Supplementary Information for "Measuring and Suppressing Quantum State Leakage in a Superconducting Qubit"

Zijun Chen,[1] Julian Kelly,[2] Chris Quintana,[1] R. Barends,[2] B. Campbell,[1] Yu Chen,[2] B. Chiaro,[1] A. Dunsworth,[1] A. G. Fowler,[2] E. Lucero,[2] E. Jeffrey,[2] A. Megrant,[1,3] J. Mutus,[2] M. Neeley,[2] C. Neill,[1] P. J. J. O'Malley,[1] P. Roushan,[2] D. Sank,[2] A. Vainsencher,[1] J. Wenner,[1] T. C. White,[1] A. N. Korotkov,[4] and John M. Martinis[1,2,*]

[1]*Department of Physics, University of California, Santa Barbara, California 93106-9530, USA*
[2]*Google Inc., Santa Barbara, California 93117, USA*
[3]*Department of Materials, University of California, Santa Barbara, California 93106, USA*
[4]*Department of Electrical and Computer Engineering, University of California, Riverside, California 92521, USA*

(Dated: December 8, 2015)

RATE EQUATION FOR $|2\rangle$ STATE POPULATION

In this section we discuss the rate equation which describes the $|2\rangle$ state population in the randomized benchmarking (RB) procedure.

Neglecting the population of the $|3\rangle$ state and higher levels, it is natural to describe (phenomenologically) the *average* population $p_{|2\rangle}(m)$ of the state $|2\rangle$ after m Cliffords using the evolution equation

$$p_{|2\rangle}(m+1) = p_{|2\rangle}(m) + \gamma_\uparrow[1 - p_{|2\rangle}(m)] - \gamma_\downarrow p_{|2\rangle}(m), \quad (1)$$

where $\gamma_\uparrow$ is the probability of the $|2\rangle$ state excitation per Clifford, averaged over Cliffords and also over the initial state in the qubit subspace, while $\gamma_\downarrow$ is the probability of returning from the state $|2\rangle$ to the qubit subspace, averaged over Cliffords. We emphasize that Eq. (1) would be invalid for a particular RB sequence, but we apply it only assuming averaging over the RB sequences: $p_{|2\rangle}(m)$, $\gamma_\uparrow$, and $\gamma_\downarrow$ are all the averaged values. So far we have introduced Eq. (1) phenomenologically; we will discuss the applicability of this equation later.

The solution to Eq. (1) is

$$p_{|2\rangle}(m) = C(1-\Gamma)^m + p_\infty, \quad p_\infty = \frac{\gamma_\uparrow}{\Gamma}, \quad \Gamma = \gamma_\uparrow + \gamma_\downarrow, \quad (2)$$

where C is a constant, determined by the initial condition, $C = p_{|2\rangle}(0) - p_\infty$. In the case $\Gamma \ll 1$ this solution can be replaced with

$$p_{|2\rangle}(m) = [p_{|2\rangle}(0) - p_\infty]\, e^{-\Gamma m} + p_\infty, \quad (3)$$

that obviously corresponds to the standard rate equation

$$dp_{|2\rangle}(m)/dm = \gamma_\uparrow[1 - p_{|2\rangle}(m)] - \gamma_\downarrow p_{|2\rangle}(m), \quad (4)$$

to which Eq. (1) reduces when m is considered as a quasicontinuous variable ($m \gg 1$). Thus, m plays the role of the dimensionless time, while $\gamma_\uparrow$ and $\gamma_\downarrow$ are the excitation and relaxation rates in this dimensionless time. Note that if $p_{|2\rangle}(0) = 0$, then Eq. (3) becomes $p_{|2\rangle}(m) = p_\infty(1 - e^{-\Gamma m})$.

