PHYSICAL REVIEW A, VOLUME 64, 032111

Sensitivity optimization in quantum parameter estimation

F. Verstraete,[1,2] A. C. Doherty,[1] and H. Mabuchi[1]
[1]*Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125*
[2]*SISTA/ESAT, Katholieke Universiteit Leuven, K. Mercierlaan 94, Leuven, Belgium*
(Received 24 April 2001; published 17 August 2001)

We present a general framework for sensitivity optimization in quantum parameter estimation schemes based on continuous (indirect) observation of a dynamical system. As an illustrative example, we analyze the canonical scenario of monitoring the position of a free mass or harmonic oscillator to detect weak classical forces. We show that our framework allows the consideration of *sensitivity scheduling*, as well as estimation strategies for nonstationary signals, leading us to propose corresponding generalizations of the standard quantum limit for force detection.

DOI: 10.1103/PhysRevA.64.032111 PACS number(s): 03.65.Ta, 03.65.Yz, 42.50.Lc

The primary motivation for work presented in this paper has been to contribute to the continuing integration of quantum measurement theory with traditional (engineering) disciplines of measurement and control. Various researchers engaged in this endeavor have found that the concepts and methods of theoretical engineering provide a fresh perspective on how differences and relationships between quantum and classical metrology can be most cleanly understood. This approach has been especially fruitful in scenarios involving *continuous* measurement, for which a number of important physical insights and results of practical utility follow simply from the formal connections between quantum trajectory theory and Kalman filtering [1–7].

Here we describe a general formalism for parameter estimation via continuous quantum measurement, whose equations are amenable to analytic and numerical optimization strategies. In addition to being useful for practical design of quantum measurements, we find that this approach sharpens our understanding of the significance and origin of standard quantum limits (SQL's) in precision metrology. Following the basic notion that the "standard limit" for any measurement scenario should be derivable by optimization over some parametric family of "standard" measurement strategies, we present results that generalize the SQL for force estimation through continuous monitoring of the position of a test mass. Our analysis shows that the canonical expression for the force SQL in continuous position measurement stems from a rather arbitrary limitation of the set of allowable measurement strategies to those with constant sensitivity, and we find that a lower expression (by a factor of 3/4) can be obtained when time variations are allowed. It follows that further expansions of the optimization space (such as adaptive measurements with real-time feedback [1]) should be considered in order to arrive at an SQL that consistently accounts for a natural set of measurement strategies that are "practically equivalent" in terms of inherent experimental difficulty.

For clarity, the main results of this paper are presented in the first and third sections within the concrete context of force estimation via continuous position measurement. In order to emphasize the general nature of our formalism and the conclusions we derive from it, the second section provides a more abstract development that arrives at all the equations needed for sensitivity optimization in a broad class of continuous measurement scenarios. As this general treatment is rather technical, we note that it is not crucial to the overall logical flow of the paper. Very recently, Gambetta and Wiseman have discussed a similar approach to parameter estimation for resonance fluorescence of a two-level atom, paying particular attention to how information about the unknown parameter, and also about the quantum state, changes with different kinds of measurements [8].

I. FORCE ESTIMATION BY CONTINUOUS MEASUREMENT OF POSITION

The aim of this section is to present a formalism for continuous parameter estimation in the specific context of a harmonic oscillator subject to an unknown force linear in $\hat{x}$. This section gives a rigorous and a more general treatment of the ideas previously worked out by one of us [4]. We first derive the conditional evolution equations for the oscillator under continuous position measurement, then discuss their control-theoretic interpretation as Kalman filtering equations. We then show how a Bayesian parameter estimator can be obtained from the Kalman filter in this scenario.

A. Conditional evolution equations

We will derive the equations of motion of a continuously observed system conditioned on the measurement record. Our treatment is based on the model of continuous measurement of Caves and Milburn [9], which in turn was based on work of Barchielli *et al.* [10]. Their derivation is solely based on the standard techniques of operations and effects in quantum mechanics, which makes it very transparent. Similar results could have been obtained by making use of the quantum-stochastic calculus of Hudson and Parthasarathy [11] as was done by Belavkin and Staszewski [12].

