Quantum System Identification: Hamiltonian Estimation using Spectral and Bayesian Analysis

Sophie G Schirmer*, Frank C Langbein†
* Department of Applied Maths & Theoretical Physics, Univ. of Cambridge,
Wilberforce Rd, Cambridge, CB3 0WA, UK; Email: sgs29@cam.ac.uk
† School of Computer Science, Cardiff University,
5 The Parade, Cardiff, CF24 3AA, UK; Email: F.C.Langbein@cs.cardiff.ac.uk

Abstract—**Identifying the Hamiltonian of a quantum system from experimental data is considered. General limits on the identifiability of model parameters with limited experimental resources are investigated, and a specific Bayesian estimation procedure is proposed and evaluated for a model system where a-priori information about the Hamiltonian's structure is available.**

I. INTRODUCTION

At a fundamental level nature is governed by the laws of quantum mechanics, but until recently such phenomena were mostly a curiosity studied by physicists. However, significant advances in theory and technology are increasingly pushing quantum phenomena into the realm of engineering, as building blocks for novel technologies and applications from chemistry to computing. E.g., advances in laser technology enable ever more sophisticated coherent control of atoms, molecules and other quantum systems. Recent advances in nanofabrication have made it possible to create nanostructures such as quantum dots and quantum wells that behave like artificial atoms or molecules and exhibit complex quantum behaviour. Cold-atom systems and the creation of Bose condensates demonstrate that even macroscopic systems can exhibit quantum coherence.

Harnessing the potential of quantum systems is a challenging task, requiring exquisite control of quantum effects and system designs that are robust with regard to fabrication imperfection, environmental noise and loss of coherence. Although significant progress has been made in designing effective controls, most control design is model-based, and available models for many systems do not fully capture their complexity. Model parameters are often at best approximately known and may vary, in particular for engineered systems subject to fabrication tolerances. Experimental system identification is therefore crucial for the success of quantum engineering. While there has been significant progress in quantum state identification and quantum process tomography, we require dynamic models if we wish to control a system's evolution. Furthermore, effective protocols must take into account limitations on measurement and control resources for initial device characterization. This presents many challenges, from determing how much information can be obtained in a given setting to effective and efficient protocols to extract this information. Here we illustrate some problems and solutions for the case of identifying the dynamics of a three-level system.

II. IDENTIFIABILITY OF MODEL PARAMETERS

One of the first questions to consider before attempting to find explicit protocols for experimental system identification is clearly what information we can hope to extract about a given system with a certain limited set of resources. For instance, given a system with a Hilbert space of dimension N, it is well known that the ability to prepare and measure the system in a set of computational basis states $\{|n\rangle : n = 1 \ldots, N\}$ is insufficient for quantum process tomography, even if the process is unitary [1], [2]. However, recent work shows that a substantial amount of information about the generators of the dynamics can be obtained for Hamiltonian [3]–[6] and even dissipative systems [7]–[10] at least generically, by mapping the evolution of the computational basis states stroboscopically over time. More precisely, this is done by determining the probabilities that a measurement of the observable $M = \mathrm{diag}(m_1, \ldots, m_N)$ produces the outcome m_ℓ after the system was initialized in the computational basis state $|k\rangle$ and allowed to evolve for time t for a number of different times t_n. This begs the question how much information we can hope to obtain in general from such experiments. In this paper we consider Hamiltonian systems, whose evolution is governed by the Schrodinger equation $i\hbar \dot{U}(t, t_0) = HU(t, t_0)$ with a fixed Hamiltonian H and $U(t_0, t_0) = \mathbb{I}$, for which we have $p_{k\ell}(t) = |\langle \ell | U(t, t_0) | k \rangle|^2$.

