Practical Experimental Certification of Computational Quantum Gates Using a Twirling Procedure

Osama Moussa,[1,*] Marcus P. da Silva,[2,3,†] Colm A. Ryan,[1,3] and Raymond Laflamme[1,4,‡]

[1]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

[2]*Département de Physique, Université de Sherbrooke, Sherbrooke, Quebec J1K 2R1, Canada*

[3]*Raytheon BBN Technologies, Disruptive Information Processing Technologies Group, Cambridge, Massachusetts 02138, USA*

[4]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2J 2W9, Canada*

(Received 16 November 2011; published 17 August 2012)

Because of the technical difficulty of building large quantum computers, it is important to be able to estimate how faithful a given implementation is to an ideal quantum computer. The common approach of completely characterizing the computation process via quantum process tomography requires an exponential amount of resources, and thus is not practical even for relatively small devices. We solve this problem by demonstrating that twirling experiments previously used to characterize the average fidelity of quantum memories efficiently can be easily adapted to estimate the average fidelity of the experimental implementation of important quantum computation processes, such as unitaries in the Clifford group, in a practical and efficient manner with applicability in current quantum devices. Using this procedure, we demonstrate state-of-the-art coherent control of an ensemble of magnetic moments of nuclear spins in a single crystal solid by implementing the encoding operation for a 3-qubit code with only a 1% degradation in average fidelity discounting preparation and measurement errors. We also highlight one of the advances that was instrumental in achieving such high fidelity control.

DOI: 10.1103/PhysRevLett.109.070504 PACS numbers: 03.67.Lx, 03.65.Wj, 03.67.Ac

Introduction.—Because of the technical challenges of building quantum computers, only small building blocks of such devices have been demonstrated so far in a number of different physical systems. In order to quantify how closely these demonstrations come to the desired ideal operations, the experiments are fully characterized via quantum process tomography (QPT) [1,2], and, often, the average fidelity [3,4] between the experiment and the ideal operator is calculated from the description of the estimated process. The main drawback of this approach is that QPT fundamentally requires an exponential number of experiments. Moreover, the classical postprocessing of the data is nontrivial, as the raw experimental data do not lead to a physical description, and approaches such as maximum likelihood or Bayesian estimation on an exponentially large parameter space are needed to find the most appropriate physical description [5]. Therefore, approaches based on QPT to estimate the average fidelity are not practical, and cannot be reasonably expected to be used even in systems that are only moderately larger than the current experimental state of the art. Here we solve this problem by showing that, for an important class of quantum operations, the average fidelity can be estimated efficiently, requiring a number of experiments which are independent of the system size. This new proposal is also practical, and enables the demonstration of processes which would not have been possible due to the complexity of QPT.

Twirling.—It has been recently shown that if one wishes to compare an experimental implementation of a quantum process to the identity process (e.g., in the case of ideal quantum memories), then it is possible to estimate the average fidelity via a technique known as twirling [6–10], with a number of experiments which depend only on the desired accuracy of the estimate, not on the system size—moreover, these experiments are simple to implement, requiring only local operations and measurements [8]. The twirling procedure consists of applying a random unitary before the process to be characterized, followed by the inverse of this randomly chosen unitary. When these unitaries obey certain symmetry properties, the resulting invariant information about the noise under this symmetry can be extracted by repeating the experiment with different random choices. For example, if the twirling gates are random permutations followed by tensor products of single-qubit Cliffords, then information about the weights of the noise terms can be determined [8]. The schematic for such an experiment is depicted in Fig. 1(a), where the process we would like to compare to the identity process, $\mathcal{E}$, is conjugated with $\mathcal{C}_i$, an appropriately chosen randomizing local operation, and $\hat{M}_j$ is the measurement of the parity of a subset of qubits in the computational basis.