Also note that if observed probabilities $\tilde{p}_{|2\rangle}$ are different from actual probabilities $p_{|2\rangle}$ due to measurement infidelity in a linear way, $\tilde{p}_{|2\rangle}(m) = A p_{|2\rangle}(m) + B[1 - p_{|2\rangle}(m)]$ (here $A \approx 1$ is the fidelity of the state $|2\rangle$ measurement, while $B \ll 1$ is the average probability of misidentifying a state within the qubit subspace as the $|2\rangle$ state), then Eqs. (1)–(4) remain valid for $\tilde{p}_{|2\rangle}(m)$, but with the slightly changed rates: $\gamma_\uparrow \to \tilde{\gamma}_\uparrow = A\gamma_\uparrow + B\gamma_\downarrow$, $\gamma_\downarrow \to \tilde{\gamma}_\downarrow = \Gamma - \tilde{\gamma}_\uparrow$, $\tilde{\Gamma} = \Gamma$. Therefore, the rates $\tilde{\gamma}_\uparrow$ and $\tilde{\gamma}_\downarrow$ extracted from the RB results, may slightly differ from the actual rates $\gamma_\uparrow$ and $\gamma_\downarrow$.

Next we discuss the applicability of the rate equation (1) for the $|2\rangle$ state population. A rate equation usually assumes incoherent processes. However, in our case both coherent and incoherent processes are important: while the rate $\gamma_\downarrow$ is mostly determined by incoherent energy relaxation, the rate $\gamma_\uparrow$ is mostly determined (at least for short gates) by a unitary evolution, though with possibly fluctuating pulse shapes. Therefore, it is not obvious if the simple rate equation is applicable. Note that we do not apply random ± 1 pulses for the $|2\rangle$ state as was suggested [1–3] for formal randomization of the coherent processes. In our opinion, for practical purposes it is not necessary because of different transition frequencies ω_{21} and ω_{10}. To illustrate this argument, let us assume only coherent excitations of the $|2\rangle$ state and consider the evolution of the wavefunction $c_0|0\rangle + c_1|1\rangle + c_2|2\rangle$ in the rotating frame based on ω_{10}. Then for a particular sequence of Cliffords (assuming $|c_2|^2 \ll 1$)

$$c_2(m) = c_2(0) + \sum\nolimits_k g_{\uparrow,k}\, e^{-i(\omega_{21}-\omega_{10})t_k}, \quad (5)$$

where the complex number $g_{\uparrow,k}$ is the contribution from kth Clifford in the sequence ($\gamma_\uparrow = \langle |g_{\uparrow,k}|^2 \rangle$) and t_k is the start time of kth Clifford. For $\omega_{21} - \omega_{10} = 2\pi \times -212\,\mathrm{MHz}$ and elementary gate time $\gtrsim 10\,\mathrm{ns}$, it is unlikely that the phase shifts in Eq. (5) are close to exact integers of 2π. Therefore, even if averaging over Cliffords and initial states does not provide full randomization in the sense that $\langle g_{\uparrow,k} \rangle \neq 0$, the extra phase factor (accumulating with k) helps to average the contributions to zero, so that in this example $|c_2(m)|^2 \propto m$ (from two-dimensional random walk), as would also be expected from a simple rate-equation model. Thus, we expect that the rate

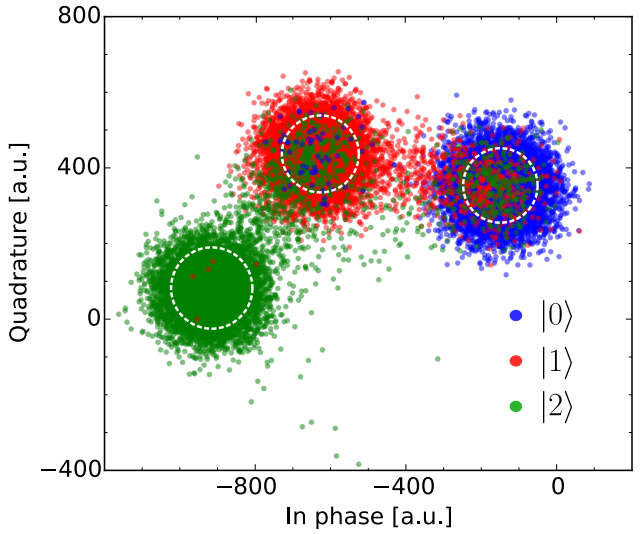

Figure S1. Phase space points corresponding to the qubit being prepared in the $|0\rangle$ (blue), $|1\rangle$ (red), and $|2\rangle$ (green) states. Out of a total of 50,000 preparations of each state, 5000 are shown here. The states are discriminated based on their distance from the center of the cloud corresponding to each state. Points of one color positioned in a cloud of a different color indicate readout errors. The white circles in each cloud have radii corresponding to one standard deviation of the complex data in each cloud.

equation should work well for coherent contributions to the leakage, and since it also works for incoherent processes, we expect the rate equation to be well applicable to our RB procedure. Experimental results presented in the main text confirm this expectation.