Theorem 1: Let H and M be Hermitian operators representing the Hamiltonian and the measurement, respectively, and let ρ_0 be a positive operator with $\mathrm{Tr}(\rho_0) = 1$ representing the initial state of the system. If M, H and ρ_0 are simultaneously blockdiagonalizable, i.e., there exists a decomposition of the Hilbert space $\mathcal{H} = \oplus_{s=1}^{S>1} \mathcal{H}_s$ such that

$$M = \mathrm{diag}(M_s), \quad H = \mathrm{diag}(H_s), \quad \rho_0 = \mathrm{diag}(\rho_s), \quad (1)$$

where M_s, H_s and ρ_s are operators on the Hilbert spaces $\mathcal{H}_s$, then we can at most identify H up to $\sum_s \lambda_s \mathbb{I}_s$, where $\mathbb{I}_s$ is the identity on the subspace $\mathcal{H}_s$.

Proof: If H is block-diagonal then any initial state ρ_0 starting in a subspace $\mathcal{H}_s$ must remain in this subspace. Thus, the dynamics on each subspace is independent, $U(t) = \otimes_s U_s(t)$ with $U_s(t) = e^{-itH_s}$. Per hypothesis M and ρ_0 are also blockdiagonal, so $\mathrm{Tr}[MU(t)\rho_0 U(t)^\dagger] = \sum_s \mathrm{Tr}[M_s U_s(t)\rho_s U_s(t)^\dagger]$. If $\tilde{H} = H + \sum_s \lambda_s \mathbb{I}_s$ then

$\tilde{U}(t) = \otimes_s \tilde{U}_s(t)$ with $\tilde{U}_s(t) = e^{-it\lambda_s}U_s(t)$. Thus, $\text{Tr}[M\tilde{U}(t)\rho_0\tilde{U}(t)^\dagger] = \sum_s \text{Tr}[M_s e^{-it\lambda_s}U_s(t)\rho_s e^{it\lambda_s}U_s(t)^\dagger] = \sum_s \text{Tr}[M_s U_s(t)\rho_s U_s(t)^\dagger]$ shows that H and $\tilde{H}$ are indistinguishable. ∎

Thus, there are some limitations on the maximum amount of information we can obtain about the system by initializing and measuring the system in a fixed computational basis. In particular, if H and M commute, we can infer that H and M are simultaneously diagonalizable, and assuming the eigenvalues m_ℓ of M are distinct, this fixes the Hamiltonian basis, i.e., we have $H = \sum_\ell \lambda_\ell \Pi_\ell$, where Π_ℓ is the projector on the eigenspace of M corresponding to m_ℓ, i.e., the computational basis state $|\ell\rangle$. However, no information about the eigenvalues λ_m or the transition frequencies $\omega_{k\ell} = \lambda_\ell - \lambda_k$ can be obtained by measuring $p_{k\ell}(t)$, all of which are constant in this case.

Maximum information about the Hamiltonian can be obtained if H and M are not simultaneously block-diagonalizable. This is the generic case, and in this case we can identify H at most up to a diagonal unitary matrix $D = (1, e^{i\phi_2}, \ldots, e^{i\phi_N})$ and a global energy shift $\lambda_0 \mathbb{I}$, i.e., $\tilde{H} \simeq H = D^\dagger \tilde{H} D + \lambda_0 \mathbb{I}$, as was noted in [6]. The term $\lambda_0 \mathbb{I}$ is generally physically insignificant as it gives rise only a global phase factor $\tilde{U}(t,0) = e^{-it(H+\lambda_0\mathbb{I})} = e^{-i\lambda_0 t}e^{-itH}e^{-i\lambda_0 t}U(t,0)$, which is generally unobservable, as the abelian phase factors cancel, $\rho(t) = U(t,0)\rho_0 U(t,0)^\dagger = \tilde{U}(t,0)\rho_0\tilde{U}(t,0)^\dagger$ for any ρ_0. The diagonal unitary matrix D represents the freedom to redefine the measurement basis states, $|n\rangle \mapsto e^{i\phi_n}|n\rangle$ as $\Pi_n = |n\rangle\langle n| = e^{i\phi_n}|n\rangle\langle n|e^{-i\phi_n}$. The phases ϕ_n cannot be ignored in general but in certain special cases they can be effectively eliminated. For example, if H is known to be real-symmetric, a common case in physics, then we can choose all basis vectors to be real and restrict $e^{i\phi_n}$ to ± 1. Moreover, if the off-diagonal elements in the computational basis are known to be real and positive, $H_{k\ell} = \langle k|H|\ell\rangle = |\langle k|H|\ell\rangle|$, then $|\langle k|H|\ell\rangle| = |\langle k|\tilde{H}|\ell\rangle|$ with $\tilde{H}$ as above. Hence, with this additional constraint the Hamiltonian is effectively uniquely determined (up to a global energy level shift and global inversion of the energy levels).

