Randomized Benchmarking and Process Tomography for Gate Errors in a Solid-State Qubit

J. M. Chow,[1] J. M. Gambetta,[2] L. Tornberg,[3] Jens Koch,[1] Lev S. Bishop,[1] A. A. Houck,[1] B. R. Johnson,[1] L. Frunzio,[1] S. M. Girvin,[1] and R. J. Schoelkopf[1]

[1]*Departments of Physics and Applied Physics, Yale University, New Haven, Connecticut 06520, USA*
[2]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*
[3]*Chalmers University of Technology, SE-41296 Gothenburg, Sweden*

(Received 26 November 2008; published 5 March 2009; corrected 11 March 2009)

We present measurements of single-qubit gate errors for a superconducting qubit. Results from quantum process tomography and randomized benchmarking are compared with gate errors obtained from a double π pulse experiment. Randomized benchmarking reveals a minimum average gate error of $1.1 \pm 0.3\%$ and a simple exponential dependence of fidelity on the number of gates. It shows that the limits on gate fidelity are primarily imposed by qubit decoherence, in agreement with theory.

DOI: 10.1103/PhysRevLett.102.090502 PACS numbers: 03.67.Lx, 42.50.Pq, 85.25.−j

The success of any computational architecture depends on the ability to perform a large number of gates and gate errors meeting a fault-tolerant threshold. While classical computers today perform many operations without the need for error correction, gate error thresholds for quantum error correction are still very stringent, with conservative estimates on the order of 10^{-4} [1,2].

Gate fidelity is the standard measure of agreement between an ideal operation and its experimental realization. Beyond the gate fidelity, identifying the nature of the dominant errors in a specific architecture is particularly important for improving performance. While NMR, linear optics, and trapped ion systems are primarily limited by systematic errors such as spatial inhomogeneities and imperfect calibration [3–5], for solid-state systems decoherence is the limiting factor. The question of how to measure gate errors and distinguish between various error mechanisms has produced different experimental metrics for gate fidelity, such as the double π metric employed in superconducting qubits [6], process tomography as demonstrated in trapped ions, NMR, and superconducting systems [3–5,7], and randomized benchmarking, as performed in trapped ions and NMR [8,9].

Here we present measurements of single-qubit gate fidelities where the three metrics mentioned above are implemented in a circuit QED system [10,11] with a transmon qubit [12]. We compare the results for the different metrics and discuss their respective advantages and disadvantages. We find single-qubit gate errors at the 1%–2% level consistently among all metrics. These low gate errors reflect recent improvements in coherence times [13,14], systematic microwave pulse calibration, and accurate determination of gate errors despite limited measurement fidelity. In circuit QED, measurement fidelity can be as high as 70%, though in this experiment it is $\sim$5%, as readout is not optimized. The magnitude of errors and their dependence on pulse length are consistent with the theoretical limits imposed by qubit relaxation and the presence of higher qubit energy levels, with only small contributions from calibration errors.

We first discuss the double π metric (π-π). Similar to the "bang-bang" technique [15], two π pulses are applied in succession, which ideally should correspond to the identity operation $\mathbb{1}$. The aim of π-π is to determine the deviations from $\mathbb{1}$ by measuring the residual population of the excited state following the pulses. Despite its simplicity, this metric captures the effects of qubit relaxation and the existence of levels beyond a two-level Hilbert space. However, in general, it is merely a rough estimate of the actual gate fidelity as it does not contain information about all possible errors. In particular, errors that affect only eigenstates of σ_x or σ_y and deviations of the rotation angle from π are not well captured by this measure.

A second metric that, in principle, completely reveals the nature of all deviations from the ideal gate operation is quantum process tomography (QPT) [16]. Ideally, QPT makes it possible to associate deviations with specific error sources, such as decoherence effects or nonideal gate pulse calibration. However, in systems with imperfect measurement, it is difficult to assign the results from QPT to a single gate error. Moreover, the number of measurements that are necessary for QPT scales exponentially with the number of qubits.

