

Frequency tracking and parameter estimation for robust quantum state estimationJason F. Ralph,^{1,*} Kurt Jacobs,^{2,†} and Charles D. Hill^{3,‡}¹*Department of Electrical Engineering and Electronics, The University of Liverpool, Brownlow Hill, Liverpool L69 3GJ, United Kingdom*²*Department of Physics, University of Massachusetts at Boston, 100 Morrissey Blvd, Boston, Massachusetts 02125, USA*³*Centre for Quantum Computation and Communication Technology, School of Physics, University of Melbourne, Victoria 3010, Australia*

(Received 15 September 2011; published 28 November 2011)

In this paper we consider the problem of tracking the state of a quantum system via a continuous weak measurement. If the system Hamiltonian is known precisely, this merely requires integrating the appropriate stochastic master equation. However, even a small error in the assumed Hamiltonian can render this approach useless. The natural answer to this problem is to include the parameters of the Hamiltonian as part of the estimation problem, and the full Bayesian solution to this task provides a state estimate that is *robust* against uncertainties. However, this approach requires considerable computational overhead. Here we consider a single qubit in which the Hamiltonian contains a single unknown parameter. We show that classical *frequency estimation* techniques greatly reduce the computational overhead associated with Bayesian estimation and provide accurate estimates for the qubit frequency.

DOI: [10.1103/PhysRevA.84.052119](https://doi.org/10.1103/PhysRevA.84.052119)

PACS number(s): 03.65.Wj, 03.65.Yz

I. INTRODUCTION

It is well known that the key to large scale quantum computing is the ability to maintain and manipulate a pure quantum state, minimizing any sources of possible error. Errors, either due to environmental fluctuations or inherent in the quantum system itself, are liable to produce decoherence, such as dissipation or dephasing, and lead to initially pure quantum states becoming mixed. In many of the candidate quantum technologies (such as quantum optics, atoms, and ions in traps) the systems being used as the qubits can be reasonably well characterized. Laser frequencies can be measured accurately. Atomic and ionic transition frequencies are consistent from one qubit to another. In the solid state, where devices need to be fabricated individually, there will always be some uncertainty in the parameters that determine the Hamiltonians of the individual qubits, the coupling between qubits, and the coupling to any external controls or environment.

Any attempt to apply feedback control to solid-state qubits will therefore have to deal with uncertainties in their Hamiltonians. Feedback control is based upon continuous state estimation, in which the observer tracks the system via a continuous measurement [1–4], and uses this knowledge to correct the evolution via control inputs [5–8]. It is therefore essential that continuous state estimation handles uncertainties in the Hamiltonian. An estimation process that does this is referred to as being *robust*. An effective approach to achieving this robustness is to estimate continuously both the state of the system and the uncertain parameters in the Hamiltonian as the measurement proceeds. This is the approach we consider here, comparing the optimal Bayesian technique and classical estimation techniques to determine an unknown parameter.

A number of authors have examined continuous state estimation in which the Hamiltonian is uncertain. There has been work on parameter estimation by Gambetta and Wiseman

[9], Verstraete *et al.* [10], Stockton *et al.* [11], Chase and Geremia [12], and Tsang [13–15]. Gambetta and Wiseman have examined a situation similar to the one we consider here, and derived the full Bayesian estimation equations. They found that the ability to determine the unknown parameter depended strongly on the way that the system was measured. This is a result of the way in which the measured observable is affected by errors in the parameter, and the noise introduced into the dynamics by the measurement. The goal of the work by Verstraete *et al.* and Stockton *et al.* was to estimate continuously a signal that appeared as a (possibly time dependent) parameter in the Hamiltonian of a measured system. This problem arises naturally in the context of measuring classical fields, such as gravity-wave detection [16] and magnetometry [11]. In tracking the classical parameter it was necessary to track the state of the system, and thus the authors also derived the full Bayesian estimation equations. More recently, Chase and Geremia have developed a quantum analog of a classical tracking filter called a particle filter [17], which estimates the evolution of continuous probability density function for a parameter by considering a finite set of weighted sample points [12], and Tsang has developed a generalisation of classical smoothing filters [13–15], where measurements are processed forward and backward in time to improve state estimates and the probability distributions for classical parameters. In addition, related work has been done by Yamamoto [18], who considered the problem of robust state estimation for linear quantum systems, using results from classical control theory. Issues of robustness in quantum control have also been discussed in Refs. [19,20].

We consider what is possibly the simplest nonlinear quantum state-estimation problem, that of tracking the state of a single qubit. We first examine what happens when the observer uses the usual stochastic master equation to estimate the state of the system, but with an incorrect Hamiltonian. We find that this procedure is, as expected, very sensitive to errors in the Hamiltonian, failing quickly when the observer's Hamiltonian differs only a little from the true Hamiltonian. We then describe how this problem is solved by extending the estimation equations to include the Hamiltonian parameters.

*jfralph@liverpool.ac.uk

†kurt.jacobs@umb.edu

‡cdhill@unimelb.edu.au

The resulting estimation equations are those derived in Refs. [9] and [10]. (We also include a simple and concise derivation of these equations in an Appendix A.) Since the estimation equations simultaneously estimate both a quantum state and a set of classical parameters, we will refer to this set of equations as the *hybrid* stochastic master equation (HME).

The HME, while an optimal solution to the robust state-estimation problem, does generate a considerable computational overhead. Our main purpose here is to show that this overhead can be reduced, with a relatively small reduction in performance, by using “frequency tracking” methods that were originally developed for classical signal processing applications (such as radar frequency tracking, sonar processing, and communications) [21]. In particular, we examine three frequency tracking techniques: one Fourier-based method and two that do not require a Fourier transform (thereby avoiding a large computational overhead). The Fourier-based technique is the maximizer of the periodogram, finding the frequency that maximizes the correlation between the noisy measurement signal and a complex Fourier kernel. It is optimal for a single sinusoidal signal in noise, but it is computationally expensive.

The first non-Fourier technique (Quinn-Fernandes [22]) is iterative and can be applied entirely in the time domain. It uses notch filtering and relies on the iterative construction of a filter that cancels an instability at the desired frequency. It can be shown to approach the Cramér-Rao bound for the convergence of the estimated frequency, in which the variance of the estimator asymptotes to $O(N^{-3})$, where N is the number of data points in the measurement record [21].

The second non-Fourier technique is common in signal processing and is called “multiple signal characterisation” or MUSIC [23,24]. This technique uses the properties of the eigenvectors of the autocovariance matrix of the data. These are related to the corresponding eigenvalues, which are in turn related to the frequencies present in the data. The variance of this technique scales as $O(N^{-1})$ rather than $O(N^{-3})$, but unlike the Quinn-Fernandes method it does not require an initial frequency estimate and can be more stable for large, appropriately filtered data sets. We compare the accuracy of these techniques with that of the full Bayesian estimator and find that they perform very well in general, although some do require additional filtering to obtain a stable parameter estimate.

We begin by reviewing the formalism used to describe continuous weak measurement, and define the relevant measurement record upon which the parameter estimates will be based. We then describe the optimal Bayesian estimator (the HME) for updating the parameter estimates and discuss several approximate methods for the estimation of frequencies from noisy data, with particular emphasis on computational complexity. We then analyze the performance of the Fourier, Quinn-Fernandes, and MUSIC techniques for a range of measurement strengths, and find that these perform well in comparison with the Bayesian method. We conclude by describing how these techniques can be included in a quantum feedback system to estimate an unknown (possibly time-dependent) Hamiltonian.

II. CONTINUOUS WEAK MEASUREMENT AND STATE ESTIMATION

Measurement processes in quantum mechanics have always been the subject of intense theoretical interest. Traditionally, measurements have been projective and the quantum evolution has been reconstructed from a stochastic series of results, averaged over an ensemble of similarly prepared quantum systems. The main problem with this approach is the restrictive nature of the projective measurement process. There is no alternative to performing a series of projective measurements and building up a probabilistic model for the average evolution. Continuous weak measurement models were developed as a means to describe the effect of weak (i.e., nonprojective) measurement processes. This is achieved by coupling the quantum system of interest to environmental degrees of freedom, and then performing a series of projective measurements on the environment. Because the system becomes correlated with the environment, the projective measurements provide a stream of information about the system of interest. As long as the coupling strength between the system and the environment is sufficiently weak, then the effect of the projective measurements will be small compared to the underlying quantum evolution of the system, and the result is a (weak) continuous measurement.

The evolution of the system now depends on the stream of measurement results, called the *measurement record*, and since these are necessarily random, the evolution of the system is given by a stochastic equation. If one averages over all possible measurement records, then one obtains the usual Lindblad-form master equation for a system coupled to an environment. Each possible measurement record corresponds to a particular realisation of an experiment on a single quantum system. This is often referred to as an “unravelling” of the master equation, with different unravellings corresponding to different kinds of measurements that one could make on the environment [25]. Weak measurements have a relatively small effect on the underlying quantum evolution, but increasing the strength of the coupling between the quantum system and the environment will also allow stronger measurements to be made, including projective-like measurements, where the nonunitary terms dominate the evolution, along with the quantum Zeno effect [26].