In continuous measurement—often an accurate description of experimentally realizable measurements—projective collapse of the wave function, and hence also the Zeno effect, can be avoided by continually performing infinitesimally weak measurements. A weak measurement consists of weakly coupling the system under interest to a (quantum-mechanical) meter, followed by a von Neumann measurement of the meter state. As there was only a weak coupling,

 ©2001 The American Physical Society

only very little information about the system of interest is revealed and there will only be a limited amount of back action. At first, we will introduce the concept of weak measurements in the framework of position measurement. Then we will show how to derive the equations of motion for a quantum particle subject to a whole series of weak measurements. The treatment of continuous measurements will then be obtained by taking appropriate limits.

The aim of a weak position measurement is to get some information out of the system, although without disturbing it too much. This can be done by applying an operation valued measure $\{\hat{A}_\xi(\hat{x})\}$ where there is a lot of overlap between the $\hat{A}_\xi(\hat{x})$ associated with different measurement results ξ. This overlap is proportional to the variance of the measurement outcome, but inversely proportional to the variance of the back-action noise. As shown by Braginsky and Khalili [14], the product of those variances always exceeds $\hbar^2/4$. Equality is achieved if and only if $\hat{A}_\xi(\hat{x})$ is Gaussian in $\hat{x}$. As we are interested in the ultimate limits imposed by quantum mechanics, we will assume our measurement device is optimally constructed so as to yield a Gaussian $\hat{A}_\xi(\hat{x})$:

$$\hat{A}_\xi(\hat{x})=\frac{1}{(\pi D)^{1/4}}\exp\left(-\frac{(\xi-\hat{x})^2}{2D}\right).$$

This is equivalent to the model of Barchielli and also of Caves and Milburn [9] who obtained it by explicitly working out the case of linear coupling between a (Gaussian) meter and the particle followed by a von Neumann measurement on the meter.

We will now assume that the wave function of the observed particle is also Gaussian. This is a reasonable assumption as we will soon take the limit of many Gaussian measurements, each of which effects a Gaussian "conditioning" of the particle's wave function. Ultimately, the wave function itself will become Gaussian, whatever its original shape. We furthermore assume that the Hamiltonian of the unobserved particle would be given by

$$H_0=\frac{\hat{p}^2}{2m}+\frac{m\omega^2}{2}\hat{x}^2+\theta\hat{x}, \tag{1}$$

where θ is the (eventually time-dependent) force to be estimated. It will turn out to be very useful to parametrize the Gaussian wave function of the particle by a complex mean $\tilde{x}=\tilde{x}_r+i\tilde{x}_i$ and complex variance $\tilde{\sigma}=\tilde{\sigma}_r+i\tilde{\sigma}_i$ (throughout the paper, the notation σ instead of σ^2 will be used to denote the variance):

$$|\psi\rangle=|\tilde{x}(t),\tilde{\sigma}(t)\rangle, \tag{2}$$

$$\langle x|\psi\rangle=\left(\frac{\tilde{\sigma}_r}{\pi|\tilde{\sigma}|^2}\right)^{1/4}\exp\left(-\frac{(x-\tilde{x})^2}{2\tilde{\sigma}}-\frac{\tilde{x}_i^2}{2\tilde{\sigma}_r}\right),$$

$$\overline{x}=\tilde{x}_r+\frac{\tilde{\sigma}_i}{\tilde{\sigma}_r}\tilde{x}_i,\quad \overline{p}=\hbar\frac{\tilde{x}_i}{\tilde{\sigma}_r},$$

$$\overline{\Delta x^2}=\frac{|\tilde{\sigma}|^2}{2\tilde{\sigma}_r},\quad \overline{\Delta p^2}=\frac{\hbar^2}{2\tilde{\sigma}_r},\quad \overline{\Delta x\Delta p+\Delta p\Delta x}=\frac{\hbar\tilde{\sigma}_i}{\tilde{\sigma}_r}.$$