For more general processes, in order to compare a given process to a desired unitary evolution one could in principle apply the physical process under consideration, and then apply the inverse of the unitary evolution we would like to compare it against, finally measuring the overlap between the initial state and the resulting state for a set of initial states—this is, in essence, the definition of the average fidelity. The obstacle to implementing such a protocol is that often one is attempting to demonstrate or certify the implementation of a unitary, and a noiseless implementation of its inverse cannot be assumed to be available. One way [6,11,12] to address this problem is to estimate the average fidelity over a set of quantum processes that form a group by considering random sequences of such processes chosen to result in the identity process—examples of such sets include the group of all unitary processes as well as the Clifford group [13]. Such motion-reversal benchmarking schemes suffer from two shortcomings: they apply only to noise satisfying certain strength conditions [12], and they only provide information about the average over a set of processes instead of specific information about a particular process. While this information is useful, what is experimentally most useful is to diagnose coherent control implementation errors, which, in our experience, are highly process dependent. Therefore, it is critical for the experimentalist to be able to characterize a particular process.

Certification procedure.—The result we report here, which sidesteps many of the shortcomings listed above, is that the average fidelity between any physical process on multiple qubits and any particular element of the Clifford group can be estimated efficiently by a simple modification to the twirling protocol, leading to the same favorable scaling as experiments which compare a physical process to the identity. If we define U to be the desired element of the Clifford group, with the superoperator $\mathcal{U}(\rho) = U\rho U^\dagger$, then the noisy implementation represented by the superoperator $\tilde{\mathcal{U}}$ can be thought of as some noisy process $\mathcal{E}$ followed by the application of $\mathcal{U}$, i.e., $\tilde{\mathcal{U}} = \mathcal{U} \circ \mathcal{E}$. Unitaries in the Clifford group include operations needed to encode and decode quantum information to protect it from noise [14]—in current approaches to fault tolerance these operations comprise the vast majority of (if not all) operations. Clifford group operations can also be used to achieve universal fault-tolerant quantum computation with the aid of specially prepared resource states [13,15], so these operations are of great importance and utility for quantum computation.

In order to see why the average fidelity can be estimated efficiently for these operations, consider Fig. 1(b), which modifies the original twirling protocol [8] by inserting the identity process—in this case written as $\mathcal{U}^\dagger \circ \mathcal{U}$. One can in principle combine all processes after the first application of $\mathcal{U}$ in Fig. 1(b) into a new measurement. For a general unitary process, this new measurement will be as hard to implement as performing the process $\mathcal{U}$ itself. However, if $\mathcal{U}$ is an element of the Clifford group [16], this results in the measurement of the parity of a different set of qubits in a different local basis (or equivalently, the measurement of a different Pauli operator [13]) which can be precomputed efficiently given $\mathcal{U}$, the local randomizing Clifford operation $\mathcal{C}_i$, and the original measurement $\hat{M}_j$, as depicted in Fig. 1(c), where

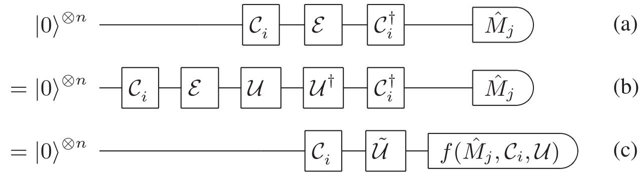


FIG. 1. The relationship between (a) comparing a physical process $\mathcal{E}$ to the identity, and (c) comparing a physical process superoperator $\tilde{\mathcal{U}} = \mathcal{U} \circ \mathcal{E}$ to a unitary superoperator $\mathcal{U}$ can be seen by (b) the appropriate insertion of the identity $\mathcal{U}^\dagger \circ \mathcal{U}$. When $\mathcal{U}$ and all $\mathcal{C}_i$ are elements of the Clifford group, and $\hat{M}_j$ is a Pauli operator, $f(\hat{M}_j, \mathcal{C}_i, \mathcal{U})$ is also a Pauli operator.

$f(\hat{M}_j, \mathcal{C}_i, \mathcal{U}) = \mathcal{U}(C_i \hat{M}_j C_i^\dagger) = U C_i \hat{M}_j C_i^\dagger U^\dagger$. In essence, the protocol in Fig. 1(c) is the experiment, but the data are analyzed according to Fig. 1(a) as described in [8,9], which separates the noise $\mathcal{E}$ from the unitary $\mathcal{U}$. Because C_i and U are elements of the Clifford group, and $\hat{M}_j$ is a Pauli observable, $f(\hat{M}_j, \mathcal{C}_i, \mathcal{U})$ can be computed efficiently using the Gottesman—Knill theorem [17].

