Benchmarking Quantum Computers: The Five-Qubit Error Correcting Code

E. Knill,* R. Laflamme,† R. Martinez, and C. Negrevergne

Los Alamos National Laboratory, MS B265, Los Alamos, New Mexico 87545
(Received 29 January 2001)

The smallest quantum code that can correct all one-qubit errors is based on five qubits. We experimentally implemented the encoding, decoding, and error-correction quantum networks using nuclear magnetic resonance on a five spin subsystem of labeled crotonic acid. The ability to correct each error was verified by tomography of the process. The use of error correction for benchmarking quantum networks is discussed, and we infer that the fidelity achieved in our experiment is sufficient for preserving entanglement.

Robust quantum computation requires that information be encoded to enable removal of errors unavoidably introduced by noise [1]. Every currently envisaged scalable quantum computer has encoding, decoding, and error-correction procedures among its most frequently used subroutines. It is therefore critical to verify the ability to implement these procedures with sufficient fidelity. The experimental fidelities achieved serve as useful benchmarks to compare different device technologies and to determine to what extent scalability can be claimed.

Liquid state nuclear magnetic resonance (NMR) is currently the only technology that can be used to investigate the dynamics of more than four qubits [2,3]. Although it is not practical to apply it to more than about ten qubits [4], it can be used to investigate the behavior of quantum networks on representative physical systems to learn more about and begin to solve the problems that will be encountered, more scalable device technologies. In this Letter, we describe an experimental implementation using NMR of a procedure for benchmarking the one-error-correcting five-qubit code. This is the shortest code that can protect against depolarizing one-qubit errors [5,6]. The effect of depolarizing errors is to randomly apply one of the Pauli matrices or the identity to the state of the system. The experiment is one of the most complex quantum computations implemented so far and the first to examine the behavior of a quantum error-correcting code (QECC) that protects against all one-qubit errors. We discuss the principles underlying error-correction benchmarks and offer a sequence of specific challenges to be met by this and future experiments. Our experiment shows an average polarization preservation of 67% corresponding to an entanglement fidelity of 0.75. This easily achieves the goal of demonstrating the preservation of entanglement in principle.

The five-qubit code.—A QECC for encoding a qubit is a two-dimensional subspace of the state space of a quantum system. In the case of interest, the quantum system consists of five qubits. The code, C_5, can be specified as one of the 16 two-dimensional joint eigenspaces of the four commuting operators

$$\sigma_z^{(2)}\sigma_y^{(3)}\sigma_y^{(4)}\sigma_x^{(5)}, \qquad \sigma_z^{(1)}\sigma_y^{(2)}\sigma_y^{(3)}\sigma_x^{(4)}, \qquad \sigma_y^{(2)}\sigma_z^{(3)}\sigma_z^{(4)}\sigma_z^{(5)}, \qquad \sigma_x^{(1)}\sigma_z^{(2)}\sigma_x^{(3)}\sigma_z^{(4)}. \tag{1}$$

Here, $\sigma_x^{(k)}, \sigma_y^{(k)}, \sigma_z^{(k)}$ are the Pauli spin operators acting on qubit k. This is an instance of a stabilizer code [7,8].

QECCs can be used to protect quantum states against noise. In our experiment, the procedure begins with qubit 2 containing the state to be protected and *syndrome* qubits 1, 3, 4, 5 in the initial state $|1\rangle$. A unitary *encoding* transformation maps the two-dimensional input state space to the code C_5. The five qubits would then be stored in a noisy memory. In our experiment, we explicitly applied one of the correctable errors. The information is retrieved by *decoding* the state. The properties of the code guarantee that the syndrome qubits' state determines which error was applied. To complete the process, conditional on the syndrome qubits' state, it is necessary to *correct* the state of qubit 2 by applying a Pauli operator. Quantum networks for encoding, decoding, and error correction are shown in Fig. 1.

Quantum codes as benchmarks.—The purpose of a benchmark is to compare different devices on the same task. Since quantum codes will be used to maintain information in future quantum computing devices, they are excellent candidates for benchmarking the reliability of proposed quantum processors. A quantum code benchmark consists of a sequence of procedures that implement encoding, evolution, decoding, and error-correction networks. In the simplest cases, the networks are applied to one qubit's state and use several ancilla qubits. An experimental implementation measures the reliability (see below) with which the qubit's state is processed. It is necessary to include a means for verifying that a code with the desired properties was indeed implemented.