While QPT provides information about a single gate, randomized benchmarking (RB) [8,17] gives a measure of the accumulated error over a long sequence of gates. This metric hypothesizes that with a sequence of randomly chosen Clifford group generators ($R_u = e^{\pm i\sigma_u \pi/4}$, $u = x, y$) the noise can behave as a depolarizing channel, such that an average gate fidelity can be obtained. In contrast to both π-π and QPT, RB is approximately independent of errors in the state preparation and measurement. Also, while the other metrics measure a single operation and extrapolate the performance of a real quantum computation, RB tests the concatenation of many operations (here up to $\sim$200), just as would be required in a real quantum algorithm.

The gate error metrics are performed in a circuit QED sample consisting of a transmon qubit coupled to a coplanar waveguide resonator [10–12]. The theory and discussion, however, extend generally to all qubit systems including ions and spins. The sample fabrication and measurement techniques are similar to those in Refs. [13,14,18]. Experimentally measured parameters include the qubit-cavity coupling strength given by $g_0/\pi = 94.4$ MHz, the resonator frequency $\omega_r/2\pi = 6.92$ GHz, the photon decay rate of $\kappa/2\pi = 300$ kHz, and the qubit charging energy $E_C/2\pi = 340$ MHz. The qubit is detuned from its flux sweet spot by $\sim$1.5 GHz with a resonant frequency of $\omega_{01}/2\pi = 5.96$ GHz and coherence times of $T_1 = 2.2$ μs and $T_2^* = 1.3$ μs.

In analogy to the NMR language, our single-qubit operations are rotations about the x, y, and z axes of the Bloch sphere [19]. Rotations about any axis in the x-y plane are performed using microwave pulses. The carrier frequency is resonant with the qubit transition frequency, and the pulse amplitudes and phases define the rotation angle and axis orientation, respectively. In all experiments, the pulse shape is Gaussian with standard deviation σ between 1 and 12 ns. The pulses are truncated at 2σ on each side, and a constant buffer time of 8 ns is inserted after each pulse to ensure complete separation of the pulses. Using tune-up sequences similar to those used in NMR [20], each pulse amplitude is calibrated by repeated application of the pulse and matching the measurement outcome to theory. (See supplementary material [21] for details.)

Double π.—After calibration, we perform the π-π experiments with $\sigma = 2$ ns and varying separation time $t_{\rm sep}$ between the two π gates. Subsequently, the excited state probability P_1 is measured, as shown in Fig. 1(a). Because of the decay of the excited state following the first π pulse, P_1 increases as a function of $t_{\rm sep}$. This can be accurately captured in simulations with a simple theoretical model consisting of the dynamics from a master equation for a driven three-level atom subject to relaxation and dephasing, with corresponding time scales T_1 and T_ϕ. The coherent evolution is governed by the Hamiltonian

$$H = \hbar \sum_{j=1,2} [\omega_{0j}\sigma_j^\dagger \sigma_j + \varepsilon_j(t)(\sigma_j^\dagger + \sigma_j)], \qquad (1)$$

where $\sigma_j = |j-1\rangle\langle j|$ is the lowering operator for the multilevel atom with eigenenergies $\hbar\omega_j$. The corresponding transition energies are denoted $\hbar\omega_{ij} = \hbar(\omega_j - \omega_i)$. Drive strength and pulse shapes are determined by

$$\varepsilon_j(t) = \frac{g_j^2}{\omega_r - \omega_{j-1,j}}[X(t)\cos(\omega_d t) + Y(t)\sin(\omega_d t)]. \qquad (2)$$

Here $g_j \sim \sqrt{j} g_0$ is the transmon coupling strength [12], $\omega_d/2\pi$ is the frequency of the drive, and $X(t)$ and $Y(t)$ are the pulse envelopes in the two quadratures.

The inset in Fig. 1(a) shows the experiment with $t_{\rm sep}$ varying between 0 and 30 ns repeated 2.5×10^6 times. We measure $P_1 = 0.014 \pm 0.008$ at $t_{\rm sep} = 0$ ns. Dividing this probability by 2 as in Ref. [6] gives a single gate error of $0.7 \pm 0.4\%$.