MEASUREMENT SETUP

The measurement setup is largely as described in the supplementary information for Ref. 4, with two primary differences. First, the qubits are no longer statically biased with a programmable voltage source separate from the Z-control DAC. Instead they are operated by internally adding a DC offset to the output of the control DAC. As such, the bias tees and attenuators on the Z-control lines at the 20 mK stage were removed. Second, the thermalization of all lines was improved by clamping the lines to all stages from 4 K to 20 mK using copper thermal anchors [5].

STATE DISCRIMINATION

Readout parameters for this device have previously been detailed in Ref. 4. At the operating point used for the experiment, we find the dispersive shift to be about 1 MHz. We readout using a 1 μs pulse. To characterize our readout fidelity, we prepare the qubit in each of the three states 50,000 times and measure. The raw IQ points of the demodulated signal [6] are shown in Fig. S1. The probability of measuring the qubit in each state given preparation in another is as follows:

$$\begin{pmatrix} 0.993 & 0.0069 & 5\times10^{-5} \\ 0.055 & 0.945 & 5\times10^{-4} \\ 0.0246 & 0.083 & 0.892 \end{pmatrix}$$

where the row indicates the state prepared and the columns indicate the state measured. The primary source of error is T_1 decay of the excited states. The readout frequency was chosen to maximize the separation between the $|2\rangle$ state and the $|1\rangle$ state, resulting in a separation error between the two clouds of phase space points of around 1×10^{-4}. However, the actual probability of preparing $|1\rangle$ and measuring $|2\rangle$ is greater, at around 5×10^{-4}. This is consistent with the heating rate of 4×10^{-7} per nanosecond as measured in the main paper, multiplied by the readout time of 1 μs.

In Ref. 4, where fast and accurate state discrimination was a requirement for the experiment, readout was noted to be a possible source of leakage when the readout resonator is strongly driven. In this work, we avoid the leakage problem and achieve high separation fidelity by reading out with low drive power and a long pulse because we do not require the readout to be fast. However, the length of the pulse leads to increased measurement error due to T_1 decay. For the purposes of this work, we prefer to have minimize measurement leakage. Optimizing measurement fidelity and minimizing leakage from measurement is an outstanding problem.

In general, we do not correct for measurement fidelity except in the thermalization measurement shown in Fig. 4 of the main article. As noted above, the extraction of leakage rates from RB data is affected by readout fidelity. Thus, the leakage rates we quote are about 10% lower than the actual leakage rates.

DEPENDENCE OF OPTIMAL PULSE DETUNING ON DRAG WEIGHT AND PULSE LENGTH

In Fig. S2(a) we show the dependence of the optimal pulse detuning on the DRAG weight α for three different π-pulse lengths. For each pulse length, the dependence is linear, and the slope becomes more shallow with longer pulse length. In Fig. S2(b), we plot the dependence of this slope on pulse length. We find that the slope between optimal detuning and DRAG is proportional to the inverse square of the pulse length. Equivalently stated, the slope depends quadratically on the drive strength, which we expect because the AC Stark shift scales quadratically with the strength of the driving field.