The values of these quantities will in general depend on the value of θ. In this section, we will supress this dependence, but in the following, we will denote the mean position conditioned on a particular value of θ by $\overline{x}_\theta$ and likewise for the other expectation values. We will now derive the dynamics of this state if a measurement takes place at time τ. From time 0 to τ^-, just before the measurement, the equations of motion are governed by the Schrödinger equation:

$$\frac{d\tilde{\sigma}}{dt}=\frac{i\hbar}{m}\left(1-\frac{m^2\omega^2}{\hbar^2}\tilde{\sigma}(t)^2\right),\quad \frac{d\tilde{x}}{dt}=\frac{\tilde{\sigma}(t)}{i\hbar}(\theta+m\omega^2\tilde{x}). \tag{3}$$

The corresponding $\overline{x}$, $\overline{p}$, and second-order moments can easily be derived. The equation for $\tilde{\sigma}$ indicates the expanding and contracting of the wavepacket induced by the harmonic oscillation. At time τ, the operation valued measure $\{\hat{A}_\xi(\hat{x})\}$ is performed. ξ will be a Gaussian-distributed random variable with expectation value $\overline{x(\tau^-)}$ and variance $D+\overline{\Delta x^2}(\tau^-)$. Straightforward calculations show that the post-measurement wave function, conditioned on the result ξ, is parametrized by

$$\frac{1}{\tilde{\sigma}(\tau)}=\frac{1}{\tilde{\sigma}(\tau^-)}+\frac{1}{D},\quad \tilde{x}_\xi(\tau)=\frac{\tilde{\sigma}(\tau^-)\xi+D\tilde{x}(\tau^-)}{\tilde{\sigma}(\tau^-)+D}. \tag{4}$$

The equation for $\tilde{\sigma}$ now indicates the contracting effect of the position measurement. The expectation values $\overline{x}$ and $\overline{p}$ become

$$\overline{x(\tau)}=\overline{x(\tau^-)}+\frac{|\tilde{\sigma}(\tau)|^2}{\tilde{\sigma}_r(\tau)D}[\xi-\overline{x(\tau^-)}],$$

$$\overline{p(\tau)}=\overline{p(\tau^-)}+\frac{\hbar\tilde{\sigma}_i(\tau)}{D\tilde{\sigma}_r(\tau)}[\xi-\overline{x(\tau^-)}]. \tag{5}$$

Note that the wave function collapses manifest themselves by periodically shifting the center of the wave packet through the white noise terms proportional to $\xi-\overline{x(\tau^-)}$.

It is trivial to write down the dynamical equations in the case of a finite number (N) of measurements: we just have to repeat the previous two-stage procedure N times. However, we are interested in taking the limit of infinitesimal time intervals dt between two measurements. This will only make sense if at each infinitesimal time step the wave function is only subject to an infinitesimal disturbance. Referring to Eq. (4), this implies that the measurement accuracy D has to scale as $1/dt$. Therefore, we define the finite sensitivity k by the relation $D=1/(kdt)$, implying that only an infinitesimal amount of information is obtained in each measurement. In this limit, the random zero-mean variable $[\xi-\overline{x(\tau^-)}]/D$ has a standard deviation given by $\sqrt{kdt/2}$. This is very convenient as a Gaussian random variable with zero mean and variance $\sqrt{dt}$ is by definition a Wiener increment, and therefore, we can make use of the theory of Ito calculus. Defining $d\Xi(t)=\xi_t dt$ as being the measurement record, and using the notation of Ito calculus, the complete equations of motion conditioned on the measurement result for a Gaussian particle subject to continuous observation of the position can be written as

