Direct Fidelity Estimation from Few Pauli Measurements

Steven T. Flammia and Yi-Kai Liu

Direct Fidelity Estimation from Few Pauli Measurements

Steven T. Flammia[1] and Yi-Kai Liu[2]

[1]*Institute for Quantum Information, California Institute of Technology*
[2]*Computer Science Department, University of California, Berkeley*

We describe a simple method for certifying that an experimental device prepares a desired quantum state ρ. Our method is applicable to any pure state ρ, and it provides an estimate of the fidelity between ρ and the actual (arbitrary) state in the lab, up to a constant additive error. The method requires measuring only a *constant* number of Pauli expectation values, selected at random according to an importance-weighting rule. Our method is faster than full tomography by a factor of d, the dimension of the state space, and extends easily and naturally to quantum channels.

In recent years there has been substantial progress in preparing many-body entangled quantum states in the laboratory [1]. A key step in such experiments is to *verify* that the state of the system is the desired one. This can be done using quantum state tomography, or techniques such as entanglement witnesses [2]. However, in many cases these solutions are not fully satisfactory. Tomography gives complete information about the state, but it is very resource-intensive, and has difficulty scaling to large systems. Entanglement witnesses can be much easier to implement, but are not a generic solution since known constructions only work for special quantum states.

Here we propose a new method, *direct fidelity estimation*, that is much faster than tomography, is applicable to a large class of quantum states, and requires minimal experimental resources. Let us first describe the setting of the problem. Consider a system of n qubits, with Hilbert space dimension $d = 2^n$, and let ρ be the desired state, i.e. the state we hope to accurately prepare. We make two basic assumptions. First, we assume that ρ is *pure*. However, we do not assume any additional structure or symmetry, so our method goes beyond previous work [3, 4] to encompass nearly all of the states of interest in experimental quantum information science (e.g., the GHZ and W states, stabilizer states, cluster states, matrix product states, projected entangled pair states, etc.) in a unified framework. Second, we assume that we can measure n-qubit Pauli observables, that is, tensor products of single-qubit Pauli operators; we do not need to perform any other operations. Thus our method is applicable to any system that is capable of single-qubit gates and readout, without needing to rely on 2-qubit gates or entangled measurements.

Our method works by measuring a random subset of Pauli observables chosen according to an "importance-weighting" rule. Roughly, we select Pauli operators that are most likely to detect deviations from the desired state ρ. We use the resulting measurement statistics to estimate the fidelity $F(\rho, \sigma)$, where σ is the actual state in the lab. Surprisingly, although there are 4^n distinct Pauli operators, we only need to sample a *constant* number of them to estimate $F(\rho, \sigma)$ up to a constant additive error, for *arbitrary* σ. That is, for every possible state σ, with high probability over the choice of Pauli measurements, we get an accurate estimate of $F(\rho, \sigma)$.

Although we measure only a constant number of Pauli observables, we need to repeat each measurement many times in order to estimate the corresponding expectation value. The number of repetitions depends on the desired state ρ. In the worst case, it is $O(d)$, but in many cases of practical interest, it is much smaller. For example, for stabilizer states, the number of repetitions is *constant*, independent of the size of the system; and for the W state, it is only quadratic in the number of qubits n.

Even in the worst case, our method requires far fewer resources than full tomography, both in theory and in practice. We demonstrate this by proving lower bounds on the sample complexity of full tomography, and by numerical simulations.

Finally, we show an analogous method for certifying any unitary quantum channel by estimating the entanglement fidelity. We discuss applications to benchmarking quantum circuits — as a special case, our method can certify Clifford circuits in *constant* time, independent of the number of qubits and gates.