To allow for unbiased comparison of devices, the reliability measurement and the verification steps of the benchmark need to be standardized. There are many different
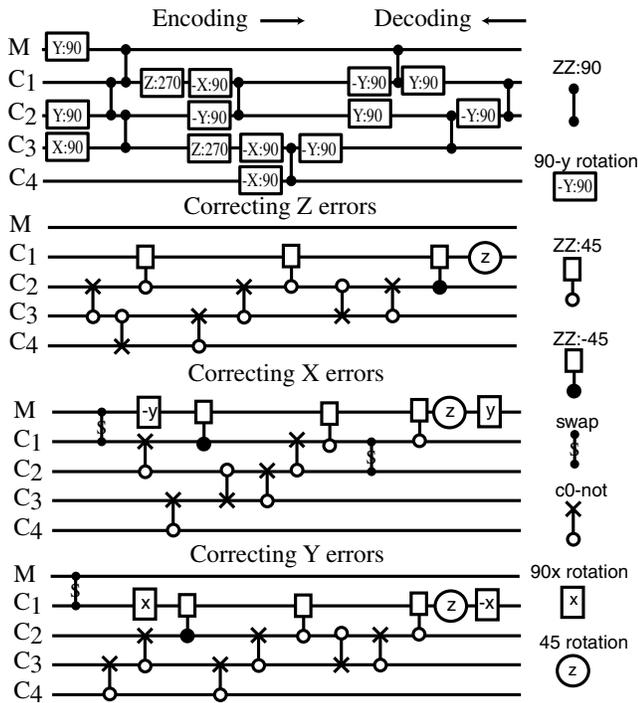
FIG. 1. Networks for the five-qubit code. Top: the encoding network using 90° rotations. Except for refocusings required to eliminate unwanted couplings, these are directly implementable with pulses. The decoding network is the inverse of the encoding network. Angles are all give in units of degrees and the gates denoted by ZZ:ϕ implement $e^{-i\sigma_z\sigma_z\phi/2}$. Bottom: the three steps of the error-correction procedure, which implements a rotation on $C1$ conditional to the syndrome state. The controlled operations can be translated to sequences of 90° rotations using standard quantum network methods [18]. The full encoding, decoding, and the three error-correction networks were implemented in their entirety. An evolution consisting of applying no error or a Pauli operator on one of the qubits was inserted between encoding and decoding.

ways of quantifying reliability. The best known such measure is *fidelity*. If the input state is $|\psi\rangle$ and the output density matrix is ρ, then the fidelity of the output is given by $F(|\psi\rangle, \rho) = \langle\psi|\rho|\psi\rangle$, the probability of measuring $|\psi\rangle$ in a measurement that distinguishes this state from the orthogonal states. In our case we are interested in an arbitrary state of one qubit. One quantity of interest would be the worst case fidelity for pure states. However, an easier to use quantity is the *entanglement fidelity F_e* [9], the fidelity with which a Bell state on the qubit and a perfect reference qubit is preserved. Entanglement fidelity does not depend on the choice of Bell state and has the property that $F_e = 1$, if and only if the process perfectly preserves every input state. F_e can be determined from the fidelities of pure states. Define $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$ (the eigenstates of σ_x and σ_y, respectively). Let F_s be the fidelity of the process for input $|s\rangle$. Then

$$F_e = (F_0 + F_1 + F_+ + F_- + F_{+i} + F_{-i})/4 - 1/2.$$
(2)

The pure state fidelities can be inferred by measurements involving any set of states whose density operators form a complete set. A formula in terms of polarization preservation useful for NMR is given later.

The standard verification procedure for a quantum code benchmark needs to satisfy that good fidelity is a demonstration of having implemented a code with the desired properties. For codes defined as the common eigenspace of a commuting set of products of Pauli operators (stabilizer codes), it is, in principle, enough to verify that applying a product P of Pauli operators during the evolution has the expected effect on the output, namely, that it differs from the input by the application of a Pauli operator $\sigma(P)$ determined by the code and the applied product. A single fidelity measure may be obtained by averaging the entanglement fidelities measured for each P after applying $\sigma(P)$ (to cancel the effect of P) to the output states. To make this procedure experimentally feasible, one may randomize the choice of P and use statistical methods to estimate the desired average.

