Estimation of many-body quantum Hamiltonians via compressive sensing

A. Shabani,[1] M. Mohseni,[2] S. Lloyd,[2,3] R. L. Kosut,[4] and H. Rabitz[1]
[1]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[2]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[4]*SC Solutions, Sunnyvale, California 94085, USA*
(Received 1 March 2010; revised manuscript received 21 February 2011; published 11 July 2011)

We develop an efficient and robust approach for quantum measurement of nearly sparse many-body quantum Hamiltonians based on the method of compressive sensing. This work demonstrates that with only $O(s\ln(d))$ experimental configurations, consisting of random local preparations and measurements, one can estimate the Hamiltonian of a d-dimensional system, provided that the Hamiltonian is nearly s sparse in a known basis. The classical postprocessing is a convex optimization problem on the total Hilbert space which is generally not scalable. We numerically simulate the performance of this algorithm for three- and four-body interactions in spin-coupled quantum dots and atoms in optical lattices. Furthermore, we apply the algorithm to characterize Hamiltonian fine structure and unknown system-bath interactions.

DOI: 10.1103/PhysRevA.84.012107 PACS number(s): 03.65.Wj, 02.30.Zz, 03.67.−a

I. INTRODUCTION

The dynamical behavior of multipartite quantum systems is governed by the interactions amongst the constituent particles. Although the physical or engineering considerations may specify some generic properties about the nature of quantum dynamics, the specific form and the strength of multiparticle interactions are typically unknown. Additionally, quantum systems usually have an unspecified interaction with their surrounding environment. In principle, one can characterize quantum dynamical systems via "quantum process tomography" (QPT) [1–5]. However, the relationship between relevant physical properties of a system to the information gathered via QPT is typically unknown. Alternatively, knowledge about the nature of inter- and intra-many-body interactions within the system and/or its environment can be constructed by identifying a set of (physical or effective) Hamiltonian parameters generating the dynamics [6–17]. Currently, a scalable approach for efficient estimation of a full set of Hamiltonian parameters does not exist.

The dynamics of a quantum system can be estimated by observing the evolution of some suitable test states. This can be achieved by a complete set of experimental configurations consisting of appropriate input states and observables measured at given time intervals. Knowledge about the dynamics may then be reconstructed via inversion of the laboratory data by fitting a set of dynamical variables to the desired accuracy. Estimating Hamiltonian parameters from such a procedure faces three major problems: (1) The number of required physical resources grows exponentially with the degrees of freedom of the system [1,3–5]. (2) There are inevitable statistical errors associated with the inversion of experimental data [1,3–5]. (3) The inversion generally involves solving a set of nonlinear and nonconvex equations, since the propagator is a nonlinear function of Hamiltonian parameters [6–17]. The first two problems are always present with any form of quantum tomography, but the last problem is specific to the task of Hamiltonian identification as we wish to reconstruct the generators of the dynamics. Many quantum systems involve two-body local interactions, so the goal is often to estimate sparse Hamiltonians with effectively a polynomial number of unknown parameters. Unfortunately, quantum state and process tomography cannot readily exploit this potentially useful feature.

The highly nonlinear feature in the required inversion of laboratory data was studied in Ref. [6] in which closed-loop learning control strategies were used for the Hamiltonian identification. In that approach one estimates the unknown Hamiltonian parameters by tailoring shaped laser pulses to enhance the quality of the inversion. Identification of time-independent (or piecewise constant) Hamiltonians have been studied for single-qubit and two-qubit cases [10,11] to verify the performance of quantum gates. Estimation of these Hamiltonians is typically achieved via monitoring the expectation values of some observable, e.g., concurrence, which are time periodic functions. Through Fourier transform of this signal the identification task is reduced to finding the relative location of the peaks and heights of the Fourier spectrum [10,11]. Bayesian analysis is another method proposed for robust estimation of a two-qubit Hamiltonian [12]. The difficulty with these methods is then scalability with the size of the system. A symmetrization method for efficient estimation of the magnitude of effective two-body error generators in a quantum computer was studied in [13] by monitoring quantum gate average fidelity decay. Recently, it was demonstrated that direct or selective QPT schemes could be used for efficient identification of short-time behavior of sparse Hamiltonians [14] assuming controllable two-body quantum correlations with auxiliary systems and the exact knowledge of the sparsity pattern. Alternative schemes for the indirect estimation of the coupling parameters in spin networks with restricted controllability have been recently introduced [15–18].

In this work, inspired by recent advances in classical signal processing known as *compressive sensing* [19], we use random *local* input states and measurement observables for efficient Hamiltonian identification. We show how the difficulties with the nonlinearity of the equations can be avoided by either a short time or a perturbative treatment of the dynamics. We demonstrate that randomization of the measurement observables enables compressing the extracted Hamiltonian

information into an exponentially smaller set of outcomes. This is accomplished by a generalization of compressive sensing to utilize random matrices with correlated elements. This approach is applicable for Hamiltonians that are nearly sparse in a known basis with an arbitrary unknown sparsity pattern of parameters. The laboratory data can then be inverted by solving a convex optimization problem. This algorithm is tolerant to noise and experimental imperfections. The power of this procedure is illustrated by simulating three- and four-body Hamiltonians for neutral atoms in an optical lattice and spin-coupled quantum dot systems, respectively. Furthermore, we apply the algorithm to estimate Hamiltonian fine structure and to characterize unknown system-bath interactions for open quantum systems.