Fidelity Estimation. The fidelity between our desired pure state ρ and the actual state σ is given by [5]:

$$F(\rho, \sigma) = \left(\mathrm{tr}\left[(\sqrt{\rho}\sigma\sqrt{\rho})^{1/2}\right]\right)^2 = \mathrm{tr}(\rho\sigma). \qquad (1)$$

We can write $\mathrm{tr}(\rho\sigma)$ in terms of the Pauli expectation values of ρ and σ. Let W_k $(k = 1, \ldots, d^2)$ denote all possible Pauli operators (n-fold tensor products of I, σ_x, σ_y and σ_z). Define the characteristic function $\chi_\rho(k) = \mathrm{tr}(\rho W_k/\sqrt{d})$, and note that

$$\mathrm{tr}(\rho\sigma) = \sum_k \chi_\rho(k)\chi_\sigma(k). \qquad (2)$$

In general, Eq. (2) involves the expectation values of all d^2 Pauli operators. However, it is easy to see that in certain cases fewer Pauli operators are required. For example, if ρ is a stabilizer state, $\chi_\rho(k)$ takes on values of $\pm 1/\sqrt{d}$ at the d points in the stabilizer group of ρ, and vanishes everywhere else. So the sum in (2) contains only d terms, and one can compute $\mathrm{tr}(\rho\sigma)$ by measuring only d Pauli operators. Furthermore, to merely estimate $\mathrm{tr}(\rho\sigma)$

one only needs to measure a small random subset of these Pauli operators. We will now generalize this strategy to work with an arbitrary pure state ρ.

We will construct an estimator for $\mathrm{tr}(\rho\sigma)$ as follows. Select $k \in \{1, \ldots, d^2\}$ at random with probability [6]

$$\Pr(k) = (\chi_\rho(k))^2. \tag{3}$$

By measuring the expectation value of the Pauli observable W_k, we can estimate $\chi_\sigma(k)$, up to some finite precision which we will discuss later. For the time being, let us suppose we can measure $\chi_\sigma(k)$ perfectly. We then construct the estimator

$$X = \chi_\sigma(k)/\chi_\rho(k). \tag{4}$$

It is easy to see that $\mathbb{E}X = \mathrm{tr}(\rho\sigma)$ (where $\mathbb{E}$ denotes the expected value over the random choice of k).

Now say we want to estimate $\mathrm{tr}(\rho\sigma)$ with some fixed additive error ε and failure probability δ. We repeat the above process $\ell = \lceil 1/(\varepsilon^2\delta) \rceil$ times: we choose $k_1, \ldots, k_\ell$ independently, which give independent estimates $X_1, \ldots, X_\ell$, and we let $Y = \frac{1}{\ell}\sum_{i=1}^{\ell} X_i$. By Chebyshev's inequality [7], Y satisfies

$$\Pr[|Y - \mathrm{tr}(\rho\sigma)| \geq \varepsilon] \leq \delta. \tag{5}$$

To complete the description of our method, we show how the ideal "infinite-precision" estimator Y can be approximated by an estimator $\tilde{Y}$ that uses a finite number of copies of the state σ. Given any choice of $k_1, \ldots, k_\ell$, we proceed as follows. For each $i = 1, \ldots, \ell$, we will use m_i copies of σ, where we set

$$m_i = \left\lceil \frac{2}{d\chi_\rho(k_i)^2\ell\varepsilon^2} \log(2/\delta) \right\rceil. \tag{6}$$

(Note that m_i depends on k_i.) We measure the Pauli observable W_{k_i} on each of these copies of σ, and get measurement outcomes $A_{ij} \in \{1, -1\}$ ($j = 1, \ldots, m_i$). Note that $\mathbb{E}A_{ij} = \sqrt{d}\chi_\sigma(k_i)$ (taking the expectation over the random measurement outcomes). Let

$$\tilde{X}_i = \frac{1}{m_i\sqrt{d}\chi_\rho(k_i)} \sum_{j=1}^{m_i} A_{ij}. \tag{7}$$

Finally, we let $\tilde{Y} = \frac{1}{\ell}\sum_{i=1}^{\ell} \tilde{X}_i$. This is our estimate for Y. (Note that $\mathbb{E}\tilde{Y} = Y$.) By Hoeffding's inequality [7], $\tilde{Y}$ has additive error ε and failure probability δ:

$$\Pr[|\tilde{Y} - Y| \geq \varepsilon] \leq \delta. \tag{8}$$

We can then conclude that, with probability $\geq 1 - 2\delta$, the fidelity $F(\rho, \sigma)$ lies in the range $[\tilde{Y} - 2\varepsilon, \tilde{Y} + 2\varepsilon]$.

