Compressed sensing for Hamiltonian reconstruction

Kenneth Rudinger* and Robert Joynt
University of Wisconsin—Madison, Madison, Wisconsin 53706, USA
(Received 28 September 2014; revised manuscript received 23 April 2015; published 19 November 2015)

In engineered quantum systems, the Hamiltonian is often not completely known and needs to be determined experimentally with accuracy and efficiency. We show that this may be done at temperatures that are higher than the characteristic interaction energies, but not too much higher. The condition for this is that there are not too many multiparticle interactions: the Hamiltonian is sparse in a well-defined sense. The protocol that accomplishes this is related to compressed sensing methods of classical signal processing, in this case applied to sparse rather than low-rank matrices.

DOI: 10.1103/PhysRevA.92.052322 PACS number(s): 03.67.Lx, 03.65.Wj, 75.10.Dg

I. INTRODUCTION

In quantum physics, the standard method for understanding a large system has long been to make an approximate model Hamiltonian that captures the essential physics of the material in question. More recently, this situation is often turned on its head: a quantum system of n qubits is constructed and we need to find its Hamiltonian from experimental data. To do quantum information processing of any kind, accurate control of the Hamiltonian is always a prerequisite. One needs to be able to apply external controls to guide the desired time-dependent Hamiltonian, but it is usually also the case that there are "always-on" terms, generally time independent or nearly so, in the Hamiltonian that need to be determined at a precise quantitative level [1]. This is a particularly pressing issue for a quantum memory or in cold-atom systems that are specifically constructed in order to simulate many-body Hamiltonians. For electron spin qubits in semiconductor quantum dots [2], for example, the single-qubit energy-level splittings are subject to unknown random hyperfine fields, and there are also dipole-dipole interactions. These are one- and two-qubit interactions, but there is also the more challenging case of multiqubit interactions. In this paper we propose an efficient way to determine these always-on terms.

For $n = 1$ and $n = 2$, considerable work has been done, since these cases are relevant to the performance of gates [3–5]. Process tomography is the usual tool for problems with $n > 2$, but standard methods [6,7] require a number of measurements that scales exponentially with n. Other methods that pertain particularly to spin systems require only a small number of measurements, but they appear to involve full simulation of the system, a task that again scales exponentially [8–11]. Another proposal aims to characterize large-scale quantum simulators; this procedure requires access to a trusted small-scale quantum simulator [12].

Several authors have investigated the use of techniques from compressed sensing (CS) [13], which would give an efficient solution to this problem when the process matrix χ is s sparse (has only s nonzero elements) in some basis [14–16]. The number of measurements needed to determine χ is then $O(sn)$. However, this scheme requires prior knowledge of the basis in which χ is sparse. Thus it is useful for verifying quantum gates but cannot be used to determine processes (or Hamiltonians) when an approximation of the true dynamics is not known in advance, which is the case we consider.

As pointed out in Ref. [17], it makes sense to take advantage of the fact that, to a very good approximation, almost all qubit Hamiltonians H have only one- and two-qubit interactions, so that the number of parameters to be determined scales only as n^2. These authors suggest a sequence of randomly chosen measurements on randomly prepared states. If the time interval t between preparation and measurement is short enough, $||H||t \ll 1$, then the density matrix is simply related to H. Here $||H||$ is the operator norm (largest eigenvalue) of H. CS techniques can then come into play and the number of experimental configurations required to determine H is $O(n^3)$. However, $||H||$ grows with the size of the system, which limits the usefulness of this scheme.

II. METHOD

Here we propose a different approach for the experimental determination of H. The most general Hamiltonian for an array of n qubits is