II. QUANTUM DYNAMICAL EQUATIONS

The time evolution of a quantum system in a pure state is governed by the Schrödinger equation $\dot{\rho}(t) = -i[H,\rho(t)]$. The solution of this equation for a time-independent Hamiltonian can be simply expressed as $\rho(t) = \exp(-itH)\rho(0)\exp(itH)$. In principle, the Hamiltonian of the system H can be estimated by preparing an appropriate set of test states $\{\rho_k\}$ and measuring the expectation value of a set of observables $\{M_j\}$ after the system has evolved for a certain period of time. The expectation value of these observables can be expressed as

$$p_{jk} = \langle M_j \rangle_{\rho_k} = \mathrm{Tr}(e^{itH} M_j e^{-itH} \rho_k). \tag{1}$$

Equation (1) implies that the experimental outcomes $\{p_{jk}\}$ are nonlinear functions of the Hamiltonian parameters. To avoid the difficulties of solving a set of coupled nonlinear equations we consider the short-time behavior of the system. Monitoring the short-time dynamics of the system is valid when the relevant time scales of the system evolution satisfy $t \ll K^{-1}$, where for positive operator-valued measure (POVM) operators $\{M_j\}$, the constant K equals $||\mathbf{H} = -i[H,.]||_{\mathrm{spec}}$. In this paper, for an operator O, we represent the superoperator $-i[O,.]$ by $\mathbf{O}$. The general expression of K is given in Appendix B; also see Appendix A for definition of the superoperator norm $||.||_{\mathrm{spec}}$. This yields the linearized form of Eq. (1),

$$p_{jk} = \mathrm{Tr}(M_j\rho_k) + it\,\mathrm{Tr}([H,M_j]\rho_k) + O(K^2t^2). \tag{2}$$

The linear approximation contains enough information to fully identify the Hamiltonian and the higher-order terms do not provide additional information. The short-time approximation implies prior knowledge about the system dynamical time scale or the order of magnitude of $||H||_{\mathrm{spec}}$. This prior knowledge can be available from generic physical and engineering considerations. For example, in solid-state quantum devices the time scale of single-qubit rotations is typically on the order of 1–10 ns. The switching time for exchange interactions varies among different solid-state systems from 1 to 100 ps (for more details, see Appendix B).

We expand the Hamiltonian in an orthonormal basis $\{\Gamma_\alpha\}$, where $\mathrm{Tr}(\Gamma_\alpha^\dagger\Gamma_\beta) = d\delta_{\alpha,\beta}$: $H = \sum_\alpha h_\alpha\Gamma_\alpha$. Here d is the dimension of the Hilbert space. In this representation the Hamiltonian parameters are the coefficients h_α. The expanded form of the above affine equation (2) is

$$\bar{p}_{jk} = it\sum_\alpha \mathrm{Tr}([\Gamma_\alpha,M_j]\rho_k)h_\alpha. \tag{3}$$

Here we introduce the experimental outcomes as $\bar{p}_{jk} = p_{jk} - \mathrm{Tr}(M_j\rho_k)$, since $\mathrm{Tr}(M_j\rho_k)$ is *a priori* known. The relation (3) corresponds to a single experimental configuration (M_j,ρ_k). For a d-dimensional system, the total number of Hamiltonian parameters h_α is d^2. Thus, one requires the same number of experimental outcomes, p_{jk} that leads to d^2 linearly independent equations. For a system of n qubits, this number grows exponentially with n as $d = 2^{2n}$.

A Hamiltonian H is considered to be s sparse if it only contains s nonzero parameters $\{h_\alpha\}$. More generally, a Hamiltonian H is termed nearly s sparse, for a threshold η, if at most s coefficients h_α ($H = \sum h_\alpha\Gamma_\alpha$) have magnitudes greater than $\eta h_{\max}$, where $h_{\max} = \max(h_\alpha)$. By definition, the sparsity is basis dependent. However, for local interactions, the basis in which the Hamiltonian is sparse is typically known from physical or engineering considerations.

III. COMPRESSIVE HAMILTONIAN ESTIMATION

Our algorithm is based on general methods of so-called compressive sensing that recently have been developed in signal processing theory [19]. Compressive sensing allows for condensing signals and images into a significantly smaller amount of data, and recovery of the signal becomes possible from far fewer measurements than required by traditional methods.

Compressive sensing has two main steps: encoding and decoding. The information contained in the signal is mapped into a set of laboratory data with an exponentially smaller representation. This compression can be achieved by randomization of data acquisition. The actual signal can be recovered via an efficient algorithm based on convex optimization methods. Compressive sensing has been applied to certain quantum tomography tasks. Standard compressive sensing has been used for efficient pseudothermal ghost imaging [20,21] and quantum state tomography [22]. Recently a compressive sensing theory for quantum process tomography has been developed and experimentally tested in tomography of a photonic two-qubit controlled-phase gate [23].

Here, we first describe how the Hamiltonian information is compressed into the experimental data. The output of a single measurement is related to the unknown signal (Hamiltonian parameters) through the relation (3). Suppose we try m different experimental configurations [i.e., m different pairs of (M_j,ρ_k)]. This yields a set of linear equations

$$\vec{p'} = \Phi\vec{h}, \tag{4}$$

where Φ is a $m \times d^2$ matrix with elements $\Phi_{jk,\alpha} = (it/\sqrt{m})\mathrm{Tr}([\Gamma_\alpha,M_j]\rho_k)$ (a factor $1/\sqrt{m}$ is included for simplifying the proofs, Appendix C) and $\vec{p'} = \vec{p}/\sqrt{m}$. In general, m has to be greater than or equal to d^2 in order to solve Eq. (4). A Hamiltonian estimation attempt with $m < d^2$ seems impossible as we face an underdetermined system of linear equations with an infinite number of solutions.

However, any two s-sparse Hamiltonians h_1 and h_2 still can be distinguished via a properly designed experimental setting, if the measurement matrix Φ preserves the distance between h_1 and h_2 to a good approximation:

$$(1-\delta_s)||h_2-h_1||_{l_2}^2 \leqslant ||\Phi(h_2-h_1)||_{l_2}^2 \leqslant (1+\delta_s)||h_2-h_1||_{l_2}^2 \qquad (5)$$

for a constant $\delta_s \in (0,1)$. A smaller δ_s ensures higher distinguishability of s-sparse Hamiltonians. The inequality relation (5) is termed a *restricted isometry property* (RIP) of the matrix Φ [24]. We now discuss how to construct a map Φ satisfying this inequality, and how small the value of m can be made.

The RIP (5) for a matrix Φ can be established by employing the measure concentration properties of random matrices. In each experiment the test state and the measurement observable can be drawn randomly from a set of configurations $\{M_j,\rho_k\}$ realizable in the laboratory. The independent selection of ρ_k and M_j leads to a matrix Φ with independent rows but correlated elements $\Phi_{jk,\alpha}$ in each row. Thus the standard results from compressive sensing theory are not applicable here (Appendix C).