Our method uses $\ell = \lceil 1/(\varepsilon^2\delta) \rceil$ Pauli observables, independent of the size of the system. It requires m copies

of the state σ, where $m = \sum_{i=1}^{\ell} m_i$. Though this depends on the random choices k_i, we have

$$\mathbb{E}(m_i) = \sum_{k_i} (\chi_\rho(k_i))^2 m_i \leq 1 + \frac{2d}{\ell\varepsilon^2} \log(2/\delta), \tag{9}$$

and hence the expected number of copies satisfies

$$\mathbb{E}(m) \leq 1 + \frac{1}{\varepsilon^2\delta} + \frac{2d}{\varepsilon^2} \log(2/\delta). \tag{10}$$

By Markov's inequality, m is unlikely to exceed its expectation by much: $\Pr(m \geq t \cdot \mathbb{E}(m)) \leq 1/t$, for all $t \geq 1$.

Example: the W state. Suppose our desired state ρ is the W state, i.e. the uniform superposition over computational basis states where a single qubit is $|1\rangle$ and the rest are $|0\rangle$, as previously considered in [3, 4]. To apply our method, we need to sample Pauli operators from the probability distribution (3). It is straightforward to give a short formula for these probabilities, and an explicit algorithm that does the sampling in $\mathrm{poly}(n)$ time [7].

The distribution for a W state is quite different from what one would expect for a Haar-random quantum state. For a random state, one expects most of the Pauli matrices to occur with probability $\sim 1/d^2$; but for the W state, most of the Pauli matrices have probability 0, and all the nonzero probabilities are at least $1/n^2d$. This is an example of a *well-conditioned* state. As we now show, our method requires fewer resources for such states.

Well-conditioned states. We say that a state ρ is well-conditioned with parameter α if for all k, either $\mathrm{tr}(\rho W_k) = 0$ or $|\mathrm{tr}(\rho W_k)| \geq \alpha$. For example, stabilizer states (including the GHZ state) and the W state are well-conditioned with $\alpha = 1$ and $\alpha = 1/n$, respectively, and Dicke states with k excitations have $\alpha = O(1/n^k)$. When ρ is well-conditioned, our method requires a smaller number of measurement settings, as well as fewer copies of the actual state σ. Note first that the estimator X is bounded: $|X| \leq 1/\alpha$. Now we can use the stronger Hoeffding inequality for Eq. (5), and we can choose the number of measurement settings to be $\ell = O\left(\frac{\log(1/\delta)}{\alpha^2\varepsilon^2}\right)$. Thus, the dependence on δ is exponentially better, at a cost of a factor of $1/\alpha^2$.

The total number of copies used in the procedure, m, is bounded in expectation by (10). For well-conditioned states, we can prove a much stronger bound that holds with certainty: $m_i \leq 1 + \frac{2\log(2/\delta)}{\alpha^2\ell\varepsilon^2}$, and hence $m \leq O\left(\frac{\log(1/\delta)}{\alpha^2\varepsilon^2}\right)$. In particular, when ρ is a stabilizer state, m is *independent* of the size of the system; when ρ is the W state, m is only quadratic in the number of qubits n.

Truncating bad events. For an arbitrary pure state ρ, it is possible to modify our protocol so that m is always bounded by $O\left(\frac{1}{\varepsilon^2\delta} + \frac{d\log(1/\delta)}{\varepsilon^2}\right)$. The idea is to construct a nearby ρ' which is well-conditioned with $\alpha = O(1/\sqrt{d})$, by truncating small values of $\chi_\rho(k)$. This eliminates the bad choices of k that cause m to be large, at the expense of introducing a small bias into the fidelity estimate [7].

Dephasing and depolarizing noise. Our method also performs better if one makes some mild assumptions about the noise in the system. For an arbitrary pure state ρ, suppose the actual state σ is given by $\sigma = \mathcal{E}(\rho)$, where $\mathcal{E}$ is some quantum process that shrinks the characteristic function, i.e., for all k, $|\chi_{\mathcal{E}(\rho)}(k)| \leq |\chi_\rho(k)|$. For example, dephasing and depolarizing noise both do this. Again, this implies that $|X| \leq 1$, hence we can use a smaller number of measurement settings, $\ell = O\left(\frac{\log(1/\delta)}{\varepsilon^2}\right)$.