$$H = -\eta \sum_{a=1}^{4^n-1} J_a \lambda_a, \tag{1}$$

where a is an n-digit base-4 number $a = a_1 a_2 \ldots a_n$ and the λ_a are tensor products of Pauli matrices: $\lambda_a = \sigma_{a_1} \otimes \sigma_{a_2} \otimes \ldots \otimes \sigma_{a_n}$. $\sigma_{1,2,3} = \sigma_{x,y,z}$ and σ_0 is the identity matrix. For notational convenience we have defined the energy scale η, set by the condition that the dimensionless variables J_a satisfy $|J_a| \leqslant 1$. We assume that only s of the $4^n - 1$ possible J_a values are 0 and $s \ll 4^n$. The system is placed in a bath and comes to thermal equilibrium. The density matrix is $\rho = \exp(-\beta H)/Q$, where Q is the partition function: $Q = \mathrm{Tr}\exp(-\beta H)$ and $\beta = 1/k_B T$. If $T = 0$, ρ reduces to $\rho = |0\rangle\langle 0|$, where $|0\rangle$ is the ground state so that the density matrix has rank 1. We work in the opposite, high-temperature, limit $\eta\beta \ll 1$, where $\rho = 2^{-n}(I - \beta H + \beta^2 H^2/2 + \ldots)$, and we may truncate the expansion. In general there are a macroscopic number of energy eigenstates that enter ρ, and ρ represents a high-rank state. It is important to note that the application of CS proposed here is opposite to others in the literature, which focus primarily on the determination

*Current address: Sandia National Laboratories, Albuquerque, New Mexico 87185-1322, USA; kmrudin@sandia.gov

of states of low rank [18,19]. In fact the density matrix is technically of *full* rank at any finite temperature and the naive (but inefficient) procedure to determine the J_a would be to measure the observables λ_a. For $\eta\beta \ll 1$ this gives $\eta J_a = -2^{-n}\mathrm{Tr}(\lambda_a H) \approx \beta^{-1}\mathrm{Tr}(\lambda_a \rho)$. However, we use the fact that only s values of J_a are nonzero to reduce the number of observables that need to be measured, when s is sufficiently small. [Given that many-body systems typically have few many-body interactions (resulting in small s), it is immediately natural to represent the Hamiltonian in the "Pauli basis" of Eq. (1), guaranteeing sparsity in a known basis.]

The measurement and processing protocol is as follows. After the system reaches equilibrium, its state is given by $\rho = 2^{-n} I + 2^{-n} \sum_{a=1}^{4^n-1} v_a \lambda_a$, where $\vec{v}$ is the equilibrium polarization vector of the system. We then subject the system to a random unitary transformation U so that the new state of the system is $\rho' = U\rho U^{-1}$. The procedure for generating random U's that are efficiently implementable with a small gate set is a modified version of one proposed for quantum data hiding by DiVincenzo, Leung, and Terhal [20], using work by Harrow and Low on random quantum circuits [21]. The U's are not selected uniformly from the Haar distribution but our results indicate that they provide usable compression matrices. (Details for generating each U are provided in Appendix B.)

The new polarization vector $\vec{v}'$ is linearly related to the previous one: $v'_a = \sum_{b=1}^{4^n-1} C_{ab} v_b$, with $C_{ab} = 2^{-n}\mathrm{Tr}(\lambda_a U \lambda_b U^{-1})$. C is an orthogonal matrix and $\vec{v}$ is a long but approximately sparse vector, the "signal vector." We now choose M unique observables in the set $\{\lambda_k\}_{k=1}^{4^n-1}$. Measuring the M chosen observables yields the results $\{y_k\}_{k=1}^{M}$, with the y_k satisfying $-1 \leqslant y_k \leqslant 1$; $|y_k|$ will, in general, be of order $\eta\beta$.

(Additionally, we note that, in practice, each observable λ_i will be repeatedly measured to increase the accuracy of the observed value y_i. If we denote by R_i the number of times that λ_i is measured on identical copies of ρ, then y_i approaches $\mathrm{Tr}(\lambda_i U \rho U^\dagger)$ as R_i approaches infinity. For the time being, we make the idealizing assumption that the expectation value of each observable is exactly known; we address the point of finite measurement statistics at the end of Sec. III.)