In contrast, here we derive a concentration inequality for a matrix with independent rows and correlated columns as the backbone for the RIP of our quantum problem in Appendix C. Using Hoeffding's inequality, we show that for any Hamiltonian h and a random matrix Φ, with column-only correlations, the random variable $||\Phi h||^2$ is concentrated around $||h||^2$ with a high probability, i.e., $\forall\, 0<\delta<1$,

$$\text{Prob}\left\{\left|||\Phi h||_{l_2}^2-||h||_{l_2}^2\right| \geqslant \delta||h||_{l_2}^2\right\} \leqslant 2e^{-mc_0(\delta+c_1)^2} \qquad (6)$$

for some constants c_0 and c_1.

Using the above inequality, now we can show how an exponential reduction in the minimum number of the required configurations can be achieved for Hamiltonian estimation. The inequality (6) is defined for any h, while the inequality in the definition of RIP, Eq. (5), is for any s-sparse h. As shown in Ref. [25], there is an inherent connection between these two inequalities. It is proved that any matrix Φ satisfying Eq. (6) has RIP with probability greater than $1-2\exp\{-mc_0(\delta_{s/2}+c_1)^2+s[\ln(d^4/s)+\ln(12e/\delta_{s/2})]\}$. In addition, whenever $m \geqslant c_2 s\,\ln(d^4/s)$, for a sufficiently large constant c_2 one can find a constant $c_3 \geqslant 0$ such that the likelihood of the RIP to be satisfied converges exponentially fast to unity as $1-2\exp(-c_3 m)$.

The set of experimental configurations defined by Eq. (4) and the concentration properties given by Eqs. (5) and (6) can be understood as encoding the information of a sparse Hamiltonian into a space with a lower dimension. Next we need to provide an efficient method for decoding in order to recover the original Hamiltonian. The decoder is simply the minimizer of the l_1 norm of the signal h. Implementing this decoder is a special convex optimization problem, which can be solved via fast classical algorithms, yet not strictly scalable. Furthermore, the encoding/decoding scheme is robust to noisy data as $||p'-\Phi h||_{l_2} \leqslant \epsilon$, where ϵ is the noise threshold. Note that ϵ includes the error of linearization [see Eq. (2)] that is in the order of K^2t^2. Denote h_0 as the true representation of the Hamiltonian. For a threshold η, $h_0(s)$ is an approximation to h_0 obtained by selecting the s elements of h_0 as those that are larger than $\eta h_{\max}$ and setting the remaining elements to zero. Now we state our main result.

IV. ALGORITHM EFFICIENCY

If the measurement matrix $\Phi \in \mathbb{C}^{m\times d^4}$ is drawn randomly from a probability distribution that satisfies the concentration inequality in Eq. (5) with $\delta_s < \sqrt{2}-1$, then there exist constants $c_2,c_3,d_1,d_2>0$ such that the solution $h^\star$ to the convex optimization problem,

$$\begin{aligned} &\text{minimize } ||h||_{l_1}, \\ &\text{subject to } ||p'-\Phi h||_{l_2} \leqslant \epsilon, \end{aligned} \qquad (7)$$

satisfies

$$||h^\star-h_0||_{l_2} \leqslant \frac{d_1}{\sqrt{s}}||h_0(s)-h_0||_{l_1}+d_2\epsilon, \qquad (8)$$

with probability $\geqslant 1-2e^{-mc_3}$ provided that

$$m \geqslant c_2 s\,\ln(d^4/s), \qquad (9)$$

where the performance of a l_1 minimizer, Eq. (8), and the necessary bound $\delta_s < \sqrt{2}-1$ are derived by Candés in Ref. [26].

As an example, for a system consisting of n interacting qubits, the exponential number of parameters describing the dynamics, 2^{2n}, can be estimated with a linearly growing number of experiments $m \geqslant c_2 s[8\,\ln(2)n-\ln(s)]$. The second term, $d_2\epsilon$, indicates that the algorithm performance is bounded by the experimental uncertainties. Consequently, for fully sparse Hamiltonians and $\epsilon=0$ the exact identification of an unknown Hamiltonian is achievable. The properties of the ensemble from which the states and measurement observables are chosen would determine the parameter δ_s, and consequently, the performance of the algorithm. The exponentially small required number of experimental configurations gives an important feature to the presented method for Hamiltonian identification: the initial states can be selected from product states only and the measurements can be limited to single-body observables. Such restriction on the configuration set becomes possible since this set has $c_4 n 2^n$ number of independent elements, for some constant c_4, which is large enough for random selection of required $c_2 s[8\,\ln(2)n-\ln(s)]$ configurations. In this way we can avoid measuring many-body (k-body) observables whose expectation values are exponentially small in k and require exponential accuracy in the measurements. The linear independency of the Φ matrix rows for a random set of local state preparations and observables can be guaranteed by a polynomial level of computational overhead before conducting the experiments.

A certification for the near-sparsity assumption can be obtained from Eqs. (8) and (9) as follows: Suppose $h_m^\star$ is the algorithm's outcome for m configurations. The near-sparsity assumption is certified on the fly during the experiment, if the estimation improvement $||h_{m+1}^\star-h_m^\star||$ converges to zero for a polynomially large total number of configurations.

V. PHYSICALLY NEARLY SPARSE HAMILTONIAN

Although physical systems at the fundamental level involve local two-body interactions, many-body Hamiltonians often

describe quantum dynamics in a particular representation or in well-defined approximate limits. The strength of the nonlocal k-body terms typically is much smaller than the two-body terms with strength J and decreases with the number k. For a fixed sparsity threshold η, k_η is defined as the largest number k for which k-body terms have strength larger than ηJ. Then the number of elements of an s-sparse approximation of an n-body Hamiltonian grows linearly, approximately as $ng(k_\eta)$, where the $g(k_\eta)$ is determined by the geometry of the system. Consequently, for the case of an n-qubit system, the minimum number of measurements m grows quadratically with the number of qubits that is $m \geqslant c_2 g(k_\eta)\{8\ln(2)n^2 - \ln[g(k_\eta)]n\}$.