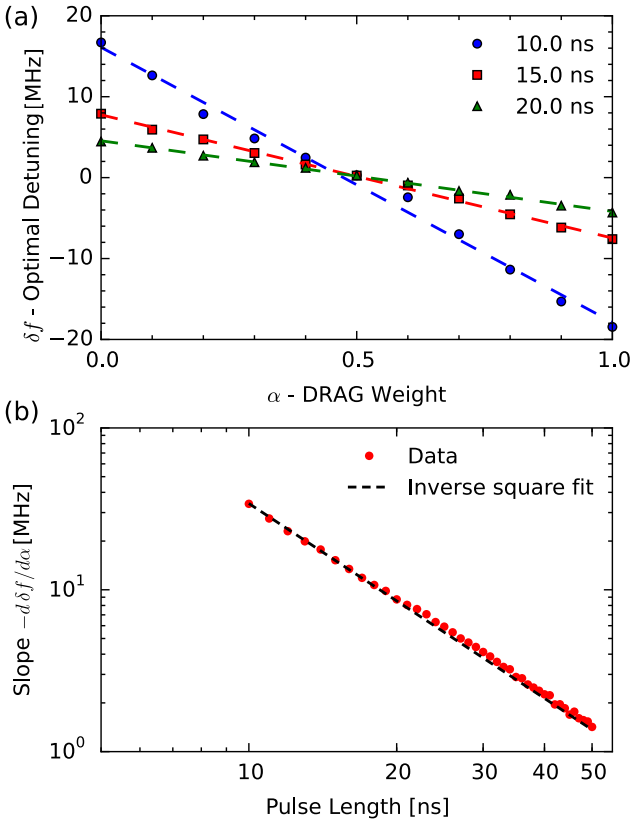

Figure S2. (a) Dependence of the optimal detuning on α. Three different pulse lengths are shown. The dashed lines are linear fits. (b) The slopes from the linear fits as shown in (a), for a range of pulse lengths. The dashed line is a fit to the inverse square of the pulse length, as expected from the AC Stark shift.

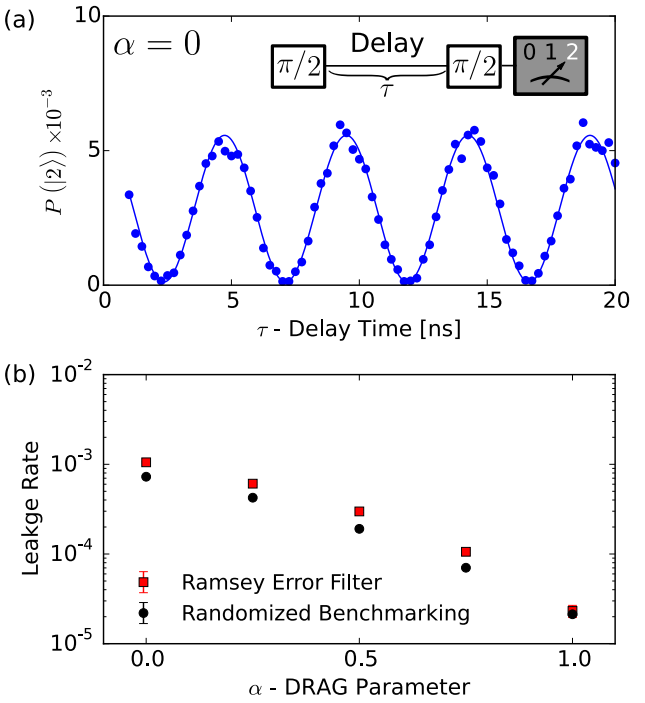

Figure S3. (a) Ramsey error filter experiment. The pulse sequence shows a typical Ramsey experiment, and the coherent population of the $|2\rangle$ state due to the two gates causes an interference fringe at the anharmonicity frequency. The leakage is estimated to be one quarter of the fringe amplitude. The data is for a $\pi/2$ pulse, but the experiment can also be done for a π pulse. (b) Comparison of leakage measured using randomized benchmarking and Ramsey error filter.

COMPARISON BETWEEN RANDOMIZED BENCHMARKING AND RAMSEY ERROR FILTERING

State leakage in superconducting qubit gates has previously been measured using a technique known as the Ramsey error filter [7, 8]. The pulse sequence for this technique is shown in the inset of Fig. S3(a). The two gates coherently populate the $|2\rangle$ state, and by varying the delay, we observe an interference fringe in the $|2\rangle$ state population because the phase of the $|2\rangle$ state evolves at a frequency different from ω_{10} by the anharmonicity η. We then estimate the leakage of the gate as a quarter of the peak to peak amplitude of the fringe.

We perform the Ramsey error filter experiment for both π and $\pi/2$ pulses, and take the weighted average of the two leakage rates ($1.5 \times \pi/2$ leakage + $0.375 \times \pi$ leakage) to estimate the leakage per Clifford gate. We then compare the leakage obtained using Ramsey error filtering with randomized benchmarking over a number of DRAG weights, with the results shown in Fig. S3(b). We find that leakage rates measured using the two techniques are generally within 50%, and follow the same trend as a function of DRAG weight. We note that the data shown is from a different device than the one used for the main paper, and the pulse lengths for a π and $\pi/2$ pulse were 15 ns and 10 ns respectively.