As the parity measurement is equivalent to local measurement followed by simple data postprocessing, and the initial states required are product states locally equivalent to the all-zeros state, it is important that precise local operation be available. In other words, the problem of implementing the inverse of a multibody Clifford unitary $\mathcal{U}$ can be translated into the problem of implementing classical data processing and local (single body) quantum operations reliably. These operations are often readily available at high fidelities, as randomizing benchmarking results have demonstrated [11,18,19]. Thus, the average fidelity of any implementation of a Clifford group operation can be estimated using a number of experiments that depend only on the desired accuracy, as is the case for twirling experiments with quantum memories [8].

Because of this connection to twirling protocols, our proposal also enables the estimation of other parameters beyond the gate fidelity, such as the probability of errors of a given weight. Recent proposals for Monte Carlo estimation of state and gate fidelity have the same scaling as the protocol we describe here (in the case of Clifford gates) [20,21]. However, the probabilities of errors of a given weight are not natural parameters to be considered in the Monte Carlo sampling proposals, demonstrating the advantage of considering twirling protocols in this context. The simplicity of the experiments also shows that our proposal is of practical significance in the benchmarking of these important operations. Moreover, because the estimation of the average fidelity in the twirling protocol corresponds to the estimation of the probability of no errors having occurred (a single parameter that is accessible with an accuracy that does not depend on the number of qubits [8]), Bayesian estimation of such a probability is straightforward, as is the calculation of uncertainties associated with these estimates.

Experiment.—A common task for an experimentalist is to optimize and tweak the performance of a particular gate on the system. The experimenter has many potential knobs to adjust and he or she needs a reliable robust method for certifying whether any changes actually improved the performance. A trivial example is calibrating the power of a pulse but here we demonstrate how we can easily quantify the improvement from more subtle and sophisticated control improvements.

Methods inspired by optimal control theory have been successful in aiding pulse design for small systems. However, for these pulses to achieve the designed fidelity, it is important that the implemented control fields at the sample match the designed ones. That is to say, any systematic deviations from the designed pulses, in pulse generation or amplification, need to be accounted for or rectified. To this end, a feedback system can be employed to correct for these systematic imperfections [22]. We use an antenna to measure the fields in the vicinity of the sample, then this data is fed back for comparison with the target pulse, and a new pulse form that attempts to compensate for the imperfections is computed and sent back to the signal generation unit. This loop is repeated a number of times to reach a satisfactory pulse form [23,24]. Figure 2 shows a typical example of the measured pulse forms of the initial and corrected attempts to match a target pulse shape. The development of this feedback pulse rectification protocol has led to a great improvement in the fidelity of coherent control of nuclear spins in the solid state—the certification scheme is used herein to demonstrate and quantify the typical improvement in fidelity resulting from using the feedback system.

The computational register used in this demonstration is an ensemble of molecular nuclear spins in a macroscopic single crystal of malonic acid ($C_3H_4O_4$). A small fraction ($\sim 3\%$) of the molecules are triply labeled with (spin-$\frac{1}{2}$)

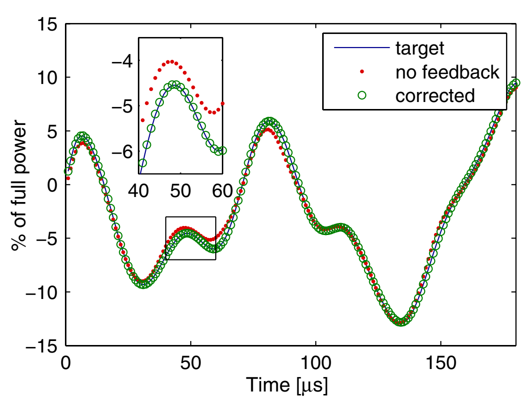


FIG. 2 (color online). Portion of a typical pulse shape showing the designed target shape (solid blue line), the initial attempt at implementing the pulse including systematic imperfections due to nonlinearities in pulse generation and amplification as well as finite bandwidth effects from the probe's resonant circuit (red dots), and the corrected shape after the feedback protocol (green circles). Full power corresponds to nutation frequency of 80 kHz.

