Two-qubit Hamiltonian tomography by Bayesian analysis of noisy data

Sonia G. Schirmer[1,3,*] and Daniel K. L. Oi[2,3]

[1]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

[2]*Department of Physics, SUPA, University of Strathclyde, Glasgow G4 0NG, United Kingdom*

[3]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore*

We present an empirical strategy to determine the Hamiltonian dynamics of a two-qubit system using only initialization and measurement in a single fixed basis. Signal parameters are estimated from measurement data using Bayesian methods from which the underlying Hamiltonian is reconstructed, up to three unobservable phase factors. We extend the method to achieve full control Hamiltonian tomography for controllable systems via a multistep approach. The technique is demonstrated and evaluated by analyzing data from simulated experiments including projection noise.

PACS number(s): 03.67.Lx, 03.65.Wj

I. INTRODUCTION

Using quantum phenomena to perform new modes of computation is a daunting challenge [1]. Significant achievements in the theory of quantum computation include the development of error correction, fault tolerance [2], and scalability of quantum circuits [3]. However, in order to build large-scale quantum processors, many individual quantum systems must be manipulated with extraordinary precision and accuracy. A prerequisite for this level of quantum control is the precise characterization of the underlying dynamics and its response to control fields and what is sometimes referred to as Hamiltonian engineering (see [4,5] and references therein). This is especially crucial for manufactured devices such as solid-state quantum bits (qubits), e.g., quantum dots (Fig. 1) or superconducting quantum interference devices. Any manufacturing process will introduce variations, so it is important to empirically identify the control relationship for each component. In a large-scale quantum computer, it is desirable to be able to achieve this using *in situ* resources, i.e., initialization, control actuators, and measurement capabilities already present for performing computation.

The canonical method for assessing quantum dynamics is quantum process tomography (QPT) [7–9]. This involves initialization of a quantum system in a (complete) set of states, allowing it to evolve under the dynamics under consideration, and then performing an informationally complete measurement on the output state for each input. From this set of input-output data, the superoperator, or the completely positive map, governing the quantum evolution of the system can be reconstructed. This may then be repeated for different evolution times to obtain an estimate of the Lindblad operators (generators of the dynamics) [10]. For control purposes, QPT would be performed for a variety of actuator settings to build up a map of the control space.

A potential disadvantage of QPT is the need for *ab initio* initialization and measurement outside of the computational

basis, a capability which may not exist in the absence of characterization in the first place. It is usually argued that initialization and measurement in an arbitrary basis can be achieved by unitary rotation of a fixed basis; however, this presupposes that the system response to control fields has already been characterized, a vicious circle. Previous work has addressed this issue for the case of a single qubit subjected to multiple control Hamiltonians, decoherence, and imperfect subspace confinement [11–16]. Following similar ideas of using limited tools and accessibility to extract as much information as possible about a quantum system, the response of a single qubit to external driving forces has been used to extract information about the qubit's environment and the qubit-environment coupling mechanism [17]. Here, we extend the basic idea of Hamiltonian characterization to two coupled qubits with an unknown generic internal Hamiltonian and control Hamiltonian response.
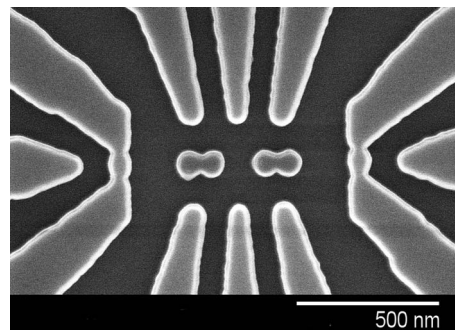

FIG. 1. Manufactured qubit system. A pair of horizontally aligned double quantum dots (center) can act as two qubits. A qubit can be defined in each double quantum dot by two different charging states, e.g., a single excess electron located on the left or the right dot of each pair. Electrodes (top and bottom) control the potentials and electron tunneling rates. Single-electron transistors (left and right) measure the locations of the excess electrons which defines the measurement basis or logical states of the qubit. Due to finite manufacturing precision, the placement of the control and the measurement structures may not be exactly as calculated; hence, the Hamiltonian dependence upon control signals will have to be determined empirically [6].

*sgs29@cam.ac.uk

This paper is organized as follows. In Sec. II we discuss the basic principles of Hamiltonian tomography for a two-qubit system with a fixed but unknown Hamiltonian, assuming only the ability to measure the system at specific times in a fixed measurement basis, but no control or *a priori* knowledge of the system. We also deliberately exclude the ability to perform local operators on either qubit or the ability to initialize the system in states other than the measurement basis states. Our approach differs in this regard from related works on two-qubit Hamiltonian identification using concurrence spectroscopy [18,19] or optimal experiment design [20]. These approaches may be preferable for certain types of systems but have some limitations as they presume that the single-qubit dynamics can be fully characterized independently of the interqubit coupling, which is required to prepare the two-qubit system in superposition states by applying local rotations. Using concurrence also limits us to reconstructing the nonlocal part of the two-qubit Hamiltonian. Thus, this approach may be well suited for some systems e.g., with weak-coupling and nonlocal interaction Hamiltonians, but may be problematic for other systems. The approach in this paper should be seen as complementary to these works.

In Sec. III we discuss how to extract the relevant system parameters from the noisy measurement data accurately and robustly. The difficulty of this task is greatly magnified compared to the single-qubit case due to the number of parameters to be determined, as well as the increased signal complexity. A naive approach using straightforward least-squares error minimization failed completely when applied to noisy data from simulated experiments. The power spectrum of the signal (which was sufficient for the single-qubit case) is still useful, but no longer an optimal frequency estimator in the presence of multiple frequencies, and obtaining accurate estimates of the amplitudes of different frequency components is very difficult. For these reasons Bayesian analysis is employed to determine the signal parameters, which is shown to result in significant improvements in the accuracy and the robustness of the procedure.

In Sec. IV we show how to reconstruct the total Hamiltonian or, more precisely, its matrix representation with respect to the fixed measurement basis, from the estimated parameters. Unlike the single-qubit case, calculating the 16 matrix elements of the two-qubit Hamiltonian from the 214 parameters estimated from the 16 measured signals is nontrivial and requires several optimization steps, from identifying the most likely level structure from the set of transition frequencies to determining the magnitudes and phases of the Hamiltonian matrix elements that provide the best fit with the estimated parameters. The analysis also shows that the fixed Hamiltonian can be determined only up to a global phase and sign, as well as three phases, which define U(1) transformations of the measurement basis states. If there are no other measurements or control available then these U(1) transformations of the basis states have no observable effect. Modulo these unobservable parameters, we demonstrate that we can reconstruct the overall Hamiltonian with a very good accuracy from noisy data.

In Sec. V we consider the more general case of control Hamiltonian tomography. In particular, we are interested in characterizing Hamiltonians $H=H(\mathbf{f})$ that depend on a number of external parameters $\mathbf{f}=(f_1,\ldots,f_M)$ that can be varied experimentally, such as voltages applied to certain gate electrodes that allow us to vary confinement potentials, tunneling rates, etc. By varying these parameters over time, we can engineer complicated effective Hamiltonians and efficiently achieve a wide range of control tasks from quantum state preparation to gate implementation [21] using powerful optimal control techniques [22,23]. However, effective control requires knowledge of the dependence of the Hamiltonian on these parameters $H(\mathbf{f})$. When applying different Hamiltonians, the previously unobservable phase factors now have practical effects and are critical for full control Hamiltonian tomography. We show how to determine these phases, relative to a reference Hamiltonian, using a simple two-step experiment, and how to use this information to achieve full control Hamiltonian tomography. Finally, in Sec. VI we discuss applications of the results, as well as future improvements and generalizations to our method.

II. FIXED HAMILTONIAN TOMOGRAPHY

Throughout this paper we assume that we are given a two-qubit system with an unknown Hamiltonian and a measurement apparatus that enables us to perform a fixed projective measurement on each qubit, including the ability to perform effectively simultaneous measurements on both qubits [24]. We denote the measurement basis states of the resulting four-outcome measurement by $|1\rangle=|00\rangle$, $|2\rangle=|01\rangle$, $|3\rangle=|10\rangle$, and $|4\rangle=|11\rangle$. We then perform the following simple experiment:

(1) Initialize the system in one of the four measurement basis states $|k\rangle$ by performing simultaneous measurements on both qubits.

(2) Let the system evolve for some time t.

(3) Perform simultaneous measurements on both qubits, projecting the system back into one of the four measurement basis states.

By repeating this experiment many times for a fixed evolution time t, we can estimate the probabilities $p_{k\ell}(t)=|\langle\ell|\Psi_k(t)\rangle|^2$, where $|\Psi_k(t)\rangle$ is the time-evolved state and $|\Psi_k(0)\rangle=|k\rangle$. By further repeating the experiment for different times t_n for $n=0,\ldots,N-1$, we can stroboscopically capture the evolution of the probabilities $p_{k\ell}(t)$ for $k,\ell=1,2,3,4$, yielding 16 noisy signals as shown in Fig. 2 [25]. What information about the Hamiltonian can we extract from this data, and what is the most effective way to extract this information?

Assume that the evolution of the system is governed by a fixed Hamiltonian according to the Schrödinger equation

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = H|\Psi(t)\rangle. \tag{1}$$

Expanding the Hamiltonian H with respect to its orthonormal eigenbasis $\{|\xi_\nu\rangle:\nu=1,\ldots,4\}$,
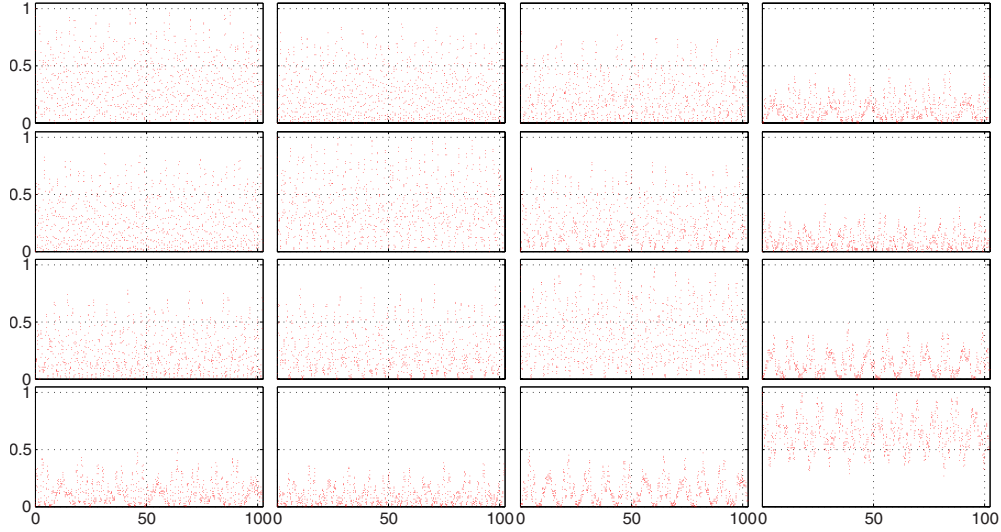
FIG. 2. (Color online) Simulated measurements of system 1 with $2^{10}+1$ data points per trace sampled at $\Delta t = 0.1$, signal length $T = 102.4$ (arbitrary units), number of experiment repetitions per data point $N_e = 250$. Each graph is the probability $p_{k\ell}(t)$ at time t of detecting the system in state $|\ell\rangle$ ($\ell = 1, 2, 3, 4$ left to right) if initialized in state $|k\rangle$ ($k = 1, 2, 3, 4$ top to bottom).