Of course, in a real experiment, all one has is the measurement record rather than the underlying processes. In such cases it is possible to reconstruct aspects of the underlying quantum evolution by solving the relevant stochastic master equation (SME) for the specific measurement process and unravelling using the measurement record obtained from the experiment. The resultant evolution provides estimates of the instantaneous state of the quantum system *given* the measurement record [5]. This type of state estimation process is common in classical systems, where the classical evolution of a system is estimated from a noisy series of measurements, often using the Kalman filter or its variants [27]. The parallels between classical state estimation and quantum state estimation have been widely discussed in the literature [28]. In this context, the estimated quantum state represents the best guess of what the actual state of the system is, based upon the set of measurements that are currently available. As the system evolves and more

information becomes available via the measurements, the estimate of the quantum state should improve and the purity of the state should increase.

For a quantum system with a Hamiltonian \hat{H} , subject to a measurement corresponding to an operator \hat{c} , the stochastic master equation for the unravelling is given by Ref. [29]

$$\begin{aligned} \hbar d\rho_c = & -i[\hat{H}, \rho_c]dt + [\hat{c}\rho_c\hat{c}^\dagger - \frac{1}{2}(\hat{c}\hat{c}^\dagger\rho_c + \rho_c\hat{c}\hat{c}^\dagger)]dt \\ & + dz^\dagger(\hat{c}\rho_c - \langle\hat{c}\rangle\rho_c) + (\rho_c\hat{c}^\dagger - \rho_c\langle\hat{c}^\dagger\rangle)dz, \end{aligned} \quad (1)$$

where dt is an infinitesimal time increment and dz is an infinitesimal complex Wiener increment. The complex Wiener increments obey the relations $E[dz] = 0$ (where $E[\dots]$ represents an expectation value), $dzdz^\dagger = \hbar\Theta dt$, and $dzdz^T = \hbar\Upsilon dt$. In general, dz is a complex vector and Θ and Υ are matrices, but if we restrict consideration to a single measurement interaction, Θ is a real number (which we will set to one, corresponding to efficient detection [29]) and we set $\Upsilon = 1$ so that the Wiener increments are real. The density matrix ρ_c is conditioned upon the measurement record, which is given by Ref. [29]

$$dy(t) = \langle\hat{c}^T\Theta + \hat{c}^\dagger\Upsilon\rangle_c dt + dz^T = \langle\hat{c} + \hat{c}^\dagger\rangle_c dt + dz, \quad (2)$$

where the expectation values are taken with respect to the conditional state, $\langle\hat{c}\rangle_c = \text{Tr}[\rho_c\hat{c}]$.

Unlike projective measurements, weak measurement processes are not restricted to Hermitian operators but, for the purposes of the current paper, we will assume that the measurement operator is Hermitian. This means that the sole action of the measurement is to extract information about a single physical observable. (Measurement operators that are not Hermitian describe damping in addition to the extraction of information, a common example is the damping of a lossy optical cavity [30]). For a Hermitian weak measurement $\hat{c} = \sqrt{2k\hbar}\hat{y} = \hat{c}^\dagger$, where k is the strength of the measurement interaction, the SME reduces to

$$\begin{aligned} d\rho_c = & -\frac{i}{\hbar}[\hat{H}, \rho_c]dt - k[\hat{y}\hat{y}, \rho_c]dt \\ & + \sqrt{2k}(\hat{y}\rho_c + \rho_c\hat{y} - 2\langle\hat{y}\rangle_c\rho_c)dW, \end{aligned} \quad (3)$$

where dW is a real Wiener increment (such that $dW^2 = dt$), and the measurement record is

$$\frac{dy(t)}{\sqrt{\hbar}} = \sqrt{8k}\langle\hat{y}\rangle_c dt + dW. \quad (4)$$

At each point in time, the measurement record (from an experiment or a simulated process) will determine the associated dW , given the current estimate of the state ρ_c . The state of the system may then be updated using the value found for dW and Eq. (3). In classical state estimation, the value dW is often referred to as the *innovation* [27], in that it represents the new information provided by the measurement. The innovation is normally the difference between the actual measurement [in this case, $dy(t)$] and the current estimate of what the measurement should be given the current state of knowledge of the system [which is $\sqrt{8k}\langle\hat{y}\rangle_c dt$]. The main difference between classical and quantum measurement processes is that if the measured observable does not commute with the Hamiltonian, then the measurement actually feeds noise into the observable that is being measured. Because of this, no

matter how long we measure for, we will always be left with some uncertainty about the observable.

The main potential problem with the above approach to state estimation is that the stochastic master equation requires that the system is well defined, in the sense that the Hamiltonian and the measurement interaction need to be known very accurately. If the Hamiltonian and interaction are not well defined, the measurement record that is used to update the system will produce incorrect state estimates and assign too much confidence to the results of the SME. This is also common in classical state estimation problems, where too much confidence is assigned to the dynamical processes and valid measurements are often discounted because they conflict with the (inaccurate) state estimates [27]. Ideally we would like our state estimates to be insensitive to small errors in the Hamiltonian, or to allow the parameters in the Hamiltonian to be estimated from the measurement record so that errors can be corrected online.

To examine the sensitivity of the state estimates to errors in the Hamiltonian, we consider two SMEs: one is used to produce a measurement record (playing the role of an experimental system) and the other is used to estimate the underlying state of the system based on a Hamiltonian that contains a small error. Initially, we select a simple single-qubit Hamiltonian,

$$H = \hbar\omega_x\sigma_x/2, \quad (5)$$

where σ_x is the Pauli x matrix. This Hamiltonian generates a rotation of the single qubit Bloch vector about the Bloch x axis. We then apply an interaction that generates a measurement along the Bloch z axis. This maximizes the size of the output signal relative to the noise by choosing a measurement that is orthogonal to the axis of rotation. The Hamiltonian and measurement are selected because they reflect a common situation for solid state charge qubits, where the σ_x term corresponds to a tunneling interaction and the σ_z interaction corresponds to a measurement of charge (see Ref. [31] for example). Rewriting the SME in the Bloch vector representation [32] with

$$r_i = \text{Tr}[\sigma_i\rho], \quad i = X, Y, Z,$$

where σ_i is a Pauli matrix and the density operator is

$$\rho = \frac{I + r_X\sigma_x + r_Y\sigma_y + r_Z\sigma_z}{2},$$

the SME becomes three coupled stochastic equations

$$\begin{aligned} dr_X = & -4kr_X dt - \sqrt{8k}r_X r_Z dW, \\ dr_Y = & -\omega_x r_Z dt - 4kr_Y dt - \sqrt{8k}r_Y r_Z dW, \\ dr_Z = & \omega_x r_Y dt + \sqrt{8k}(1 - r_Z^2)dW. \end{aligned} \quad (6)$$

The corresponding measurement record is

$$\frac{dy(t)}{\sqrt{\hbar}} = \sqrt{8k}r_Z dt + dW. \quad (7)$$

We use one SME with a fixed frequency ω_x to generate a measurement record, and then feed this measurement record to a second SME. This second SME is used by an observer to estimate the state of the system. The first SME is initialized in a pure state. The second SME is initialized in a completely mixed state ($r_X = r_Y = r_Z = 0$), reflecting the fact that the

observer does not initially know what state the qubit is in. If the Hamiltonian used in the second SME is perfect (that is, is exactly the same as that used in the first SME), then the initial mixed state will gradually purify as information is extracted. It will converge to the correct underlying state. (The process of purification or state reduction of single qubits subject to weak measurements and the use of feedback to increase the rate of purification have been discussed in detail by the current authors and others [31,33–35].) If the Hamiltonian used in the second SME contains an error, the estimated or conditioned state will purify, but the errors in the predicted evolution due to the inaccurate Hamiltonian will cause the estimated state and the underlying state from the first SME to diverge. The question is how accurate the Hamiltonian needs to be so that the estimated state purifies before the prediction errors accumulate and cause the states to diverge. A robust parameter estimation process would allow a state to be purified to the required level before they diverge significantly. Preferably this should be achieved without affecting the underlying evolution of the state, thereby allowing the Hamiltonian parameters to be estimated and updated. The alternative, a measurement strong enough to completely dominate the evolution, projecting the state on a fast time scale compared to the Hamiltonian, would certainly give the correct estimated state, but it is not the situation we are interested in here.