^{13}C to form an ensemble of 3-qubit processor molecules, spatially buffered from one another by molecules of the same compound but with naturally abundance ($\sim 1\%$) carbon nuclei [25–27]. The carbon control pulses are numerically optimized to implement the required unitary gates using the gradient ascent pulse engineering [28] algorithm, and are typically designed [29] to have an average Hilbert—Schmidt fidelity of 99.8% over appropriate distributions of Zeeman-shift dispersion and control-fields inhomogeneity.

Preparation and measurement.—The certification protocol calls for the preparation of a number of input states each with a known nonzero projection on some arbitrarily chosen 3-qubit Pauli operator, $\hat{M}_k$. To each of these preparations, the experimental implementation to be certified is applied, followed by a measurement of the corresponding Pauli operator, $\hat{U}\hat{M}_k\hat{U}^\dagger$, where $\hat{U}$ is the target Clifford operation. These measurements are averaged for $\hat{M}_k$ with the same Pauli weight, and linearly transformed [8,27] to find the average fidelity of the experimental implementation. That is to say, the certification protocol follows the twirling protocol for quantum memories [8], with the one exception that the required measurements are transformed by the target operation $\hat{U}$.

In the current implementation, the first step in the initial preparation procedure is a selective polarization transfer from one of the methylene protons (H_{m_1}) to the methylene carbon (C_m). This is realized using a short [30] Hartman—Hahn cross-polarization sequence [31] after tipping the proton polarization to the transverse plane, and is sufficient because the coupling strength between these two nuclei is more than an order of magnitude larger than any other coupling. The state of the three carbon nuclei after this polarization can be described as $\rho_i = \hat{I}^{\otimes 3} + \alpha I\hat{I}X$, where α quantifies the amount of polarization transferred from the proton, and is on the order of $\sim 10^{-5}$ for protons in 7.1 T at room temperature. A free induction decay is collected for this initial state to establish a reference for α, against which all subsequent experiments are compared. Simple coherence-transfer pulses can then be used to prepare all states of the form $\rho_w = \hat{I}^{\otimes 3} + \alpha\hat{X}^{\otimes w}\hat{I}^{\otimes 3-w}$, and their permutations over the 3 qubits, for all possible Pauli weights, $w = 1, 2, 3$. From these states, pulses realizing single-qubit $\frac{\pi}{2}$ rotations are all that is required for preparing a state with nonzero projection on any arbitrary 3-qubit Pauli operator, $\hat{M}_k$. The same set of pulses are sufficient to map the required measurement, $\hat{U}\hat{M}_k\hat{U}^\dagger$, unto an observable in a nuclear magnetic resonance (NMR) experiment.

These single-qubit $\frac{\pi}{2}$ rotations can be realized with very high fidelity, which we now demonstrate using single-qubit randomized benchmarking [12] on each of the qubits—the average fidelity of randomized sequences of pulses that compose to the identity is measured for varying sequence lengths, and assuming that the implementation errors do not depend on which gate is being applied, the average

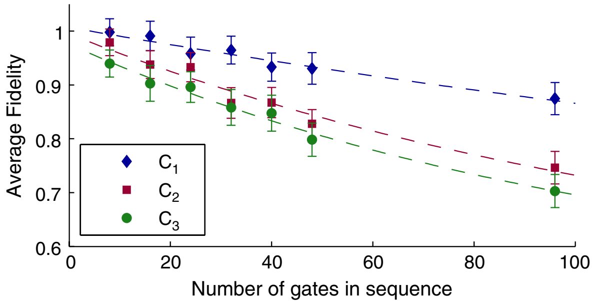


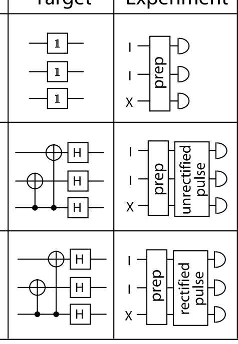
FIG. 3 (color online). Randomized benchmarking of single-qubit $\frac{\pi}{2}$ rotations required for state preparation and measurement—shown is the average fidelity decay of randomized sequences of $\frac{\pi}{2}$ pulses on each of the 3 qubits. Each data point is the average fidelity of 24 sequences. Fitting the data to [12] $\log(F - \frac{1}{2}) = \log A_0 + m \log p$, we extract an average error per gate of $1.6 \pm 0.4 \times 10^{-3}$ for C_1 (blue diamonds), $3.8 \pm 0.7 \times 10^{-3}$ for C_2 (red squares), and $4.4 \pm 0.6 \times 10^{-3}$ for C_3 (C_m) (green circles).