$$H = \sum_{\nu=1}^{4} \lambda_\nu |\xi_\nu\rangle\langle\xi_\nu|, \qquad (2)$$

where λ_ν are the (real) eigenvalues, and setting $\langle k|\xi_\nu\rangle = r_{k\nu} e^{i\phi_{k\nu}}$, we obtain

$$\langle\ell|H|k\rangle = \sum_{\nu=1}^{4} \lambda_\nu \langle\ell|\xi_\nu\rangle\langle\xi_\nu|k\rangle = \sum_{\nu=1}^{4} \lambda_\nu r_{\ell\nu} r_{k\nu} e^{i(\phi_{k\nu} - \phi_{\ell\nu})}. \quad (3)$$

Further defining $s_{k\ell;\nu} = r_{k\nu} r_{\ell\nu}$, $\delta_{k\ell;\nu} = \phi_{k\nu} - \phi_{\ell\nu}$, and $\Delta_{k\ell;\mu\nu} = -\delta_{k\ell;\mu} + \delta_{k\ell;\nu}$ we obtain

$$\langle\ell|H|k\rangle = \sum_{\nu=1}^{4} \lambda_\nu s_{k\ell;\nu} e^{i\delta_{k\ell;\nu}} = e^{i\delta_{k\ell;1}} \sum_{\nu=1}^{4} \lambda_\nu s_{k\ell;\nu} e^{i\Delta_{k\ell;1\nu}}, \quad (4)$$

where the phase terms satisfy $\delta_{\ell k;1} = -\delta_{k\ell;1}$ and

$$\delta_{23;1} = \delta_{13;1} - \delta_{12;1}, \qquad (5a)$$

$$\delta_{24;1} = \delta_{14;1} - \delta_{12;1}, \qquad (5b)$$

$$\delta_{34;1} = \delta_{14;1} - \delta_{13;1}. \qquad (5c)$$

If the system is initialized in one of the measurement basis states $|\Psi_k(0)\rangle = |k\rangle$, its time-evolved state $|\Psi_k(t)\rangle$ under the action of H is given by

$$|\Psi_k(t)\rangle = \sum_{\nu=1}^{4} e^{-i\lambda_\nu t} |\xi_\nu\rangle\langle\xi_\nu|k\rangle \qquad (6)$$

and since $\langle\xi_\nu|\xi_\mu\rangle = \delta_{\nu\mu}$, its projection onto the measurement basis state $|\ell\rangle$ at time t is

$$\langle\ell|\Psi_k(t)\rangle = \sum_{\nu=1}^{4} e^{-i\lambda_\nu t} \langle\ell|\xi_\nu\rangle\langle\xi_\nu|k\rangle = \sum_{\nu=1}^{4} s_{k\ell;\nu} e^{-i(\lambda_\nu t - \delta_{k\ell;\nu})}. \qquad (7)$$

Hence, the probability $p_{k\ell}(t) = |\langle\ell|\Psi_k(t)\rangle|^2$ of the outcome $|\ell\rangle$ for a projective measurement of $|\Psi_k(t)\rangle$ is

$$\left[\sum_{\nu=1}^{4} s_{k\ell;\nu} e^{-i(\lambda_\nu t - \delta_{k\ell;\nu})} \right] \left[\sum_{\mu=1}^{4} s_{k\ell;\mu} e^{i(\lambda_\mu t - \delta_{k\ell;\mu})} \right]$$

$$= \sum_{\mu=1}^{4} s_{k\ell;\mu}^2 + 2 \sum_{\nu > \mu} s_{k\ell;\mu} s_{k\ell;\nu} \cos(\omega_{\mu\nu} t - \Delta_{k\ell;\mu\nu}),$$

where $\omega_{\mu\nu} = \lambda_\nu - \lambda_\mu$, and using $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$,

$$p_{k\ell}(t) = c_{k\ell} + 2 \sum_{\nu > \mu} a_{k\ell;\mu\nu} \cos(\omega_{\mu\nu} t) + b_{k\ell;\mu\nu} \sin(\omega_{\mu\nu} t),$$

$$(8)$$

where the coefficients are

$$a_{k\ell;\mu\nu} = s_{k\ell;\nu} s_{k\ell;\mu} \cos(\Delta_{k\ell;\mu\nu}), \qquad (9a)$$

$$b_{k\ell;\mu\nu} = s_{k\ell;\nu} s_{k\ell;\nu} \sin(\Delta_{k\ell;\mu\nu}), \qquad (9b)$$

$$c_{k\ell} = \sum_\nu s_{k\ell;\nu}^2. \qquad (9c)$$

Equations (9) show that the observed dynamics are completely determined by the transition frequencies $\omega_{\mu\nu}$; the phase differences $\Delta_{k\ell;\mu\nu}$; and the (real) coefficients $a_{k\ell;\mu\nu}$, $b_{k\ell;\mu\nu}$, and $c_{k\ell}$, from which we can reconstruct the Hamiltonian $\tilde{H}$ defined by

$$\langle \ell | \widetilde{H} | k \rangle = \sum_{\nu=1}^{4} \widetilde{\lambda}_\nu s_{k\ell;\nu} e^{i\Delta_{k\ell;1\nu}}, \tag{10}$$

where $\widetilde{\lambda}_\nu = \omega_{1\nu} - \frac{1}{4}(\omega_{12} + \omega_{13} + \omega_{14})$, which is related to the actual Hamiltonian H by

$$H = D^\dagger \widetilde{H} D + \text{const } \mathbb{I}, \tag{11}$$

where $D = \text{diag}(1, e^{i\delta_{12}}, e^{i\delta_{13}}, e^{i\delta_{14}})$. The last term is simply a global energy shift, which has no observable consequences in general. The diagonal operator D represents the U(1) degree of freedom for redefining each measurement basis state. With only a single constant Hamiltonian, and preparation and measurement in a single fixed basis only, we cannot completely determine the Hamiltonian.

III. PARAMETER ESTIMATION

The first task is to analyze the measurement traces (8) to extract the t signal parameters (9), and the frequencies $\omega_{\mu,\nu}$. For convenience we label the transition frequencies of the system ω_m for $m = 1, \ldots, 6$, assuming $\omega_{m+1} > \omega_m > 0$, and define the vectors $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_6)$, $\mathbf{a}_{k\ell} = (a_{k\ell;m})$, and $\mathbf{b}_{k\ell} = (b_{k\ell;m})$ for $k, \ell = 0, 1, 2, 3$. The first step toward identifying the Hamiltonian $\widetilde{H}$ is to extract the six transition frequencies $\boldsymbol{\omega}$ and 13 linear coefficients $\mathbf{a}_{k\ell}$, $\mathbf{b}_{k\ell;m}$, and $c_{k\ell}$ for each of the 16 signals. Although there are $6 + 13 \times 16 = 214$ parameters, the problem would be relatively simple if $p_{k\ell}(t)$ was known with infinite precision for a set of sample times t_n. In practice, the accuracy with which the $p_{k\ell}(t_n)$ can be determined is limited by noise. In our case the acuracy is fundamentally limited by projection noise due to the finite number of measurement repetitions N_e. This renders the problem one of parameter estimation for a harmonic signal with multiple frequencies and phases from noisy data. Problems of this type are common in engineering from acoustics to image processing, and many techniques have been developed, but our parameter estimation problem is nontrivial due to the large number of parameters involved.

According to Eq. (9) the traces $p_{k\ell}(t)$ should be linear combinations of the 13 basis functions $g_{2m-1}(t) = \cos(\omega_m t)$, $g_{2m}(t) = \sin(\omega_m t)$ for $m = 1, \ldots, 6$, and $g_{13}(t) = 1$, i.e.,

$$p_{k\ell}(t_n) = \sum_{m=1}^{6} a_{k\ell,m} g_{2m-1}(t) + b_{k\ell,m} g_{2m}(t) + c_{k\ell} \tag{12}$$

and our objective is to find parameters ω_m, $a_{k\ell;m}$, $b_{k\ell;m}$, and $c_{k\ell}$ that maximize the likelihood of the measured data. Setting $\mathbf{d}_{k\ell} = (d_{k\ell;1}, \ldots, d_{k\ell;N})$, where $d_{k\ell;n}$ denotes the *approximate* value of $p_{k\ell}(t_n)$ derived from the measurement data; one way to proceed is to try to fit the parameters to minimize the squared L^2 norm of the error

$$\sum_{k,\ell} \|\mathbf{e}_{k\ell}\|_2^2 = \sum_{k,\ell} \|\mathbf{p}_{k\ell} - \mathbf{d}_{k\ell}\|_2^2, \tag{13}$$

where $\mathbf{p}_{k\ell} = (p_{k\ell;1}, \ldots, p_{k\ell;N})$ with $p_{k\ell;n} = p_{k\ell}(t_n)$ and $\|\mathbf{e}\|_2^2 = \Sigma_{n=1}^{N} e_n^2$ as usual. However, for problems with a large number of noisy data points and a large number of parameters, as in our case, finding a solution close to the (unknown) global

minimum of the error using brute-force optimization over all system parameters at once is difficult at best. We tested this strategy and in most cases achieved only poor results.

Instead of minimizing the global error, we can alternatively try to maximize the related likelihood function

$$L(\mathbf{a}_{k\ell}, \mathbf{b}_{k\ell}, c_{k\ell}, \boldsymbol{\omega}, \sigma) = \prod_{k,\ell=1}^{4} \sigma_{k\ell}^{-N} \exp\left[-\frac{\|\mathbf{p}_{k\ell} - \mathbf{d}_{k\ell}\|_2^2}{2\sigma_{k\ell}^2} \right]. \tag{14}$$

Note that we have implicitly assumed here that the signals $p_{k\ell}(t)$ are independent and subjected to Gaussian white noise with variance $\sigma_{k\ell}^2$, assumptions that are not strictly valid in our case. Hermitian symmetry of the Hamiltonian requires $\mathbf{a}_{k\ell} = \mathbf{a}_{\ell k}$ and $\mathbf{b}_{k\ell} = -\mathbf{b}_{\ell k}$ but we will enforce this symmetry later by averaging the estimated coefficients

$$\mathbf{a}_{k\ell} \mapsto \tfrac{1}{2}(\mathbf{a}_{k\ell} + \mathbf{a}_{\ell k}), \tag{15a}$$

$$\mathbf{b}_{k\ell} \mapsto \tfrac{1}{2}(\mathbf{b}_{k\ell} - \mathbf{b}_{\ell k}). \tag{15b}$$

The Gaussian noise model is not strictly valid; if the measurements are projection-noise limited, then a Poissonian error model would be more accurate, but we shall see that this is nonetheless a good approximation.