For benchmarks involving a QECC implemented with the complete error-correction procedure, the emphasis is on having corrected the set of errors $\mathcal{E}$ for which it was designed, and verification involves applying the errors in $\mathcal{E}$ during the evolution and observing the extent to which they are indeed corrected. Ideally, the errors occur naturally in the course of evolution, and one would like to see that information is preserved better by encoding it. In order to investigate the code in a controlled way, it is easier to apply different errors explicitly and observe the effect on the reliability of the process. The experiment described here involves measuring the entanglement fidelity for each of the one-qubit Pauli operators applied during the evolution.

When implementing a benchmark, it is useful to have some goals in mind. For benchmarks involving codes designed to correct independent errors on qubits, we offer a sequence of four such goals, depending on how well the implementation succeeds at protecting against various error models. Most involve comparing the entanglement fidelity for two situations involving a specific error process $\mathcal{E}_i$. In the first, the information is stored in any one of the qubits, giving an optimum $F_{e,1}(\mathcal{E}_i)$. In the second, the information is stored by using the implemented code, giving an experimentally determined $F_{e,C}(\mathcal{E}_i)$. A sufficiently good implementation satisfies an improvement in the fidelity: $F_{e,C}(\mathcal{E}_i) > F_{e,1}(\mathcal{E}_i)$. Specific goals are given next. Numerical values are given under the assumption that for each correctable error, the experimental implementation induces depolarizing errors. The goals are as follows: (i) An improvement in fidelity for the noise model $\mathcal{E}_1$ that independently for each qubit depolarizes it with probability p: Assuming that the error behavior of the implementation is depolarizing with equal fidelity F_{e,C_5} for each possible Pauli-product error, then this goal requires $F_{e,C_5} > 0.97$, giving an improvement when $p = 0.08713$ (see Fig. 2). (ii) Improvement in fidelity for the noise model $\mathcal{E}_2$ that first randomly chooses a qubit and then depolarizes it: For our

code this requires $F_{e,C_5}(\mathcal{E}_2) > 0.85$. (iii) Preservation of some entanglement for $\mathcal{E}_2$: This requires $F_{e,C}(\mathcal{E}_2) > 0.5$. This is due to the fact that if one-half of a pair of qubits in a Bell state is transmitted through a depolarizing channel with entanglement fidelity ≤ 0.5, no entanglement remains in the pair [5]. (iv) Improvement in fidelity for the *demonic* error process $\mathcal{E}_4$ that, knowing the method for storing the qubit, chooses the worst possible one-qubit depolarizing error and applies it: In this case we need $F_{e,C}(\mathcal{E}_4) > 0.25$. The ultimate goal is to demonstrate that the code can be implemented sufficiently well to permit long-term preservation of information.

Experimental implementation.—A standard 500 MHz NMR spectrometer (DRX-500, Bruker Instruments) with a triple resonance probe was used with a sample of ^{13}C labeled transcrotonic acid synthesized as described in [10], but with deuterated acetone as a solvent. The chemical shifts (CS) and primary couplings (J) for the nuclei along the backbone of crotonic acid are given by

$$H_3 - {}^{13}C - {}^{13}C = {}^{13}C - {}^{13}C \text{ OOH}$$

| CS (khz): | 0.9 | | 2.2 | | 18.4 | | 15.4 | | 21.2 |
| J (hz): | | 127 | | 42 | | 70 | | 72 | |

The five spin-$1/2$ systems used for the code are the methyl group (M) and the four ^{13}C nuclei, labeled $C1$, $C2$, $C3$, and $C4$, starting from M. The methods of [10] were used to prepare the methyl group as an effective spin-$1/2$ system and to initialize the labeled pseudopure state $1\sigma_z 11111$ on the active nuclei with the gradient echo sequence. Here, $1 = |1\rangle\langle 1|$ and the last two nuclei are the protons $H1$, $H2$ adjacent to $C2$, $C3$. The pseudopure state was subjected to a "crusher" gradient. To absolutely guarantee the pseudopure state and eliminate the possibility of contributions from zero coherences, more randomization is required (see [10]) but we did not implement this. $H1$ and $H2$ were not used and were only affected by some hard pulse refocusings on the protons. The states of $H1$ and $H2$ (up or down)
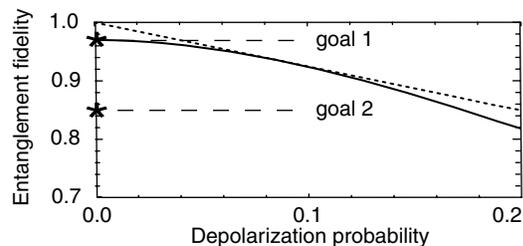