fidelity decay is fit to [12]: $F = A_0 p^m + B_0$, where m is the sequence length, A_0 and B_0 encompass initialization and measurement errors, and p is a parameter related to the average error per gate, $r = \frac{1-p}{2}$. Assuming the gate errors are unital, we set $B_0 = \frac{1}{2}$. In Fig. 3, the average fidelity of 24 sequences each for up to 96 pulses per sequence is plotted, and the average error per single-qubit $\frac{\pi}{2}$ pulse is estimated to be less than 0.5%. Furthermore, to get an estimate of the average combined fidelity of the state preparation and measurement processes, we certify the do nothing operation against the target identity evolution. The results are summarized in Fig. 4(a).

Certifying the 3-qubit encoding.—Next, we choose to certify the (1.5 ms) pulse [32] designed to perform the encoding operation of the phase variant of the 3-qubit quantum error correcting code against the ideal gate [33], which is a 3-qubit Clifford gate that decomposes into two controlled-NOTs (CNOTs) followed by transversal single-qubit Hadamards. As shown in Fig. 4, the average fidelity of the implemented pulse, before and after rectification—including preparation and measurement errors—is estimated to be 86.3% and 97.3%, respectively. Under an assumption that the errors from preparation and measurement are factorable, we estimate the average fidelity of the rectified implementation to be 99%.

Discussion.—We have shown how it is possible to certify individual Clifford group operations efficiently using a modified twirling protocol. As an illustrative example, we demonstrated the certification of the encoding operation for a 3-qubit error correction code, and the improvements on the performance of this operation via feedback of measurements of the control field at the NMR sample. The high fidelity achieved in this demonstration marks state-of-the-art coherent control of nuclear spins in the solid state.

It is worth emphasizing that what we have shown here is the efficiency of the certification algorithm, irrespective of the method used to obtain the pulse form to be certified. Indeed, the problem of designing control sequences to perform a given task remains one of the formidable challenges in quantum control, at least as challenging as simulating the full quantum dynamics. Precisely because it avoids the inefficiency of the simulation, this certification



	w	k_w	λ_w	$\bar{F}$
a	1	6	0.967 ± 0.010	$0.983^{+0.007}_{-0.006}$
	2	21	1.000 ± 0.009	
	3	7	0.978 ± 0.017	
b	1	8	0.848 ± 0.022	$0.863^{+0.013}_{-0.012}$
	2	21	0.883 ± 0.017	
	3	8	0.799 ± 0.023	
c	1	6	0.959 ± 0.014	$0.973^{+0.009}_{-0.008}$
	2	21	0.989 ± 0.013	
	3	8	0.964 ± 0.016	

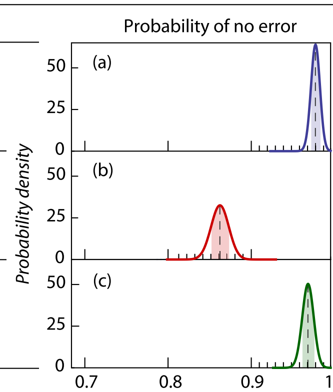


FIG. 4 (color online). Summary of the experimental parameters and results for the three sets of certification experiments—the Target column shows the quantum circuit representation of the ideal process; the Experiment column represents the experimental setup to certify the corresponding implementation, including state preparation and measurement using local readout pulses as described in the text; k_w and λ_w are, respectively, the number of performed experiments, and the average surviving polarization, partitioned by the Pauli weight, w, of the input preparation [8,27]. Shown also are the Bayesian estimated probability density functions for the probability of no error in the experimental implementation of the target gate, as well as the estimated average fidelity. The three sets of experiments are (a) state preparation and measurement compared to the identity operation—this can be thought of as a calibration for the certification procedure; (b) the target is the encoding operation for the 3-qubit phase quantum error correcting code, and the experimental implementation is a numerically designed pulse using gradient ascent pulse engineering; and (c) is the same as (b) but the pulse is corrected for implementation errors using the feedback procedure described in the text.

scheme may be extended to an efficient *in situ* pulse design protocol, in which individual parameters can be optimized iteratively without assumptions about experimental imperfections.