The main advantage of the latter formulation is that we can eliminate the explicit dependence on the linear coefficients $\mathbf{a}_{k\ell}$, $\mathbf{b}_{k\ell}$, $c_{k\ell}$ and the noise variances $\sigma_{k\ell}$ by integration over suitable priors to obtain an explicit expression for the probability of a particular model given the observed data $\mathbf{d}_{k\ell}$ that depend only on the six transition frequencies $\boldsymbol{\omega}$, rather than the >200 parameters in the full model. Following the standard Bayesian analysis, [26] we obtain

$$P(\boldsymbol{\omega}|\mathbf{d}) \propto \prod_{k,\ell=1}^{4} \left[1 - \frac{13\langle \mathbf{h}_{k\ell}^2 \rangle}{N\langle \mathbf{d}_{k\ell}^2 \rangle} \right]^{(13-N)/2}, \tag{16}$$

where the averages are defined by

$$\langle \mathbf{d}_{k\ell}^2 \rangle = \frac{1}{N} \sum_{n=1}^{N} d_{k\ell;n}^2, \tag{17a}$$

$$\langle \mathbf{h}_{k\ell}^2 \rangle = \frac{1}{13} \sum_{m=1}^{13} h_{k\ell;m}^2. \tag{17b}$$

The components $h_{k\ell;m}$ are essentially the orthogonal projections of the data onto a set of orthonormal basis vectors $H_m(t_n)$

$$h_{k\ell;m} = \sum_{n=1}^{N} H_m(t_n) d_{k\ell;n}. \tag{18}$$

The orthonormal basis vectors are derived from the (nonorthogonal) basis functions $g_m(t)$ defined above, evaluated at the respective sample times t_n, via

$$H_m(t_n) = \frac{1}{\sqrt{\alpha_m}} \sum_{m'=1}^{13} e_{m'm} g_{m'}(t_n), \tag{19}$$

where $e_{m'm}$ is a 13×13 matrix whose columns $\mathbf{e}_m$ are the normalized eigenvectors $G\mathbf{e}_m = \alpha_m \mathbf{e}_m$ of the 13×13 matrix $G = (G_{m_1 m_2})$ with

$$G_{m_1 m_2} = \sum_{n=1}^{N} g_{m_1}(t_n) g_{m_2}(t_n). \tag{20}$$

The objective is to find $\boldsymbol{\omega}$ that maximizes $P(\boldsymbol{\omega}|\mathbf{d}_{k\ell})$ or, equivalently, the logarithmic-likelihood function

$$\log_{10} P(\boldsymbol{\omega}|\mathbf{d}_{k\ell}) = \frac{13 - N}{2} \sum_{k,\ell=1}^{4} \log_{10}\left[1 - \frac{13\langle \mathbf{h}_{k\ell}^2 \rangle}{N\langle \mathbf{d}_{k\ell}^2 \rangle}\right]. \tag{21}$$

Note that N and $\langle \mathbf{d}_{k\ell} \rangle$ are constants, while $\mathbf{h}_{k\ell}$ indirectly depends on $\boldsymbol{\omega}$ via the basis functions $g_m(t)$. It can be shown that the corresponding optimal coefficients are

$$\mathbf{a}_{k\ell} = (\langle x_{k\ell;1} \rangle, \langle x_{k\ell;3} \rangle, \ldots, \langle x_{k\ell;11} \rangle), \tag{22a}$$

$$\mathbf{b}_{k\ell} = (\langle x_{k\ell;2} \rangle, \langle x_{k\ell;4} \rangle, \ldots, \langle x_{k\ell;12} \rangle), \tag{22b}$$

$$c_{k\ell} = \langle x_{k\ell;13} \rangle, \tag{22c}$$

where $\langle x_{k\ell;m} \rangle$ is the shorthand notation for the expectation values $E(x_{k\ell;m}|\boldsymbol{\omega}, \mathbf{d}_{k\ell})$ of the linear coefficients of the basis functions, given the optimal frequencies $\boldsymbol{\omega}$ and the data $\mathbf{d}_{k\ell}$. Furthermore [26],

$$\langle x_{k\ell;m} \rangle = \sum_{m'=1}^{13} \frac{e_{mm'} h_{k\ell;m'}}{\sqrt{\alpha_{m'}}}. \tag{23}$$

We can similarly derive expressions for second moments

$$\langle x_{k\ell;m_1} x_{k\ell;m_2} \rangle - \langle x_{k\ell;m_1} \rangle \langle x_{k\ell;m_2} \rangle = \sigma_{k\ell}^2 \sum_{m'=1}^{13} \frac{e_{m_1 m'} e_{m_2 m'}}{\alpha_{m'}}, \tag{24}$$

where $\sigma_{k\ell}^2$ is the noise variance of the (k,ℓ)th signal, which can be approximated by its estimated expectation

$$\langle \sigma_{k\ell}^2 \rangle = \frac{1}{N-15}[N\langle \mathbf{d}_{k\ell} \rangle - 13\langle \mathbf{h}_{k\ell} \rangle]. \tag{25}$$

Note that for $m_1 = m_2$, Eq. (24) is simply the variance of the parameter $x_{k\ell;m}$, which gives an estimate of the uncertainty $\Delta x_{k\ell;m}$ of the coefficient $x_{k\ell;m}$,

$$\Delta x_{k\ell;m}^2 \approx \mathrm{var}(x_{k\ell;m}) = \langle \sigma_{k\ell}^2 \rangle \sum_{m'=1}^{13} \frac{e_{mm'}^2}{\alpha_{m'}}. \tag{26}$$

Figure 3 shows that for a sufficiently large number of data points N and experiment repetitions per data point, N_e, these uncertainties can be made very small indeed. For N and/or N_e small, the uncertainties are much larger, but simulations for our specific problem suggest that the estimated values are generally still very close to the actual values even for small
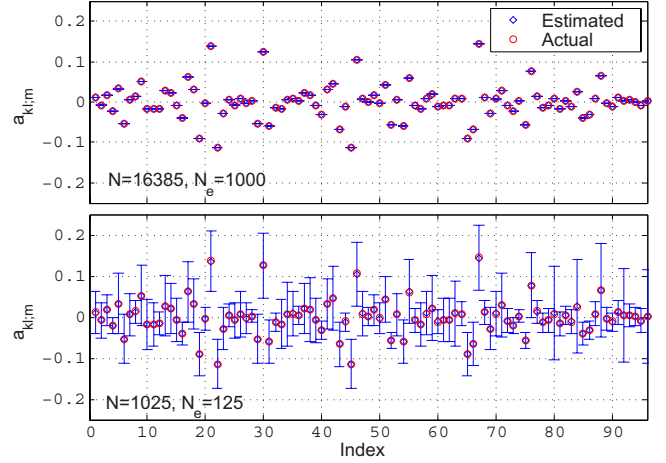

FIG. 3. (Color online) Estimated and actual values of the coefficients $a_{k\ell;m}$ with estimated error bars of system 1. When both N and N_e are large (top) the error bars are nearly invisible, and the estimated and the actual values are almost indistinguishable. When N and N_e are both small (bottom), the error bars are significantly larger, mainly due to increased noise variances $\sigma_{k\ell}^2$; yet the actual and the estimated values for the coefficients are still almost indistinguishable. This suggests that the estimated coefficients are in fact much more accurate than the uncertainty estimates suggest.

N and/or N_e, much closer than the uncertainty estimates would suggest.

Although the logarithmic-likelihood function (21) depends explicitly only on the six frequencies $\boldsymbol{\omega} \in \mathbb{R}^6$ rather than the full 214 model parameters, finding its (global) maximum is not trivial as the logarithmic likelihood is sharply peaked with many local extrema, and thus computationally efficient gradient-based optimization algorithms are likely to get trapped in local extrema if the starting point $\boldsymbol{\omega}_0$ is chosen randomly. An alternative is to use global search algorithms such as pattern search or genetic algorithms, but these are computationally expensive and the results for our problem proved inaccurate. To circumvent this problem, we adopt a combination strategy.

We can first estimate the resonant frequencies by looking for peaks in the power spectra

$$C_{k\ell}(\omega) = \left| \frac{1}{N} \sum_{n=1}^{N} d_{k\ell;n} e^{i\omega t_n} \right|^2. \tag{27}$$

Using spectral filtering combined with a basic peak finding routine, we locate (up to) six peaks ω_m in the combined power spectrum

$$C(\omega) = \sum_{k,\ell=1}^{4} C_{k\ell}(\omega) \tag{28}$$

as illustrated in Fig. 4, which are then used as input $\boldsymbol{\omega}^{(0)} = (\omega_1, \ldots, \omega_6)$ to an optimization routine based on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method with cubic line search [27–30] to find the maximum of the logarithmic likelihood (21). Although the discrete Fourier transform is *not* an optimal frequency estimator for a signal with multiple frequencies, it proved generally effec-
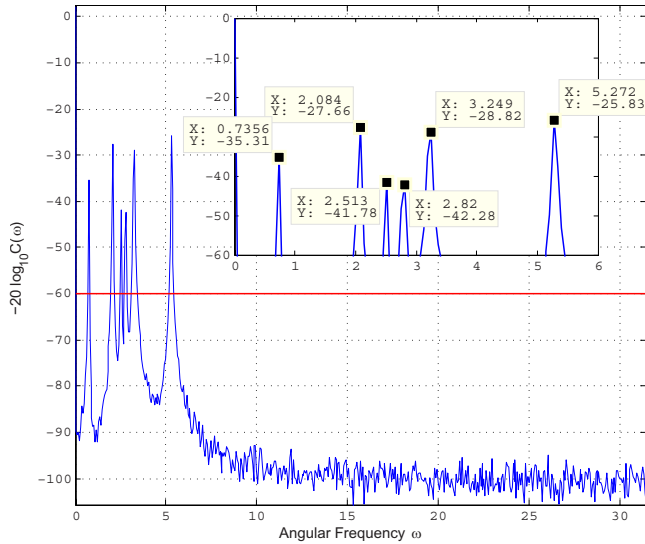
FIG. 4. (Color online) Power spectrum $C(\omega)$ of system 1. Although the power spectrum is noisy, the logarithmic plot of $C(\omega)$ of the measured signals shown in Fig. 2 shows six well-defined peaks for $\omega_m > 0$ in addition to the peak at $\omega = 0$. The inset shows the filtered power spectrum $C(\omega) > C_0$, from which the six peaks ω_m can easily be identified using standard peak detection.

tive in providing good starting values for the logarithmic-likelihood optimization routine, provided that the total sampling time (signal length) T was sufficiently long to resolve the resonant peaks.