We begin by examining the divergence of the underlying state, given by the first SME, from that of the estimated state, given by the second SME. We do this by calculating the average fidelity between the two density matrices as a function of time for several different measurement strengths. The question relating to the extraction of the parameters in the Hamiltonian will be dealt with in the next section. The fidelity F for two density matrices ρ_0 and ρ_c is a common measure of the difference between two (possibly mixed) density matrices and it is given by Ref. [36]

$$F = F(\rho_0, \rho_c) = |\text{Tr}[\sqrt{\sqrt{\rho_c}\rho_0\sqrt{\rho_c}}]|^2.$$

In Fig. 1 we show the average fidelity for a completely mixed initial state as a function of time for a number of different expected frequency errors (the oscillation frequency ω_x being the parameter that we will be trying to estimate and the errors being characterized by their standard deviation $\sigma_{\delta\omega_x}$, which will normally be given as a percentage) and for a measurement strength $k = 0.07\omega_x/(2\pi)$. The fidelity is calculated by comparing the estimated density matrix ρ_c and the underlying (pure state) density matrix ρ_0 . Both SMEs are integrated numerically using Milstein stochastic integration [37] and at least 4000 time steps per oscillation cycle. The estimated states' purify as information is extracted from the measurement record, thereby increasing their fidelity. The average purity of ρ_c will increase with time (the estimated purity tends to one) but the average fidelity eventually saturates as the information gain from the measurement record is balanced by the information loss due to the inaccuracies in the estimated Hamiltonian parameter. In the inset of Fig. 1, the saturation value of the average fidelity is given as a function of measurement strength for three different values of the error ($\sigma_{\delta\omega_x} = 1\%$, 2% , and 5%). Figure 2 shows $1 - \bar{F}$ ($1 -$ average fidelity, sometimes called the “infidelity”) as a function of the

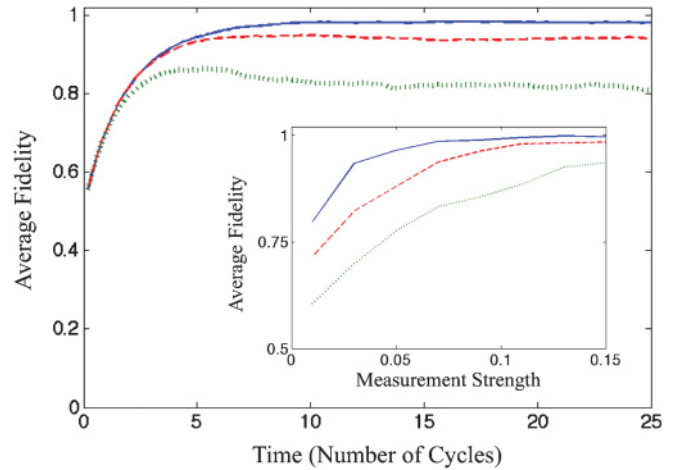


FIG. 1. (Color online) Average fidelity of the estimated state vs time for different values of expected errors $\sigma_{\delta\omega_x} = 1\%$ (blue solid line), 2% (red dashed line), and 5% (green dotted line) for $k = 0.07\omega_x/(2\pi)$. Inset shows the saturated value for the average fidelity as a function of measurement strength $2\pi k/\omega_x$ for $\sigma_{\delta\omega_x} = 1\%$ (blue solid line), 2% (red dashed line), and 5% (green dotted line).

percentage error in the estimate of the oscillation frequency for three different measurement strengths. It is clear from the data shown in Figs. 1 and 2 that errors of the order of a few percent can have a significant effect on the fidelity of the estimation process. To obtain a fidelity value in excess of 0.99 the frequency error must be less than about 0.2%–0.3% if the measurement strength is around $k = 0.03\omega_x/2\pi$, but the error can be as large as 1% if the measurement strength is greater than about 0.07–0.08 $\omega_x/2\pi$. Therefore, it is crucial that the measurement strength is chosen to reflect the expected accuracy of the frequency estimation process, although the stronger the coupling between the measurement environment and the qubit, the greater the effect of the backaction of the measurement on the qubit evolution. In the next section we consider how an optimal Bayesian inference method can be

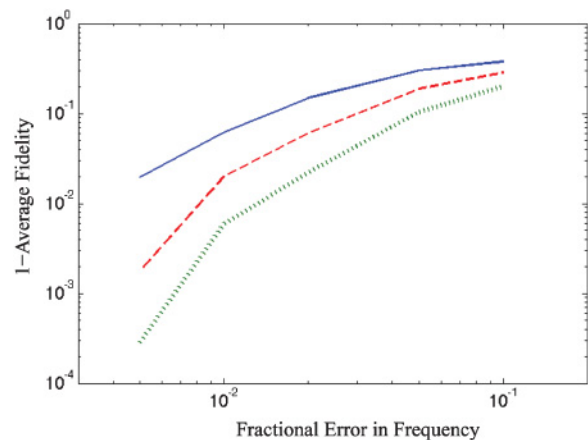


FIG. 2. (Color online) Average fidelity error ($1 -$ average fidelity) of the estimated state vs percentage frequency errors for $2\pi k/\omega_x = 0.03$ (blue solid line), 0.07 (red dashed line), and 0.12 (green dotted line) after 25 oscillation cycles.

used to update the parameter estimates online and consider several alternatives taken from classical signal processing and frequency tracking.

III. QUANTUM PARAMETER ESTIMATION AND FREQUENCY TRACKING

A. Bayesian inference

Making a continuous measurement will, in general, provide information not only about the state of a system, but also about the dynamics, and thus the parameters that determine the Hamiltonian. We will denote the parameters by the vector λ . Our task is to process the measurement results in order to extract that information. The full solution to this problem is given by applying Bayesian inference [38] (classical measurement theory) in combination with quantum measurement theory (being the quantum version of Bayesian inference [39]). To begin with, our state of knowledge regarding the system and the parameters is described by a density matrix for the state of the system ρ_c , and a probability density for the parameters $P(\lambda)$. Here we are concerned only with a single parameter (the frequency of the oscillations), so we have a scalar parameter λ . The solution to the continuous estimation problem tells us how to update both the density matrix and $P(\lambda)$ in each infinitesimal time-step dt , that is, it gives differential equations for ρ_c and $P(\lambda)$, and these are stochastic because they are driven by the measurement record.

We now present the stochastic differential equations that solve the combined state and parameter estimation problem. While they look somewhat involved, they are quite simple to derive, and we give this derivation in Appendix A. Since the Hamiltonian is a function of λ , we will write it as $H(\lambda)$. The stochastic master equation for the density matrix is

$$d\rho_c = \frac{-i}{\hbar} \left[\int P(\lambda) H(\lambda) d\lambda, \rho_c \right] dt - k[\hat{y}, [\hat{y}, \rho_c]] dt + 4k(\hat{y}\rho_c + \rho_c\hat{y} - 2\langle\hat{y}\rangle_c\rho_c)(dr - \langle\hat{y}\rangle_c dt), \quad (8)$$

where dr is the measurement record, $dr = dy(t)/\sqrt{8k\hbar}$, \hat{y} is the measured observable, and $\langle\hat{y}\rangle_c = \text{Tr}[\hat{y}\rho_c]$. The equation of motion for $P(\lambda)$ is

$$dP(\lambda) = 8k(\langle\hat{y}\rangle - \langle\hat{y}\rangle_\lambda)(dr - \langle\hat{y}\rangle dt)P(\lambda). \quad (9)$$

Here we have introduced a new quantity $\langle\hat{y}\rangle_\lambda$. This is the expectation value that \hat{y} would have if we knew the value of λ . To be able to calculate $\langle\hat{y}\rangle_\lambda$ we must continually update our state-of-knowledge of the system, assuming that we know λ for every value of λ . Denoting our state-of-knowledge of the system, given λ , by ρ_λ , the update equation for each value of λ is

$$d\rho_\lambda = \frac{-i}{\hbar} [H(\lambda), \rho_\lambda] dt - k[\hat{y}, [\hat{y}, \rho_\lambda]] dt + 4k(\hat{y}\rho_\lambda + \rho_\lambda\hat{y} - 2\langle\hat{y}\rangle_\lambda\rho_\lambda)(dr - \langle\hat{y}\rangle_\lambda dt), \quad (10)$$

where $\langle\hat{y}\rangle_\lambda = \text{Tr}[\hat{y}\rho_\lambda]$. We will refer to Eqs. (8)–(10) together as the *hybrid master equation*, since it involves both quantum and classical states-of-knowledge. Unlike the smoothing filters developed by Tsang [13–15], the hybrid master equation propagates forwards in time only.

It is because we must propagate the density matrices ρ_λ for every value of our parameter λ that the full state estimator is so numerically intensive. If λ is a continuous variable, as is the case here, then it is impossible to have a density matrix for all values of λ , and we therefore use a discrete grid of values. Note that the spacing of values on this grid must be at least as small as the accuracy with which we wish to track λ . In addition, the length of the grid (the range of values of λ) must be large enough to encompass our initial uncertainty. If we use a fixed grid, these two requirements determine how many values of λ require, which can be large. One approach to reducing this number would be to use a moving grid, updating the points on the grid along with $P(\lambda)$. Even this procedure would require at least three copies of the density matrix (likely more), and certainly more if we had two or more parameters. In addition, it requires the rather involved task of interpolating a new density matrix each time a grid point is moved. Rather than exploring dynamical grids as a numerical method for implementing the full estimator, in the next section we consider replacing the Bayesian updater for $P(\lambda)$ with a classical frequency estimation technique.