A constructive procedure for reconstructing a generic unknown Hamiltonian from stroboscopic measurements of the observables $p_{k\ell}(t)$ at fixed times $t = t_n$ using Bayesian parameter estimation techniques was also given in [6].

III. IDENTIFICATION WITH A-PRIORI INFORMATION

The previous section shows that when essentially no a-priori information about the Hamiltonian is available then even measurement of all the observables $p_{k\ell}(t)$ is not sufficient to uniquely determine the Hamiltonian. However, in many cases some a-priori knowledge about the system is available. For instance, the transition frequencies $\omega_{\mu\nu} = \lambda_\nu - \lambda_\mu$ of the system, where λ_ν are the eigenvalues of the Hamiltonian H, may be known from available spectroscopic data, and we may be able to infer basics such as the level structure and allowed transitions from fundamental physical principles. In such cases the identification problem can be substantially simplified and far less information may be required.

As a specific simple example, consider a three-level system with known transition frequencies ω_{12} and ω_{23} and no direct transitions between states $|1\rangle$ and $|3\rangle$ subject to external fields driving the $(1,2)$ and $(2,3)$ transitions, respectively. If our computational/measurement basis coincides with the eigenbasis of the undriven system, then we know that the Hamiltonian of the driven system must be of the form $H = H_0 + f(t)H_1$ with $H_0 = \text{diag}(0, \omega_{12}, \omega_{12} + \omega_{23})$ and $H_1 = \begin{bmatrix} 0 & d_1 & 0 \\ d_1 & 0 & d_2 \\ 0 & d_2 & 0 \end{bmatrix}$, i.e., we have only two unknowns, d_1 and d_2. If we take the field to be of the form $f(t) = A_1 \cos(\omega_{12}t) + A_2 \cos(\omega_{23}t)$, i.e., consisting of two components that resonantly drive the $(1,2)$ and $(2,3)$ transition, then transforming to a rotating frame and making the rotating wave approximation (RWA), we obtain an effective Hamiltonian $H_{\text{eff}} = \begin{bmatrix} 0 & \Omega_1 & 0 \\ \Omega_1 & 0 & \Omega_2 \\ 0 & \Omega_2 & 0 \end{bmatrix}$, where $\Omega_k = d_k A_k / 2\hbar$ for $k = 1, 2$. If the field amplitudes A_k are constant, this Hamiltonian is constant and we could use the general protocol in [6] to fully characterize the dynamics by stroboscopically measuring the probabilities $p_{k\ell}(t)$ for $k, \ell = 1, 2, 3$ at sufficiently many times t_n. This requires the ability to initialize the system in all three basis states $|k\rangle$ and measure the populations of all three states. Due to conservation of probability $\sum_\ell p_{k\ell} = 1$ and symmetry $p_{k\ell} = p_{\ell k}$, the requirements can be reduced to initialization and measurement in two basis states, e.g., $|1\rangle$ and $|3\rangle$, as the remaining probabilities can be inferred from the other two, but we can do even better by using all the information available.