Since the frequency resolution of the power spectrum is limited by the signal length T, $\Delta \omega = \frac{\pi}{T}$, if there are two or more closely spaced transition frequencies then it may not be possible to resolve six peaks in the power spectrum without increasing the signal lengths significantly. But this is generally not necessary as we can improve the frequency reso-
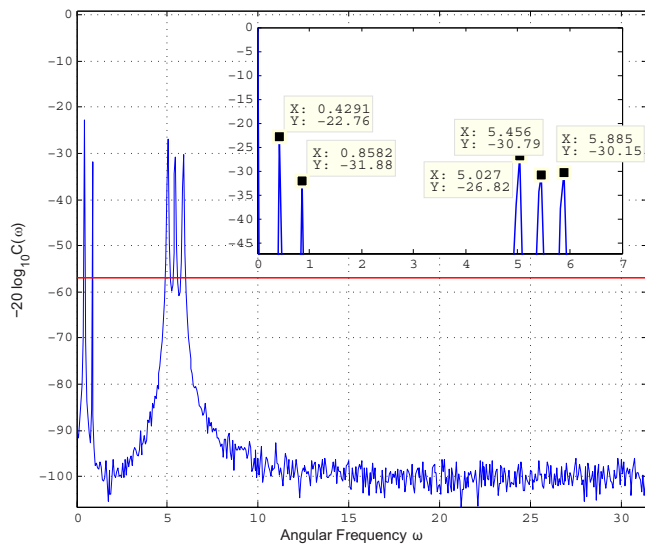

FIG. 5. (Color online) The power spectrum of system 12. The power spectrum $C(\omega)$ has only five peaks $\omega_m > 0$ in addition to the peak at $\omega = 0$. This could mean that the system has only five distinct transition frequencies, or that the measured signals are not sufficient to resolve two (closely spaced) transition frequencies.
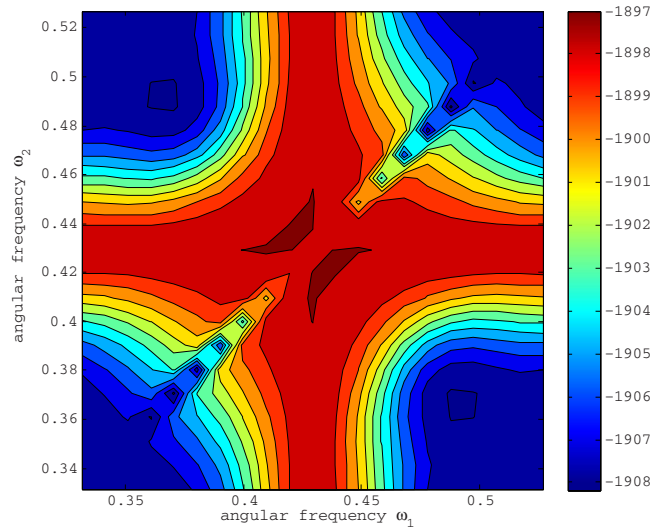

FIG. 6. (Color online) Logarithmic likelihood on $I_1 \times I_1$ for system 12 with five-peak power spectrum shown in Fig. 5 shows symmetry about $y = x$ as $\log_{10} P[(\omega_1, \omega_2, \ldots)|\mathbf{d}_{k\ell}] = \log_{10} P[(\omega_2, \omega_1, \ldots)|\mathbf{d}_{k\ell}]$ with twin peaks for $y \neq x$ indicating that the most probable model on this subspace of the parameter space is a six-frequency model.

lution as follows. Suppose there are five identifiable peaks, $\omega_1 - \omega_5$, in the power spectrum, as shown in the example in Fig. 5. Then we proceed as before, using the five-peak frequencies in the power spectrum as input $\boldsymbol{\omega}^{(0)}$ for the optimization routine to find the most likely five-frequency model $\boldsymbol{\omega}^{(*)}$. To ascertain whether there is a more probable six-frequency model, we choose an interval I_m about each $\omega_m^{(*)}$, $m = 1, \ldots, 5$, and investigate the logarithmic-likelihood function (21) on the two-dimensional (2D) parameter space $I_m \times I_m$, keeping the other four frequencies fixed in each case. For example, for $m = 1$ in the example above we find the maximum of $\log_{10} P(\boldsymbol{\omega}|\mathbf{d}_{k\ell})$ for $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_2^{(*)}, \omega_3^{(*)}, \omega_4^{(*)}, \omega_5^{(*)})$ with $(\omega_1, \omega_2) \in I_1^2$ and $I_1 = [\omega_1^{(1)} - \frac{10}{T}, \omega_1^{(1)} + \frac{10}{T}]^2$ by calculating $\log_{10} P$ on a coarse 2D grid, finding the maximum on the grid, and using the resulting $\boldsymbol{\omega}$ as a starting point for the BFGS optimization routine as before. A contour plot
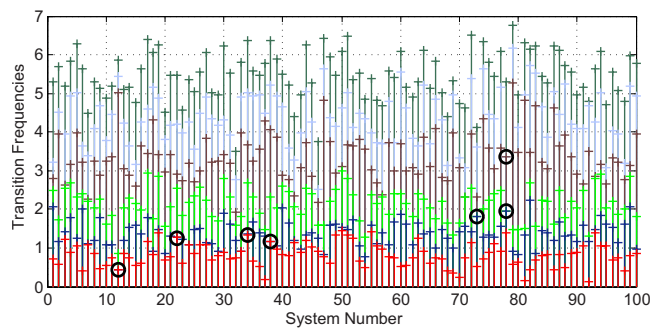

FIG. 7. (Color online) The transition frequency diagram for each of the 100 test systems shows that the transition frequencies range from 0.3 to 7, and there are six systems (12, 22, 34, 38, 73, and 78) with two transition frequencies that differ by less than 0.01 (circled), which are difficult to resolve, including one system (78) with two such cases.

TABLE I. Logarithmic likelihood for different five- and six-frequency models and actual transition frequencies for system 12 with five-peak power spectrum shown in Fig. 5.

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	$\log_{10} P$
$\boldsymbol{\omega}^{(0)}$	0.4293	0.8586	4.9983	5.4276	5.8569		924.4486
$\boldsymbol{\omega}^{(*)}$	0.4291	0.8558	5.0046	5.4282	5.8604		938.2960
$\boldsymbol{\omega}^{(1)}$	0.4235	0.4323	0.8558	5.0046	5.4282	5.8604	943.3509
$\boldsymbol{\omega}^{(2)}$	0.4291	0.7631	0.8558	5.0046	5.4282	5.8604	938.3099
$\boldsymbol{\omega}^{(3)}$	0.4291	0.8558	5.0046	5.1023	5.4282	5.8604	938.2977
$\boldsymbol{\omega}^{(4)}$	0.4291	0.8558	5.0046	5.4282	5.5063	5.8604	938.2993
$\boldsymbol{\omega}^{(5)}$	0.4291	0.8558	5.0046	5.4282	5.8604	5.9287	938.2975
$\boldsymbol{\omega}_{\mathrm{act}}$	0.4236	0.4322	0.8558	5.0046	5.4282	5.8604	

showing the maxima in the logarithmic likelihood on $I_1 \times I_1$ is shown in Fig. 6.

We repeat this procedure for each m in turn. The results, summarized in Table I, show that the six-frequency model $\boldsymbol{\omega}^{(1)}$ is most likely, more than the five-frequency model, and the other five six-frequency models. Indeed, the frequencies of the most likely six-frequency model are very close to the actual transition frequencies of the system simulated. However, the relative flatness of the peak corresponding to the global maximum of the logarithmic-likelihood function and the relatively small differences between the likelihoods of the most likely model and the less likely models suggest that more data would be desirable to improve the resolution of the parameter estimates, and our confidence that the model is indeed the correct choice. If there are fewer than five peaks in the power spectrum, the procedure described can be iterated to sequentially resolve peaks in the power spectrum until the most probable model has been found.

To test the effectiveness and the accuracy of this parameter estimation technique, we test the method for 100 randomly generated Hamiltonians, sampled at $\Delta t = 0.1$ (arbitrary units) for different signal lengths $T = (N-1)\Delta t = 0.1(2^d)$ for $d = 10, 11, 12, 13, 14$ and different levels of projection noise, with the number of measurements per data point, $N_e \in \{125, 250, 500, 1000\}$. The test Hamiltonians have transition frequencies in the range of $[0.3, 7]$ and include cases with very closely spaced transition frequencies, as shown in Fig. 7. To assess the quality of the models found, we calculate the transition frequencies ω_m and the corresponding parameters $\mathbf{a}_{k\ell}$, $\mathbf{b}_{k\ell}$, and $c_{k\ell}$ for each Hamiltonian and consider the relative errors of the parameters identified from the noisy data with the parameter estimation technique described.

Tables II and III show the means and the medians, respectively, over 100 systems, of the maximum relative error (in percent)

$$\epsilon_{\max}(\boldsymbol{\omega}^{(0)}) = 100 \max_{m \in 1, \ldots, 6} \left| 1 - \frac{\omega_m^{(0)}}{\omega_m} \right| \tag{29}$$

of the estimated transition frequencies for each system, where ω_m are the exact transition frequencies. Comparison of the errors for the initial frequency estimates obtained from the power spectrum, labeled $\boldsymbol{\omega}^{(0)}$, and the optimal values $\boldsymbol{\omega}^{\mathrm{opt}}$ obtained by maximizing the logarithmic likelihood shows that the optimized frequencies are generally about two orders of magnitude more accurate than the estimates obtained from the power spectrum.

The linear coefficients $a_{k\ell;m}$, $b_{k\ell;m}$, and $c_{k\ell}$ are then estimated from the maximization of Eq. (21) and from Eq. (22). Taking the median of the relative errors (again, in percent)

$$\epsilon_{\mathrm{med}}(a_{k\ell;m}) = 100 \underset{k, \ell, m}{\mathrm{median}} \left| 1 - \frac{a_{k\ell;m}^{\mathrm{est}}}{a_{k\ell;m}} \right|, \tag{30}$$

where k, ℓ range from 1 to 4 and $m = 1, \ldots, 6$, as a general measure of the quality of the fit, Table IV shows that the average errors in the coefficients $a_{k\ell;m}$, $b_{k\ell;m}$, and to a lesser extent $c_{k\ell}$, are generally at least one order of magnitude larger than the error in the frequency estimates. Overall, the quality is still good, however, with the (average) errors ranging from a fraction of a percent to less than 2.5% for $a_{k\ell}$, and much less for $c_{k\ell}$, depending on the number of data points N and the accuracy of the data points determined by the number of experiment repetitions per data point, N_e. Figure 8 shows the distribution of the errors for both the least and the

TABLE II. The percentage relative errors $\langle \epsilon_{\max}(\boldsymbol{\omega}^{(0)}) \rangle$ (left) and $\langle \epsilon_{\max}(\boldsymbol{\omega}^{\mathrm{opt}}) \rangle$ (right) show that the logarithmic-likelihood optimization improves the accuracy of the frequency estimates by at least two orders of magnitude compared to the estimates obtained from the power spectrum.

$N \backslash N_e$	125	250	500	1000	125	250	500	1000
16385	0.093	0.094	0.094	0.094	0.0002	0.0002	0.0001	0.0001
8193	0.231	0.226	0.231	0.231	0.0006	0.0006	0.0004	0.0003
4097	0.432	0.432	0.432	0.432	0.0018	0.0019	0.0009	0.0009
2049	0.696	0.685	0.680	0.685	0.0065	0.0040	0.0030	0.0024
1025	1.646	1.650	1.646	1.650	0.0272	0.0184	0.0085	0.0108

TABLE III. The medians of percentage relative errors $\epsilon_{\max}(\boldsymbol{\omega}^0)$ (left) and $\epsilon_{\max}(\boldsymbol{\omega}^{\mathrm{opt}})$ (right) show the same accuracy improvements of the logarithmic-likelihood estimates. Median errors lower than the averages indicate that the error distribution is peaked toward the origin.

$N\backslash N_e$	125	250	500	1000	125	250	500	1000
16385	0.068	0.068	0.068	0.068	0.0001	0.0001	0.0001	0.0001
8193	0.167	0.164	0.167	0.167	0.0005	0.0003	0.0002	0.0002
4097	0.327	0.327	0.327	0.327	0.0011	0.0012	0.0006	0.0005
2049	0.5100	0.493	0.493	0.493	0.0035	0.0023	0.0019	0.0011
1025	1.164	1.142	1.164	1.142	0.0126	0.0089	0.0052	0.0036

greatest numbers of experiments. Apart from a few outliers, the distribution follows a roughly exponential form with most estimates being within a fraction of a percent of the true values, even for the least number of experimental samples.