We present in Fig. 3 results for using the hybrid master equation to estimate the frequency of a single qubit. For these simulations we used a grid of 301 points for λ . The simulations are highly intensive: running the estimator for 500 periods, and averaging over 5000 realizations, takes ~ 33 hours on a 232 processor cluster. In Fig. 3 we plot the average error of the estimator (the average rms deviation from the true frequency) as a function of time, and for different values of the measurement strength k . While a stronger measurement provides a faster rate of information extraction, the laws of quantum mechanics mean that it also introduces more noise into the system. As a result there is an optimal value for the measurement strength for a given frequency. We see from Fig. 3 that this is approximately $k = 0.035\omega/2\pi$, at least for a measurement time of 150 cycles.

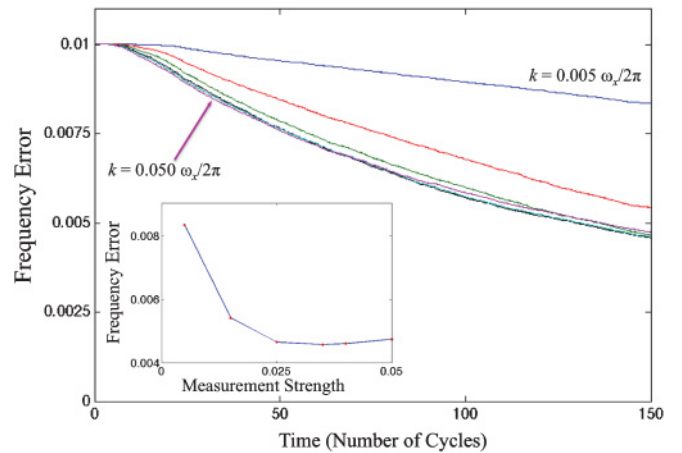


FIG. 3. (Color online) Expected frequency errors (1 standard deviation) vs time for Bayesian inference method and different measurement strengths, $2\pi k/\omega_x = 0.005, 0.015, 0.025, 0.035, 0.040$, and 0.050 . Inset figure shows the expected frequency error after 150 qubit cycles for the data shown in the main figure.

B. Classical frequency tracking

The ability to determine which frequencies are present in a noisy signal is extremely important in a large number of practical engineering applications. In communication systems, analog signals are often transmitted by the modulation of the frequency of a carrier signal or digital data bits can be encoded in a frequency-shift key modulation scheme [40]. In classical signal processing, the frequencies present in sonar signals provide information about the engine and gearing mechanisms in sea vessels [21] and tracking the Doppler frequency shifts in radar signals provides information about the relative (radial) motion of the radar and the target [41]. A large number of techniques have been developed to extract the dominant frequencies from a noisy signal. Some rely on calculating the Fourier transform of the data record and estimating the likelihood that frequency component corresponds to a signal rather than to noise, but these are computationally expensive and are similar in nature to the Bayesian method discussed above. However, other techniques have been developed which do not require explicit calculation of a Fourier transform and some can be applied in the time domain alone. It is natural then to ask whether these techniques can be used to extract useful information from a quantum measurement record of the form given in Eq. (7). From the wide variety of techniques that are available, we have chosen three for further study: the maximizer of the periodogram, the Quinn-Fernandes method [22], and the *multiple signal characterization* method or MUSIC [23,24]. The frequency that maximizes the periodogram provides the best correlation between the measurement record and a sine wave, it is a computationally inefficient technique, but it provides the most accurate frequency estimate for cases where a single frequency signal is corrupted by white noise. The Quinn-Fernandes method is selected for its relative simplicity, it works in the time domain, removing the need to perform computationally expensive Fourier transforms, and for the fact that its theoretical accuracy approaches the optimal Cramer-Rao bound. The MUSIC method is slightly more complicated, using an estimate of the signal's autocorrelation matrix, but it does not require the initial frequency estimate used by Quinn-Fernandes and it can be used for signals containing more than one frequency component.

The main advantage of classical techniques is that the measurement record can be integrated over much longer time steps, thereby reducing the computational requirements. The optimal Bayesian inference technique requires that the SME is integrated using very small time steps. For the purity value of the quantum state to be well behaved (i.e., not significantly larger than one), several thousand time steps are required per period of the oscillation even when Milstein stochastic integration techniques are used. By contrast, the classical tracking techniques only require on the order of 50 sample points per period, a significant saving in computational terms. The measurement record used by the classical trackers is a discrete sequence of values sampled at regular intervals from the continuous measurement record. As such, each sample y_n is an integral of $dy(t)$ over a finite time-step $t_n \rightarrow t_n + \Delta t$ where $n = 0, \dots, N - 1$:

$$y_n = \int_{t_n}^{t_n + \Delta t} dy(t). \tag{11}$$

When the coupling between the qubit and the measured environment is very weak, the discrete signal y_n can be approximated by a single sinusoid and a noisy background signal,

$$y_n \approx A \cos[\omega_x(n\Delta t) + \phi] + \epsilon_n, \tag{12}$$

where A and ϕ are the amplitude and phase of the sinusoidal signal and ϵ_n is the noise. For a single sinusoid in Gaussian white noise, the minimum variance for an unbiased estimate of the frequency given by the asymptotic Cramer-Rao bound [21]

$$\sigma_{\text{CRB}}^2 = \frac{48\pi f_x(\omega_x)}{A^2(N\Delta t)^3},$$

where $f_x(\omega)$ is the spectral density of the noise at a frequency ω , and the signal-to-noise ratio at the qubit frequency is given by $\text{SNR} = 10 \log_{10}\{A^2/[4\pi f_x(\omega_x)]\}$ [21].

1. Maximizing the periodogram

The periodogram for N discrete points is formed by calculating

$$I_y(\omega) = \left| \sum_{n=0}^{N-1} y_n \exp(-i\omega t_n) \right|^2 \tag{13}$$

and the estimate of the qubit frequency is given by the value of ω that maximizes $I_y(\omega)$. For the cases considered in this paper, the number of points will be large and the signal-to-noise ratio for the signal (i.e. the measurement record) will be very low when compared to most signal processing applications. Typically there will be 50 points per oscillation cycle of the qubit, and we will consider the performance over several hundred cycles. A rough estimate of the signal-to-noise ratio is given by the ratio between the variance of the two terms in the $dy(t)$ measurement record,

$$\text{SNR} = 10 \log_{10} \left[\frac{(\sqrt{8k} \Delta t)^2}{2\Delta t} \right] = 10 \log_{10}(4k \Delta t),$$

which, for the discrete time examples that we will be considering, will provide signal-to-noise ratios between -30 and -20 dB. However, this simple calculation ignores the backaction of the measurement on the qubit, which induces slightly higher levels of noise in the output signal and broadens the peak associated with the sinusoidal oscillation of the qubit. In such cases the accuracy of the Fourier-based technique is limited by the broadening of the peak because a broad peak has many maxima and a large number of points can generate a maximum when subject to a large amount of noise. Other techniques based on the Fourier transform, such as the interpolation of Fourier coefficients [42], offer better performance than simply maximizing the discrete Fourier transform of the discrete time signal (which was found to provide very poor performance for all the cases studied in this paper). However, these techniques will always underperform when compared to maximizing the periodogram. An alternative approach based on fitting the whole frequency response of the broadened measurement peak could be envisaged, and could offer more accurate frequency estimates, but this is likely to require even more computational power than calculating the periodogram.

2. The Quinn-Fernandes technique

The Quinn-Fernandes technique belongs to a class called *adaptive notch filters* [22,43]. The general idea is to introduce a filter that removes a particular frequency from the signal, the “notch” frequency. If the signal is made up of white noise and a sinusoid, finding the notch frequency that minimizes the variance of the filtered signal is equivalent to finding the frequency of the sinusoid. The background to the method and the implementation of the technique are described fully in Appendix B. The main point to note is that this technique requires an initial estimate of the qubit frequency $\omega_1 = \tilde{\omega}_x$. The accuracy of the initial estimate will have an effect on the accuracy of the final frequency estimate ω_{QF} because it affects the convergence properties of the minimization procedure used by the technique. For sinusoidal signals with low to moderate noise levels the Quinn-Fernandes technique is fairly insensitive to inaccuracies in this initial estimate, but for the situation considered here, where the signal-to-noise level is very low and the qubit oscillation is affected by the backaction of the measurement, the accuracy of the initial estimate is very important. Figure 4 shows the evolution of the Quinn-Fernandes frequency estimates for the qubit signal for an example measurement record, and for several different initial frequency estimates. On short time scales there are very few stable solutions to the iterative procedure outlined above, but as time increases the accuracy of the initial estimate is required to be better because the number of stable solutions and estimates increases. This multiplicity of solutions is partly due to the very low signal-to-noise level of the measurement record and partly due to the broadening of the corresponding peak in the Fourier domain (one effect of measuring the qubit is to reduce the effective quality factor of the oscillation). The Quinn-Fernandes technique performs much better when the noise level is (artificially) lowered. Reducing the size of the dW term in the measurement record Eq. (7) relative to the size of the signal reduces the number of solutions, and the accuracy of the technique approaches the theoretical