$$d\Xi(t)=\overline{x(t)}dt+v_{\xi}(t)dW, \tag{6}$$

$$d\overline{x(t)}=\frac{\overline{p(t)}}{m}dt+v_x(t)dW, \tag{7}$$

$$d\overline{p(t)}=-m\omega^2\overline{x(t)}dt-\theta(t)dt+v_p(t)dW, \tag{8}$$

$$\dot{\tilde{\sigma}}(t)=\frac{i\hbar}{m}\left(1-\frac{m^2\omega^2}{\hbar^2}\tilde{\sigma}(t)^2\right)-k(t)\tilde{\sigma}(t)^2, \tag{9}$$

$$v_x(t)=\sqrt{\frac{k(t)}{2}}\frac{|\tilde{\sigma}(t)|^2}{\tilde{\sigma}_r(t)},\quad v_p(t)=\sqrt{\frac{k(t)}{2}}\frac{\hbar\tilde{\sigma}_i(t)}{\tilde{\sigma}_r(t)},$$

$$v_{\xi}(t)=\frac{1}{\sqrt{2k(t)}}. \tag{10}$$

If the sensitivity k is kept constant during the whole observation $[\forall t, k(t)=k(0)]$, Eq. (9) can be solved exactly. Given initial condition $\tilde{\sigma}_0$, the solution is

$$\tilde{\sigma}(t)=\tilde{\sigma}_\infty\left(\frac{[(\tilde{\sigma}_\infty+\tilde{\sigma}_0)/(\tilde{\sigma}_\infty-\tilde{\sigma}_0)]\exp(2i\Omega t)-1}{[(\tilde{\sigma}_\infty+\tilde{\sigma}_0)/(\tilde{\sigma}_\infty-\tilde{\sigma}_0)]\exp(2i\Omega t)+1}\right),$$

$$\Omega=\sqrt{\omega^2-\frac{i\hbar k}{m}},\quad \tilde{\sigma}_\infty=\frac{\hbar/m}{\Omega}. \tag{11}$$

This shows that the position variance of the wave function evolves at least exponentially fast to a steady state. The damping is roughly proportional to the square root of the sensitivity, while the steady-state solution has a variance inversely proportional to it. This result means that a continuously observed particle is localized, although not confined, in space. It is interesting to note that this localization increases with the mass of the particle, such that it is very difficult to localize a light particle. Indeed, the steady-state position variance can be understood from the point of view of standard quantum limits for position measurement [14]. For example if $\omega^2 \gg \hbar k/m$ then $\overline{\Delta x^2}_\infty \simeq \hbar/2m\omega$. Similarly, if we take $t=1/\mathrm{Re}[\Omega]$ to be the time for an effectively complete measurement, then for a free particle $\overline{\Delta x^2}_\infty=\hbar t/m$ and so the steady-state position variance is the same as the SQL for ideal position measurements separated by time intervals of length $1/\mathrm{Re}[\Omega]$.

B. Kalman filtering interpretation

Let us now try to give a "signal processing" interpretation to Eqs. (6)–(10). The Wiener increment was defined as the difference between the actual and the expected measurement result. As it is white noise, it is clear that the expected measurement result was actually the best possible guess for the result. This is reminiscent to the innovation process in classical control theory: the optimal filtering equations of a classical stochastic process can be obtained by imposing that the difference between the actual and expected (i.e., filtered) measurement be white noise. Indeed, in a previous paper [5], one of us noticed that Eqs. (6)–(10) have exactly the structure of the Kalman filtering equations associated with a classical stochastic linear system. This is in complete accordance with the dynamical interpretation of quantum mechanics as describing the evolution of our knowledge about the system.