FIG. 2. Entanglement fidelities for independent depolarization. The fidelities for an unencoded qubit (dotted line) and a qubit encoded with C_5 are shown as a function of the depolarization probability. The implementation of the code is assumed to be imperfect, with an additional error that is syndrome independent and has entanglement fidelity $F_e = 0.97$. The lower curve was obtained in two independent ways: by direct simulation and by a combinatorial analysis of the error operators for the stabilizer code. The two curves are tangent at $p = 0.08713$, so if $F_e > 0.97$ (which raises the lower curve), encoding preserves information better for an interval about $p = 0.08713$. The first two fidelity benchmarking goals are indicated.

induce an effective frequency shift on the other nuclei, depending on the coupling constants, and was compensated for in phase calculations. To greatly reduce the effect of radio frequency (rf) inhomogeneity, we used the nutation-based selection scheme of [10], applied to both the proton and the carbon transmitters. The quantum networks of Fig. 1 were directly translated to pulse sequences, again using the methods described in [10]. The only significant use of manual intervention was to place the refocusing pulses. The evolution period between encoding and decoding was carefully isolated from both the preceding and the following pulses: It implemented the identity unitary operator (corresponding to having no error), or one of the one qubit 180° rotations (corresponding to a one-qubit Pauli error) by refocusing the molecule's internal Hamiltonian and applying an extra inversion or by shifting the phase by 180°. The qubit's output state appeared on $C1$ at the end of the experiment. The peak group associated with $C1$ was observed in each experiment. Spectra were analyzed by comparing the spectrum of the pseudopure state $1\sigma_x 11111$ to the output, using the knowledge of the peak positions and shapes to compute relative intensities and phases. No phase adjustment was made after phasing the pseudopure state spectrum. This was possible since the relative phase is precomputed by the pulse compiler and integrated into the acquisition. Excluding the state preparation steps, the experimentally implemented pulse sequence involved 368 rf pulses applied over a period of 0.38 s.

We performed one experiment for each of the 16 possible evolutions with one-qubit or no Pauli error, each of the three initial states σ_x, σ_y, or σ_z on $C1$, and each of three observations (no pulse, 90° X pulse, or 90° Y pulse) on $C1$. This resulted in a total of 144 experiments, each of which was repeated sufficiently often to obtain better than 8% error in the inferred state of $C1$. For each evolution E and input σ_u, we determined the amount of signal $P(E, u)$ in the correct direction in the output relative to the input signal. This requires "tracing out" the other spins, which was done by adding the intensities of each peak in the $C1$ spectrum that is associated with the **11** state on $H1$ and $H2$. (The spectrum of $C1$ resolves all its couplings.) Thus, except for noise, $-1 \leq P(E, u) \leq 1$. Under the assumption that input $1(|0\rangle\langle 0| + |1\rangle\langle 1|)11111$ results in no observable signal, the entanglement fidelity for a given evolution E is given by $F_e(E) = [P(E, x) + P(E, y) + P(E, z) + 1]/4$ [see Eq. (2)]. We did not verify the assumption in this experiment, but note that it has been verified in related experiments [11], and could be enforced by modifying the process with random pairs of canceling 180° pulses before and after the implemented pulse sequence. This would also enforce the depolarizing noise model for the implementation while preserving the observed polarization.

Results.—Typical spectra compared to the spectrum of the input pseudopure state are shown in Fig. 3. The relative polarization after the error-correction procedure in the correct output state varies between 48% and 87%. The distribution is shown in Fig. 3. The inferred entanglement
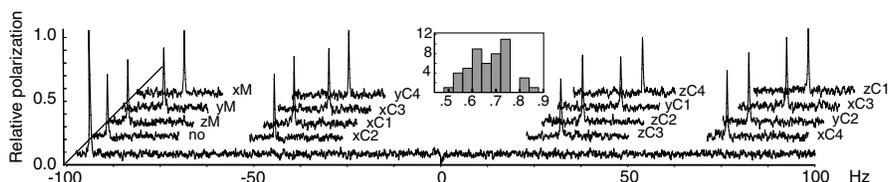
FIG. 3. Experimental input and output spectra. The reference spectrum for the pseudopure input is at the bottom, and the partial spectra for each one-qubit error are shown above it using the same scale. The labels indicate which error was applied to which nucleus. One peak is observed for each possible error input. Its position corresponds to the error syndrome. Its phase reflects the error correction procedure and corresponds to the input state up to a small error. Signal in the wrong locations or phase was consistently small and comparable to the estimated noise. The bar graph inserted in the middle shows the distribution of relative polarizations. There are a total of 48 polarization measurements. Each bar represents the number of measurements with relative polarization in the bar's interval. The distribution strongly suggests some syndrome-dependent effects on the implementation error.