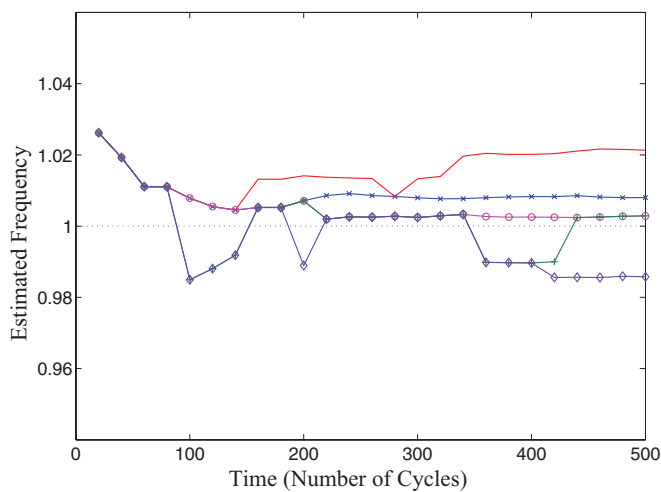


FIG. 4. (Color online) Example outputs from Quinn-Fernandes frequency estimation algorithm as a function of time with different initial frequency estimates: $\tilde{f}_x = 1.02$ (red line), $\tilde{f}_x = 1.01$ (blue x), $\tilde{f}_x = 1.00$ (magenta circle), $\tilde{f}_x = 0.99$ (green plus), and $\tilde{f}_x = 0.98$ (purple diamond).

best case. (The best case corresponds to the output of the Quinn-Fernandes estimate when the initial estimate is perfect.) If the qubit signal is a spin- $\frac{1}{2}$ system, then moving to a larger value of angular momentum would boost the relative size of the signal compared to the dW noise and may give the Quinn-Fernandes technique a significant advantage over the others. Increasing the signal-to-noise ratio by a factor of about 3 to 4 is enough to increase the accuracy of this technique significantly, and the Quinn-Fernandes technique is the least expensive computationally of the different techniques considered in this paper. Even without an increased signal-to-noise ratio, the Quinn-Fernandes does provide a slight improvement in accuracy over the initial estimate, as discussed in the next section.

3. Multiple signal characterization (MUSIC)

Multiple signal characterization, or MUSIC, is a widely used frequency estimation technique [23,24]. MUSIC uses properties of the eigenvectors of the autocovariance matrix of the measurement record to decompose the data into a signal subspace and a noise subspace. The autocovariance matrix is formed by constructing a symmetric Toeplitz matrix

$$C = \begin{pmatrix} C_0 & C_1 & \dots & C_{M-1} \\ C_1 & C_0 & \dots & C_{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{M-1} & \dots & C_1 & C_0 \end{pmatrix} \quad (14)$$

from the autocovariances

$$C_j = N^{-1} \sum_{n=j}^{N-1} (y_n - \bar{y})(y_{n-j} - \bar{y}), \quad (15)$$

where \bar{y} is the mean of the measurement record. The size of the autocorrelation matrix $M > 2$ can be varied to optimize the performance of the technique and for the examples shown here $M = 5$, which provides a good level of performance and does not require large computational resources. MUSIC can be used to find more than one frequency, however, we will only use it to find one frequency in this paper.

Once the autocovariance matrix has been constructed, the eigenvectors of the matrix v_m ($m = 1, \dots, M$) are found and ordered so that the eigenvalues γ_m are in ascending order $\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_M$. If there are p sinusoids present in the data, the first $M - p$ eigenvectors corresponding to the lowest $M - p$ eigenvalues are used to find the MUSIC “spectrum” $S(\omega)$ from

$$S(\omega) = \frac{1}{\sum_{m=1}^{M-p} |e^*(\omega)v_m|^2}, \quad (16)$$

where $e(\omega) = \{1 \exp(i\omega\Delta t) \exp(2i\omega\Delta t) \dots \exp[i(M - 1)\omega\Delta t]\}^T$ and the eigenvectors v_m are normalized. The MUSIC spectrum contains a number of maxima, one for each of the p sinusoids present in the data. In our case, $p = 1$ and $\omega \simeq \omega_x$. The maximum value of the MUSIC spectrum therefore provides our estimate for the qubit frequency.

The signal-to-noise level in the measurement record is sufficiently high that the MUSIC technique applied on its own proves to be unstable. The frequency estimates generated occasionally contain very large errors, which unduly affect

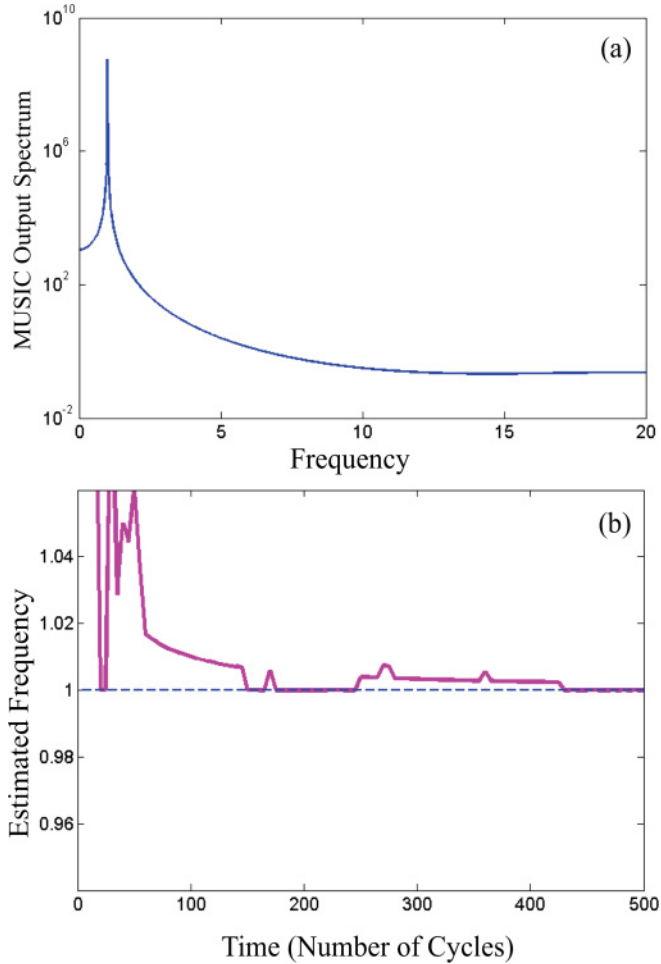


FIG. 5. (Color online) (a) Example MUSIC spectrum and (b) corresponding frequency estimate as a function of time.

the variance of the errors and the resultant accuracy of the technique. To remove these very large errors, it is necessary to prefilter the measurement record to remove frequency components too far away from the initial estimate of the qubit frequency. This is not as restrictive as the initial estimate used in the Quinn-Fernandes technique: the prefilter used here is a bandpass filter (a Butterworth filter [44]) with a pass band which is $\pm 10\%$ of the initial frequency estimate. Figure 5(a) shows an example of a MUSIC spectrum calculated for an example measurement record and the corresponding MUSIC frequency estimates are shown in Fig. 5(b) as a function of time.

C. Comparison of performance

The principal aim of the frequency estimation algorithm is to improve the final fidelity of the estimated state in the stochastic master equation. Figures 1 and 2 show that the expected fidelity of the state, after the initial purification phase, is dependent on the measurement strength and the expected error in the qubit frequency. The Bayesian method gives the best estimates for the qubit frequency for relatively weak measurements $2\pi k/\omega_x \simeq 0.035$. However, such weak measurements require very accurate frequency estimates to obtain high fidelity values: Fig. 2 indicates that the frequency

error should be less than about 0.2%–0.3% for the average fidelity to be in excess of 0.99. Using stronger measurements, $2\pi k/\omega_x \simeq 0.070$ –0.120, reduces the accuracy required from the frequency estimation algorithm to around 0.8%–1%. The strength of the measurement interaction is an important design consideration for solid state circuits because it is normally more difficult to decouple or otherwise modify the interactions between fabricated circuits than it would be for an atomic or ionic qubit system. As such, the measurement strength would be an important design parameter in the construction of solid-state qubits and the associated control circuitry.

The accuracy of the frequency estimation algorithms for the different measurement strengths is summarized in Fig. 6. For weak measurement values around $2\pi k/\omega_x \simeq 0.030$, the accuracy of each of the three classical techniques is found to be approximately equal in the long time limit, although maximizing the periodogram is more accurate over short time scales. (Two values are given for the Quinn-Fernandes technique, representing the upper and lower bounds to the accuracy, which is a function of the accuracy of the initial estimate.) The baseline accuracy of all three classical techniques is limited to around 0.8%, which would be accurate enough for a maximum average fidelity of around $F = 0.95$ (see Fig. 2). Only the Bayesian technique approaches the accuracy that would be required for fidelities approaching 0.99 at the expense of significantly increased computational loading.