Given that the two techniques show similar results, randomized benchmarking has two primary advantages over Ramsey error filtering when characterizing the general leakage performance of a gateset. First, Ramsey error filtering only amplifies leakage errors by a factor of 4, whereas randomized benchmarking can amplify leakage errors by orders of magnitude, as seen in Fig. 1(d) of the main paper. More importantly, Ramsey error filtering can only detect coherent leakage errors, while the main paper established that incoherent leakage can be a dominant source of leakage when control has been well optimized. However, Ramsey error filtering is still useful for measuring the coherent leakage errors of a single gate. Comparing Ramsey error filtering to interleaved randomized benchmarking to isolate single gates would be a logical next step to understanding leakage metrology.

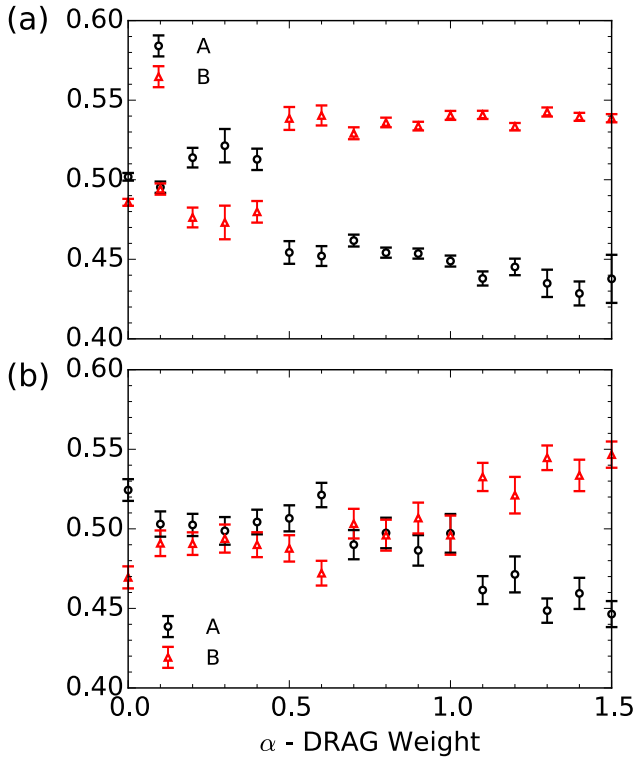

Figure S4. Parameters A and B from fitting randomized benchmarking sequence fidelity decay curves. (a) and (b) correspond to the data in Fig. 3(a) and Fig. 3(b) of the main paper.

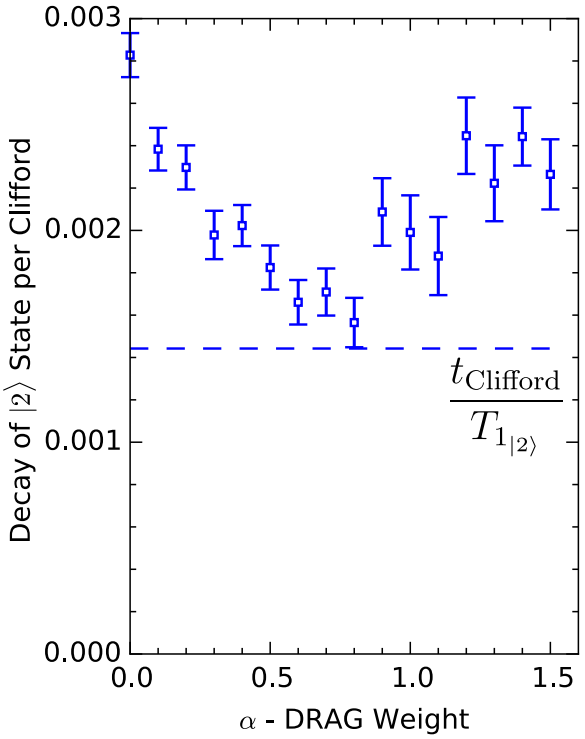

Figure S5. Decay probability of the $|2\rangle$ state per Clifford measured using RB, corresponding to Fig. 3(a) of the main paper. The dashed line indicates the expected incoherent decay from the measured T_1.