Table II shows that increasing N_e and thus the accuracy of the data points does not improve the accuracy of the initial frequency estimates obtained from the power spectrum at all, while doubling N tends to reduce the error by more than half. This is what we expect as once N_e is large enough to permit discrimination of the resonant peaks from the noise floor, little is gained by increasing N_e. Doubling N_e does reduce the error for the optimized frequencies obtained from our Bayesian analysis, although if the accuracy of frequency estimates alone is considered, doubling the number of data points is preferable to doubling N_e. Increasing the accuracy (by doubling N_e) is more effective in reducing the errors in the coefficients a, b, c, but the contour plots in Fig. 9 show that the errors decrease faster with N, i.e., increasing the number of data points is generally still preferable.

IV. HAMILTONIAN RECONSTRUCTION

Once the frequencies $\boldsymbol{\omega}$ and amplitudes $\mathbf{a}_{k\ell}$, $\mathbf{b}_{k\ell}$, and $c_{k\ell}$ have been extracted from the measured data using parameter estimation, reconstructing the Hamiltonian (up to equivalence) requires at least two further steps: identification of the

TABLE IV. Relative errors $\epsilon_{\mathrm{med}}(a_{k\ell;m})$, $\epsilon_{\mathrm{med}}(b_{k\ell;m})$, and $\epsilon_{\mathrm{med}}(c_{k\ell})$ (in %) and estimated error variances $\langle \sigma^2 \rangle$, averaged over 100 test systems for different signal lengths $T=0.1(N-1)$ and numbers of experiment repetitions N_e per data point.

|	$N\backslash N_e$	125	250	500	1000
$\langle \epsilon_{\mathrm{med}}(a_{k\ell;m}) \rangle$	16385	0.3825	0.2671	0.1912	0.1454
	8193	0.5538	0.3598	0.2857	0.1923
	4097	0.7711	0.5516	0.4075	0.2786
	2049	1.0630	0.7940	0.5755	0.3762
	1025	1.5817	1.1210	0.7880	0.5573
$\langle \epsilon_{\mathrm{med}}(b_{k\ell;m}) \rangle$	16385	0.2417	0.1739	0.1174	0.0846
	8193	0.3333	0.2519	0.1755	0.1144
	4097	0.4860	0.3470	0.2394	0.1733
	2049	0.6715	0.5098	0.3436	0.2485
	1025	1.0194	0.7197	0.4691	0.3523
$\langle \epsilon_{\mathrm{med}}(c_{k\ell}) \rangle$	16385	0.0734	0.0525	0.0378	0.0279
	8193	0.1002	0.0751	0.0538	0.0372
	4097	0.1463	0.1037	0.0770	0.0518
	2049	0.2007	0.1483	0.1148	0.0751
	1025	0.2817	0.2258	0.1555	0.1047
$\langle \sigma^2 \rangle$	16385	0.0012	0.0006	0.0003	0.0001
	8193	0.0971	0.0959	0.0953	0.0950
	4097	0.2896	0.2873	0.2861	0.2855
	2049	0.6763	0.6717	0.6692	0.6681
	1025	1.4580	1.4487	1.4437	1.4414
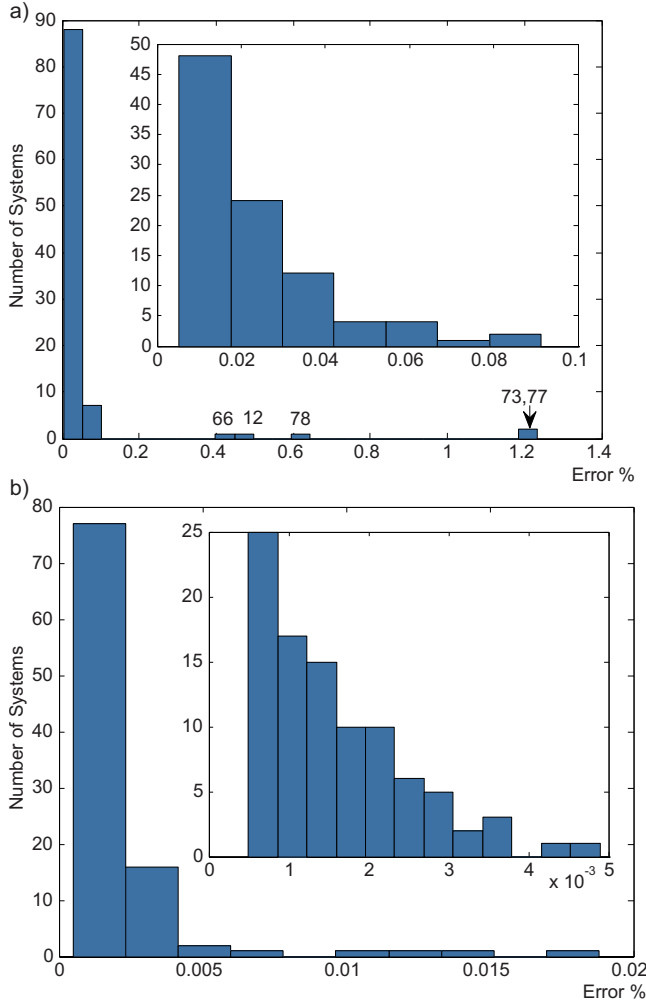
FIG. 8. (Color online) Histogram of the relative % error for the test 100 systems for sampling numbers of (a) $N=1025$, $N_e=125$ and (b) $N=16\,385$, $N_e=1000$. Inset graphs magnify the region around the origin showing the general distribution of errors, which is roughly exponential. The numbers 66, 12, 78, 73, and 77 in (a) refer to outlier systems.

resonant frequencies with transitions (μ, ν) between eigenstates $|\xi_\mu\rangle$ and $|\xi_\nu\rangle$ of the system and computation of the parameters $s_{k\ell;\nu}$ and $\Delta_{k\ell;\mu\nu}$ in Eq. (10) from the coefficients $\mathbf{a}_{k\ell}$, $\mathbf{b}_{k\ell}$, and $c_{k\ell}$. For a four-level system we have three primary transitions $\{\omega_{12}, \omega_{23}, \omega_{34}\}$ between adjacent energy levels and three other transitions $\{\omega_{13}, \omega_{24}, \omega_{14}\}$, which must satisfy

$$\omega_{13} = \omega_{12} + \omega_{23}, \tag{31a}$$

$$\omega_{24} = \omega_{23} + \omega_{34}, \tag{31b}$$

$$\omega_{14} = \omega_{12} + \omega_{23} + \omega_{34}. \tag{31c}$$

We identify the possible level structure (up to inversion) by examining the relationships between the frequencies. In the generic case, i.e., when there are six distinct transition frequencies, $0 < \omega_1 < \omega_2 < \cdots < \omega_6$, it follows immediately from Eqs. (31) that $\omega_6 = \omega_{14}$, and the primary transitions are
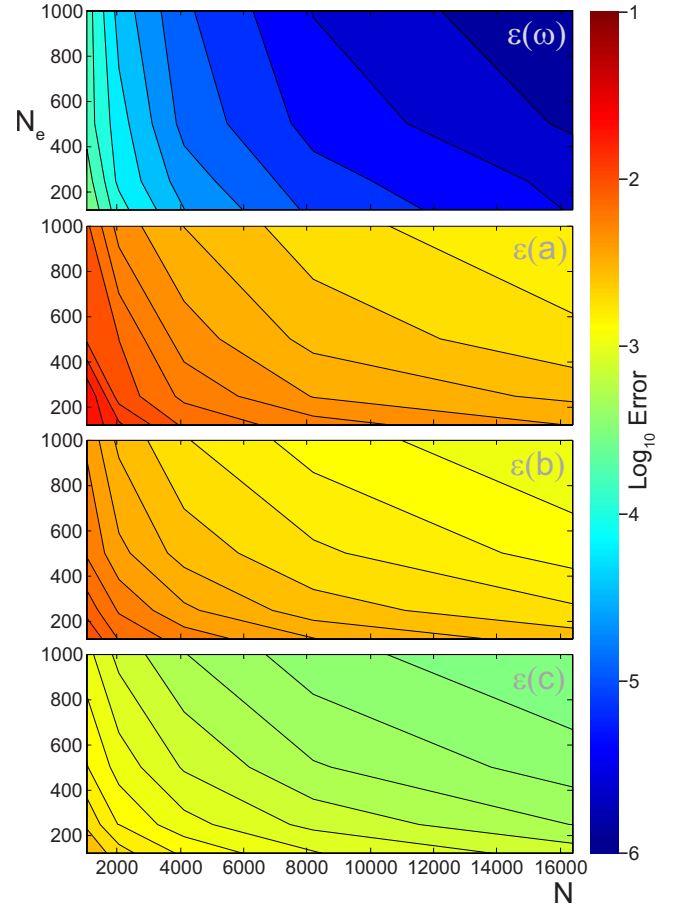

FIG. 9. (Color online) Contour plots of the $\log_{10}$ mean (relative) errors for the frequencies ω and coefficients a, b, and c. The frequencies show the smallest errors (down to 10^{-6}), while the a coefficients show errors up to a few percent ($\sim 10^{-2}$) for the shortest signal lengths and greatest projection noise.

$\{\omega_1, \omega_2, \omega_6 - \omega_1 - \omega_2\}$. Closer inspection shows that there are ten possible arrangements of the six transition frequencies as shown in Fig. 10, and the exact transition frequencies $\boldsymbol{\omega}$ must satisfy $A_s \boldsymbol{\omega} = \mathbf{0}$ for one of the following matrices:

$$A_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 \end{bmatrix}, \quad A_2$$

$$= \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}, \quad A_4$$

$$= \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & -1 & 0 \end{bmatrix},$$
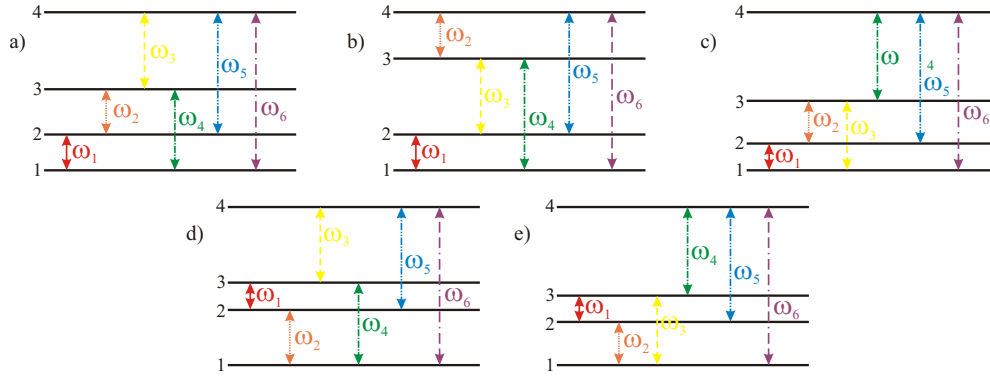
FIG. 10. (Color online) Possible arrangements for a generic four-level system with six distinct transition frequencies. Not shown are the other five configurations that correspond to a reflection of the energies, which merely flips the above level structures.

$$A_5 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 \end{bmatrix}. \quad (32)$$

Given the estimated frequencies $\boldsymbol{\omega}^{opt}$ the most likely case is that for which $\|A_s \boldsymbol{\omega}^{opt}\|_2^2$ assumes its minimum, which should be close to zero, and significantly smaller than the errors for the other cases. A larger minimum error indicates and none of the possibilities is likely, suggesting that the system may not be a Hamiltonian four-level system. Similarly, if we have two cases for which the error is close to the minimum, this would be an indication that further data are required to resolve the ambiguity.