Note that the norm of a sparse Hamiltonian $||\mathbf{H}||_{\text{spec}}$ increases with s and therefore with the size of the system. Considering the experimental limitation of fine tuning the measurement time t limits the size of the system for which the linearization assumption holds. Therefore the experimental time resolution and the local interaction strength determines the largest size of the system that our algorithm can handle.

A general class of many-body interactions arises when we change the basis for a bosonic or fermionic system expressed by a (typically local) second-quantized Hamiltonian to a Pauli basis, e.g., via a Jordan-Wigner transformation. For fermionic systems the interactions are imposed physically from Coulomb's force and Pauli exclusion principle. The second-quantized Hamiltonian for these systems can be generally written as

$$\hat{H} = \sum_{p,q} b_{pq}\hat{a}_p^\dagger \hat{a}_q + \sum_{p,q,r,s} b_{pqrs}\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s, \qquad (10)$$

where the annihilation and creation operators ($\hat{a}_j$ and $\hat{a}_j^\dagger$, respectively) satisfy the fermionic anticommutation relations: $\{\hat{a}_i,\hat{a}_j^\dagger\} = \delta_{ij}$ and $\{\hat{a}_i,\hat{a}_j\} = 0$ [27]. For example, in chemical systems the coefficients h_{pq} and h_{pqrs} can be evaluated using the Hartree-Fock procedure for N single-electron basis functions. The Jordan-Wigner transformation can then be used to map the fermionic creation and annihilation operators into a representation in terms of Pauli matrices $\hat{\sigma}^x,\hat{\sigma}^y,\hat{\sigma}^z$. This allows for a convenient implementation on a quantum computer, as was demonstrated recently for the efficient simulation of chemical energy of molecular systems [28]. An important example of a Coulomb-based Hamiltonian is the spin-coupled interaction in quantum dots which has the following Pauli representation:

$$H = \sum_{i,j,k,\ldots} b_{i,j,k,\ldots}\sigma_A^i \otimes \sigma_B^j \otimes \sigma_C^k \ldots, \qquad (11)$$

where $A,B,C,\ldots$ indicate the location of the quantum dots, σ^i's are Pauli operators, and $b_{i,j,k,\ldots}$ generally represents a many-body spin interacting term. In practice, these Hamiltonians are highly sparse or almost sparse due to symmetry considerations associated with total angular momentum [29]. For example, the Hamiltonian for the case of four quantum dots (A,B,C,D) takes the general form [29]

$$H_{\text{exchange}} = J \sum_{A\leqslant i<j\leqslant D} \sigma_i\sigma_j + J'[(\sigma_A\sigma_B)(\sigma_C\sigma_D) + (\sigma_A\sigma_C)(\sigma_B\sigma_D) + (\sigma_A\sigma_D)(\sigma_B\sigma_C)]. \qquad (12)$$

Another class of effective many-body interactions often emerge in a perturbative and/or short-time expansion of dynamics, such as effective three-body interactions between atoms in optical lattices [30] that we study in this work.

Next, we simulate the performance of our algorithm for estimation of such sparse many-body Hamiltonians in optical lattices [30] and quantum dots [29].

A. Three-body interactions in optical lattices

An optical lattice is a periodic potential formed from interference of counterpropagating laser beams where neutral atoms are typically cooled and trapped one per site. Consider four sites in two adjacent building blocks of a triangular optical lattice filled by two species of atoms [30]. The interaction between atoms is facilitated by the tunneling rate J between neighboring sites and collisional couplings U when two or more atoms occupy the same site. For each site an effective spin is defined by the presence of one type of atom as the up state $\uparrow$ and the presence of the other type as the down state $\downarrow$. Three-body interactions between atoms in a triangular optical lattice can be significant. The effective Hamiltonian for this system is studied in Ref. [30]. The on-site collisional interaction U and the tunneling rates $J = J^\uparrow = 2J^\downarrow$ are taken to be the same in all sites; also $U = U_{\uparrow\uparrow} = U_{\downarrow\downarrow} = 2.12U_{\uparrow\downarrow} =$ 10 kHz. The effective Hamiltonian of the four-spin system is

$$H_{\text{opt-latt}} = \sum_{j,\alpha=x,y,z} b_1^\alpha \sigma_j^\alpha \sigma_{j+1}^\alpha + b_2^\alpha \sigma_j^\alpha \sigma_{j+1}^\alpha \sigma_{j+2}^\alpha, \qquad (13)$$

where $\{b_1^\alpha, b_2^\alpha\}$ are functions of $\{J,U\}$ and their explicit forms are given in Appendix D. The ratio $|J/U|$ quantifies the sparsity level. For a fixed value of U, a smaller J leads to weaker three-body interactions and therefore a higher level of sparsity. As expected, this enhances the algorithm performance. A four-qubit Hamiltonian has 255 parameters, while the above Hamiltonian (13) has only 15 nonzero parameters for $J \ll 1$, or 21 otherwise. More precisely, for a threshold η the sparsity level is $15 + 6\,\Theta(|J/U| - \eta)$, where $\Theta(\cdot)$ is the Heaviside step function.

We assume that the system can be initialized in a random pure product state $\rho_k = |\psi_k\rangle\langle\psi_k|$, where $|\psi_k\rangle = |\psi_k^1\rangle \otimes \cdots \otimes |\psi_k^4\rangle$ and $|\psi_k^i\rangle$'s are drawn from the Fubini-Study metric-induced distribution. The required observables for the algorithm are uniformly selected from single-qubit Pauli operators $\{\sigma_i^x,\sigma_i^y,\sigma_i^z\}$. This choice of states and observables allows for $\delta \approx 0.37 < \sqrt{2} - 1$. Let us denote the extracted Hamiltonian and the true Hamiltonian by H^* and H_{true}, respectively. Here, the performance of the algorithm is defined by the relative error $1 - ||H^* - H_{\text{true}}||_{\text{Fro}}/||H_{\text{true}}||_{\text{Fro}}$. Notice that a smaller error can be assigned as the true Hamiltonian H_{true} can be shifted by $e_g I_{d\times d}$ for the ground-state energy free parameter e_g. The results for different numbers of configurations are depicted in Fig. 1, for various values of J. As is evident in Fig. 1, performance accuracy of above 94% can be obtained with only 80 settings significantly smaller than the approximately 6×10^4 configurations required in QPT. The robustness of this scheme was also investigated for 10% random error in simulated experimental data leading to about a 5% reduction in the overall performance.