The classical stochastic system that has exactly the same filtering equations as our continuously observed quantum system is given by

$$d\begin{pmatrix}x_\theta\\ p_\theta\end{pmatrix}=\begin{pmatrix}0 & \frac{1}{m}\\ -m\omega^2 & 0\end{pmatrix}\begin{pmatrix}x_\theta\\ p_\theta\end{pmatrix}dt+\begin{pmatrix}0\\ 1\end{pmatrix}\theta(t)dt$$
$$+\begin{pmatrix}0\\ \hbar/2\end{pmatrix}\sqrt{2k}dV_1,$$

$$d\Xi=(1\quad 0)\begin{pmatrix}x_\theta\\ p_\theta\end{pmatrix}dt+\frac{1}{\sqrt{2k}}dV_2. \tag{12}$$

dV_1 and dV_2 are two independent Wiener increments and correspond to the process noise and measurement noise, respectively. It is very enlightening to look at the corresponding weights of these noise processes: the higher the sensitivity, the more accurate the measurements, but the more noise is introduced into the system. Moreover, measuring the position only introduces noise into the momentum. This clearly is a succinct manifestation of the Heisenberg uncertainty relation. Indeed, the product of the amplitude of the measurement noise process and the back-action noise is independent of the sensitivity k and exactly given by $\hbar/2$. The close relation between the quantum mechanical and classical problems becomes even more evident when one realizes that the first system of equations for a classical position and momentum has precisely the same form as the quantum langevin equations for this system (see, for example, [13,12]). The equation for the measurement process is then seen to have the same form as the input-output relations [13] for such a position measurement. The quantum equations are obtained simply by reading x, p, dV_1, and dV_2 as operators. The quantum stochastic increments dV_1 and dV_2 arise from the coupling of the quantum system to the meter environment and are noncommuting, $[dV_2, dV_1]=dt$.

The equations for the means $\overline{x_\theta}$ and $\overline{p_\theta}$ are now given by the Kalman filter equations of this classical system, and the equations for the variances $\overline{\Delta x_\theta^2}, \overline{\Delta p_\theta^2}, \overline{\Delta x_\theta \Delta p_\theta+\Delta p_\theta \Delta x_\theta}$ are given by the associated Riccati equations. This is very convenient as this will allow us to use the language of classical control theory to solve the estimation problem.

C. Continuous parameter estimation

Let us now consider the basic question of this paper: how can we get the best estimates of the unknown force $\{\theta(t)\}$ acting on the system, given the measurement record $\{d\Xi_t\}$? The natural way to attack this problem is the use of Bayes' rule. As we have a linear system with $\{d\Xi_t\}$ a linear function of $\{\theta(t)\}$, and the noise in the system is Gaussian, this will lead to a Gaussian distribution in $\{\theta(t)\}$. Moreover, due to the linearity, the second-order moments of this distribution will be independent of the actual measurement record. Therefore, the accuracy of our estimates will only be a function of the sensitivity chosen during the observation process and of the prior knowledge we have about the signal $\{\theta(t)\}$ (for example, that it is constant). This will allow us to devise optimal measurement strategies.

The formalism that we have developed is particularly useful in the case where we parametrize $\{\theta(t)\}$ as a linear combination of known time-dependent functions $\{f_i(t)\}$, but with unknown weights $\{\theta_i\}$:

$$\theta(t)=\sum_{i=1}^{n}\theta_i f_i(t). \tag{13}$$

The estimation, based on Bayes' rule, will lead to a joint Gaussian distribution in the parameters $\{\theta_i\}$. Indeed, we have the relations

$$\begin{aligned}&p(\{\theta_i\}|\{\xi(t+dt)\})\\&\quad\sim p(d\Xi(t)|\{\theta_i\},\{\xi(t)\})p(\{\theta_i\}|\{\xi(t)\})\\&\quad\sim p(d\Xi(t)|\overline{x_{\{\theta_i\}}}(t)\{\xi(t)\}])p(\{\theta_i\}|\{\xi(t)\}).\end{aligned} \tag{14}$$

In the last step, we made use of the fact that the Kalman estimate $\overline{x_{\{\theta_i\}}}(t)$ is a sufficient statistic for $d\Xi(t)$. Moreover, all distributions are Gaussian, while $\overline{x_{\{\theta_i\}}}(t)$ is some linear function of $\{\theta_i\}$ due to the linear character of the Kalman filter:

$$\overline{x_{\{\theta_i\}}}(t)=\sum_i \theta_i\int_0^t dt' g(t,t')f_i(t'). \tag{15}$$

