit. The decay rates for these channels were determined by measured coherence times (shown in Table I) and the duration of the preceding gate (30ns single-qubit gates and 60ns two-qubit gates, with a 5ns gap between pulses).

Our state tomography analysis does not discriminate between $|1\rangle$ and $|2\rangle$, so leakage from $|11\rangle$ to $|20\rangle$ looks like incoherent population transfer to $|10\rangle$. Accordingly, we used the following Kraus operators used to model leakage:

$$L_{1,\lambda}(t) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sqrt{1-\lambda} \end{pmatrix}, \quad L_{2,\lambda\text{amba}}(t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \sqrt{\lambda} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The channel that composes amplitude damping and dephasing is given by:

$$\mathcal{E}_{qk}(t) : \rho_{qk} \mapsto \sum_{i=1,2} A_{i,\Gamma_1}(t)D_{j,\Gamma_\phi}(t)\rho_{qk}D_{j,\Gamma_\phi}^*(t)A_{i,\Gamma_1}^*(t),$$

where $A_{i,\Gamma_1}(t)$ is the amplitude damping process (with $\Gamma_1 = 1/T_1$), and $D_{j,\Gamma_\phi}(t)$ is the dephasing process ($\Gamma_\phi = 1/T_{2R} - 1/2T_1$). $\Gamma_1,qk$ and $\Gamma_\phi,qk$ are the appropriate coherence parameters for qubit $k$, and $t$ is the time of the
preceeding single- or two-qubit gate on that qubit. The amplitude damping and dephasing Kraus operators are given by:

\[
A_{1, \Gamma_1}(t) = \begin{pmatrix} 1 & 0 \\ 0 & e^{-\Gamma_1 q_k t / 2} \end{pmatrix}
\] (3)

\[
A_{2, \Gamma_1}(t) = \begin{pmatrix} 0 & \sqrt{1 - e^{-\Gamma_1 q_k t}} \\ 0 & 0 \end{pmatrix}
\] (4)

\[
D_{1, \Gamma_\phi}(t) = \begin{pmatrix} e^{-\Gamma_\phi q_k t / 2} & 0 \\ 0 & e^{-\Gamma_\phi q_k t / 2} \end{pmatrix}
\] (5)

\[
D_{2, \Gamma_\phi}(t) = \begin{pmatrix} \sqrt{1 - e^{-\Gamma_\phi q_k t}} & 0 \\ 0 & 0 \end{pmatrix}
\] (6)

\[
D_{3, \Gamma_\phi}(t) = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{1 - e^{-\Gamma_\phi q_k t}} \end{pmatrix}
\] (7)