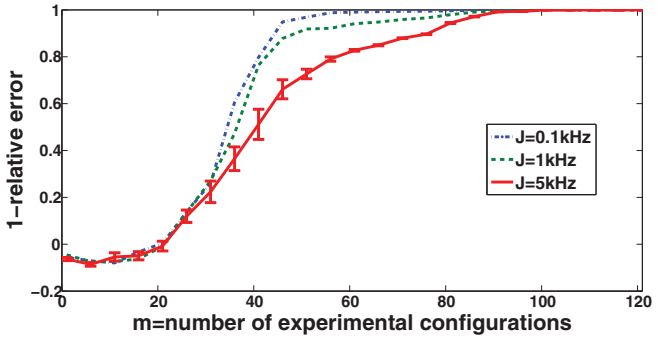

FIG. 1. (Color online) The Hamiltonian estimation average performance is illustrated for a system of four adjacent sites in an optical lattice for different tunneling rates J and collisional coupling $U = 10$ kHz. The error bars demonstrate the standard deviation of the performance due to the random and independent selection of m configurations (shown only for $J = 5$ kHz). Performance accuracy of above 90% with only 60 settings is achievable for $J = 1$ kHz, which is significantly smaller than the approximately 6×10^4 required experimental configurations in QPT.

B. Four-body interactions in quantum dots

Another important class of effective many-body Hamiltonians can be obtained for electrons in quantum dots coupled through an isotropic (Heisenberg) or anisotropic exchange interaction. For example, the Hamiltonian for the case of four quantum dots (A,B,C,D) takes the general form of Eq. (12). The first term in the summation is a two-body Heisenberg exchange interaction and the last three terms are four-body spin interactions. In certain regimes, the ratio $|J'/J|$ can reach up to 16%. The amplitude of $|J'/J|$ determines the sparsity level of the Hamiltonian. A four-qubit Hamiltonian has 255 parameters, while the above Hamiltonian (12) has only 18 nonzero parameters for $J' \ll J$, or 39 otherwise. More precisely, for a threshold η the sparsity level is $18 + 21\,\Theta(|J'/J| - \eta)$.

Here we use an efficient modification of signal recovery referred to as "reweighted l_1 minimization," which is described in Appendix E. The performance of this algorithm is demonstrated in Fig. 2, which shows a significant reduction of the required number of settings in contrast to the standard QPT.

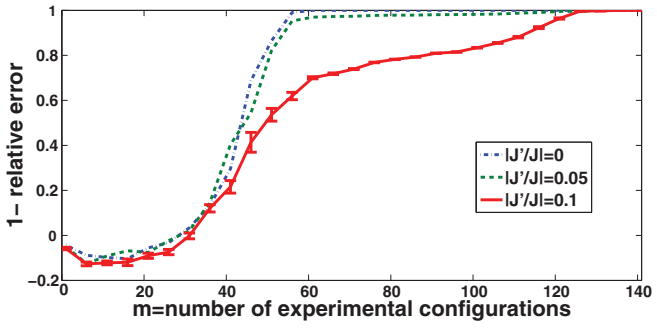

FIG. 2. (Color online) Estimation of the exchange interaction Hamiltonian for four electrons in quantum dots. The average performance of the procedure is illustrated for different values of $|J'/J|$ with 50 iterations of the l_1-reweighted minimization. The standard deviations are shown only for $|J'/J| = 0.1$ It is demonstrated that only 60 different configurations are sufficient for estimating the unknown Hamiltonian with an accuracy above 95% for $|J'/J| = 0.05$, instead of the approximately 6×10^4 required settings via QPT.

VI. CHARACTERIZATION OF HAMILTONIAN FINE STRUCTURES AND SYSTEM-BATH INTERACTIONS

A. Hamiltonian fine estimation

In many systems a primary model of the interactions is often known through physical and/or engineering considerations. Starting with such an initial model we seek to improve our knowledge about the Hamiltonian by random measurements. Let's assume the initial guess about the Hamiltonian H_0 is close to the true form H_{true} that is for $\Delta = H_{\text{true}} - H_0$, $||\Delta||_{\text{spec}} \ll ||\mathbf{H}_{\text{true}}||_{\text{spec}}$. Therefore for a perturbative treatment we demand $t||\Delta||_{\text{spec}} \ll 1$, which is a much weaker requirement compared to $t||\mathbf{H}_{\text{true}}||_{\text{spec}} \ll 1$. We can approximate Eq. (1) in the paper by using the matrix identity $e^{(A+A')t} = e^{At} + \int_0^t e^{A(t-s)} A' e^{(A+A')s} ds$ for any matrices A and A'. If matrix A' has a norm much smaller than the norm of A, we can approximately write

$$e^{(A+A')t} \approx e^{At} + \int_0^t e^{A(t-s)} A' e^{As} ds. \tag{14}$$

By taking $A = H_0$ and $A' = \Delta$, we find

$$p_{jk} \approx \text{Tr}\left(M_j^0 \rho_k\right) + \text{Tr}\left(\left[i \int_0^t e^{isH_0} \Delta e^{-isH_0} ds, M_j^0\right] \rho_k\right), \tag{15}$$

where $M_j^0 = e^{itH_0} M_j e^{-itH_0}$. This equation is linear in $\triangle$, consequently, in a similar fashion as above, the compressive sensing analysis can be applied for efficient estimation of the fine structure of Hamiltonians.