The function $g(t,t')$ can easily be calculated using Eqs. (6)–(10). To obtain the variance of the optimal estimates of $\{\theta_i\}$, formula (14) has to be applied recursively. By explicitly writing out the Gaussian distributions, and making use of the fact that the product of Gaussians is still a Gaussian, it is then easy to show that the variances at time τ are given by

$$\frac{1}{\sigma_{\theta_i}}=\int_0^\tau\frac{dt}{v_\xi^2(t)}\left(\int_0^t dt' g(t,t')f_i(t')\right)^2. \tag{16}$$

A more intuitive way of obtaining the same optimal estimation, given a fixed measurement strategy, of $\{\theta_i\}$ can be obtained by a little trick: we can enlarge the state vector (x_θ,p_θ) with the unknowns, and construct the Kalman filter and Riccati equation of the new enlarged system. $\overline{x_\theta}$ and $\overline{p_\theta}$, till now the expected values conditioned on a fixed value of the force, then get the meaning of the mean of these expected values over the probability distribution of the unknown force. In other words, the new $\overline{x}$ and $\overline{p}$ become the ensemble averages over the pure states labeled by a fixed force θ. The new enlarged system, in the case of one unknown parameter θ, reads

$$\begin{aligned}d\begin{pmatrix}x\\p\\\theta\end{pmatrix}=&\underbrace{\begin{pmatrix}0&1/m&0\\-m\omega^2&0&f(t)\\0&0&0\end{pmatrix}}_{A(t)}\begin{pmatrix}x\\p\\\theta\end{pmatrix}dt\\&+\underbrace{\begin{pmatrix}0\\\hbar/2\\0\end{pmatrix}}_{B}\sqrt{2k(t)}dV_1,\end{aligned} \tag{17}$$

$$d\Xi=\underbrace{(1\ 0\ 0)}_{C}\begin{pmatrix}x\\p\\\theta\end{pmatrix}dt+\underbrace{\frac{1}{\sqrt{2k}}}_{D}dV_2. \tag{18}$$

The Kalman filter equations will give us the best possible estimation of the vector (x,p,θ) at each time, while the Riccati equation determines the evolution of the covariance matrix P:

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix}\overline{x}\\\overline{p}\\\overline{\theta}\end{pmatrix}=&A(t)\begin{pmatrix}\overline{x}\\\overline{p}\\\overline{\theta}\end{pmatrix}+2k(t)P(t)C^T\\&\times\left[d\Xi(t)-C\begin{pmatrix}\overline{x}\\\overline{p}\\\overline{\theta}\end{pmatrix}\right],\end{aligned} \tag{19}$$

$$\dot{P}=A(t)P+PA^T(t)-2k(t)PC^TCP+2k(t)BB^T. \tag{20}$$

An optimal measurement strategy, dependent on the sensitivity, will then be the one that minimizes the (3,3) element in P at time t_{final}. An analytic solution of this problem does not exist in general, as the Riccati equations are quadratic. However, in the case of constant $f(t)=f(0)$ and constant sensitivity $k(t)=k(0)$ analytical results will be derived.

Before proceeding, however, it is interesting to do a dimensional analysis to see how the variances will scale. We begin by scaling $\tilde{t}=t/\tau$ with τ the duration of the complete measurement. Introducing the matrix

$$T=\begin{pmatrix} \sqrt{\frac{\hbar\tau}{2m}} & 0 & 0 \\ 0 & \sqrt{\frac{\hbar m}{2\tau}} & 0 \\ 0 & 0 & \sqrt{\frac{\hbar m}{2\tau^3}} \end{pmatrix}, \tag{21}$$