ADDITIONAL FIT PARAMETERS FROM RANDOMIZED BENCHMARKING

To estimate gate fidelity, we fit sequence fidelity decay from randomized benchmarking to the function $Ap^m + B$. In Fig. S4, we show the parameters A and B corresponding to the fidelities in Fig. 3 of the main paper. The values of A and B are related to state preparation, measurement, and recovery error (which may vary from experiment to experiment due to interactions of the qubit with background fluctuators), as well as the Markovianity of the error [2]. Non-Markovian error can arise from leakage, or from non-Markovian phase noise which can be studied using other techniques [9].

LEAKAGE STATE DECAY

Equation 2 contains both a leakage rate and a decay rate of the $|2\rangle$ state back into the computational subspace. We show in Fig. S5 the decay rates corresponding to the data in Fig. 3(a) of the main paper. The dashed line represents the decay expected due to T_1 decay of the $|2\rangle$ given an average Clifford time of $t_{\text{Clifford}} = 18.75\,\text{ns}$. The T_1 for the $|2\rangle$ we use here is $13\,\mu$s as measured concurrently with the RB data. We note that this is a different from the $18\,\mu$s quoted in the context of Fig. 4 of the main paper because these measurements were performed many days apart. Over that time scale, the fine features of the spectrum of two-level state (TLS) defects tend to drift. In general, the decay rates are higher than expected from T_1 decay.

RAW DATA FOR SIMULTANEOUSLY OPTIMIZED FIDELITY AND LEAKAGE

In Fig. S6, we show the raw randomized benchmarking data for 10 ns pulses simultaneously optimized for fidelity and leakage, as described in Fig. 3(b) of the main article. Here, α =1.4, and $\delta f = -30\,\text{MHz}$.

THERMALIZATION AT THE $1 \leftrightarrow 2$ TRANSITION FREQUENCY

To verify the heating rate measured in Fig. 4 of the main article, we bias the qubit so that the $0\leftrightarrow1$ transition frequency is equal to the original $1\leftrightarrow2$ frequency, which was about 5.1 GHz. We measure the T_1 of the $|1\rangle$ state at this frequency to be $39\,\mu$s, roughly a factor of two greater than the measured $|2\rangle$ state T_1 of $18\,\mu$s, as expected [10]. Next, we measure the heating rate of the $0\leftrightarrow1$ transition

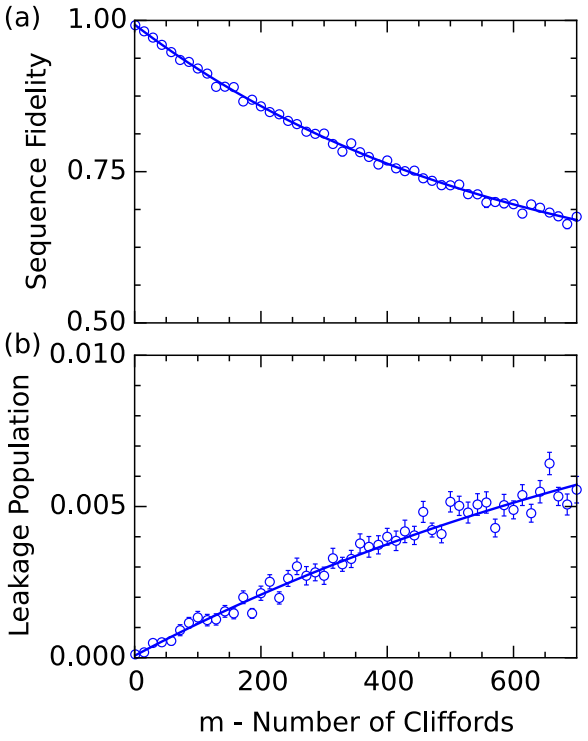

Figure S6. Raw randomized benchmarking data for pulses optimized for both gate fidelity and leakage. (a) Sequence fidelity decay. The error per Clifford is $8.7 \pm 0.4 \times 10^{-4}$. (b) Leakage accumulation. The leakage per Clifford is $1.2 \pm 0.1 \times 10^{-5}$.