B. Characterizing system-bath interactions

The identification of a decoherence process is a vital task for quantum engineering. In contrast to the usual approach of describing the dynamics of an open quantum system by a Kraus map or a reduced master equation, here we use a microscopic Hamiltonian picture to efficiently measure the system-bath coupling terms generating the overall decoherence process. However, since we consider a full dynamics of the system and bath, this method can be applied to a finite-size environment such as a spin bath, or a surrogate Hamiltonian modeling of an infinite bath. In the latter case a harmonic bath of oscillators is approximated by a finite spin bath [31].

Consider an open quantum system with a total Hamiltonian:

$$H = H_S \otimes I_B + I_S \otimes H_B + H_{SB} \tag{16}$$

and

$$H_{SB} = \sum_{p,q} \lambda_{p,q} S_p \otimes B_q, \tag{17}$$

where H_S (H_B) denotes the system (bath) free Hamiltonian, H_{SB} is the system-bath interaction with coupling strengths $\{\lambda_{p,q}\}$, and complete operator bases of the system and bath are $\{S_p\}$ and $\{B_q\}$, respectively.

We develop a formalism to estimate $\lambda_{p,q}$ parameters in the weak system-bath coupling regime and with the sparsity assumption that a few numbers of $\lambda_{p,q}$ have a significant value.

The Liouvillian dynamical equation is

$$\frac{d}{dt}\rho_{SB}(t) = \left(\mathbf{H}_0 + \sum_{pq}\lambda_{pq}\mathbf{H}_{pq}\right)\rho_{SB}(t), \tag{18}$$

where $\mathbf{H}_0 = -i[H_S \otimes I_B + I_S \otimes H_B,.]$ and $\mathbf{H}_{pq} = -i[S_p \otimes B_q,.]$. In the regime of weak coupling to a finite bath, $||\mathbf{H}_{SB}|| \ll \min\{||\mathbf{H}_S||, ||\mathbf{H}_B||\}$, the Liouvillian equation (18) can be solved perturbatively if time t satisfies $t||\mathbf{H}_{SB}|| \ll 1$. For an initial system density state ρ_k, using the matrix identity (14) with $A = H_S + H_B$ and $A' = H_{SB}$, we find the measurement outcomes as

$$p_{jk} \approx \mathrm{Tr}(\rho_k M_j) + \sum_{pq}\lambda_{pq}\,\mathrm{Tr}\left(\left[\int_0^t ds e^{(t-s)\mathbf{H}_0}\mathbf{H}_{pq}e^{s\mathbf{H}_0}\rho_k^{SB}, M_j\right]\right). \tag{19}$$

where M_j is a system-only observable and ρ_0^{SB} is the initial system-bath state. Here we assume that the system and bath are initially in an equilibrium state ρ_0^{SB} before the action on the system only that transforms $\rho_0 = \mathrm{Tr}_B(\rho_0^{SB})$ into $\rho_k = \mathrm{Tr}_B(\rho_k^{SB})$. This above affine function (19) between the outcomes p_{jk} and coupling parameters $\{\lambda_{pq}\}$ is similar to Eq. (2) in the paper for Hamiltonian estimation. Consequently, the compressive sensing algorithm can be employed for computing $\{\lambda_{pq}\}$'s.

VII. OUTLOOK

We have introduced an efficient and robust experimental procedure for the identification of nearly sparse Hamiltonians using only separable (local) random state preparations and single-body measurements. There are a number of future directions and open problems associated with this work. It is not known how the performance of the algorithm depends on the distribution of the ensemble from which the states and measurement observables are drawn. Also, a general closed-loop learning approach for updating knowledge of the sparsity basis of an arbitrary Hamiltonian is an interesting open problem that will be of importance for generic compressive system identification. The presented method for Hamiltonian estimation is promising for drastic reduction in the number of experimental configurations. However, the classical resources for postprocessing are not scalable. A fully scalable Hamiltonian estimation method might also be achievable via a hybrid of compressive sensing and DMRG (density-matrix renormalization group) methods as recently it has been applied in quantum state tomography [32].

ACKNOWLEDGMENTS

We thank DARPA [Grant No. FA9550-09-1-0710 administered through AFOSR (A.S., R.L.K., H.R.)], NSERC, and Center for Extreme Quantum Information Theory (M.M.), for funding.

APPENDIX A: VECTORS AND OPERATOR NORM

In this paper we use the following different norms. For a vector x, l_1 and l_2 norms are defined as

$$||x||_{l_1} = \sum_i |x_i|, \quad ||x||_{l_2} = \sqrt{x^\dagger x}. \tag{A1}$$

For a matrix A, the Frobenius norm is

$$||A||_{\mathrm{Fro}} = \sqrt{\mathrm{Tr}(A^\dagger A)}, \tag{A2}$$

and in the next appendix we use the spectral norm ,

$$||A||_{\mathrm{spec}} = \sqrt{\max_{|\psi\rangle}\ \langle\psi|\,A^\dagger A\,|\psi\rangle\,/\,\langle\psi|\psi\rangle}. \tag{A3}$$

APPENDIX B: ANALYSIS OF THE SHORT-TIME APPROXIMATION

The short-time monitoring of the system's dynamics requires a prior knowledge of the dynamical time scales. In the solid-state quantum devices, in particular, in the context of quantum control and quantum information processing, the time scale of single-qubit rotations is typically on the order of 1–10 ns. The switching time for exchange interactions varies among different solid-state systems. For superconducting phase qubit the duration of a swap gate is about 10 ns [33]. For electron-spin qubits in quantum dots and in donor atoms (Heisenberg models) [34–36], and also for quantum dots in cavities (anisotropic exchange interactions) [37] the coupling time is between 10 and 100 ps, while for exciton-coupled quantum dots (XY model) and Forster energy transfer in multichromophoric complexes the relevant time scale is on the order of 1 ps. Next we rigorously derive bound on the evolution time t that guarantees the validity of the short-time approximation.