it can easily be checked that $\tilde{P}=T^{-1}PT^{-1}$ is dimensionless. If we then scale the sensitivity as $k(t)=\tilde{k}(\tilde{t})\hbar\tau^2/(2m)$, the force $\theta=\tilde{\theta}\sqrt{\hbar m/2\tau^3}$, and do the appropriate transformations $B\rightarrow\tilde{B}$ and $C\rightarrow\tilde{C}$, we get the equivalent state space model

$$\tilde{A}=\begin{pmatrix} 0 & 1 & 0 \\ -\omega^2\tau^2 & 0 & f(t) \\ 0 & 0 & 0 \end{pmatrix}, \quad \tilde{B}=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \tilde{C}=(1 \quad 0 \quad 0). \tag{22}$$

The new filter equations are still given by Eqs. (19),(20) with the substitution $[A,B,C,k(t)]\rightarrow[\tilde{A},\tilde{B},\tilde{C},\tilde{k}(\tilde{t})]$. This observation has an immediate consequence if we are measuring the force acting on a free particle ($\omega=0$): the standard deviation on our estimate will always scale like $\sqrt{\hbar m/2\tau^3}$, and the chosen sensitivity will only affect the accuracy by a multiplicative prefactor.

II. GENERAL FORMALISM FOR QUANTUM PARAMETER ESTIMATION

In this section, we develop a description of the problem of estimating unknown parameters θ of the dynamics of a quantum system from the results of generalized measurements. This general problem can be addressed in essentially the same way as the specific problem of force estimation for an oscillator that was discussed in the previous section. An approach to this problem has been proposed by one of us [3] and we will formulate the theory in the language of operations and effects and consider, in particular, the case of measurement currents that are continuous in time, as in the case of homodyne detection [15] or continuous position measurement. The fundamental basis of this approach is to propagate an *a posteriori* probability distribution $p(\theta|\mathbf{I}_{[0,t)})$ for the parameter θ conditioned on the history of measurement results $\mathbf{I}_{[0,t)}$ up to time t by employing Bayes' rule and using the theory of operations and effects to calculate the relative likelihood of the known measurement record as a function of θ. Readers who are less interested in mathematical details and more interested in the application of our formalism to the force estimation problem may skip this section.

A. General theory

We will treat the quantum parameter estimation as an essentially classical parameter estimation problem coupled to the quantum measurement updating rules. For each value θ' of θ there will be a conditioned state $\rho_{\theta'}$ describing the state of the quantum system conditioned on the measurement history and a particular value of the unknown parameter θ. This density matrix would be our best description of the state if we knew the measurement record and also that θ took this particular value. However, the value of θ is not assumed to be known exactly and is described by a probability distribution $p(\theta)$. Hence, the density matrix describing the state from the point of view of the experimenter is

$$\rho=\int d\theta p(\theta)\rho_\theta. \tag{23}$$

The most general quantum evolution and measurement can be described by the theory of operations and effects. The following discussion will adapt the treatment of Wiseman and Diósi to our problem [16]. In this paper, we assume that either the dynamics or the measurement are unknown and belong to a family parametrized by θ. Thus, we consider quantum measurements characterized by a set of operators $\Omega_{\theta,r}$ where θ labels the value of the unknown parameter and r labels the measurement result. Thus, there is a separate measurement for each value of θ and the operators $\Omega_{\theta,r}$ are constrained by completeness

$$\int d\mu_{\theta,0}(r)\Omega^\dagger_{\theta,r}\Omega_{\theta,r}=1. \tag{24}$$

Here, $d\mu_{\theta,0}(r)$ is a normalized measure on the space of measurement results r. As in the standard theory, the probability of the measurement result r conditioned on θ is

$$d\mu_\theta(r)=d\mu_{\theta,0}(r)\mathrm{Tr}[\Omega^\dagger_{\theta,r}\Omega_{\theta,r}\rho_\theta]. \tag{25}$$