We shall assume Ω_1 and Ω_2 are real and positive. For notational convenience, let $\Omega = \sqrt{\Omega_1^2 + \Omega_2^2}$ and $\alpha = \arctan(\Omega_2/\Omega_1)$ be the polar coordinates of the vector (Ω_1, Ω_2), i.e., $\Omega_1 = \Omega \cos\alpha$ and $\Omega_2 = \Omega \sin\alpha$ with $\Omega \in \mathbb{R}_0^+$ and $\alpha \in [0, \pi/2]$. Then $U(t,0) = \exp(-itH_{\text{eff}})$ is

$$\begin{bmatrix} c^2\cos(\Omega t) + s^2 & -ic\sin(\Omega t) & cs[\cos(\Omega t) - 1] \\ -ic\sin(\Omega t) & \cos(\Omega t) & -is\sin(\Omega t) \\ cs[\cos(\Omega t) - 1] & -is\sin(\Omega t) & s^2\cos(\Omega t) + c^2 \end{bmatrix} \quad (2)$$

where $c = \cos\alpha$ and $s = \sin\alpha$. This shows immediately that a single measurement trace $p_{k\ell}(t) = |\langle\ell|U(t,0)|k\rangle|^2$ except $p_{22}(t)$ contains information about both parameters and thus should be sufficient to fully identify the Hamiltonian. Specifically, if we choose to measure $p_{11}(t)$ we obtain

$$p_{11}(t) = x^2\cos^2(\Omega t) + 2x(1-x)\cos(\Omega t) + (1-x)^2$$
$$= \frac{x^2}{2}\cos(2\Omega t) + 2x(1-x)\cos(\Omega t) + (1-x)^2 + \frac{x^2}{2},$$

using $\cos^2(\Omega t) = \frac{1}{2}[\cos(2\Omega t)+1]$ and setting $x = c^2 = 1-s^2$. This shows that there are three frequency components 0, Ω and 2Ω, whose amplitudes determine α.

IV. EFFICIENT PARAMETER ESTIMATION

The form of $p_{11}(t)$ suggests Fourier analysis to determine the parameters Ω and α, e.g., by identifying the non-zero Fourier components. The highest frequency peak will be at 2Ω and the corresponding peak amplitude $a_2 = x^2/2$ uniquely determines $x = \sqrt{2a_2}$. In some cases (as in the

example shown in Fig. 1) there may be only one clearly identifiable non-zero peak in the power spectrum, which could correspond to either Ω or 2Ω. This problem can in principle be overcome by estimating x from the average signal $\langle p_{11}(t) \rangle = a_0(x) = 1 - 2x + \frac{3}{2}x$, from which we can obtain estimates for the coefficients $a_2(x) = \frac{1}{2}x^2$ and $a_1(x) = 2x(1-x)$. If $a_2 \gg a_1$ then we identify the non-zero-frequency peak with 2Ω, otherwise with Ω.

Alternatively, we can estimate the base frequency Ω and the signal amplitudes using a Bayesian approach. The signal in our case is a linear combination of the basis functions $g_0 = 1$, $g_1(t) = \cos(\Omega t)$ and $g_2(t) = \cos(2\Omega t)$. Following standard techniques, we maximize the log-likelihood function [6], [11]

$$ P(\omega|\mathbf{d}) \propto \frac{m_b - N_t}{2} \log_{10}\left[1 - \frac{m_b \langle \mathbf{h}^2 \rangle}{N_t \langle \mathbf{d}^2 \rangle}\right], \qquad (3) $$

where m_b is the number of basis functions, $m_b = 3$ in our case, N_t is the number of data points, and

$$ \langle \mathbf{d}^2 \rangle = \frac{1}{N_t} \sum_{n=0}^{N_t-1} d_n^2, \quad \langle \mathbf{h}^2 \rangle = \frac{1}{m_b} \sum_{m=0}^{m_b-1} h_m^2, \qquad (4) $$

where the elements h_m of $(m_b, 1)$-vector $\mathbf{h}$ are projections of the $(1, N_t)$-data vector $\mathbf{d}$ onto a set of orthonormal basis vectors derived from the non-orthogonal basis functions $g_m(t)$ evaluated at the respective sample times t_n. Concretely, setting $G_{mn} = g_m(t_n)$, let λ_m and $\mathbf{e}_m$ be the eigenvalues and corresponding (normalized) eigenvectors of the $m_b \times m_b$ matrix $GG^\dagger$ with $G = (G_{mn})$, and let $E = (e_{m'm})$ be a matrix whose columns are $\mathbf{e}_m$. Then we have $H = VG$ and $\mathbf{h} = H\mathbf{d}^\dagger$ with $V = \mathrm{diag}(\alpha_m^{-1/2})E^\dagger$, and the corresponding coefficient vector is $\mathbf{a} = \mathbf{h}^\dagger V$.