As the measurement strength is increased, the accuracy of the periodogram–Fourier-based technique is seen to deteriorate relative to either of the other two classical techniques. For very short periods of time, it remains the most accurate technique, but the accuracy asymptotes very quickly to an error of around 1.2%. With an initial 1% error in the estimate, the Quinn-Fernandes technique provides a slight improvement in the frequency estimate after about 150 qubit cycles. However, the accuracy of the Quinn-Fernandes technique for $2\pi k/\omega_x \simeq 0.070$ is around 0.9%, which is just outside the desired range for an average fidelity of $F = 0.99$, and the accuracy of this approach can be seen to deteriorate when the measurement strength is increased. The lower bound for the Quinn-Fernandes accuracy is given in each graph in Fig. 6 and is the accuracy of the Quinn-Fernandes technique if the initial estimate is perfect (i.e., zero error). This represents the best estimate that could be achieved using this (time-domain) technique, but in practice it will not be achievable because of errors in the initial frequency.

The expected accuracy of the MUSIC frequency estimates is slightly better than the other two classical methods, the error reduces to around 0.7% at $2\pi k/\omega_x \simeq 0.070$. This is well within the range for an average fidelity greater than 0.99. The expected accuracy of this technique would normally be expected to be worse than the other two methods because it does not converge to the Cramer-Rao bound in the long time limit, however, it is found to be remarkably robust in the presence of noise and it was also found to be stable in the presence of a broader peak in the Fourier domain. In fact, the expected accuracy was found to be fairly insensitive to noise and to variations in the measurement strength. It provides a reasonable estimate of the qubit frequency and is relatively efficient in terms of computational demands, slightly longer

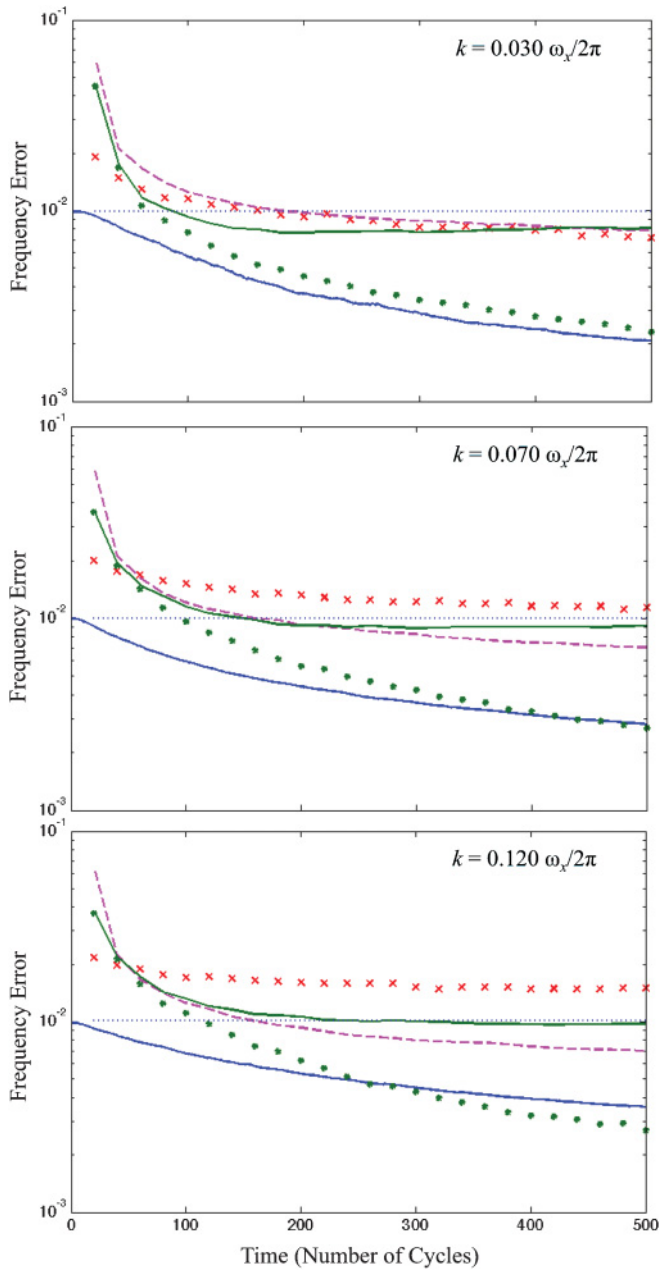


FIG. 6. (Color online) Expected frequency errors (1 standard deviation) vs time for classical frequency tracking techniques: Bayesian [blue (lower) solid line], the maximum of the periodogram error (red cross), Lower limit for Quinn-Fernades [green dotted line] error, upper limit for Quinn-Fernandes [green (upper) solid line] with 1% initial error (blue dotted line), and MUSIC (purple dashed line) for three different measurement strengths $2\pi k/\omega_x = 0.03, 0.07,$ and 0.12 .

to calculate than Quinn-Fernandes but much quicker than the periodogram or the Bayesian methods.

While the Bayesian frequency estimates are most accurate in the region of $2\pi k/\omega_x \simeq 0.035$, the accuracy of this technique after 500 cycles is always good enough to give an average fidelity above 0.99. However, comparing the trends for the fidelity errors shown in Fig. 2 with the accuracies of the frequencies given in Fig. 6, it suggests that the improvement in the average fidelity for increasing

measurement strength increases faster than the degradation in the frequency estimates. This means that the Bayesian technique should still give higher fidelity estimates for the qubit states for stronger measurements even though frequency estimates are less accurate. It should also be noted that the Bayesian method generates a state estimate to obtain an estimate of the state from any of the classical tracking methods one must integrate the SME once the qubit frequency has been determined.

IV. FREQUENCY TRACKING FOR A TIME-DEPENDENT HAMILTONIAN

So far we have considered a frequency estimation process rather than a true tracking process, but the extension of the estimation processes discussed to tracking is relatively straightforward for the case of a single qubit subject to a continuous measurement. The simplicity derives from the fact that there is only one frequency to be estimated in the measurement signal. If the qubit frequency is fixed, then more measurement data can be collected and the accuracy of the frequency estimates can be improved beyond the values given in Fig. 6. The accuracies of the Bayesian and MUSIC techniques are still improving after 500 qubit cycles. It is not clear whether this is also the case for the periodogram or Quinn-Fernandes techniques because any improvement in accuracy after a few hundred cycles is very slow in these cases.

If the qubit frequency is not fixed but varies slowly with time, slow compared with the time scales associated with the frequency estimation process, there are a number of approaches that could be taken depending upon the nature of the variations in the frequency. For solid-state qubits, the most common problems are likely to be either a gradual frequency drift or a discontinuous jump. A gradual frequency drift could arise because the qubit controls are not absolutely stable and give rise to small drifts in the applied control fields leading to slight changes in the qubit frequency. Alternatively, a control could contain infrequent, but sudden, small jumps in voltage which alter the bias point of the qubit circuit in a random way.

For gradual frequency drifts, a natural method for generating a time-dependent frequency “track” would be to generate a moving window for the measurement data so that the frequency is always estimated over a fixed length window of data where the frequency is approximately stationary. The estimates are continually updated as the window is moved. The length of the window can be varied depending upon the anticipated time dependence of the underlying signal. This can be computationally expensive, the data having to be reprocessed each time the window moves, but it provides incremental updates to the frequency and generates a natural track history for the frequency. To reduce computational demands, the measurement data can be divided up into fixed blocks with finite length, but the frequency updates are then discrete measurements associated with each block and a separate tracking algorithm, such as a Kalman filter [27], is required to generate a continuous track in between the estimates and updates. For sudden discrete frequency changes, there are a number of techniques based on hidden Markov models to generate the discontinuous transitions between different frequency values [21]. In each case, it is often helpful

to have a model for the dynamics of the time dependence of the system to reduce any residual errors in the tracked frequency values.

The frequency estimation techniques can be generalized if more than one frequency is present in the measurement signal. The MUSIC technique, in particular, is intended to allow multiple frequencies to be estimated. However, once measurements are made of more than one signal, potential problems can occur and slightly more thought is required when tracking multiple signals. When two frequencies are close together there are often problems deciding which measurement is associated with which tracked frequency. This is a well-known problem in object tracking. The misassociation of measurements with tracks can lead to significant errors in the parameter estimates. For algorithms like the Kalman filter, additional information regarding the expected errors in the parameters being tracked (the frequencies in this case) propagates with the estimates. The expected errors can then be used to minimize the probability of misassociation by finding the most likely association of measurements with tracks based on a statistically weighted distance (e.g., the Mahalanobis distance [27]).