We thank J. Emerson, A. Blais and J. Gambetta for comments on the manuscript. This work was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

*omoussa@iqc.ca
†msilva@bbn.com
‡laflamme@iqc.ca

[1] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).
[2] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).
[3] M. D. Bowdrey, D. K. L. Oi, A. J. Short, K. Banaszek, and J. A. Jones, Phys. Lett. A **294**, 258 (2002).
[4] M. A. Nielsen, Phys. Lett. A **303**, 249 (2002).
[5] R. Blume-Kohout, New J. Phys. **12**, 043034 (2010).
[6] J. Emerson, R. Alicki, and K. Zyczkowski, J. Opt. B **7**, S347 (2005).
[7] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A **80**, 012304 (2009).
[8] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Science **317**, 1893 (2007).
[9] M. Silva, Ph.D. thesis, University of Waterloo, 2008.
[10] C. C. López, A. Bendersky, J. P. Paz, and D. G. Cory, Phys. Rev. A **81**, 062113 (2010).
[11] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A **77**, 012307 (2008).
[12] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. Lett. **106**, 180504 (2011).
[13] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).
[14] D. Gottesman, Phys. Rev. A **57**, 127 (1998).
[15] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[16] Other operations can also be certified efficiently in this manner. The fundamental requirement is that the new measurement $f(\hat{M}_j, \mathcal{C}_i, \mathcal{U})$ be a measurement that is easily implemented with high fidelity.
[17] D. Gottesman, in *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, edited by S. P. Corney, R. Delbourgo, and P. D. Jarvis (International Press, Cambridge, MA, 1999), pp. 32–43.
[18] C. A. Ryan, M. Laforest, and R. Laflamme, New J. Phys. **11**, 013034 (2009).
[19] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. Lett. **102**, 090502 (2009).
[20] S. T. Flammia and Y.-K. Liu, Phys. Rev. Lett. **106**, 230501 (2011).
[21] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Phys. Rev. Lett. **107**, 210404 (2011).
[22] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. G. Cory, J. Chem. Phys. **121**, 6117 (2004).
[23] C. A. Ryan, Ph.D. thesis, University of Waterloo, 2009.
[24] O. Moussa, Ph.D. thesis, University of Waterloo, 2010.
[25] D. G. Cory, R. Laflamme, E. Knill, L. Viola, T. F. Havel, N. Boulant, G. Boutis, E. Fortunato, S. Lloyd, R. Martinez, C. Negrevergne, M. Pravia, Y. Sharf, G. Teklemariam, Y. S. Weinstein, and W. H. Zurek, Fortschr. Phys. **48**, 875 (2000).
[26] J. Baugh, O. Moussa, C. A. Ryan, R. Laflamme, C. Ramanathan, T. F. Havel, and D. G. Cory, Phys. Rev. A **73**, 022305 (2006).
[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.109.070504 for additional experimental details, and an illustrative example of the steps necessary to design a gate-twirling experiment.
[28] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).
[29] C. A. Ryan, C. Negrevergne, M. Laforest, E. Knill, and R. Laflamme, Phys. Rev. A **78**, 012328 (2008).
[30] L. Müller, A. Kumar, T. Baumann, and R. R. Ernst, Phys. Rev. Lett. **32**, 1402 (1974).
[31] S. R. Hartmann and E. L. Hahn, Phys. Rev. **128**, 2042 (1962).
[32] O. Moussa, J. Baugh, C. A. Ryan, and R. Laflamme, Phys. Rev. Lett. **107**, 160501 (2011).
[33] S. L. Braunstein, arXiv:quant-ph/9603024v1.