In our case the $P(\omega|\mathbf{d})$ is a function of a single frequency ω and Ω is the frequency for which $P(\omega|\mathbf{d})$ achieves its global maximum. If $\mathbf{a}(\Omega)$ is the corresponding coefficient vector, we can obtain the best estimate for $x = \cos^2 \alpha$ and thus α by minimizing $\|\mathbf{a}(x) - \mathbf{a}(\Omega)\|$ with $a_m(x)$ as defined above. Thus, the problem of finding the most likely model (Ω, α) is reduced to finding the global maximum of $P(\omega|\mathbf{d})$. Unfortunately, this is not an easy task as $P(\omega|\mathbf{d})$ is sharply peaked and can have many local extrema and a substantial noise floor depending on the number and accuracy of the data points. One way to circumvent this problem is to use the peaks in the discrete Fourier spectrum $\mathrm{DFT}(\mathbf{d})$ of the data $\mathbf{d}$ as input for a gradient-based optimization of $P(\omega|\mathbf{d})$. To make the peak detection simpler and more robust, especially when the data is noisy, we find the position ω_0 of the highest peak in the rescaled power spectrum $F(\omega) = 20 \log_{10}[|\mathrm{DFT}[\mathbf{d} - \langle \mathbf{d} \rangle]|^2 + 1]$, which should correspond to either Ω or 2Ω, and then find the location of the maxima ω_1 and ω_2 of $P(\omega|\mathbf{d})$ in the intervals $I_1 = [\omega_0 - \Delta\omega, \omega_0 + \Delta\omega]$ and $I_2 = [\frac{1}{2}\omega_0 - \Delta\omega, \frac{1}{2}\omega_0 + \Delta\omega]$, where $\Delta\omega$ depends on the resolution of the discrete Fourier transform, e.g., $\Delta\omega \approx 2\pi/T$ for regularly sampled data. We take the best estimate ω_3 for the system frequency Ω to be ω_1 if $P_1 > P_2$, and ω_2 otherwise, where $P_j = P(\omega_j|\mathbf{d})$ for $j = 1, 2$. If P_1 and P_2 differ by less than a certain amount we can flag the system suggesting that more data is needed for reliable discrimination.

To test this strategy 30 Hamiltonians $H(\Omega_k, \alpha_k)$ with $\Omega_k \in [0, 2\pi]$ and $\alpha_k \in [0, \frac{\pi}{2}]$ and a range of sampling time vectors $\mathbf{t} = (t_n)$ with $t_n \in [0, 100]$ were generated with the number of samples N_t ranging from 2^{10} to 2^5. Regular and irregular time vector samplings were considered, where for irregular samples a (fast) non-uniform Fourier transform was used [17]. For each test system and time vector $\mathbf{t}_\ell$, noisy data vectors $\mathbf{d}$ were generated by simulating actual experiments, noting that in a laboratory experiment each data point d_n would normally be estimated by initializing the system in state $|1\rangle$, letting it evolve for time t_n, and performing a projective measurement $P_1 = |1\rangle\langle 1|$, whose outcome is random, either 0 or 1. To estimate the probability $p_{11}(t_n)$ the experiment is repeated many times and $p_{11}(t_n)$ approximated by the relative frequency d_n of 1's. The simplest approach is to use a fixed number of experiment repetitions N_e for each time t_n, but noting that the uncertainty of the estimate d_n of $p_{11}(t_n)$ is $N_e^{-1/2}$ shows that it is advantageous to adjust the number of repetitions N_e for each time t_n to achieve a more uniform signal-to-noise ratio. Specifically, for each data point we sample until $d_n\sqrt{N_e} \approx 10$ or we reach a maximum number of repetitions (here 10^4). Although the projection noise for a single data point is Poissonian, the overall error distribution for a large number of samples is roughly Gaussian, justifying the use of a Gaussian error model in the Bayesian analysis.