Comparison with full tomography. We have shown that it is possible to estimate the fidelity of an arbitrary pure state using Pauli measurements on $O(d)$ copies of the state. (In this discussion, let us fix the accuracy ε and failure probability δ to be constant.) How good is this result? We argue that our protocol is more efficient than full tomography by a factor of d. By tomography, we mean any procedure that distinguishes arbitrary quantum states with accuracy Δ, so that for every pair of states ρ and σ with $F(\rho, \sigma) \leq 1 - \Delta$, the procedure returns different outputs for ρ and σ.

First, as a toy example, consider what is possible using *arbitrary* quantum operations. Fidelity estimation of a pure state can then be done with $O(1)$ copies using the swap test [8], while full tomography requires $\Omega(d/\operatorname{poly}\log d)$ copies, by Holevo's theorem [9] (see [7]).

In the more realistic situation where only Pauli measurements are allowed (and one cannot perform joint measurements on more than one copy of the state), fidelity estimation uses $O(d)$ copies. We now prove that full tomography requires at least $\Omega(d^2/\log d)$ copies. The idea of the proof is as follows (details in [7]). First, we construct a set of $2^{\Omega(d)}$ quantum states $|\phi_i\rangle$ that are almost orthogonal (for all $i \neq j$, $|\langle\phi_i|\phi_j\rangle|^2 < 1 - \Delta$), and whose Pauli expectation values are small (for all i and k with $W_k \neq I$, $|\langle\phi_i|W_k|\phi_i\rangle| \leq \tau\sqrt{\log d}/\sqrt{d}$). (This is done using repeated applications of Levy's lemma [7, 10].)

Now suppose there is some tomography procedure that can distinguish these states, given t copies. This implies the existence of a classical protocol for transmitting $\Omega(d)$ bits of information over a particular noisy channel $\mathcal{E}$. Intuitively, Bob encodes an $\Omega(d)$-bit message i by sending a string of ± 1 bits through the channel $\mathcal{E}$, in such a way that when Alice receives these bits, they have the same distribution as the measurement outcomes she would have obtained by measuring Pauli observables on the state $|\phi_i\rangle$. Then Alice uses the tomography procedure to reconstruct $|\phi_i\rangle$ and extract the message i. One can show that the channel $\mathcal{E}$ has capacity $O\big((\log d)/d\big)$ (even allowing feedback from Alice to Bob) [11]. Then the converse to Shannon's (classical) noisy coding theorem [11] implies that $t \geq \Omega(d^2/\log d)$.

Numerics. In order to evaluate how tight our analysis is for typical states, we simulated our protocol as follows. We sampled Haar-random states of $n = 8$ qubits and ran our protocol with $\varepsilon = \delta = .05$ (and $\ell = \frac{1}{\varepsilon^2\delta}$) where the true state was created by subjecting the ideal state to
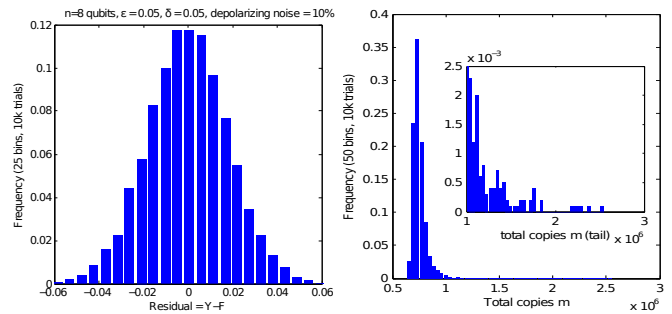

FIG. 1. *Left:* The residual error has a standard deviation of 1.8%. *Right:* Most states use only a typical number of copies, with just .1% of trials using more than four times the expected number of copies, as shown in the inset.

independent 10% depolarizing noise. The residual error $(Y - F)$ and the total number of copies m are plotted as histograms in Fig. 1. We see that the accuracy is always well-behaved, and the total number of copies, excepting a few bad events (for which our truncation procedure applies) is typically close to the average.