We can consider our collection of measurements to form a "measurement vector" $\vec{y}$, which is a subset of the elements of $\vec{v}'$. We now have

$$y_k = \sum_b C^{(M)}_{kb} v_b, \tag{2}$$

where $C^{(M)}$ consists of M rows of C, the choice of rows corresponding to the observables measured. $C^{(M)}$ is an $M \times (4^n - 1)$ matrix, the "compression matrix." The next step is to estimate the polarization vector by minimizing the L_1 norm of all possible polarization vectors that are consistent with the measurement results:

$$\vec{v}_{\mathrm{est}} = \arg\min_{\vec{w}} ||\vec{w}||_1, \quad \text{subject to} \quad \sum_b C^{(M)}_{kb} w_b = y_k. \tag{3}$$

The L_1 norm of a vector $\vec{w}$ is defined as $\|\vec{w}\|_1 = \sum_{i=1}^{d} |w_i|$. This is a convex optimization problem that can be solved efficiently. For our purposes it is important to note that this CS protocol is stable with respect to deviations from exact sparsity in the signal vector, so that, as we see below, the protocol works at moderate temperatures. Also, it can be shown that if $C^{(M)}$ is formed by choosing rows at random from C, then $C^{(M)}$ satisfies a certain restricted isometry condition which guarantees that if $M > An\ln^3 s$ for some constant A, we can recover an approximately s-sparse vector $\vec{v}$ with a high probability [22].

Once a good estimate of the polarization vector is available, we can estimate the Hamiltonian:

$$H_{\mathrm{est}} = \beta^{-1}(2^{-n}\mathrm{Tr}(\ln \rho_{\mathrm{est}})I - \ln \rho_{\mathrm{est}}). \tag{4}$$

III. RESULTS

We now turn to numerical studies of the protocol for three, four, and five qubits, for which a takes on values of $N = 63$, $N = 255$, and $N = 1023$, respectively. We input a random Hamiltonian, compute the equilibrium density matrix ρ, and measure M observables, i.e., characterize ρ by the numbers $\mathrm{Tr}(\lambda_i \rho)$, $i = 1,2,\ldots,M$. While the observables are chosen at random, they are ordered by weight; that is, all observables of weight 1 (i.e., single-qubit measurements) are measured before all observables of weight 2 (i.e., two-qubit measurements), and so on. (See Appendixes A and C for details regarding Hamiltonian generation and weight-ordering of measurements, respectively.)

The simplest case is the determination of the J_a values when we are given that only s of them are nonzero. We do not have firm guarantees of success at a finite temperature, since the density matrix is not s sparse. So the first task is to determine how high the temperature needs to be to ensure success. The temperature is quantified by the dimensionless ratio $\eta\beta$. Success is measured by the distance of H_{est}, the Hamiltonian estimated from Eq. (4), from the actual Hamiltonian H, the metric chosen as the one corresponding to the Frobenius norm: if $(||H_{\mathrm{est}} - H||_F)/\eta <$ threshold, where threshold is determined by numerical stability tests, the procedure is judged to have succeeded.

Figure 1 shows the quality of the reconstruction of H as a function of the parameters M/N, which is the number of observables measured divided by the signal length, and the sparsity ratio s/N. We assume for the moment that each observable is measured completely accurately (which would require $R_i = \infty$ for all measured observables λ_i). There are three qubits and each pixel in the plots is the result of 100 trials. Note, first, that the lower-right corner is a region where the number of nonzero entries in J_a is greater than the number of measured observables: reconstruction is impossible there. As we move away from the diagonal to the upper left, the success probability increases. As is generally observed in cases where CS works, the boundary between success and failure (that is, the Donoho-Tanner phase transition) is sharp. A high temperature is favorable for reconstruction, but even at quite moderate temperatures there is a very substantial region of parameter space where the determination of H succeeds. The red region in both panels is where H is successfully reconstructed, due to the density matrix being approximately sparse in that region.

These computations show that CS can work, in principle, and provide strong evidence that the number of measured