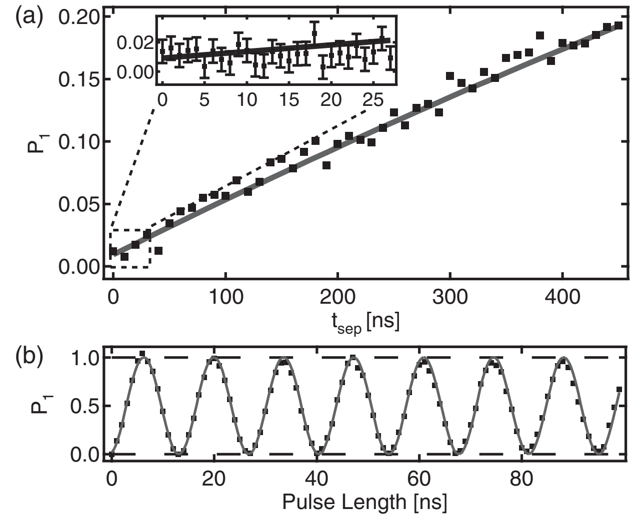

FIG. 1. (a) Excited state qubit population P_1 vs separation time $t_{\rm sep}$ between two successive π pulses ($\sigma = 2$ ns). The data agree well with the simulation (solid line) involving relaxation and decoherence. The inset shows additional data taken for $0 \le t_{\rm sep} \le 30$ ns. The residual population corresponding to the minimal separation is found to be 0.014 ± 0.008 giving a single-qubit gate error of $0.7 \pm 0.4\%$. (b) Rabi oscillations show a visibility of $100.4 \pm 1.0\%$.

Conceptually, the π-π measure is similar to the visibility measure used by Wallraff *et al.* in Ref. [22], corresponding to $(1 - \langle\sigma_z\rangle)/2$ after a single π pulse. Figure 1(b) shows Rabi oscillations made by increasing the length of a pulse resonant with the qubit transition frequency. The visibility is found to be $100.4 \pm 1.0\%$. This also agrees with our simple theoretical model taking into account the T_1, T_2, and third level at our specific operating point.

Quantum process tomography.—The idea behind QPT is to determine the completely positive map $\mathcal{E}$, which represents the process acting on an arbitrary input state ρ. The theory is detailed in Refs. [16,23] and can be summarized as follows. Any process for a d-dimensional system (for 1 qubit $d = 2$) can be written as

$$\mathcal{E}(\rho) = \sum_{m,n=0}^{d^2-1} \chi_{mn} B_m \rho B_n^\dagger, \qquad (3)$$

where $\{B_n\}$ are operators which form a basis in the space of $d \times d$ matrices and χ is the process matrix. To determine χ, we prepare d^2 linearly independent input states $\{\rho_n^{\rm in}\}$. For every input state, the output state $\rho_n^{\rm out} = \mathcal{E}(\rho_n^{\rm in})$ is determined by state tomography. The process matrix is then obtained by inverting Eq. (3). However, in general, this last step does not guarantee a completely positive map. To remedy this, we use a maximum-likelihood estimation based on Ref. [4], which is detailed in the supplementary material [21].

We perform QPT on the three processes $\mathbb{1}$, $R_x(\pi/2)$, and $R_y(\pi/2)$ using the four linearly independent input states

$|0\rangle$, $|1\rangle$, $(|0\rangle + i|1\rangle)/\sqrt{2}$, and $(|0\rangle - |1\rangle)/\sqrt{2}$. The results of this procedure are shown in Fig. 2. Here bar plots of the real and imaginary parts of χ are shown for a pulse with $\sigma = 2$ ns in the Pauli basis $\{B_n\} = \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$. We can compare our data to the ideal process matrices χ_{ideal}. For instance, for the $\mathbb{1}$ process, we expect $\chi_{\mathbb{1}\mathbb{1}} = 1$ and $\chi_{uu'} = 0$ otherwise, which is in good agreement with the measured results. Small deviations from χ_{ideal} arise from preparation and measurement errors, gate over-rotations, decoherence processes, qubit anharmonicity α [24], etc. Calibration errors of the pulses in the x axis are seen as a nonzero $\text{Im}\{\chi_{\mathbb{1}\sigma_x}\}$, and a drive detuning error is exhibited in $\text{Im}\{\chi_{\mathbb{1}\sigma_z}\}$.