Once the observed frequencies ω_m have been matched with actual transitions (μ, ν), we can associate the corresponding coefficients $a_{k\ell,m}$, $b_{k\ell,m}$ for $m = 1, \ldots, 6$ with their respective transitions, i.e., we have $a_{k\ell;\nu\ell}$ and $b_{k\ell;\nu\ell}$, and determine the phase differences

$$\Delta_{k\ell;\mu\nu} = \arctan(b_{k\ell;\mu\nu}, a_{k\ell;\mu\nu}), \quad (33)$$

where $\arctan(b, a)$ is the four-quadrant arctangent of b/a. If the estimated parameters are good, then the resulting $\Delta_{k\ell;\mu\nu}$ should satisfy $\Delta_{kk;\mu\nu} \approx 0 \pmod{2\pi}$, $\Delta_{k\ell;\mu\nu} \approx -\Delta_{k\ell;\mu\nu} \pmod{2\pi}$, and

$$\Delta_{k\ell;12} + \Delta_{k\ell;13} - \Delta_{k\ell;23} = 0 \pmod{2\pi}, \quad (34a)$$

$$\Delta_{k\ell;13} + \Delta_{k\ell;14} - \Delta_{k\ell;34} = 0 \pmod{2\pi}, \quad (34b)$$

$$\Delta_{k\ell;12} + \Delta_{k\ell;14} - \Delta_{k\ell;24} = 0 \pmod{2\pi}. \quad (34c)$$

Due to the enforced symmetrization (15) of the coefficients $\mathbf{a}_{k\ell}$ and $\mathbf{b}_{k\ell}$, the phase terms should satisfy $\Delta_{k\ell;\mu\nu} = -\Delta_{\ell k;\mu\nu}$. Minor violations of Eqs. (34) are to be expected, and can be mitigated, and the accuracy of the final reconstructed Hamiltonian improved by minimizing the constraint violations $\|\mathbf{e}_{k\ell}\|_2^2 = \Sigma_{s=1}^3 e_{k\ell;s}^2$, where

$$e_{k\ell;s} = \min\{|x_{k\ell;s}|, |x_{k\ell;s} - 2\pi|, |x_{k\ell;s} + 2\pi|\}, \quad (35)$$

with $\mathbf{x}_{k\ell} = A \boldsymbol{\Delta}_{k\ell}$ and

$$A = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 0 & -1 \end{bmatrix} \quad (36)$$

for $k, \ell = 1, \ldots, 4$ in a further refinement step, starting with the values for $\Delta_{k\ell;\mu\nu}$ obtained from Eq. (33). This refinement tries to minimize the discrepancy between the estimated signal parameters and those expected from an underlying Hamiltonian model. It must be stressed, however, that larger violations of the constraints are indicative of significant errors, which may even be exacerbated by such a refinement. In fact, Fig. 11 shows that there is a strong correlation between the maximum constraint violations prior to refinement

$$\mathcal{E}(\Delta_{k\ell;\mu\nu}) = \max_{k,\ell} \|\mathbf{e}_{k\ell}\|_2^2 \quad (37)$$

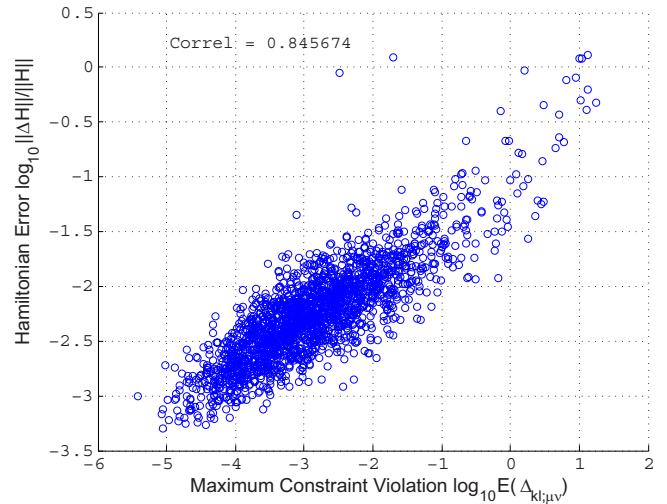and the relative error of the final estimated Hamiltonian.

FIG. 11. (Color online) A scatter plot of the relative errors $\|\Delta H\|/\|H\|$ of the estimated Hamiltonian with $\|\Delta H\|$ as defined in Eq. (41) vs the maximum constraint violation (37) on a log-log scale for our 100 systems and 20 data sets per system (total of 2000 data points) shows a strong correlation, suggesting that the maximum constraint violation (prior to refinement) is a good predictor of the accuracy of the estimated Hamiltonian.

TABLE V. Relative error $\mathcal{E}(H) = 100\|H^{\text{est}} - H\|/\|H\|$ of reconstructed Hamiltonian (with phase corrections) in %. Each table entry consists of three numbers: the median error (in %) and the number of systems (of 100) with relative error exceeding 1% and 5%, respectively.

	125			250			500			1000		
$N\backslash N_e$	$\mathcal{E}(H)$	1%	5%	$\mathcal{E}(H)$	1%	5%	$\mathcal{E}(H)$	1%	5%	$\mathcal{E}(H)$	1%	5%
16385	0.40	11	1	0.27	5	0	0.18	2	0	0.13	4	0
8193	0.57	22	0	0.41	8	0	0.31	8	1	0.19	4	0
4097	0.87	41	5	0.66	25	2	0.41	15	1	0.28	7	1
2049	1.12	60	7	0.91	45	6	0.58	19	4	0.44	12	2
1025	1.81	81	13	1.32	64	8	0.84	34	5	0.63	31	4

Once the optimal values for $\Delta_{k\ell;\mu\nu}$ have been found, we calculate the products

$$s_{k\ell;\mu}s_{k\ell;\nu} = a_{k\ell;\mu\nu}\cos(\Delta_{k\ell;\mu\nu}) + b_{k\ell;\mu\nu}\sin(\Delta_{k\ell;\mu\nu}). \quad (38)$$

Labeling the right-hand side of the previous equation $M_{k\ell;\mu\nu}$ and defining the column vector $\mathbf{s}_{k\ell}$ and the 4×4 matrix $M_{k\ell}$

$$\mathbf{s}_{k\ell} = \begin{pmatrix} s_{k\ell;1} \\ s_{k\ell;2} \\ s_{k\ell;3} \\ s_{k\ell;4} \end{pmatrix}, \quad M_{k\ell} = \begin{pmatrix} M_{k\ell;11} & \cdots & M_{k\ell;14} \\ \vdots & \ddots & \vdots \\ M_{k\ell;41} & \cdots & M_{k\ell;44} \end{pmatrix},$$

we can express Eqs. (9c) and (38) as follows:

$$\mathbf{s}_{k\ell}\mathbf{s}_{k\ell}^T = M_{k\ell}, \quad \mathbf{s}_{k\ell}^T\mathbf{s}_{k\ell} = c_{k\ell} \quad (39)$$

for $k, \ell = 1, \dots, 4$. To reconstruct the Hamiltonian (10), we must determine the coefficients $s_{k\ell;\nu}$ by solving Eqs. (39).

Each $M_{k\ell}$ is a real symmetric matrix whose off-diagonal elements $M_{k\ell;\mu\nu}$ ($\mu \neq \nu$) are determined by Eq. (38). The diagonal elements $M_{k\ell;\mu\mu}$ are unknown. However, we know that $M_{k\ell}$ should be a projector onto the one-dimensional space spanned by $\mathbf{s}_{k\ell}$, and the second equation in Eqs. (39) determines the norm of $\mathbf{s}_{k\ell}$ as well as the vector of diagonal elements $(M_{k\ell;\mu\mu})_{\mu=0}^{\mu=3}$. Thus, to determine the diagonal elements of $M_{k\ell}$ and the corresponding eigenvector $\mathbf{s}_{k\ell}$, we note that a rank-1 projector Π with matrix entries (g_{mn}) must satisfy the condition

$$dg_{mn} \equiv g_{mm}g_{nn} - g_{mn}^2 = 0, \quad \forall \; m, n. \quad (40)$$

Thus, given the off-diagonal elements of $M_{k\ell}$, we choose the diagonal elements of $M_{k\ell}$ such as to minimize the norm of the error $e = \Sigma_{m,n} dg_{mn}^2$ and take $\mathbf{s}_{k\ell}$ to be the eigenvector corresponding to the eigenvalue of $M_{k\ell}$ closest to 1, normalized to ensure $\|\mathbf{s}_{k\ell}\|_2^2 = c_{k\ell}$. It is important to carefully choose the parameters for the optimization here to ensure that we find the diagonal elements corresponding to the global minimum. Ideally, the residual error e should be 10^{-10} or less.

We implemented and tested the algorithm for our 100 Hamiltonians. We were able to correctly identify the level structures for all but one case: system 73, which has two nearly identical transition frequencies with $\omega_2 = 1.8012$ and $\omega_3 = 1.8026$, for $N = 1025$ data points sampled at $N_e = 125$, 250, and 500 experiment repetitions per data point. Even for this system, we were able to correctly identify the level structure by doubling the number of data points N, with the

exception of $N_e = 250$ where at least $N = 4097$ data points were needed. Of course, in practice more data points would be required for such a system to be confident that the identification is correct, as explained earlier.

To gauge the overall accuracy of the estimated Hamiltonians, we would like to compute the norm of the error $\|\Delta H\| = \|H^{\text{est}} - H^{\text{act}}\|$, or the relative error $\|\Delta H\|/\|H^{\text{act}}\|$, where we choose the operator norm here. However, calculating the norm of the error is complicated by the fact that we can only reconstruct the Hamiltonian up to the diagonal matrix D and the energy inversion symmetry. Thus, we must compensate for the phases that are "unobservable" in our model by setting

$$\|\Delta H\| = \|D^\dagger H^{\text{est}} D - H^{\text{act}}\| \quad (41)$$

with $D = \text{diag}(1, \delta_{12}, \delta_{13}, \delta_{14})$, where

$$\delta_{1\ell} = \text{phase}(H_{1\ell}^{\text{act}}) - \text{phase}(H_{1\ell}^{\text{est}}), \quad \ell = 2, 3, 4, \quad (42)$$

and phase(H_{1l}^{act}) is the complex phase of the $(1, l)$ matrix element of H^{act}, etc. Table V shows the results of the percentage relative errors $\|\Delta H\|/\|H^{\text{act}}\|$ for our 100 test systems, for different values of N and N_e. Medians of the relative errors range from 0.13% for $N = 16\,385$ and $N_e = 1000$ to 1.81% for $N = 1025$ and $N_e = 125$.