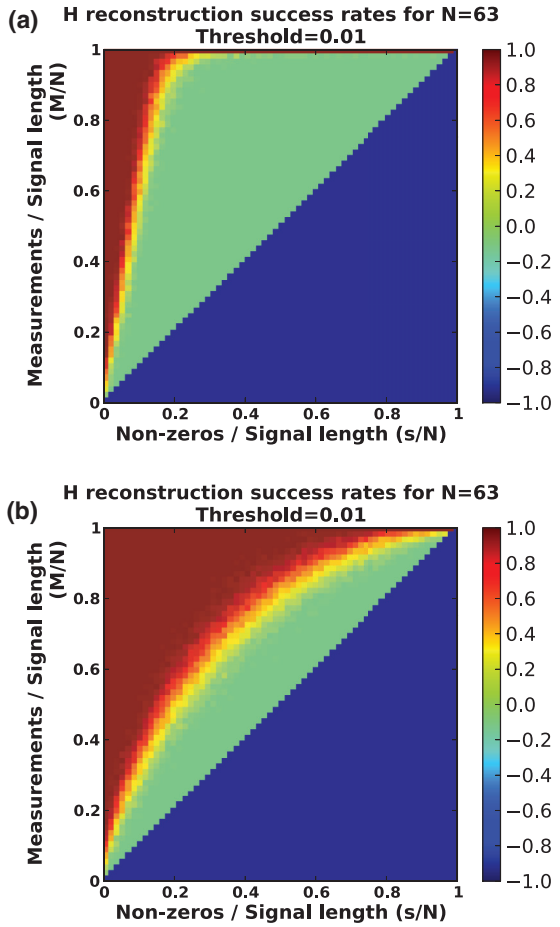

FIG. 1. (Color) Quality of Hamiltonian determination for random couplings as a function of temperature. (a) and (b) The inverse dimensionless temperature is given by $\eta\beta = 10^{-1}$ and $\eta\beta = 10^{-4}$, respectively. Red indicates a high success rate, green indicates failure, and a "negative" success rate (blue) means that reconstruction is impossible. [That is, heat map values between 0 (green) and 1 (red) are to be taken as probabilities on a linear scale, while heat map values of -1 (blue) indicate that not only would reconstruction fail for our proposed procedure, but reconstruction would fail given *any* procedure, as these blue points correspond to measuring fewer observables than there are nonzero entries to be reconstructed.] Each pixel is the average of 100 trials.

observables needed is proportional to n, the number of qubits, rather than N, the number of possible couplings, when the Hamiltonian is sparse. However, equipped with the knowledge that H is sparse, quantum state tomography can also be carried out with a reduced number of measured observables. We next examine the question of how much advantage is actually gained in practice over the straightforward method of standard tomography, stopping when H has been determined. The left column in Fig. 2 gives this comparison for $n = 3$ [Fig. 2(a)], $n = 4$ [Fig. 2(b)], and $n = 5$ [Fig. 2(c)], with small values of s, and for the moderate temperature of $\eta\beta = 10^{-1}$. The number of trials per data point is 100. The sampled M's have a spacing of 1 for $n = 3$ and $n = 4$, starting at a value of $M = 2$; due to computational constraints, every tenth value of M is used for $n = 5$, starting at a value of $M = 11$. The median value of the normalized quality $(||H_{\text{est}} - H||_F)/||H||_F$ of the estimate is plotted as a function of M, so that low values correspond to accurate estimates [23]. When the curve drops off sharply, the "phase transition" from failure to success has occurred. Thus, for example, in Fig. 2(a), the CS protocol for $n = 3$ and $s = 1$ succeeds at $M = 5$. It is shown that CS gives a large savings in the number of required observables for all cases considered, ranging (roughly) from a factor of 4 to a factor of 7 for $n = 3$, from 6 to 12 for $n = 4$, and from 12 to as high as 50 for $n = 5$. This is good evidence that the advantage of the CS protocol increases with n, as we would expect from the scaling arguments above.

In most cases of actual physical interest, we not only have some knowledge of the sparsity of H, but also have some knowledge of where the nonzeros lie. For example, for spin qubits, one- and two-body interactions are likely to be much greater in magnitude than three-body and higher interactions. We then find $s = O(n^2)$. Locality may also reduce the sparsity; for sufficiently short-range interactions $s = O(n)$. This is a very different situation than we have considered so far, where the nonzero J_a's were taken at random. Of course exponential reductions in M required to reconstruct H are now out of the question. The question is whether we can still get speedups that may be useful in real situations: even constant speedups can be important. So we perform the same numerical experiment as in the left column in Fig. 2, but now the nonzero J_a values are restricted to those corresponding to λ_a that are one- and two-qubit operators, i.e., a has at most two nonzero digits. The results are shown in the right column in Fig. 2. (The number of trials and all other parameters are the same as in the left column in Fig. 2.) In the "no-CS" (standard tomography) protocol, measurements of one- and two-body operators are made first, which improves the performance of the no-CS procedure, but not enough to overcome the advantage of the CS protocol. (See Appendix D for further details.)