From the experimentally obtained process matrix χ and its ideal counterpart χ_{ideal}, we can directly calculate the process fidelity $F_p = \text{Tr}[\chi_{\text{ideal}}\chi]$ and the gate fidelity $F_g = \int d\psi \langle\psi|U^\dagger \mathcal{E}(\psi)U|\psi\rangle$. Here the integral uses the uniform measure $d\psi$ on the state space, normalized such that $\int d\psi = 1$. F_g can be understood as how close $\mathcal{E}$ comes to the implementation of the unitary U when averaged over all possible input states $|\psi\rangle$. From Refs. [25,26], there is a simple relationship between the F_p and F_g, namely, $F_g = (dF_p + 1)/(1 + d)$. For the three processes displayed in Fig. 2, F_p is 0.96, 0.95, and 0.95 ± 0.01, respectively.

Figure 3 shows $1 - F_g$ versus pulse length. The error bars are standard deviations obtained by repeating the maximum-likelihood estimation for input values chosen from a distribution with mean and variance given by measurement. The majority of the experimental gate errors lie above the theoretical errors from a simulation incorporating T_1, T_2, and α. We attribute the higher scatter of these errors to systematic slow qubit frequency drift of ~1–3 MHz during the course of the tomography experiments.

Randomized benchmarking.—The RB protocol, described in Knill *et al.* [8], consists of the following: (i) Initialize the system in the ground state; (ii) apply a sequence of randomly chosen pulses in the pattern $\prod_i C_i P_i$, where C_i are Clifford group generators $e^{\pm i\sigma_u \pi/4}$, with $u = x, y$, and P_i are Pauli rotations, i.e., $\mathbb{1}$, σ_x, σ_y, and σ_z; (iii) apply a final Clifford or Pauli pulse to return to one of the eigenstates of σ_z; (iv) perform repeated measurements of σ_z and compare with theory to obtain the fidelity.

We choose the number of randomizations, sequences, and sequence lengths exactly as in Ref. [8] with the longest sequences consisting of 196 pulses. All 544 final pulse sequences are applied for 250 000 measurements each, taking a total time of about an hour.

The average fidelity is an exponentially decaying function of the number of gates N and approaches 0.5 for large N. Figure 4(a) plots the final state fidelity as a function of the number of computational gates for all randomized sequences with $\sigma = 3$ ns. An average error per gate of 0.011 ± 0.003 is obtained by averaging over all of the randomizations and fitting to the exponential decay. The excellent fit to a single exponential indicates a constant error per gate, consistent with uncorrelated random gate errors due to T_1 and T_ϕ, and no other mechanisms significantly affecting repeated application of single-qubit gates. The reduction of the error by a factor of $\sim 1/3$ from QPT is likely due to the overestimation of errors in QPT where gate errors cannot be isolated from measurement and preparation errors.

The benchmarking protocol is repeated for different pulse widths σ, and the average error per gate is extracted, plotted versus total gate length, and compared to theory in Fig. 4(b). At large gate lengths, experimental results agree well with theory. In this regime, errors are dominated by relaxation and dephasing. At small gate lengths, the gate fidelity is limited by the finite anharmonicity and the resulting occupation of the third level. We obtain error bars from standard deviations in error per gate having generated