V. CONTROL HAMILTONIAN TOMOGRAPHY

We have seen that our procedure can characterize a single Hamiltonian up to a (physically irrelevant) global energy shift and three relative phases δ_{1n} for $n = 2, 3, 4$, due to the freedom to redefine each of the measurement basis vectors by a U(1) phase minus an overall phase. If we can only measure the system in a fixed basis and prepare it in the measurement basis states, and the evolution is determined by a single fixed Hamiltonian, then we have determined all observable parameters. However, for the system to be controllable, we require at least two (noncommuting) Hamiltonians, or more generally we must have the ability to modify the Hamiltonian by changing control parameters $\mathbf{f}$, e.g., by applying external fields or varying applied gate voltages, etc. In this case we can still choose the phases $\delta_{1n}^{(0)}$ for one "reference" Hamiltonian $H_0 = H(\mathbf{f}_0)$ as we wish, e.g., $\delta_{1n}^{(0)} = 0$ but the phases $\delta_{1n}^{(\mathbf{f})}$ for all other Hamiltonians $H(\mathbf{f})$ are now observable and thus relevant, and the complete control Hamiltonian reconstruction therefore requires that we identify them.
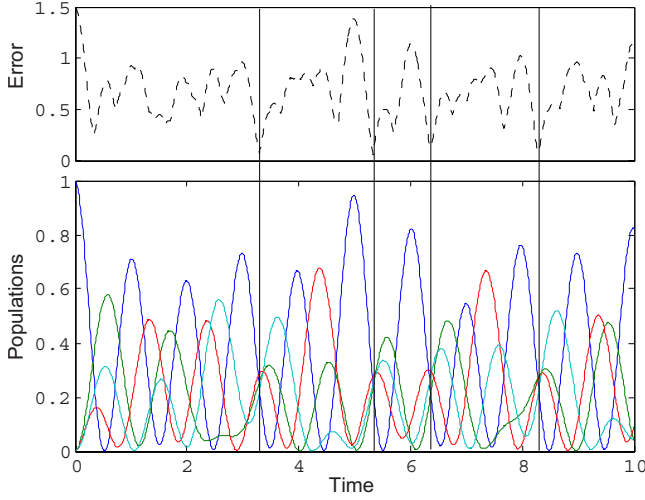
FIG. 12. (Color online) Evolution of populations $|\alpha_j(t)|$ under reference Hamiltonian (system 5) and error $\Sigma_j||\alpha_j|^2 - \frac{1}{2}|$ from the ideal balanced initial state. We selected the second minimum (which is the global minimum for $0 \leq t \leq 10$) at $t = 5.34$ as initial evolution time t_* for the δ estimation step.

FIG. 13. (Color online) Median relative error of estimated Hamiltonian after δ estimation as a function of signal lengths N and projection-noise level N_e of the measured traces $p_\ell(t_k)$. In all cases the most accurate estimates for $\tilde{H}$ from the previous step, i.e., $N = 16\,385$ and $N_e = 1000$, were used.

To achieve this, note that if can initialize the system in the superposition state $|\Phi\rangle = \Sigma_{j=1}^4 \alpha_j|j\rangle$ and measure the time-evolved state

$$|\Phi(t)\rangle = U_{\mathbf{f}}(t)|\Phi\rangle = D_{\mathbf{f}}^\dagger \tilde{U}_{\mathbf{f}}(t)D_{\mathbf{f}}|\Phi\rangle \qquad (43)$$

with $U_{\mathbf{f}}(t) = \exp[-itH(\mathbf{f})]$, $\tilde{U}_{\mathbf{f}}(t) = \exp[-it\tilde{H}(\mathbf{f})]$, then

$$p_\ell(t) = |\langle\ell|D_{\mathbf{f}}^\dagger \tilde{U}_{\mathbf{f}}(t)D_{\mathbf{f}}|\Phi\rangle|^2 = |\langle\ell|\tilde{U}_{\mathbf{f}}(t)D_{\mathbf{f}}|\Phi\rangle|^2 \qquad (44)$$

shows that the phases $\delta_{1n}^{\mathbf{f}}$ that determine $D_{\mathbf{f}} = \mathrm{diag}(1, e^{i\delta_{12}^{\mathbf{f}}}, e^{i\delta_{13}^{\mathbf{f}}}, e^{i\delta_{14}^{\mathbf{f}}})$ are now observable as $D_{\mathbf{f}}$ no longer commutes with the initial state $|\Phi\rangle$. As $\tilde{U}_{\mathbf{f}}(t)$ is fully determined by previous steps, if the initial state $|\Phi\rangle$ is known, then the only unknown parameters in Eq. (44) are δ_{1n}^B for $n = 2, 3, 4$. Given a set of measured values $d_{\ell k}$ for $p_\ell(t_k)$, we can determine the unknown parameters δ_{1n}^B by minimizing the least-squares error

$$\mathbf{e} = \sum_{\ell=1}^4 \|\mathbf{p}_\ell - \mathbf{d}_\ell\|_2^2, \qquad (45)$$

where $\mathbf{p}_\ell = (p_\ell(t_0), \ldots, p_\ell(t_K))$ and $\mathbf{d}_\ell = (d_{\ell 0}, \ldots, d_{\ell K})$ for $\ell = 1, 2, 3, 4$. An explicit expression for $p_\ell(t)$ derived in the Appendix shows that we can in principle determine all the phases if the initial state satisfies $\alpha_j \neq 0$ for all j [31]. Moreover, it is advantageous to choose a balanced initial state, $|\alpha_j|^2 \approx \frac{1}{4}$ for all j, if possible, to maximize signal to noise ratios.

To prepare such an initial state, we can use the reference Hamiltonian H_0. Unless the reference Hamiltonian is such that one or more of the measurement basis states are completely decoupled from state $|1\rangle$, it is almost certain that the time-evolved state $|\Phi_1(t)\rangle = U_0(t)|1\rangle = \Sigma_{j=1}^4 \alpha_j(t)|j\rangle$ with $U_0(t) = \exp(-itH_0)$ will satisfy $\alpha_j(t) \neq 0$ for all j for at least some $t > 0$. Thus, having characterized the Hamiltonians $H = H(\mathbf{f})$ for different control settings $\mathbf{f}$ up to the phases $\delta_{1n}^{\mathbf{f}}$, all we
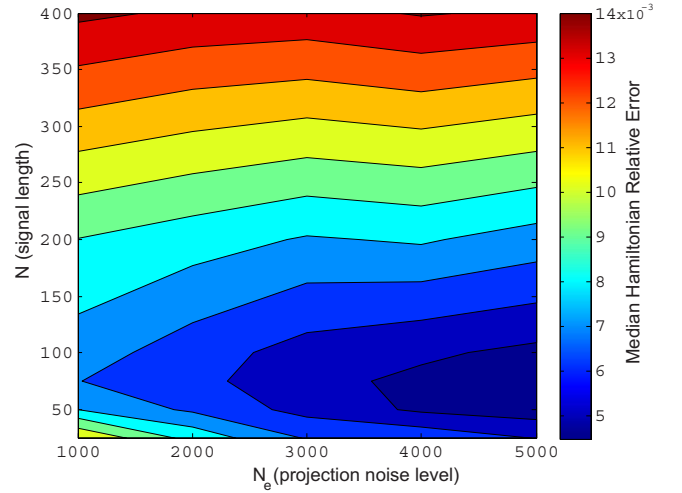
need to do is to select a suitable reference Hamiltonian $H_0 = H(\mathbf{f}_0)$ and find a time t_* such that the time-evolved state $|\Phi_1(t_*)\rangle$ satisfies $|\alpha_j(t_*)| \approx \frac{1}{2}$. This is generally not difficult. For instance, we randomly choose the Hamiltonian for test system 5 as our reference Hamiltonian. Figure 12 shows that there are several times $t \in [0, 10]$ at which the populations $|\alpha_j(t)|^2$ of all levels (in the measurement basis) are approximately equal. We pick one of these times $t_* = 5.34$, set $|\Phi\rangle = U_0(t_0)|1\rangle$, and obtain the measurement traces $p_\ell(t_k)$ as follows:

(1) Initialize system in measurement basis state $|1\rangle$.

(2) Let it evolve under Hamiltonian H_0 for time t_*.

(3) Change control settings to $\mathbf{f}$ and let system evolve for t time units under Hamiltonian $H_{\mathbf{f}}$.

(4) Perform measurement $\Rightarrow$ outcome $\ell = 1, 2, 3, 4$.

As before we repeat this experiment N_e times for a fixed t to estimate $p_\ell(t_k)$ (number of times the outcome was ℓ divided by N_e), and then repeat for different times t_k to obtain estimates for $p_\ell(t_k)$.

We tested the phase estimation procedure for the estimated Hamiltonians obtained in the previous step. For each of the 100 systems we first generated (simulated) measurement signals for $p_\ell(t_k)$ of varying length $T = (N-1)\Delta t$ and levels of projection noise N_e. The number of points ranged from $N - 1 = 25$ to 1000 data points, sampled at $\Delta t = 0.1$ fixed as before; the measurement repetitions N_e are from 1000 to 5000. In the reconstruction of the phases, we only assume that we know the estimated H_0, hence the estimated $|\Phi(0)\rangle$, and the estimated $H_{\mathbf{f}}$, determined in Sec. IV. While the most accurate estimates for the frequency and the linear coefficient estimation step (step 1) were obtained for the longest signals ($N = 16385$), we find that the accuracy of the phase estimation step peaks at around $N \approx 50$, and that longer signals are in fact highly detrimental (Fig. 13). This may seem very surprising at first but can be at least partly explained by the fact that even small inaccuracies in the initial estimates, especially for the frequencies, will accumulate over time and

TABLE VI. Relative error $\mathcal{E}(H) = 100\|H^{\text{est}} - H\| / \|H\|$ of reconstructed Hamiltonian with estimated phases δ_{1n} in % (no phase corrections). As before, each table entry consists of three numbers: the median error (in %) and the number of systems (of 100) with relative error exceeding 1% and 5%, respectively. The first row in each box are the estimates obtained for signals $p_\ell(t_k)$ of length $N=50$, while the second row are the estimates obtained for signals of length $N=200$, in both cases sampled at $N_e=5000$.

$N\backslash N_e$	125			250			500			1000		
	$\mathcal{E}(H)$	1%	5%	$\mathcal{E}(H)$	1%	5%	$\mathcal{E}(H)$	1%	5%	$\mathcal{E}(H)$	1%	5%
16385	0.89	40	2	0.70	27	1	0.56	14	1	0.45	6	1
	2.12	78	12	1.54	75	9	1.00	50	6	0.76	35	1
8193	1.09	54	5	0.84	35	0	0.69	28	1	0.53	17	1
	2.60	89	24	2.21	86	14	1.66	77	10	1.08	52	5
4097	1.48	68	7	1.12	58	5	0.91	43	5	0.61	26	2
	3.47	96	41	3.13	94	27	2.20	87	13	1.45	65	7
2049	2.24	88	15	1.45	74	7	1.12	55	7	0.78	37	4
	6.06	98	67	3.91	93	37	2.85	89	25	2.25	78	16
1025	3.14	95	29	2.44	90	18	1.60	80	8	1.22	59	6
	8.36	98	76	5.92	96	59	4.50	95	48	3.00	88	32

increase the discrepancy between the projected evolutions of the system based on our Hamiltonian estimates and the true evolution.