For an input state ρ_k, the expectation value of an observable M_j is

$$p_{jk} = \mathrm{Tr}[M_j\rho_k(t)] = \mathrm{Tr}(e^{iHt}M_j e^{-iHt}\rho_k). \tag{B1}$$

Considering the Dyson expansion of the propagator $e^{-iHt} = I - itH - \frac{1}{2}t^2H^2 + \cdots$, we find

$$p_{jk} = \mathrm{Tr}(M_j\rho_k) + it\ \mathrm{Tr}([H,M_j]\rho_k) - \frac{t^2}{2}\mathrm{Tr}([H,[H,M_j]]\rho_k) + \cdots. \tag{B2}$$

Therefore, for the linearization assumption, it is sufficient to have for the lth term,

$$t^l \times \max_j |\mathrm{Tr}(\overbrace{[H,[H,[\ldots,}^{l\ \text{times}}M_j]]]\rho_k)| \ll 1, \qquad \forall\, l. \tag{B3}$$

A tighter bound can be found for operators $\{M_j\}$ from a POVM by employing the following norm for the superoperator $\mathbf{H} = -i[H,.]$ [38]:

$$||\mathbf{H}||_{\mathrm{spec}} = \sup_X \frac{||\mathbf{H}(X)||_{\mathrm{spec}}}{||X||_{\mathrm{spec}}} = \sup_X \frac{||-i[H,X]||_{\mathrm{spec}}}{||X||_{\mathrm{spec}}}, \tag{B4}$$

$$\mathrm{Tr}[\mathbf{H}^l(M_j)\rho_k] \leqslant ||\mathbf{H}^l(M_j)||_{\mathrm{spec}} \leqslant ||\mathbf{H}^l||_{\mathrm{spec}}\cdot||M_j||_{\mathrm{spec}} \leqslant ||\mathbf{H}^l||_{\mathrm{spec}}. \tag{B5}$$

The last inequality is deduced from the fact that M_j is a POVM observable. It is easy to prove by induction that $||\mathbf{H}^l||_{\text{spec}} \leqslant ||\mathbf{H}||^l_{\text{spec}}$. Applying this result to the desired inequality (B3), we can find a single bound sufficient for linearization: $t \ll ||\mathbf{H}||^{-1}_{\text{spec}}$.

APPENDIX C: RIP FROM A CONCENTRATION INEQUALITY

In this work, we generalize the standard compressive sensing algorithm such that the necessity for independent randomness in all elements of the measurement matrix, ϕ, can be avoided. A common approach to establish RIP ([25]) for a matrix Φ is by introducing randomness in the elements of this matrix. This approach benefits from measure concentration properties of random matrices. In classical signal processing each element $\Phi_{jk,\alpha}$ can be independently selected from a random distribution such as Gaussian or Bernoulli. Whereas in the Hamiltonian estimation formulation [Eq. (4) in the paper] there is no freedom for independent selection of the Φ matrix elements.

Here we prove the concentration inequality that we employed for establishing the restricted isometry property. Though Φ is a random matrix, because it is constructed from quantum states and observables of a finite dimensional system, it is bounded. Thus we are able to apply *Hoeffding's concentration inequality:* If $v_1,...,v_m$ are independent bounded random variables such that $\text{Prob}\{v_i \in [a_i,b_i]\} = 1$, then for $S = \sum_i v_i$,

$$\begin{aligned}\text{Prob}\{S - \mathbf{E}(S) \geqslant t\} &\leqslant \exp\left[-2t^2 \Big/ \sum_i (b_i - a_i)^2\right],\\ \text{Prob}\{S - \mathbf{E}(S) \leqslant -t\} &\leqslant \exp\left[-2t^2 \Big/ \sum_i (b_i - a_i)^2\right]\end{aligned} \tag{C1}$$

for any $t > 0$. (Here $\mathbf{E}$ denotes the expectation value.) Set $v_i = |\phi_i^\dagger x|^2$ for a row ϕ_i. Then with $S = \sum_i v_i = ||\Phi x||^2_{l_2}$, we get $\forall x$,

$$\begin{aligned} v_i = x^\dagger(\phi_i\phi_i^\dagger)x \in (1/m)[w_l,w_u]||x||^2_{l_2},\\ \mathbf{E}(S) = \mathbf{E}||\Phi x||^2_{l_2} \in [f,g]||x||^2_{l_2}\end{aligned} \tag{C2}$$

for constants w_l, w_u, f, g. Note that f and g are the min and max singular values of $\mathbf{E}(\Phi^\dagger\Phi)$. From (C2) we find that $\forall t_+,t_- > 0$ and $\forall x$,

$$\begin{aligned}\text{Prob}\left\{S - g||x||^2_{l_2} \geqslant t_+\right\} &\leqslant \text{Prob}\{S - \mathbf{E}(S) \geqslant t_+\},\\ \text{Prob}\left\{S - f||x||^2_{l_2} \leqslant -t_-\right\} &\leqslant \text{Prob}\{S - \mathbf{E}(S) \leqslant -t_-\}.\end{aligned}$$

These together with Eqs. (C1) and (C2), and the choice of $t_+ = (\delta + 1 - g)||x||^2_{l_2}$ and $t_- = (f - 1 + \delta)||x||^2_{l_2}$, yields

$$\text{Prob}\left\{\left|||\Phi x||^2_{l_2} - ||x||^2_{l_2}\right| \geqslant \delta||x||^2_{l_2}\right\} \leqslant 2e^{[-2m(\delta+\epsilon)^2]/[(w_u-w_l)^2]}, \tag{C3}$$

with $\epsilon = \min\{1 - g, f - 1\}$. To ensure that $t_+,t_- > 0$, we need $1 - \delta < f \leqslant g < 1 + \delta$. Since the observable M can be scaled by any real number, a sufficient condition is $g/f < (1 + \delta)/(1 - \delta)$. For the simulations in this paper, this ratio becomes 2.176.