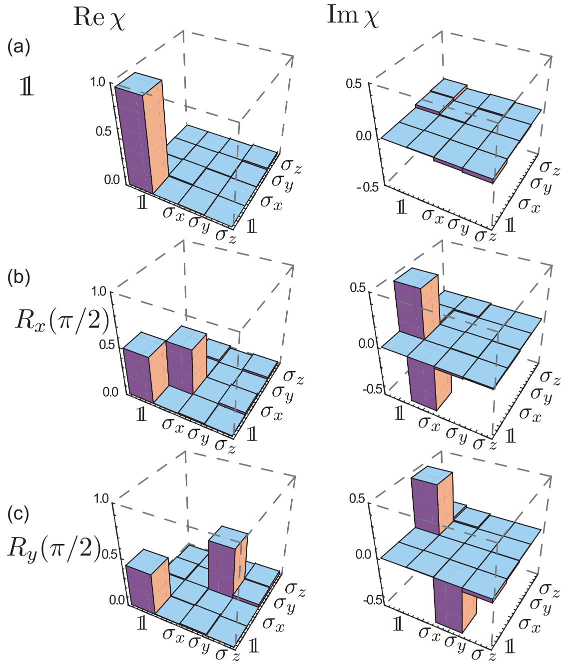

FIG. 2 (color online). Real and imaginary parts of the experimentally obtained process matrix χ for the three processes (a) $\mathbb{1}$, (b) $R_x(\pi/2)$, and (c) $R_y(\pi/2)$ for $\sigma = 2$ ns.

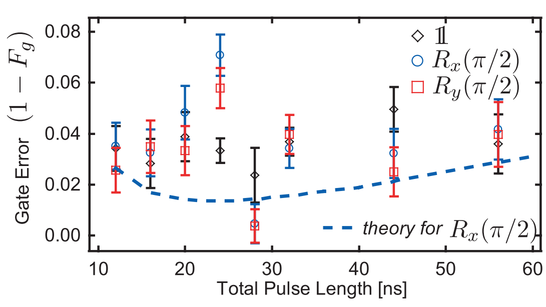

FIG. 3 (color online). Gate error vs total pulse length obtained from quantum process tomography plotted for the processes $\mathbb{1}$, $R_x(\pi/2)$, and $R_y(\pi/2)$. The dashed line is a master-equation simulation for the $R_x(\pi/2)$ process.

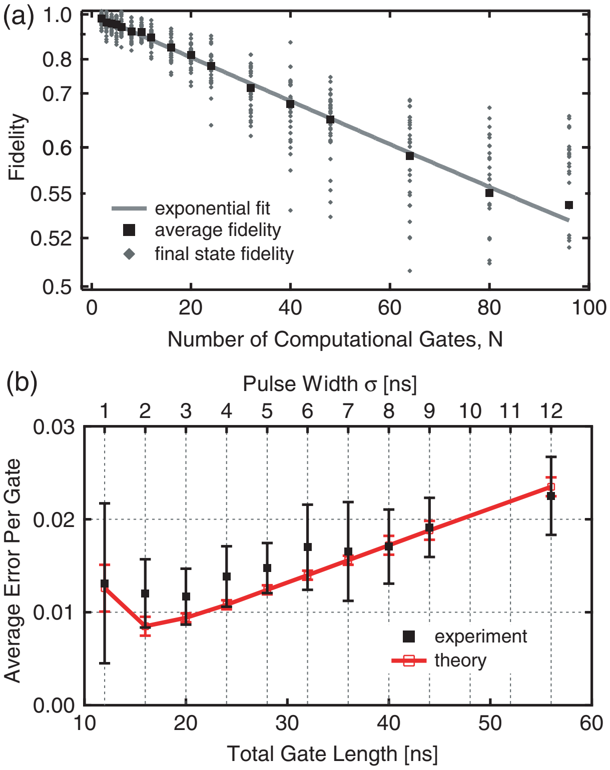

FIG. 4 (color online). (a) Average fidelity vs number of applied computational gates. Computational gates consist of a randomized Pauli with a randomized Clifford generator. For σ of 3 ns we obtain an average gate error of 1.1%. (b) Average error per gate (experimental and theoretical) at different pulse widths. The rise for $\sigma < 2$ ns corresponds to the onset of limitation by the third level of the transmon. The increase in error per gate for $\sigma > 2$ ns is due to the limitation by relaxation.

fidelity values from distributions with means and variance obtained from the experiment and theory. The optimal gate length is found to be 20 ns, as shown in Fig. 4(b), though with optimized pulse shaping we anticipate improving the gate fidelity by another order of magnitude [27].