fidelity for goals (ii) and (iii) is $F_{e,C}(\mathcal{E}_2) = 0.75$, with an estimated error of less than 0.02. The fidelity can be calculated using $F_{e,C}(\mathcal{E}_2) = 1/4 + 3/4[(1/4)w_0 + (3/4)w_1]$, where w_0 is the average polarization measured for the three input states with no error applied, and w_1 is the average polarization for the inputs and errors. The entanglement fidelity achieved is clearly sufficient for goal (iii).

The reduction in polarization is due to relaxation, incompletely refocused couplings (part of the pulse compiler's optimization trade-offs), pulse errors due to nonideal implementation of 180° and 90° pulses, and rf inhomogeneity (less than 2% after our selection procedure). At least half of the error is explained by relaxation, suggesting that this is what limits the fidelities that can be attained using liquid state NMR. The estimated phase relaxation times in our molecule are above 1 s. ($T_2 \simeq 1.4$ for $C1$, the other carbons are similar, while M is shorter based on peak shape analyses.) This is high when compared to those of nuclei in other molecules used in NMR quantum information processing experiments to date.

Discussion.—Benchmarking quantum devices for quantum information processing is crucial both for comparing different device technologies and for determining how much control over a device is achievable and how to best achieve it. Given the need for and difficulty of achieving robust quantum control, we advocate the use of quantum coding benchmarks to determine the fidelity of the implementation of standard, verifiable processes. Unlike the experimentally implemented versions (up to five qubits) of the popular quantum searching and order-finding algorithms [12–16], quantum codes offer a rich source of verifiable quantum procedures required in currently envisioned quantum computer architectures. Liquid state NMR has been used to implement several interesting, small quantum codes [10,11,17]. In this Letter, we have given specific goals for benchmarks involving error correction and implemented the shortest one-error-correcting quantum code. The fidelity achieved demonstrates preservation of entanglement in the presence of one-qubit errors. The task of demonstrating fidelity improvements in the presence of depolarizing errors remains to be accomplished. A device that achieves this challenging goal will be well on the way toward realizing robustly scalable quantum computation.

*Electronic address: knill@lanl.gov
†Electronic address: laflamme@lanl.gov

[1] For a review, see J. Preskill, Proc. R. Soc. London A **454**, 385 (1998).
[2] D. G. Cory, A. F. Fahmy, and T. F. Havel, Proc. Natl. Acad. Sci. U.S.A. **94**, 1634 (1997).
[3] N. A. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).
[4] W. S. Warren, Science **277**, 1688 (1997).
[5] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[6] R. Laflamme, C. Miquel, J.-P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).
[7] D. Gottesman, Phys. Rev. A **54**, 1862 (1996).
[8] A. Calderbank, E. Rains, P. Shor, and N. Sloane, Phys. Rev. Lett. **78**, 405 (1997).
[9] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[10] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng, Nature (London) **404**, 368 (2000).
[11] D. G. Cory, M. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo, Phys. Rev. Lett. **81**, 2152 (1998).
[12] J. A. Jones, M. Mosca, and R. H. Hansen, Nature (London) **392**, 344 (1998).
[13] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd, Nature (London) **393**, 143 (1998).
[14] I. L. Chuang, N. Gershenfeld, and M. Kubinec, Phys. Rev. Lett. **80**, 3408 (1998).
[15] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, R. Cleve, and I. L. Chuang, Phys. Rev. Lett. **85**, 5452 (2000).
[16] R. Marx, A. F. Fahmy, J. M. Myers, W. Bermel, and S. J. Glaser, Phys. Rev. A **62**, 012310/1 (2000).
[17] D. Leung, L. Vandersypen, X. L. Zhou, M. Sherwood, C. Yannoni, M. Kubinec, and I. Chuang, Phys. Rev. A **60**, 1924 (1999).
[18] S. S. Somaroo, D. G. Cory, and T. F. Havel, Phys. Lett. A **240**, 1 (1998).