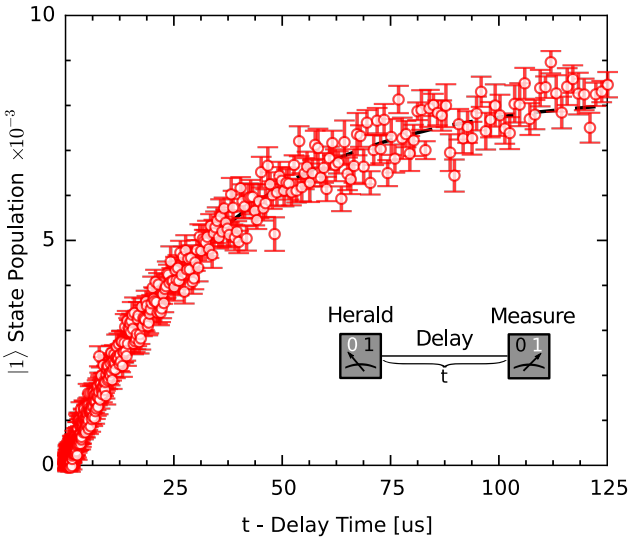

Figure S7. Heating of the qubit, measured by heralding the $|0\rangle$ state, followed by a variable delay and a second measurement. The dashed line is a fit to a rate equation, where the only free parameter is the heating rate.

by performing two measurements separated by a variable delay time, as shown in Fig. S7. The first measurement heralds the $|0\rangle$ state to ensure the qubit is in $|0\rangle$ at $t = 0$, and the second measurement probes the approach of the qubit to the equilibrium population. We fit to a rate equation with two rates, the heating rate and the T_1 decay rate; with the T_1 fixed by the previous measurement, we fit the heating rate to be $1/(4.7\,\text{ms})$. Again, we find the heating time constant to be roughly a factor of 2 larger than that of the $|2\rangle$ state, which we measured to be 2.2 ms.

DRAG WITH SECOND DERIVATIVE CORRECTION

Reference 11 notes that for long pulses and large anharmonicity, leakage can be suppressed using a DRAG-like technique with higher order derivatives. For example, DRAG correction with the second derivative takes the following form:

$$\Omega'(t) = \Omega(t) + \frac{\alpha_2}{\Delta^2}\ddot{\Omega}(t) \tag{6}$$

where α_2 is a weighting parameter. Note that unlike DRAG with first derivatives, the second derivative correction is applied to the in-phase component, which means that it does not have any effect on phase errors. We perform the same experiment as in Fig. 3(a) of the main paper to compare first and second derivative DRAG correction for 10 ns pulses without any detunings. As seen in Fig. S8(a), the second derivative correction does indeed suppress leakage, with a minimum leakage rate of 5×10^{-5} at $\alpha_2 = 1.3$. However, the first derivative correction is still more effective by about a factor of 3 when optimized. Next, we implement both first and second derivative corrections simultaneously.

$$\Omega''(t) = \Omega(t) - i\frac{\alpha_1}{\Delta}\dot{\Omega}(t) + \frac{\alpha_2}{\Delta^2}\ddot{\Omega}(t) \tag{7}$$

Because we have increased the dimension of our parameter space, performing full RB characterization by measuring leakage population versus sequence length for each set of parameters would take a prohibitively long time. Instead, we measure the leakage population for many random sequences but only for a single, large sequence length. We aim to measure the leakage state population near saturation, which is correlated with the leakage rate if the decay rate of the $|2\rangle$ state is mostly independent of the parameters under consideration. In Fig. S8(b), we show the $|2\rangle$ state population after 700 Clifford gates, averaged over 45 different random sequences, while varying both the first and second derivative DRAG weights. We see that there is a substantial parameter space over which leakage can be suppressed. We obtain a minimum leakage population after 700 Cliffords of 3×10^{-3} for $\alpha_1 = 2.8$ and $\alpha_2 = -1.8$, which is a factor of 2 improvement over using only first derivative correction (e.g. as seen in Fig. 4). However, using such a large α would also require a large