Conclusions.—We have systematically investigated gate errors in a circuit QED system by measuring gate fidelity using the π-π metric, quantum process tomography, and randomized benchmarking. Table I summarizes our results and displays consistently low gate errors across all metrics. From comparison with theory, we conclude that the observed magnitude of errors fully agrees with the limitations imposed by qubit decoherence and finite anharmonicity. Specifically, in the T_1 limited case and for moderate gate lengths t_g, we find that the gate error scales as $\sim t_g/T_1$. Once coherence times of superconducting qubits and pulse shaping are improved, the aforementioned metrics will be useful tools for characterizing gate fidelities as they approach the fault-tolerant threshold. Randomized benchmarking will be a particularly attractive option for multiqubit systems due to its favorable scaling properties as compared to QPT.

TABLE I. Gate errors for the three metrics used in this work. The measurements show consistently low gate errors of the order of 1%–2%.

Metric	Measured error in %
π-π	0.7 ± 0.4
Process tomography: $\mathbb{1}$	2.4 ± 1.1
Process tomography: $R_x(\pi/2)$	2.6 ± 0.8
Process tomography: $R_y(\pi/2)$	2.2 ± 0.7
Randomized benchmarking	1.1 ± 0.3

We acknowledge E. Knill, R. Laflamme, K. Resch, and C. Ryan for valuable discussions. This work was supported by NSA under ARO Contract No. W911NF-05-1-0365 and by the NSF under Grants No. DMR-0653377 and No. DMR-0603369. J. M. G. was supported by CIFAR, MITACS, and ORDCF. L. T. was supported by the EU through No. IST-015708 EuroSQIP and by the SRC. L. F. was partially supported by CNR-Istituto di Cibernetica.

[1] D. Gottesman, Ph.D. thesis, California Institute of Technology, Pasadena, 1997.
[2] E. Knill, Nature (London) **434**, 39 (2005).
[3] A. M. Childs, I. L. Chuang, and D. W. Leung, Phys. Rev. A **64**, 012314 (2001).
[4] J. L. O'Brien *et al.*, Phys. Rev. Lett. **93**, 080502 (2004).
[5] M. Riebe *et al.*, Phys. Rev. Lett. **97**, 220407 (2006).
[6] E. Lucero *et al.*, Phys. Rev. Lett. **100**, 247001 (2008).
[7] M. Neeley *et al.*, Nature Phys. **4**, 523 (2008).
[8] E. Knill *et al.*, Phys. Rev. A **77**, 012307 (2008).
[9] C. A. Ryan, M. Laforest, and R. Laflamme, New J. Phys. **11**, 013034 (2009).
[10] A. Wallraff *et al.*, Nature (London) **431**, 162 (2004).
[11] A. Blais *et al.*, Phys. Rev. A **69**, 062320 (2004).
[12] J. Koch *et al.*, Phys. Rev. A **76**, 042319 (2007).
[13] J. A. Schreier *et al.*, Phys. Rev. B **77**, 180502(R) (2008).
[14] A. A. Houck *et al.*, Phys. Rev. Lett. **101**, 080502 (2008).
[15] L. Viola and S. Lloyd, Phys. Rev. A **58**, 2733 (1998).
[16] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).
[17] J. Emerson *et al.*, Science **317**, 1893 (2007).
[18] J. Majer *et al.*, Nature (London) **449**, 443 (2007).
[19] C. P. Slichter, *Principles of Magnetic Resonance* (Springer, New York, 1996).
[20] R. W. Vaughan *et al.*, Rev. Sci. Instrum. **43**, 1356 (1972).
[21] See EPAPS Document No. E-PRLTAO-102-060911 for details about pulse shaping, pulse calibration, and maximum-likelihood process estimation. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[22] A. Wallraff *et al.*, Phys. Rev. Lett. **95**, 060501 (2005).
[23] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).
[24] Anharmonicity specifies the difference $\alpha = \omega_{12} - \omega_{01}$ between the fundamental qubit frequency and its next higher transition. For the transmon, $\alpha \approx -E_c$ [12].
[25] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **60**, 1888 (1999).
[26] M. A. Nielsen, Phys. Lett. A **303**, 249 (2002).
[27] F. Motzoi *et al.*, arXiv:0901.0534.