V. CONCLUSIONS

We have considered a number of frequency estimation techniques that can be used with continuous quantum measurement processes. The techniques presented have been used to estimate the Hamiltonian for a single qubit, but they are quite general and can be readily adapted to monitor more complicated continuous measurement processes. The qubit has one parameter, the oscillation frequency for rotations around the Bloch x axis. The most accurate estimates for the qubit frequency were generated by a Bayesian inference technique, but this technique is computationally expensive and typically requires thousands of integration steps per qubit cycle. By contrast, the three classical frequency estimation techniques considered—maximized periodogram, Quinn-Fernandes, and MUSIC algorithms—all performed relatively well using only 50 integration steps per qubit cycle and one (MUSIC) was found to be robust to measurement noise and produced accurate frequency estimates for a wide range of measurement strengths. The classical frequency estimation algorithms are much less computationally intensive than Bayesian inference, partly due to the fact that they can be implemented with much larger time steps. They provide frequency estimates that are accurate enough to generate robust (high fidelity) conditional state estimates for the qubit. We have also considered how these frequency estimates can be integrated into a tracking system for a time-dependent Hamiltonian, where the qubit frequency undergoes continuous drifts or sudden discontinuous changes.

ACKNOWLEDGMENTS

The authors would like to thank Professor B. G. Quinn for helpful and informative communications. This work was performed in part using the supercomputing facilities in the College of Science and Mathematics at UMass Boston. We are also grateful to Daniel Steck for allowing us to use his

parallel cluster in the Oregon Center for Optics, University of Oregon. J.F.R. and C.D.H. would like to acknowledge the support of an ESPRC Grant EP/C012674/1. K.J. is supported by the NSF under Project Nos. PHY-0902906 and PHY-1005571. This research is also partially supported by the ARO MURI Grant W911NF-11-1-0268, and was performed in part using the supercomputing facilities in the School of Science and Mathematics at the University of Massachusetts Boston. C.D.H. would also like to acknowledge the support of the Australian Research Council Centre of Excellence Scheme (CE110001027), the Australian Government, and the US National Security Agency (NSA) and the Army Research Office (ARO) under Contract No. W911NF-08-1-0527.

APPENDIX A: DERIVING THE HYBRID MASTER EQUATION

The observer starts with a state of knowledge defined by her density matrix for the system ρ and her probability density for the parameters that define the Hamiltonian of the system. For simplicity we will assume that there is only one real parameter λ , so that the probability density is $P(\lambda)$. Generalizing the result to any number of parameters is straightforward.

If the observer knew λ precisely, then upon obtaining the measurement result $dr = \langle \hat{A} \rangle dt + dW/\sqrt{8k}$, corresponding to a Hermitian operator \hat{A} , the evolution of the density matrix would be given by the stochastic master equation [1]

$$d\rho_\lambda = \frac{-i}{\hbar} [\hat{H}(\lambda), \rho] dt - k[\hat{A}, [\hat{A}, \rho]] dt + 4k(\hat{A}\rho + \rho\hat{A} - 2\langle \hat{A} \rangle \rho)(dr - \langle \hat{A} \rangle dt), \quad (\text{A1})$$

so that the observers density matrix at time $t + dt$ is $\rho_\lambda(t + dt) = \rho + d\rho_\lambda$. Since the observer does not know which value λ takes, she must average over all the possible $\rho_\lambda(t + dt)$. Thus her state of knowledge at time $t + dt$ is actually

$$\begin{aligned} \rho(t + dt) &= \int_{-\infty}^{\infty} P(\lambda) \rho_\lambda(t + dt) d\lambda \\ &= \rho + \int_{-\infty}^{\infty} [P(\lambda) d\rho_\lambda] d\lambda \\ &\equiv \rho + d\rho. \end{aligned} \quad (\text{A2})$$

The master equation for the observer's density matrix is therefore

$$d\rho = \frac{-i}{\hbar} \left[\int_{-\infty}^{\infty} P(\lambda) \hat{H}(\lambda) d\lambda, \rho \right] dt - k[\hat{A}, [\hat{A}, \rho]] dt + 4k(\hat{A}\rho + \rho\hat{A} - 2\langle \hat{A} \rangle \rho)(dr - \langle \hat{A} \rangle dt). \quad (\text{A3})$$

We now have the equation that updates the observer's density matrix, but we need also the equation that updates the observers probability density $P(\lambda)$. To derive this we must use Bayes' theorem [45]. This tells us that the observer's new state of knowledge after obtaining the measurement result dr is

$$P(\lambda|dr) = P(dr|\lambda)P(\lambda)/\mathcal{N}, \quad (\text{A4})$$

in which $P(dr|\lambda)$ is the probability density for dr given λ , and \mathcal{N} is a normalization constant. If the value of the parameter was λ , then the measurement record would be given by $dr = \langle \hat{A} \rangle_\lambda dt + dW/\sqrt{8k}$, where $\langle \hat{A} \rangle_\lambda = \text{Tr}[\hat{A}\rho_\lambda]$, and ρ_λ

is obtained by using the update equation (A1). Since dW is Gaussian distributed with mean 0 and variance dt , $P(dr|\lambda)$ is

$$P(dr|\lambda) = \frac{e^{-4k(dr - \langle \hat{A} \rangle_\lambda dt)^2/dt}}{\sqrt{2\pi dt}}. \quad (\text{A5})$$

The updated probability density is thus

$$P(\lambda|dr) \propto e^{-4k(dr - \langle \hat{A} \rangle_\lambda dt)^2/dt} P(\lambda), \quad (\text{A6})$$

where dr is the actual measurement record obtained by the observer, $dr = \langle \hat{A} \rangle_\lambda dt + dW/\sqrt{8k}$. We now expand this to first order in dt , remembering that $dW^2 = dt$ [1]. This gives

$$\begin{aligned} P(\lambda|dr) &\propto e^{-4k(\langle \hat{A} \rangle_\lambda dt - \langle \hat{A} \rangle_\lambda dt + dW/\sqrt{8k})^2/dt} P(\lambda) \\ &\propto e^{-4k(\langle \hat{A} \rangle_\lambda - \langle \hat{A} \rangle)^2 dt + \sqrt{8k}(\langle \hat{A} \rangle_\lambda - \langle \hat{A} \rangle) dW} P(\lambda) \\ &\propto [1 + \sqrt{8k}(\langle \hat{A} \rangle_\lambda - \langle \hat{A} \rangle) dW] P(\lambda). \end{aligned} \quad (\text{A7})$$

The stochastic differential equation that updates the density $P(\lambda)$ is therefore

$$\begin{aligned} dP(\lambda) &= \sqrt{8k}(\langle \hat{A} \rangle_\lambda - \langle \hat{A} \rangle) dW P(\lambda) \\ &= 8k(\langle \hat{A} \rangle_\lambda - \langle \hat{A} \rangle)(dr - \langle \hat{A} \rangle dt) P(\lambda). \end{aligned} \quad (\text{A8})$$

To be able to calculate this update to $P(\lambda)$ we must know $\langle \hat{A} \rangle_\lambda$ for *all* λ . To track the quantum state along with the parameter λ we must therefore solve the master equation for ρ [Eq. (A2)], the master equations for the ρ_λ for every value of λ [Eq. (A1)], and the update equation for $P(\lambda)$ [Eq. (A8)], which is an example of a *Kushner-Stratonovich* equation.

APPENDIX B: NOTCH FILTERING AND THE QUINN-FERNANDES METHOD

For a discrete time series x_n , $n = 0, 1, 2, \dots$, a filter y_n can be constructed using a number of values for the input signal (feed forward) and a number of previous values for the filter output (feedback). A general linear filter is given by

$$\begin{aligned} y_n &= a_0 x_n + a_1 x_{n-1} + a_2 x_{n-2} + \dots \\ &+ b_1 y_{n-1} + b_2 y_{n-2} + \dots \end{aligned} \quad (\text{B1})$$

The feed-forward coefficients a_0, a_1, \dots specify a finite duration impulse response, that is, if the input signal is a single impulse, the output signal returns to the default value (usually zero) after a finite number of time steps. Such a filter is normally referred to as a *finite impulse response* or FIR filter. The feedback coefficients b_1, b_2, \dots define an infinite impulse response or IIR filter, because the output signal only returns to zero asymptotically.

The properties of this type of filter are normally analyzed using the z transform (more precisely, the *unilateral* z transform) [44],

$$\mathcal{X}(z) = \sum_{n=0}^{\infty} x_n z^{-n}, \quad (\text{B2})$$

where z is a complex number such that $z = |z|e^{i\omega}$, where ω is the (angular) frequency of the signal. Each factor of z^{-1} represents a time delay of one time step. The z transform plays a similar role in the analysis of discrete time signals to that

of the Laplace transform for continuous and analog signals. Applying the z transform to Eq. (B1), we obtain

$$\begin{aligned} \mathcal{Y}(z) &= (a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots) \mathcal{X}(z) \\ &+ (b_1 z^{-1} + b_2 z^{-2} + \dots) \mathcal{Y}(z), \end{aligned} \quad (\text{B3})$$

which can be rearranged to find the *transfer function* for the filter

$$\mathcal{H}(z) = \frac{\mathcal{Y}(z)}{\mathcal{X}(z)} = \frac{(a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots)}{(b_1 z^{-1} + b_2 z^{-2} + \dots)} = \frac{\mathcal{A}(z)}{\mathcal{B}(z)}.$$

The two polynomials $\mathcal{A}(z)$ and $\mathcal{B}(z)$, which represent the feed-forward and feedback parts of the filter, respectively, contain all of the relevant properties for this type of linear filter. The zeros of $\mathcal{A}(z)$ determine which frequency components are suppressed by the filter if a zero of the polynomial lies on the unit circle in the z plane, $\mathcal{A}(e^{i\omega_0}) = 0$, the frequency ω_0 is removed. The zeros of $\mathcal{B}(z)$ give rise to poles in $\mathcal{H}(z)$ and determine the stability of the filter. For a stable filter, all of the poles of $\mathcal{H}(z)$ must lie within the unit circle $|z| < 1$. When a pole lies close to the unit circle, the frequencies close to the pole are enhanced.