The ratio of the number of observables required is about a factor of 2 to 4 for $n = 3$, about a factor of 3 to 6 for $n = 4$, and about a factor of 6 to 8 for $n = 5$. Thus the speedup is less when the knowledge of the locations of the nonzeros is increased, but it is still quite substantial. More importantly, it appears that the speedup still increases with the number of qubits.

It is also important to address the point of finite measurement statistics (that is, finite R_i). While all simulated data presented here were generated without finite measurement error, these findings are, in the appropriate temperature limit, robust against such error. Each observable λ_i may be thought of as a biased coin, with the bias being given by $2y_i - 1 = 2\text{Tr}(\lambda_i U\rho U^\dagger) - 1$. The accuracy in estimating this bias is given by the usual biased-coin formula; that is, the error in the estimator for λ_i will scale as $\sqrt{R_i}^{-1}$. As we are looking at thermalized states, we expect each y_i to be relatively close to 0; thus R_i should be of the order of $|y_i|^{-2}$ in order to estimate the expectation value of λ_i accurately. As $|y_i| \sim O(2^{-n}\eta\beta)$, we find that relatively hotter and larger systems would require unfeasibly large R_i to implement. (For example, $\eta\beta = 10^{-4}$, $n = 3$ would require R_i to be of the order of several billion!) However, at moderate temperatures and qubit numbers, our proposed scheme should be reasonable. For example, our main results shown in Fig. 2 assume $\eta\beta = 10^{-1}$, implying that $R_i < 7 \times 10^3$ would be sufficient for a three-qubit system.