Based on these results we settled for signals of length $N-1=50$ with $N_e=5000$ measurement repetitions per data point for the final phase estimation step. For each of the 2000 estimated Hamiltonians $\tilde{H}$ obtained in the first step—corresponding to the 100 different test systems, as well as four levels of projection noise $N_e \in \{125, 250, 500, 1000\}$ and five signal length $T=(N-1)\Delta t$ for $N-1 \in \{2^{10}, 2^{11}, 2^{12}, 2^{13}, 2^{14}\}$ with $\Delta t = 0.1$ fixed each—we estimated the phases δ_{1n} and used the results to reconstruct the total Hamiltonian $H = D^\dagger \tilde{H} D$. Table VI shows the results in terms of the median of relative errors. For comparison, we include in Table VI the results for the phase estimation using $N=200$ signals. Comparison of the numbers clearly shows that longer signals are detrimental for the phase estimation step. In addition to a substantially decreased accuracy, longer signals also slowed down the numerical optimization, making it more difficult for the routine to find the global minimum. In view of the complicated dependence of $p_\ell(t_k)$ (see the Appendix) on the parameters δ_{1n} ($n=2,3,4$), we initially explored population-based (global) optimization strategies, especially evolutionary algorithms, but found that it was substantially slower and far *less* effective in finding the global minimum of error (45) than a gradient-based (BFGS-type) local optimization algorithm. In fact, for short signals the local optimization routine generally succeeded in finding the global minimum in a single run, starting with a random guess for $\boldsymbol{\delta} = (\delta_{12}, \delta_{13}, \delta_{14})$, although the optimization was repeated with several different initial guesses to increase the probability that we had indeed found the (globally) best value for $\boldsymbol{\delta}$.

VI. CONCLUDING DISCUSSION

We have presented a method for characterizing the Hamiltonian and its dependence on external control parameters, which is a prerequisite for Hamiltonian engineering and coherent control of the system's evolution, for a generic two-qubit system, assuming only the ability of preparation and measurement in a fixed basis. Analysis of simulated measurement data shows that the task of estimating the parameters from the complex noisy measurement signals with multiple frequencies, and reconstructing the Hamiltonian is very challenging, and requires a carefully designed multistep approach, combining spectral analysis, Bayesian analysis, and several carefully designed optimization steps to reconstruct the energy-level structure and matrix representation of the Hamiltonian. In the absence of any control, the Hamiltonian can only be reconstructed up to three phases due to the freedom to redefine the measurement basis by U(1) phase rotations. This symmetry can be broken if the system can be prepared initially in a suitable superposition state, and we exploit this fact to achieve full control Hamiltonian tomography in a simple two-step procedure.

The Bayesian analysis assumes a Gaussian noise profile which, although not strictly accurate, works well especially in the large N_e limit. Any significant deviations from Gaussian noise (e.g., Poissonian statistics for small N_e for $p \approx 0, 1$) will tend to make the logarithmic-likelihood estimates worse, and thus our estimates of the confidence that the model fits the data are conservative [26]. More accurate error estimates could be obtained using Bayesian analysis with a Poissonian noise model, although our results show that even a Gaussian noise model results in a huge improvement of two orders of magnitude or more in the accuracy of the frequency estimates compared to estimates obtained from simple spectral analysis. This turned out to be crucial for a successful Hamiltonian reconstruction. The frequency estimates obtained from the power spectrum combined with a simple least-squares error minimization to find the optimal spectral amplitudes proved to be too inaccurate for Hamiltonian reconstruction, leading to inconsistent equation systems and significant errors, and any attempt to obtain estimates of the parameters by direct minimization of the least-

squares error of the measurement signals and the expected signals resulted in reconstructed Hamiltonians that were little better than random for our test systems.

Although we have implicitly assumed a Hamiltonian model, i.e., that incoherent effects will be negligible on the time scales of interest, any significant deviation from the assumed model, e.g., significant decoherence or coupling to additional states outside the two-qubit subspace would result in low likelihoods of the chosen (four-level) Hamiltonian model. Such effects can easily be incorporated into the analysis by changing the basis functions, e.g., using damped exponentials instead of sinusoids or including additional states, which we will consider in further work. Furthermore, any prior information about the structure of the Hamiltonian can be incorporated to make the Bayesian analysis more efficient. Thus, the method lends itself to adaptive protocols, as we can adaptively sample the system until certain targets for the likelihood or error estimates are met, ensuring that we perform enough measurements to get accurate estimates but no more than necessary [32]. This is especially important as the number of measurements required will vary depending on the system. For instance, for a system with well-spaced transition frequencies, a sharply peaked likelihood function with a clearly identifiable global maximum can be obtained with much less data than for a system with two almost degenerate transition frequencies.

For control Hamiltonian tomography, the small but non-zero inaccuracies in the initial estimation step lead to an optimum sampling time for the second step due to divergence of the model from the true system behavior at longer times. In principle, it should be possible to use this divergence to improve the initial estimates of the Hamiltonians, and exploring such refinements could be an interesting avenue for future research. Errors in the second step decreased with an increased signal to noise ratio (increasing N_e), as the estimate of the phase parameters does not depend on the signal length, unlike the frequency resolution. It would also be interesting to investigate the accumulation of errors in this multistep estimation, especially how uncertainties in prior steps affect the accuracy of the Bayesian estimation in subsequent stages. Finally, in this paper we have dealt with the generic case. When the Hamiltonian has exact degeneracies, then the measurement signals will contain fewer than six frequencies. In this case, the level structure reconstruction becomes harder as the number of special subcases increases and we may not be able to uniquely identify the Hamiltonian. Although the set of Hamiltonians with exact degeneracies is

of measure zero, further study of these special cases may be of interest as one may want to specifically engineer Hamiltonians with such level structures.

APPENDIX: MEASURED PROBABILITIES

If the system is initialized in the generic superposition state $|\Phi\rangle = \Sigma_{j=1}^4 \alpha_j |j\rangle$ and measured after evolving for t time units under the Hamiltonian $H_B = D^\dagger \tilde{H}_B D$, then the general expression for the probability $p_\ell(t)$ of measurement outcome ℓ is

$$
\begin{aligned}
p_\ell(t) &= \left| \langle \ell | \sum_{\mu=1}^4 e^{-i\lambda_\mu t} |\xi_\mu\rangle \langle \xi_\mu | D | \Phi \rangle \right|^2 \\
&= \sum_{\mu,\nu=1}^4 \sum_{m,n=1}^4 \alpha_m \alpha_n^* e^{-i[\omega_{\mu\nu}t - (\delta_{1m}^B - \delta_{1n}^B)]} \langle \ell|\xi_\mu\rangle \langle \xi_\mu|m\rangle \langle n|\xi_\nu\rangle \\
&\times \langle \xi_\nu|\ell\rangle = \sum_{\mu,\nu=1}^4 \sum_{m,n=1}^4 |\alpha_m| \\
&\times |\alpha_n| s_{\ell m;\mu} s_{n\ell;\nu} e^{-i[\omega_{\mu\nu}t - (\delta_{1m}^B - \delta_{1n}^B) - (\phi_m - \phi_n) - \delta_{\ell m;\mu} - \delta_{n\ell;\nu}]} \\
&= \sum_{\mu=1}^4 \left\{ \sum_{m=1}^4 |\alpha_m|^2 s_{\ell m;\mu}^2 + \sum_{m>n} 2|\alpha_m| \right. \\
&\times |\alpha_n| s_{\ell m;\mu} s_{n\ell;\mu} \cos[(\delta_{1m}^B - \delta_{1n}^B) + (\phi_m - \phi_n) + \delta_{\ell m;\mu} \\
&\left. + \delta_{n\ell;\mu})] \right\} + \sum_{\mu>\nu} \left\{ \sum_{m=1}^4 2|\alpha_m|^2 s_{\ell m;\mu} s_{m\ell;\nu} \cos(\omega_{\mu\nu}t \right. \\
&- \delta_{\ell m;\mu} - \delta_{m\ell;\nu}) + \sum_{m\neq n} 2|\alpha_m||\alpha_n| s_{\ell m;\mu} s_{n\ell;\nu} \cos[\omega_{\mu\nu}t \\
&\left. - (\delta_{1m}^B - \delta_{1n}^B) - (\phi_m - \phi_n) - \delta_{\ell m;\mu} - \delta_{n\ell;\nu}] \right\}.
\end{aligned} \tag{A1}
$$

[1] A. K. Ekert, P. Hayden, H. Inamori, and D. K. L. Oi, Int. J. Mod. Phys. A **16**, 3335 (2001).

[2] *Quantum Error Correction and Fault Tolerant Quantum Computing*, edited by Frank Gaitan (CRC Press, Boca Raton, 2008).

[3] *Scalable Quantum Computers: Paving the Way to Realization*, edited by Samuel L. Braunstein and Hoi-Kwong Lo (Wiley, New York, 2001).

[4] S. G. Schirmer, in *Lagrangian and Hamiltonian Methods for Nonlinear Control*, edited by F. Bullo and K. Fujimoto, Lecture Notes in Control and Information Sciences (Springer, Berlin, 2006), Vol. 366, p. 293.

[5] S. G. Schirmer, G. Kandasamy, and S. J. Devitt, *IEEE Proceedings of the ISCCSP, Malta, 2008* (IEEE, St. Julians, Malta, 2006), p. 966.

[6] Image courtesy of Hitachi Cambridge Laboratories, Hitachi Europe Ltd.

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

[8] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).

[9] R. Kosut, I. Walmsley, and H. Rabitz, e-print arXiv:quant-ph/0411093.

[10] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).

[11] S. G. Schirmer, A. Kolli, and D. K. L. Oi, Phys. Rev. A **69**, 050306(R) (2004).

[12] S. G. Schirmer, A. Kolli, D. K. L. Oi, and J. H. Cole, *Quantum Communication, Measurement and Computing*, edited by S. M. Barnett, E. Andersson, J. Jeffers, P. Öhberg, and O. Hirota, AIP Conf. Proc. No. 734 (AIP, Melville, New York, 2004), p. 79–82.

[13] 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).

[14] J. H. Cole, A. D. Greentree, D. K. L. Oi, S. G. Schirmer, C. J. Wellard, and L. C. L. Hollenberg, Phys. Rev. A **73**, 062333 (2006).

[15] S. J. Devitt, S. G. Schirmer, D. K. L. Oi, J. H. Cole, and L. C. L. Hollenberg, New J. Phys. **9**, 384 (2007).

[16] S. G. Schirmer, D. K. L. Oi, and S. J. Devitt, J. Phys.: Conf. Ser. **107**, 012011 (2008).

[17] J. H. Cole, S. J. Devitt, and L. C. L. Hollenberg, New J. Phys. **8**, 103 (2006).

[18] J. H. Cole, S. J. Devitt, and L. C. L. Hollenberg, J. Phys. A **39**, 14649 (2006).

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

[20] K. Young, M. Sarovar, R. Kosut, and K. Birgitta Whaley, Phys. Rev. A **79**, 062301 (2009).

[21] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2005).

[22] S. G. Schirmer, J. Mod. Opt. **56**, 831 (2009).

[23] N. Khaneja *et al.*, J. Magn. Reson. **172**, 296 (2005).

[24] Formally, the method presented will work for any four-dimensional Hamiltonian system.

[25] We consider only noise due to finite sampling for each time t_n, which is Poissonian. In the large N_e limit, and for $p_{kl} \neq 0, 1$, this tends to Gaussian noise.

[26] G. Larry Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation* (Springer, Berlin, 1998).

[27] C. G. Broyden, J. Inst. Math. Appl. **6**, 76 (1970).

[28] R. Fletcher, Comput. J. **13**, 317 (1970).

[29] D. Goldfarb, Math. Comput. **24**, 23 (1970).

[30] D. F. Shanno, Math. Comput. **24**, 647 (1970).

[31] In this case, the small number of parameters to be estimated means that a simple least-squares fit performs adequately. A Bayesian analysis is made more difficult by the complicated dependence of the basis functions on the undetermined parameters, although this approach may lead to better estimates than the ones presented here, but at a cost of greater computational complexity. This is a topic for further research.

[32] Composite pulses or robust optimal control can be used to relax the need for extremely accurate Hamiltonian characterization.