As the resolution of the discrete Fourier transform and hence the scaled power spectrum is approximately $2\pi/T$, and generally somewhat less for irregular sampling, the uncertainty in the peak positions of the power spectrum will generally be at least π/T, limiting the accuracy of the frequency estimates, in our case to ≈ 0.0314, regardless of the number of data points. This is evident in Fig. 1, which shows that the peak in power spectrum is relatively broad, compared to the peak in the likelihood function. Furthermore, the frequency range covered by the power spectrum depends on the sampling frequency, or the number of data points N_t, with the largest discernible frequency approximately $N_t\pi/T$. If the system frequency Ω is outside this range covered by the power spectrum, we are unable to detect it. For example, for a system with $\Omega = 4.0484$, we require $N_t\pi/T > \Omega$ and thus $N_t > 128$ data points (see Fig. 1). If T and N_t are sufficiently large to avoid such problems, the location ω_0 of the global maximum of the power spectrum usually provides a good starting point for finding the global optimum of the log-likelihood function but we can generally substantially improve the frequency estimates using the likelihood. Of 14440 data sets analyzed (30 test systems sampled at different times) ω_0 differed by less than 1% from the true system frequency Ω, or 2Ω, i.e., $E(\omega_0) < 0.01$ with $E(\omega_0) = \min\{|\omega_0 - \Omega|/\Omega, |\omega_0 - 2\Omega|/2\Omega\}$ in about half (7321) the cases. For almost all failed cases the number of data points was too small and Ω outside the range of the power spectrum. Even when restricted to the successful cases as defined above, the median of $E(\omega_0)$ was 0.0035, while the median of the relative error $E_1(\omega_3) = |\omega_3 - \Omega|/\Omega$ of the final estimate ω_3 obtained by maximizing the likelihood was 6.9×10^{-6}.

We also considered finding the global maximum of the

likelihood by other means, especially in those cases for which the power spectrum does not provide a useful initial frequency estimator. Since we have a function of a single parameter and evaluation of the likelihood, especially when the number of data points is small, is not expensive, it is possible to find the global maximum simply by exhaustive search. Interestingly, we found that log-likelihood still had a clearly identifiable global maximum in many cases even when the number of data points N_t was far below the minimum number of sample points required to detect a peak in the power spectrum. E.g., for the system shown in Fig. 1, the likelihood function still has a sharp peak around the system frequency Ω even if the number of samples is reduced to 32, while the peak is no longer detectable in the power spectrum even for $N_t = 128$ samples. However, as we reduce the number of samples additional peaks in the likelihood function tend to emerge at multiples or fractions of Ω, as shown in the top inset of Fig. 1. This means that we can no longer unambiguously identify the true frequency Ω. Such aliasing problems leading to sampling artefacts in the data analysis can be sustantially reduced by avoiding uniform sampling at equally spaced times (cf Fig. 1, top inset). In particular low-discrepancy sequences have been introduced with the aim to create a sampling with minimal regular patterns causing sampling artefacts, but also minimising the average gap between the samples for a fixed number of samples [15]. Here in particular we use a stratified sampling strategy, where a point is placed in each stratum of a regular grid according to a uniform probability distribution. This *may* be improved further using other low-discrepancy sequences [16]. The results are relevant as a significant reduction in the number of data points required reduces experimental overheads substantially. This comes at additional computational costs, as finding the global maximum of the likelihood function for irregular samplings with very few data points forms a hard optimization problem. Several standard optimization algorithms (simple pattern search and stochastic gradient decent) failed to reliably detect the global optimum, and exhaustive search had to be used.