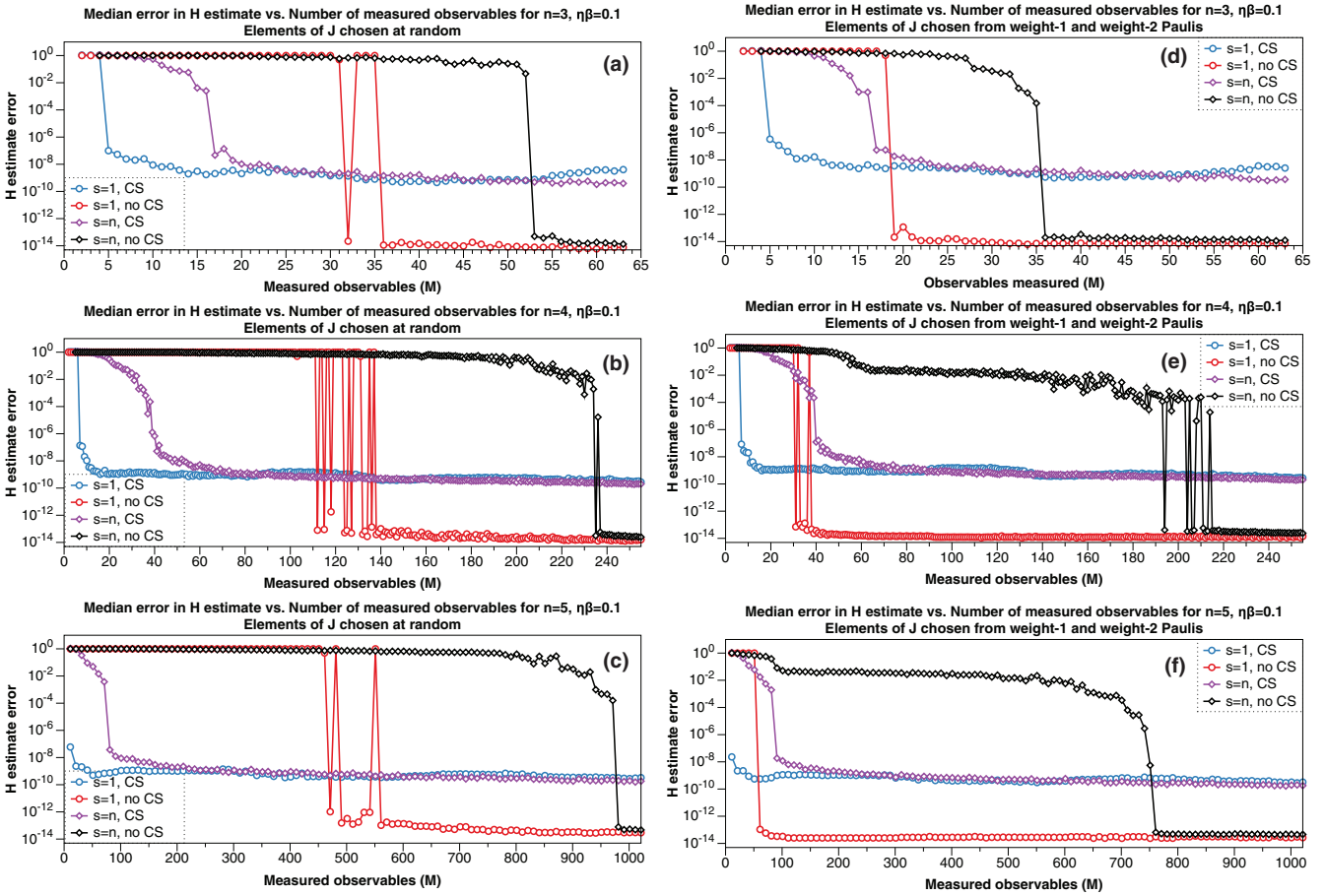

FIG. 2. (Color) Quality of Hamiltonian determination as a function of number of qubits and measured observables. (a)–(c) For Hamiltonians with entirely random couplings; (d)–(f) for Hamiltonians with random one- and two-qubit couplings only. (a) and (d) $n = 3$; (b) and (e) $n = 4$; (c) and (f) $n = 5$. Each plot gives the error in the estimated Hamiltonian as a function of the number of observables measured (M), with compressed sensing (CS) and without (no CS), for varying small values of s (number of nonzero couplings). [While the minimum error obtained by the no-CS protocol appears to be lower than the minimum error obtained with the CS protocol ($\sim 10^{-14}$ and $\sim 10^{-9}$, respectively), these differences are simply artifacts of the different noise floors of different numerical methods. For each algorithm, achieving its respective noise floor indicates that the Hamiltonian has been successfully reconstructed.] In each case, the CS protocol substantially decreases the M required to accurately reconstruct the Hamiltonian; this improvement is more significant when the couplings are totally random. Additionally, we see that for both coupling schemes, the improvement increases with the number of qubits. Each data point is the median value of 100 trials.

Thus we see that the key to success is striking a balance between the need for approximate sparsity (corresponding to smaller $\eta\beta$) and the need for a feasibly low value of R_i (corresponding to larger $\eta\beta$). This exact regime will depend on the constraints of the experimentalist wishing to implement this protocol.

Finally, we observe that one potential problem with this method is that if there exists an always-on unknown Hamiltonian, it may be hard or impossible to implement high-precision measurements via high-fidelity unitaries, as the presence of the unknown Hamiltonian terms can ruin the implementation of the desired unitaries. However, there exist several potential work-arounds. First, even if the terms in the Hamiltonian are "always on," if the experimentalist can temporarily turn off the entire Hamiltonian, such terms would not affect the applied unitaries. Second, even if the Hamiltonian is truly always on (i.e., can never be turned off), if the applied unitaries are sufficiently strong, then the applied unitaries will only be slightly perturbed by the always-on terms of the Hamiltonian.

More generally, the problem rectifying the effects of always-on Hamiltonian terms on desired unitaries and measurements indicates the importance of ensuring self-consistency in the tomographic protocol. The exploration of self-consistent tomographic protocols is a relatively new development, with several notable proposals appearing recently [24,25]. It is certainly worth exploring the possibility of adapting one of these frameworks to fit our protocol or developing an entirely new such framework to implement tomography of sparse high-rank states in a self-consistent manner.

IV. CONCLUSION

Previous improvements in the efficiency of quantum state tomography have shown the usefulness of CS techniques by focusing on the reconstruction of states of low rank. This work, by contrast, uses this technique to reconstruct states of high rank. This is not useful for validation of gate quality, but

it can be used to determine the parameters in a many-body Hamiltonian.