We also compared our method to a recent ion trap experiment, in which an 8-qubit W state was verified using full tomography [12]. Under the plausible assumption that dephasing noise is dominant, we would use our protocol with $\varepsilon = .03$, $\delta = .10$, and $\ell = \lceil \log(1/\delta)/\varepsilon^2 \rceil$. Assuming the realistic parameters of 20 ms to perform one measurement and 400 ms to reconfigure a new measurement basis, we would obtain a fidelity estimate accurate to within $\pm 1.2\%$ using just 80 minutes of experiments and a few seconds of classical processing; this compares very favorably with the 10 *hours* of experiments and one *week* of post-processing carried out in [12].

Extension to channels. We now extend our method to unitary quantum channels. Let $\mathcal{U}$ be the desired channel corresponding to some unitary evolution U, i.e., $\mathcal{U} : \rho \mapsto U\rho U^\dagger$. Let $\mathcal{E}$ be the actual channel. We will estimate the *entanglement fidelity*, given by $F_e = \operatorname{tr}(\mathcal{U}^\dagger\mathcal{E})/d^2$ (with $\mathcal{U}$ and $\mathcal{E}$ treated as matrices acting via left multiplication).

Most of the analysis for channels is exactly analogous to the case of states. The main difference is that we may also *input* a state to the channel as well as choose how to measure at the output. Thus, the characteristic function for a channel $\mathcal{E}$ is defined by $\chi_{\mathcal{E}}(k, k') = \frac{1}{d}\operatorname{tr}\big(W_k\mathcal{E}(W_{k'})\big)$, which depends on two indices. The probability distribution from which we sample indices is analogous: $\Pr(k, k') = \frac{1}{d^2}\big[\chi_{\mathcal{U}}(k, k')\big]^2$, and so is our primary estimator: $X = \chi_{\mathcal{E}}(k, k')/\chi_{\mathcal{U}}(k, k')$, for which we have $\mathbb{E}X = F_e$. Now given ℓ independent samples from our probability distribution $(k_1, k'_1), \ldots, (k_\ell, k'_\ell)$, we compute $X_1, \ldots, X_\ell$, and let $Y = \frac{1}{\ell}\sum_{i=1}^{\ell} X_i$. Then choosing $\ell = \lceil 1/(\varepsilon^2\delta) \rceil$ means that Y is an estimate of F_e which is accurate to within ε with a failure probability at most δ.

The main difference between states and channels comes in how we estimate X_i for a given sample (k_i, k'_i). We will

still measure W_{k_i} at the output, but how can we simulate inputting $W_{k_i'}$ into the channel? The key insight is that we can simply sample from states in the eigenbasis of $W_{k_i'}$ and put these states into the channel. Note that these states can always be chosen to be tensor products of local Pauli eigenstates, so no entangling gates are required.

The total number of uses of the channel is bounded in expectation by $\mathbb{E}(m) = O\big(\frac{1}{\varepsilon^2\delta} + \frac{d^2}{\varepsilon^2}\log(1/\delta)\big)$. Statements about well-conditioned channels and truncation also hold in analogy with states [7].

Benchmarking quantum circuits. One application of the above protocol is to evaluate experimental implementations of large quantum circuits: our method allows one to directly measure the entanglement fidelity and average fidelity of the entire circuit, rather than inferring it from tomography performed on individual gates. This is important because as circuits scale up, correlated noise potentially becomes an issue (c.f. Ref. [13]).

The relationship between F_e and the Haar-average fidelity is captured by the formula [14]

$$F_{\mathrm{avg}} = \int \mathrm{d}\psi\, F\big(\mathcal{U}(\psi), \mathcal{E}(\psi)\big) = \tfrac{d}{d+1}F_e + \tfrac{1}{d+1}. \qquad (11)$$

Thus, our method also gives us a direct measure of the typical performance of the channel, similar to what is achieved in other random benchmarking schemes [15–17]. Moreover, one can also prove that the worst-case behavior (as quantified by the diamond norm [18]) is bounded by $4d\sqrt{1-F_e}$ [19], so that for small high-fidelity gates, average and worst-case behavior nearly coincide.