V. CONCLUDING DISCUSSION

We have considered Hamiltonian identification using stroboscopic measurement data of a fixed observable. If the system can only be initialized in the measurement basis states then a completely unknown Hamiltonian cannot be uniquely identified even if we can measure the population of all basis states as a function of time. If a-priori information is available, however, complete identification of the system parameters is often possible with substantially reduced resources. We have illustrated this for the case of a three-level system where we can only monitor the population of state $|1\rangle$ over time, starting in $|1\rangle$, without the possibility of dynamic control or feedback as was considered in [12]. The results may be applicable to improve the efficiency of identification schemes for other systems. E.g., recent work on system identification for spin networks [13], [14] has shown that the relevant Hamiltonian parameters of a spin chain can be identified by mapping the
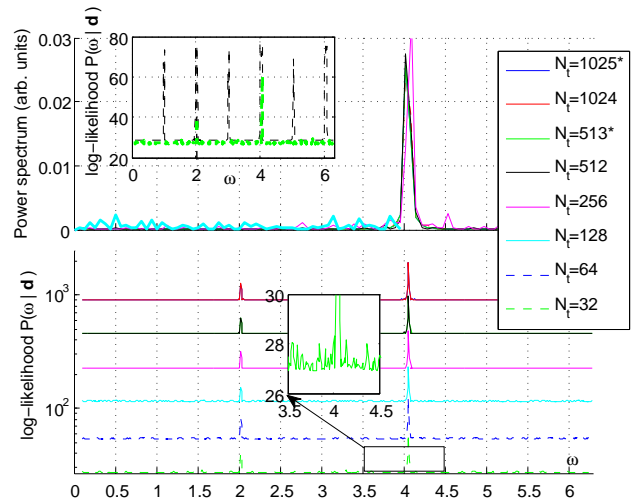

Fig. 1. Power spectra and log-likelihood for a test system with $\Omega = 4.0484$ for data sampled at different times **t** in $[0, 100]$. For $N_t \geq 128$ the power spectra have a single peak in the plotted range, which is a reasonable estimate for Ω. For $N_t = 64$ and below, the main peak is outside the range of the power spectrum and the former no longer contains any useful information. Yet, the log-likelihood still has a clearly identifiable global maximum at Ω even for data vectors with as few as 32 data points, provided a non-uniform sampling is used. For uniform sampling with $N_t = 32$ the top inset shows that $P(\omega|\mathbf{d})$ has many peaks of approximately equal height due to aliasing effects (dashed black line).

evolution of the first spin and Fourier analysis, but the scheme requires repeated quantum state tomography of the first spin for many times t_n, which is experimentally expensive.

REFERENCES

[1] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455-2467 (1997).
[2] J. F. Poyatos, J. I. Cirac and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).
[3] S. G. Schirmer, A. Kolli, D. K.L. Oi, Phys. Rev. A **69**, 050306(R) (2004).
[4] S. G. Schirmer, A. Kolli, D. K. L. Oi, J. H. Cole, In: Proc. 7th Int. Conf. QCMC, Glasgow 25-29 July 2004 (AIP 2004).
[5] J. H. Cole *et al.* Phys. Rev. A **71**, 062312 (2005).
[6] S. G. Schirmer and D. K. L. Oi, Phys. Rev. A **80**, 022333 (2009)
[7] J. H. Cole *et al.* Phys. Rev. A **73**, 062333 (2006).
[8] S. J. Devitt *et al.* New J. Phys. **9**, 384 (2007).
[9] S. G. Schirmer, D. K. L. Oi and S. J. Devitt, J. Phys.: Conf. Series **107** 012011 (2008)
[10] S. G. Schirmer and D. K. L. Oi, arXiv:0911.1367 (2009)
[11] G. Larry Bretthorst, Bayesian Spectrum Analysis and Parameter Estimation (Springer, Berlin, 1998)
[12] Z. Leghtas, M. Mirrahimi, P. Rouchon, arXiv:0903.1011 (2009)
[13] D. Burgarth, K. Maruyama, F. Nori, Phys. Rev. A **79**, 020305(R) (2009)
[14] Daniel Burgarth, Koji Maruyama, New J. Phys. **11**, 103019 (2009)
[15] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods (SIAM Review, 1992)
[16] J. A. Quinn, F. C. Langbein, R. R. Martin, G. Elber, Springer LNCS **4077**, 465-484 (2006).
[17] L. Greengard, J. Lee, SIAM Review **46**(3), 443-454 (1993).