As CS reduces the number of real-valued system parameters that must be measured, at the cost of increased postprocessing, CS is only of value for systems in which measurements are expensive but signal processing and postprocessing are cheap. This trade-off is highly attractive for many classical applications, but the trade-offs vary from case to case. In nonqubit systems, one would need to examine whether there is a clear way to determine the candidate observables. Our quantum protocol will be useful and attractive when measurement settings are expensive but quantum gate operations are cheap. Otherwise, straightforward tomography will be better. The competition between the two is greatly affected by how much advance knowledge we have about the system. It is when we do not have a very good idea in advance about the shape of the Hamiltonian that our method is useful.

ACKNOWLEDGMENTS

We thank John Aidun, Eric Bach, Charles Baldwin, Robin Blume-Kohout, S. N. Coppersmith, Daniel Crow, Adam Frees, Mark Friesen, John Gamble, Kevin Jamieson, Amir Kalev, Robert Nowak, and Nathan Wiebe for useful discussions.

APPENDIX A: GENERATION OF H

We generate random sparse Hamiltonians in the following manner. For an energy scale η [as defined in Eq. (1)] and fixed sparsity s, we generate s random pairs (J,λ), where each J is a real number chosen uniformly from the interval $[-\eta,\eta]$ and each λ_α is an n-qubit Pauli operator, chosen without replacement from the set of all appropriate Pauli operators. By "appropriate," we mean that each λ satisfies the desired weight constraints for the particular Hamiltonian we are constructing. (That is, if we are simulating a Hamiltonian with arbitrary interactions, each λ may be of any weight; if we are simulating a Hamiltonian that contains only one- and two-qubit interactions, then each λ may only be of weight 1 or weight 2.)

APPENDIX B: GENERATION OF U

To choose a random unitary map that is efficiently implementable with a small gate set, we use the following procedure, inspired by a technique for quantum data hiding proposed by DiVincenzo, Leung, and Terhal [20], along with work by Harrow and Low on random quantum circuits [21].

For an n-qubit system, we consider the set $\mathcal{G}$ of quantum gates

$$\mathcal{G} = \left\{ H_p, P_q, P_r^\dagger, R_s\left(\frac{\pi}{8}\right), \mathrm{CNOT}_{tu} \right\}, \tag{B1}$$

where H is the Hadamard gate, P is the phase gate, $R(\frac{\pi}{8})$ is the $\frac{\pi}{8}$ gate, and CNOT is the controlled-not gate. The subscripts label the qubit (or qubits) that each gate is acting on; that is, $\mathcal{G}$ contains all single-qubit copies of $\{H, P, P^\dagger, R(\frac{\pi}{8})\}$ and all two-qubit copies of CNOT.

To form the unitary map U, we simply select (with replacement) n^8 elements of $\mathcal{G}$ uniformly at random. Letting g_i denote the ith selection from $\mathcal{G}$, we define U to be given by

$$U = \prod_{i=1}^{n^8} g_i. \tag{B2}$$

Note that this gives us a random unitary operation on n qubits which, while not selected uniformly from the Haar distribution, is sufficiently random as to successfully generate a compression matrix which can be used for CS. Additionally, we note that it is an open question whether or not a smaller set of gates and/or a shorter gate sequence could yield equally successful results. The recent results in [26] indicate that the gate sequence could potentially be linear (up to additional logarithmic factors) in n.

Finally, we note that perfect knowledge of U may be unobtainable. However, this does not affect our overall point, as compressed sensing methods are robust against errors in the measurement record, provided these errors are not too large [13]. Thus, perfect knowledge of U is not required, although it should be known to reasonable approximation.

APPENDIX C: WEIGHT-ORDERING OF MEASUREMENTS

It may be of some benefit to the experimentalist for whom lower weight observables are easier to measure to be able to prioritize low-weight measurements over high-weight measurements. Therefore, we show here that the order the observables are chosen in should not affect the accuracy of the Hamiltonian or the density matrix reconstructions, allowing for the measurements to be chosen according to weight. (That is, all single-qubit measurements may be performed before any two-qubit measurement, which in turn may precede all three-qubit measurements, and so on.) This ordering by weight is justified in the following manner.