APPENDIX D: FOUR-SITES OPTICAL LATTICE HAMILTONIAN

Let us consider four sites in two adjacent building blocks of a triangular optical lattice filled by two species of atoms, $\uparrow$ and $\downarrow$. Atoms interact by tunneling between neighboring sites, $J^\uparrow$ and $J^\downarrow$, and through collisional couplings in the same site, U. The Hamiltonian for such a system can be written as [30]

$$\begin{aligned} H_{\text{opt-latt}} = &\sum_j \left(0.03\frac{J^{\uparrow 2} + J^{\downarrow 2}}{U} - 0.27\frac{J^{\uparrow 3} + J^{\downarrow 3}}{U^2}\right)\sigma_j^z\sigma_{j+1}^z\\ &- \left(\frac{2.1(J^\uparrow + J^\downarrow)J^\uparrow J^\downarrow}{U^2} + \frac{J^\uparrow J^\downarrow}{U}\right)\left(\sigma_j^x\sigma_{j+1}^x\right.\\ &\left.+ \sigma_j^y\sigma_{j+1}^y\right) + \sum_j 0.14\frac{J^{\uparrow 3} - J^{\downarrow 3}}{U^2}\sigma_j^z\sigma_{j+1}^z\sigma_{j+2}^z\\ &- 0.6\frac{J^\uparrow J^\downarrow(J^\uparrow - J^\downarrow)}{U^2}\left(\sigma_j^x\sigma_{j+1}^z\sigma_{j+2}^x + \sigma_j^y\sigma_{j+1}^z\sigma_{j+2}^y\right),\end{aligned} \tag{D1}$$

where $\sigma_j^{x,y,z}$ are Pauli operators.

APPENDIX E: REWEIGHTED l_1-MINIMIZATION

In order to simulate our algorithm performance for estimating the above Hamiltonian we use an iterative algorithm that outperforms the standard l_1 norm minimization [39]. This procedure entails initializing a weight matrix $W = I_{d^2}$ and a weight factor $\sigma > 0$, and repeating the following steps until convergence is reached: (1) Solve for h, minimize $||Wh||_{l_1}$, subject to $||p' - \Phi h||_{l_2} \leqslant \epsilon$. (2) Update weights,

$$W = \text{diag}[1/(|h_1| + \sigma), \ldots, 1/(|h_{d^2}| + \sigma)]. \tag{E1}$$

where $h = \text{vec}(h_i)$ is the Hamiltonian vectorized form. Φ is the measurement matrix and p' is the experimental data with a noise threshold ϵ.

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

[2] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. OBrien, M. A. Nielsen, and A. G. White, Phys. Rev. Lett. **90**, 193601 (2003).

[3] M. Mohseni and D. A. Lidar, Phys. Rev. Lett. **97**, 170501 (2006).

[4] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Science **317**, 1893 (2007).

[5] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).

[6] J. M. Geremia and H. Rabitz, Phys. Rev. Lett. **89**, 263902 (2002).

[7] N. Boulant, T. F. Havel, M. A. Pravia, and D. G. Cory, Phys. Rev. A **67**, 042322 (2003).
[8] R. L. Kosut, I. A. Walmsley, and H. Rabitz, e-print arXiv:quant-ph/0411093 (2004).
[9] K. C. Young, M. Sarovar, R. Kosut, and K. B. Whaley, Phys. Rev. A **79**, 062301 (2009).
[10] 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).
[11] S. J. Devitt, J. H. Cole, and L. C. L. Hollenberg, Phys. Rev. A **73**, 052317 (2006).
[12] S. G. Schirmer and D. K. L. Oi, Phys. Rev. A **80**, 022333 (2009).
[13] B. Levi, C. C. Lopez, J. Emerson, and D. G. Cory, Phys. Rev. A **75**, 022314 (2007).
[14] M. Mohseni and A. T. Rezakhani, Phys. Rev. A **80**, 010101 (2009).
[15] D. Burgarth, K. Maruyama, and F. Nori, Phys. Rev. A **79**, 020305 (2009).
[16] D. Burgarth and K. Maruyama, New J. Phys. **11**, 103019 (2009).
[17] C. Di Franco, M. Paternostro, and M. S. Kim, Phys. Rev. Lett. **102**, 187203 (2009).
[18] D. Burgarth, K. Maruyama, and F. Nori, New J. Phys. **13**, 013019 (2011).
[19] E. J. Candés and M. B. Wakin, IEEE Signal Proccess Mag. **25**, 21 (2008).
[20] O. Katza, Y. Bromberg, and Y. Silberberg, Appl. Phys. Lett. **95**, 131110 (2009).
[21] W. Gong and S. Han, e-print arXiv:0910.4823v1 (2009).
[22] D. Gross, Y. K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).
[23] 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).
[24] The common definition of RIP is $(1-\delta_{s/2})||h||_{l_2}^2 \leqslant ||\Phi h||_{l_2}^2 \leqslant (1+\delta_{s/2})||h||_{l_2}^2$ for an s-sparse h [19], which is equivalent to the definition (5) in the paper.
[25] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, Constructive Approx. **28**, 253 (2008).
[26] E. J. Candés, C. R. Acad. Sci., Ser. I: Math. **346**, 589 (2008).
[27] G. D. Mahan, *Many-Particle Physics* (Springer, New York, 2000).
[28] B. P. Lanyon *et al.*, Nat. Chem. **2**, 106 (2009).
[29] A. Mizel and D. A. Lidar, Phys. Rev. B **70**, 115310 (2004).
[30] J. K. Pachos and E. Rico, Phys. Rev. A **70**, 053620 (2004).
[31] D. Gelman, C. P. Kochb, and R. Kosloffc, J. Chem. Phys. **121**, 661 (2004).
[32] M. Cramer, M. B. Plenio, S. T. Flammia, D. Gross, S. D. Bartlett, R. Somma, O. Landon-Cardinal, Y. K. Liu, and D. Poulin, Nat. Commun. **1**, 149 (2010).
[33] R. C. Bialczak *et al.*, Nature Phys. **6**, 409 (2010).
[34] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).
[35] B. E. Kane, Nature (London) **393**, 133 (1998).
[36] R. Vrijen, E. Yablonovitch, K. Wang, H. W. Jiang, A. Balandin, V. Roychowdhury, T. Mor, and D. DiVincenzo, Phys. Rev. A **62**, 012306 (2000).
[37] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. **83**, 4204 (1999).
[38] J. Watrous, Quantum Inf. Comput. **5**, 58 (2005).
[39] E. J. Candes, M. Wakin, and S. Boyd, J. Fourier Anal. Appl. **14**, 877 (2007).