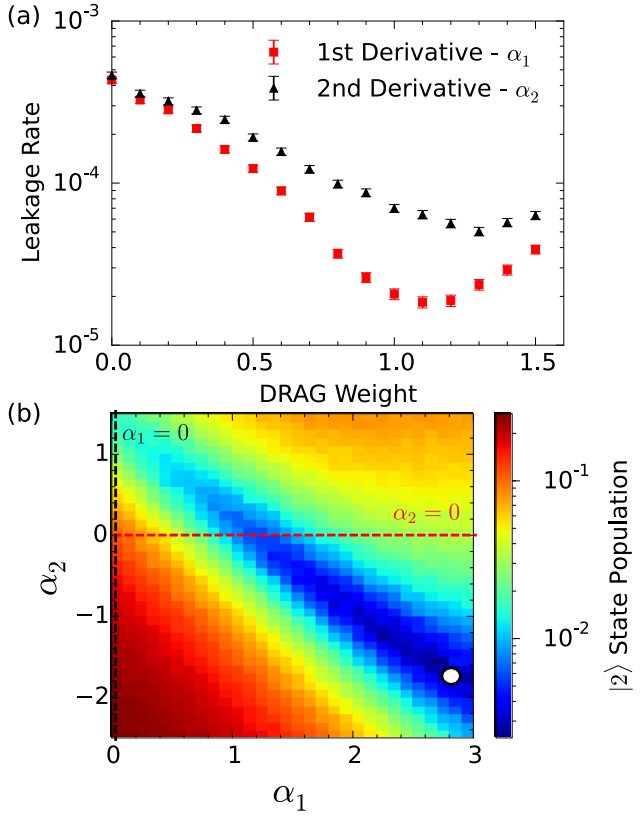

Figure S8. Suppressing leakage using second derivative DRAG. (a) Leakage rate extracted from full Clifford based RB *vs* DRAG weighting (α_1 and α_2), using first derivative correction (red) and second derivative correction (black). Data is for 10 ns pulses. (b) Leakage performance when using both first and second derivative DRAG. The color corresponds to the $|2\rangle$ state population after 700 Cliffords, and is the average of 45 different random sequences. The scale of the color is logarithmic. The dashed, horizontal red line corresponds to first derivative correction only while the vertical black line corresponds to second derivative correction only. The open circle highlights the minimum leakage population, which was 3×10^{-3}.

detuning to compensate for phase errors, which will increase leakage. Thus, while our data suggests that there are still gains to be made in leakage performance, simultaneously optimizing for fidelity and leakage while using second derivative DRAG is non-trivial and an ongoing topic of research.

* jmartinis@google.com

[1] J. M. Epstein, A. W. Cross, E. Magesan, and J. M. Gambetta, Phys. Rev. A **89**, 062321 (2014).

[2] J. J. Wallman, M. Barnhill, and J. Emerson, Phys. Rev. Lett. **115**, 060501 (2015).

[3] T. Chasseur and F. Wilhelm, arXiv preprint arXiv:1505.00580 (2015).

[4] J. Kelly, R. Barends, A. Fowler, A. Megrant, E. Jeffrey, T. White, D. Sank, J. Mutus, B. Campbell, Y. Chen, et al., Nature **519**, 66 (2015).

[5] M. Fang, B. Campbell, Z. Chen, B. Chiaro, A. Dunsworth, J. Kelly, A. Megrant, C. Neill, P. O'Malley, C. Quintana, et al., in *APS Meeting Abstracts* (2015), vol. 1, p. 39002.

[6] E. Jeffrey, D. Sank, J. Mutus, T. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, et al., Phys. Rev. Lett. **112**, 190504 (2014).

[7] E. Lucero, M. Hofheinz, M. Ansmann, R. C. Bialczak, N. Katz, M. Neeley, A. D. O'Connell, H. Wang, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **100**, 247001 (2008).

[8] E. Lucero, J. Kelly, R. C. Bialczak, M. Lenander, M. Mariantoni, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, M. Weides, et al., Phys. Rev. A **82**, 042339 (2010).

[9] P. O'Malley, J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, I.-C. Hoi, et al., Phys. Rev. Applied **3**, 044009 (2015).

[10] M. J. Peterer, S. J. Bader, X. Jin, F. Yan, A. Kamal, T. J. Gudmundsen, P. J. Leek, T. P. Orlando, W. D. Oliver, and S. Gustavsson, Phys. Rev. Lett. **114**, 010501 (2015).

[11] F. Motzoi and F. K. Wilhelm, Physical Review A **88**, 062318 (2013).