We note that if the kth Pauli measured is λ_k, then the kth element of our measurement vector $\vec{y}$ is given as

$$y_k = \mathrm{Tr}(\lambda_k U^\dagger \rho U), \tag{C1}$$

where ρ is the initial density matrix and U is the random unitary map. However, due to the cyclic property of the trace, we may re-express Eq. (C1) as

$$y_k = \mathrm{Tr}[(U\lambda_k U^\dagger)\rho]. \tag{C2}$$

That is, we may consider our kth observable measured to correspond to measuring the expectation value of a Pauli subjected to a random unitary transformation with respect to the fixed and original density matrix. Therefore, as U effectively randomizes each λ_k, choosing them in order of their weights should not affect the reconstruction algorithm. (Indeed, we have performed numerical tests which demonstrate this.)

APPENDIX D: STATE RECONSTRUCTION VIA "NO-CS" PROTOCOL

The no-CS protocol for reconstructing the density matrix ρ is as follows. For an estimate of ρ in which M Pauli observables are measured, the M expectation values are input as the appropriate v_i's; the remaining v_i's are set to 0. While this estimation procedure could theoretically produce a nonphysical ρ_{est} with negative eigenvalues, in practice,

this is not a concern, as any state we are estimating has a polarization vector with a small L_2 norm, while a nonphysical density matrix with one or more negative eigenvalues has a polarization vector with a large L_2 norm.

[1] M. D. Shulman, S. P. Harvey, J. M. Nichol, S. D. Bartlett, A. C. Doherty, V. Umansky, and A. Yacoby, Nat. Commun. **5**, 5156 (2014).

[2] D. P. DiVincenzo and D. Loss, Superlatt. Microstruct. **23**, 419 (1998).

[3] J. H. Cole, S. G. Schirmer, A. D. Greentree, C. J. Wellard, D. K. L. Oi, and L. C. L. Hollenberg, Phys. Rev. A **71**, 062312 (2005).

[4] S. J. Devitt, J. H. Cole, and L. C. L. Hollenberg, Phys. Rev. A **73**, 052317 (2006).

[5] S. G. Schirmer and D. K. L. Oi, Phys. Rev. A **80**, 022333 (2009).

[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[7] I. L. Chuang and M. A. Nielsen, J. Modern Opt. **44**, 2455 (1997).

[8] D. Burgarth, K. Maruyama, and F. Nori, Phys. Rev. A **79**, 020305 (2009).

[9] D. Burgarth and K. Maruyama, New J. Phys. **11**, 103019 (2009).

[10] D. Burgarth, K. Maruyama, and F. Nori, New J. Phys. **13**, 013019 (2011).

[11] C. Di Franco, M. Paternostro, and M. S. Kim, Phys. Rev. Lett. **102**, 187203 (2009).

[12] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Phys. Rev. Lett. **112**, 190501 (2014).

[13] E. Candès and M. Wakin, Signal Process. Mag. IEEE **25**, 21 (2008).

[14] 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).

[15] C. H. Baldwin, A. Kalev, and I. H. Deutsch, Phys. Rev. A **90**, 012110 (2014).

[16] A. V. Rodionov, A. Veitia, R. Barends, J. Kelly, D. Sank, J. Wenner, J. M. Martinis, R. L. Kosut, and A. N. Korotkov, Phys. Rev. B **90**, 144504 (2014).

[17] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut, and H. Rabitz, Phys. Rev. A **84**, 012107 (2011).

[18] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).

[19] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, New J. Phys. **14**, 095022 (2012).

[20] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, IEEE Trans. Inf. Theory **48**, 580 (2002).

[21] A. W. Harrow and R. A. Low, Comm. Math. Phys. **291**, 257 (2009).

[22] R. Vershynin, in *Compressed Sensing*, edited by Y. C. Eldar and G. Kutyniok (Cambridge University Press, Cambridge, 2012), pp. 210–268.

[23] The median error was chosen as the figure of merit, as a small number of large outliers made plotting the mean error less helpful, particularly when the errors are best compared on a logarithmic scale. It should be noted that, nonetheless, the error distributions were not terribly far from normal.

[24] R. Blume-Kohout, J. K. Gamble, E. Nielsen, P. Maunz, T. Scholten, and K. Rudinger, Sandia Technical Report SAND2015-0224 (Sandia National Laboratories, Albuquerque, NM, 2015).

[25] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, Phys. Rev. A **87**, 062119 (2013).

[26] R. Cleve, D. Leung, L. Liu, and C. Wang, arXiv:1501.04592v1.