Clifford circuits. Clifford circuits (those consisting of controlled-NOT, Hadamard and phase gates) are key components in many schemes for quantum error-correction, and become universal for quantum computation when augmented with certain state preparations [20, 21]. For a Clifford circuit $\mathcal{U}$, the characteristic function is given by $\chi_{\mathcal{U}}(k,k') = 1$ (when $W_k = \mathcal{U}(W_{k'})$) and 0 otherwise. Sampling only requires that we pick $k' \in \{1,\ldots,d^2\}$ uniformly at random, then use the Gottesman-Knill theorem to efficiently compute the k such that $W_k = \mathcal{U}(W_{k'})$. Clifford circuits are well-conditioned, so our method needs fewer measurement settings and uses of the channel, namely $\ell \le O\big(\frac{1}{\varepsilon^2}\log(1/\delta)\big)$ and $m \le O\big(\frac{1}{\varepsilon^2}\log(1/\delta)\big)$, which is *independent* of the number of qubits and gates (see also Ref. [22]).

Outlook. We have presented a general method for certifying pure states and unitary quantum channels, which requires only Pauli measurements and is faster than full tomography by a factor of d. In common cases such as stabilizer states, the W state, and Clifford circuits, our method requires even fewer resources (constant or polynomial in the number of qubits), and it provides an easy recipe to generalize beyond these examples.

Looking beyond fidelity estimation, it would be interesting to directly estimate and bound an entanglement measure [23], which would obviate the need for an entanglement witness. One may also compare our method with recent proposals for tomography for restricted classes of quantum states [24–27]. Another important direction is to find better techniques for sampling the importance-weighting distribution $\Pr(k)$ for different classes of states.

We would also like to note that M. da Silva, O. Landon-Cardinal and D. Poulin [28] have recently and independently obtained similar results.

---

[1] Nature **453** (2008), (*Nature Insight* supplement).
[2] O. Gühne and G. Toth, Physics Reports **474**, 1 (2009).
[3] R. D. Somma, J. Chiaverini, and D. J. Berkeland, Phys. Rev. A **74**, 052302 (2006).
[4] O. Gühne, C.-Y. Lu, W.-B. Gao, and J.-W. Pan, Phys. Rev. A **76**, 030305 (2007).
[5] Some authors define the square root of this as the fidelity.
[6] These probabilities are normalized since $\mathrm{tr}(\rho^2) = 1$.
[7] We defer some details to the appendices.
[8] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).
[9] A. S. Holevo, Problems of Inform. Transm. **9**, 177 (1973).
[10] P. Lévy, *Problèmes concrets d'analyse fonctionelle* (Gauthier Villars, Paris, 1951).
[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).
[12] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chekal kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, Nature **438**, 643 (2005).
[13] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).
[14] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **60**, 1888 (1999).
[15] 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).
[16] E. Magesan, J. M. Gambetta, and J. Emerson, arXiv:1009.3639 (2010).
[17] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Science **317**, 1893 (2007).
[18] D. Aharonov, A. Kitaev, and N. Nisan, in *STOC '98* (1998).
[19] S. Beigi and R. Koenig, arXiv:1101.1065 (2011).
[20] D. Gottesman and I. L. Chuang, Nature **402**, 390 (1999).
[21] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[22] R. A. Low, Phys. Rev. A **80**, 052314 (2009).
[23] R. Blume-Kohout, J. O. S. Yin, and S. J. van Enk, Phys. Rev. Lett. **105**, 170501 (2010).
[24] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).
[25] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A.

Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, Phys. Rev. Lett. **106**, 100401 (2011).

[26] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Nature Communications **1** (2010).

[27] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, Phys. Rev. Lett. **105**,

250403 (2010).

[28] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, arXiv:1104.3835 (2011).

[29] P. Sen, in *Proc. 21st Ann. IEEE Conf. on Computational Complexity* (2006) pp. 274–287.

[30] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Comm. Math. Phys. **250**, 371 (2004).