The case considered in Sec. III B consists of a single sinusoid and a noise source,

$$x_n = A \cos(n\omega\Delta t + \phi) + \epsilon_n \quad (\text{B4})$$

for $n = 0, \dots, N-1$, where A and ϕ are the amplitude and phase of the sinusoidal signal and ϵ_n is the noise. In the Fourier domain this represents a single peak (from the sinusoid) and a large noisy background. The Quinn-Fernandes technique [22,43] introduces two filters, one IIR filter and one FIR filter. The IIR filter has a pole which enhances one particular frequency $\mathcal{B}(e^{i\omega}) = 0$. The FIR filter is designed to have zero response at the same frequency [the notch frequency, $\mathcal{A}(e^{i\omega}) = 0$]. The variance of the output signal from the combined filter is a minimum when the filter frequency matches the frequency of the sinusoid. The Quinn-Fernandes algorithm provides a means to find the sinusoid frequency and, importantly for the computational efficiency, it operates purely in the time domain.

If the signal is given by Eq. (B4), it should satisfy a filter equation,

$$x_n - 2 \cos(\omega\Delta t)x_{n-1} + x_{n-2} = \epsilon_n - 2 \cos(\omega\Delta t)\epsilon_{n-1} + \epsilon_{n-2}. \quad (\text{B5})$$

Rewriting Eq. (B5) as

$$x_n - \alpha x_{n-1} + x_{n-2} = \epsilon_n - \beta \epsilon_{n-1} + \epsilon_{n-2} \quad (\text{B6})$$

subject to the constraint $\alpha = \beta$, we start with an initial estimate ω_1 for ω , and set $\alpha_1 = 2 \cos(\omega_1 \Delta t)$ along with an index $j = 1$. The data is then passed through the IIR filter to generate a new record $\zeta_{n,j}$,

$$\zeta_{n,j} = x_n - \alpha_j \zeta_{n-1,j} + \zeta_{n-2,j}, \quad (\text{B7})$$

where $\zeta_{n,j} = 0$ for $n < 0$. An improved estimate for β is then constructed by minimizing the variance of the output from an FIR filter $e_n = \zeta_{n,j} - \beta \zeta_{n-1,j} + \zeta_{n-2,j}$ with respect to β , giving

$$\beta_j = \frac{\sum_{n=0}^{N-1} (\zeta_{n,j} + \zeta_{n-2,j}) \zeta_{n-1,j}}{\sum_{n=0}^{N-1} \zeta_{n-1,j}^2}. \quad (\text{B8})$$

If $|\alpha_j - \beta_j|$ is larger than some threshold then the β value replaces the current α value $\alpha_{j+1} = \beta_j$, and the process is applied again. If the difference between α and β is small enough then the iteration in j is stopped and the signal frequency is given as $\omega = \cos^{-1}(\frac{1}{2}\beta_j)/\Delta t$. In most situations the number of iterations in j is fairly small, typically a few to ten iterations is sufficient to obtain a good frequency estimate, and the algorithm is also relatively insensitive to errors

in the initial frequency estimate. However, these properties are normally quoted for systems with higher signal-to-noise ratios than those considered here, and that do not generate a backaction on the signal from the measurement process. In these slightly more benign situations the Quinn-Fernandes technique can be shown to generate errors which scale as $O(N^{-3})$ and approaches the Cramer-Rao bound, as long as the error in the initial estimate is fairly small, $O(N^{-1})$.

-
- [1] K. Jacobs and D. Steck, *Contemp. Phys.* **47**, 279 (2006).
- [2] T. A. Brun, *Am. J. Phys.* **70**, 719 (2002).
- [3] L. Bouten, R. van Handel, and M. James, *SIAM J. Control Optim.* **46**, 2199 (2007).
- [4] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (CUP, Cambridge, 2010).
- [5] A. C. Doherty and K. Jacobs, *Phys. Rev. A* **60**, 2700 (1999).
- [6] M. R. James, *Phys. Rev. A* **69**, 032108 (2004).
- [7] J. F. Ralph, E. J. Griffith, T. D. Clark, and M. J. Everitt, *Phys. Rev. B* **70**, 214521 (2004).
- [8] A. N. Jordan and A. N. Korotkov, *Phys. Rev. B* **74**, 085307 (2006).
- [9] J. Gambetta and H. M. Wiseman, *Phys. Rev. A* **64**, 042105 (2001).
- [10] F. Verstraete, A. C. Doherty, and H. Mabuchi, *Phys. Rev. A* **64**, 032111 (2001).
- [11] J. K. Stockton, J. M. Geremia, A. C. Doherty, and H. Mabuchi, *Phys. Rev. A* **69**, 032109 (2004).
- [12] B. A. Chase and J. M. Geremia, *Phys. Rev. A* **79**, 022314 (2009).
- [13] M. Tsang, *Phys. Rev. Lett.* **102**, 250403 (2009).
- [14] M. Tsang, *Phys. Rev. A* **80**, 033840 (2009).
- [15] M. Tsang, *Phys. Rev. A* **81**, 013824 (2010).
- [16] V. B. Braginsky, F. Y. Khalili, and K. S. Thorne, *Quantum Measurement* (CUP, Cambridge, 1995).
- [17] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, *IEEE Trans. Sig. Proc.* **50**, 174 (2002).
- [18] N. Yamamoto, *Phys. Rev. A* **74**, 032107 (2006).
- [19] A. Doherty, J. Doyle, H. Mabuchi, K. Jacobs, and S. Habib, in *Proceedings of the 39th IEEE Conference on Decision and Control*, Vol. 1 (IEEE, Washington DC, 2000), p. 949.
- [20] C. D'Helon and M. R. James, *Phys. Rev. A* **73**, 053803 (2006).
- [21] B. Quinn and E. Hannan, *The Estimation and Tracking of Frequency* (Cambridge University Press, Cambridge, UK, 2001).
- [22] B. G. Quinn and J. M. Fernandes, *Biometrika* **78**, 489 (1991).
- [23] B. M. Kaveh and A. J. Barabell, *IEEE Trans. ASSP* **34**, 331 (1986).
- [24] S. M. Kay, *Modern Spectral Estimation: Theory and Application* (Prentice-Hall, Englewood Cliffs, NJ, 1988).
- [25] H. M. Wiseman, *Quantum Semiclass. Opt.* **8**, 205 (1996).
- [26] T. P. Spiller, *Phys. Lett. A* **192**, 163 (1994).
- [27] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation: Theory Algorithms and Software* (Wiley-Interscience, New York, 2001).
- [28] A. C. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. M. Tan, *Phys. Rev. A* **62**, 012105 (2000).
- [29] H. M. Wiseman and A. C. Doherty, *Phys. Rev. Lett.* **94**, 070405 (2005).
- [30] T. P. Spiller, B. M. Garraway, and I. C. Percival, *Phys. Lett. A* **179**, 63 (1993).
- [31] E. J. Griffith, C. D. Hill, J. F. Ralph, H. M. Wiseman, and K. Jacobs, *Phys. Rev. B* **75**, 014511 (2007).
- [32] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, 1997).
- [33] K. Jacobs, *Phys. Rev. A* **67**, 030301(R) (2003).
- [34] J. Combes and K. Jacobs, *Phys. Rev. Lett.* **96**, 010504 (2006).
- [35] H. M. Wiseman and J. F. Ralph, *New J. Phys.* **8**, 90 (2006).
- [36] N. A. Peters, T.-C. Wei, and P. G. Kwiat, *Phys. Rev. A* **70**, 052309 (2004).
- [37] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).
- [38] E. T. Jaynes, *Probability Theory: The Logic of Science* (CUP, Cambridge, 2003).
- [39] K. Jacobs, *Quantum Inf. Proc.* **1**, 73 (2002).
- [40] B. P. Lathi, *Modern Digital and Analog Communication Systems* (Oxford University Press, Oxford, 1998).
- [41] W. H. Long, D. H. Mooney, and W. A. Skillman, in *Radar Handbook*, 2nd ed., edited by M. Skolnik (McGraw-Hill, New York, 1990).
- [42] D. Huang, *Stat. Sinica* **10**, 157 (2000).
- [43] T.-H. Li and B. Kedem, *J. Time Ser. Anal.* **19**, 69 (1998).
- [44] T. J. Cavicchi, *Digital Signal Processing* (Wiley, New York, 2000).
- [45] R. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis* (Addison-Wesley, Sydney, 1973).