The state of the quantum system after the measurement conditioned on the pair (θ,r) is

$$\rho'_{\theta,r}=\frac{d\mu_{\theta,0}(r)\Omega_{\theta,r}\rho_\theta\Omega^\dagger_{\theta,r}}{d\mu_\theta(r)}=\frac{\Omega_{\theta,r}\rho_\theta\Omega^\dagger_{\theta,r}}{\mathrm{Tr}[\Omega^\dagger_{\theta,r}\Omega_{\theta,r}\rho_\theta]}. \tag{26}$$

If the result of the measurement is unknown or disregarded, then the state of the system is an average over the conditioned states weighted by their probabilities

$$\rho'_\theta=\int d\mu_\theta(r)\rho'_{\theta,r}=\int d\mu_{\theta,0}(r)\Omega_{\theta,r}\rho_\theta\Omega^\dagger_{\theta,r}. \tag{27}$$

This is the state of the system conditioned on a particular value of θ but not on any measurement result.

The unconditioned probability of the measurement results is found by averaging over the probability distribution for θ and is given by the measure

$$d\mu(r)=\int_\theta d\theta p(\theta)d\mu_\theta(r)=\int_\theta d\mu(\theta)d\mu_\theta(r). \tag{28}$$

After the measurement, we will require that the state conditioned on the measurement result r but not on the value of θ may still be written in the form of Eq. (23) as an average over the states conditioned on particular values of θ, thus,

$$\rho'_r = \int d\mu_r(\theta)\rho'_{\theta,r} \tag{29}$$

for some measure $d\mu_r(\theta)$ on the space of possible θ. This new measure describes the probability of θ conditioned on the measured value of r. This conditioned probability distribution for θ is precisely what we wish to calculate. For consistency, it must be the case that if the measurement result is unknown or disregarded the appropriate state is again an average over the conditioned states

$$\rho' = \int d\mu(r)\rho'_r = \int d\mu(r)d\mu_r(\theta)\rho'_{\theta,r}. \tag{30}$$

In order to calculate $d\mu_r(\theta)$, we need to develop a Bayes' rule that relates all the probability measures we have introduced. In order to do this we note that ρ' must also be able to be expressed as an average over the probability for θ of the states ρ'_θ, thus,

$$\rho' = \int d\mu(\theta)\rho'_\theta = \int d\mu(\theta)d\mu_\theta(r)\rho'_{\theta,r}. \tag{31}$$

This leads us to the Bayes' rule

$$d\mu(r)d\mu_r(\theta) = d\mu(\theta)d\mu_\theta(r), \tag{32}$$

which allows us to calculate $d\mu_r(\theta)$ in terms of $d\mu(\theta)$, the measure that characterizes our prior knowledge about θ, and the measures $d\mu_\theta(r)$, which are part of our specification of the parameterized family of measurements.

In principle, this allows us to optimally update the probability distribution for the unknown parameter in any quantum measurement. We are most interested here in the case of measurements that are continuous in time. In this situation, we wish to derive a stochastic differential equation that updates the distribution for θ conditioned on measurement current. Since the case of photon detection measurements is considered in [3] we will consider measurements like homodyne detection where the measurement results are continuous but not differentiable functions of time, in [16] these are termed diffusive measurements. This will require that we develop stochastic differential equations to describe the measurement process.

For simplicity, we will consider the case where there is only a single measurement being made and we will describe the measurement result r in an infinitesimal time interval $[t,t+dt)$ by the complex number $I(t)$. We define the measurement operators

$$\Omega_{\theta,I} = 1 - iH(\theta)dt - \tfrac{1}{2}\hat{c}^\dagger\hat{c} + I^*\hat{c}dt. \tag{33}$$

These measurement operators may be derived, for example, as the continuous limit of a model of repeated measurements [9] or from models of quantum optical measurements such as heterodyne or homodyne detection [15]. For simplicity, we consider the case where there is only one measurement current, the general case may easily be treated following the formalism of [16]. We also assume that the specific measurement that is being made is known (that is, that the operator coupling the system to the bath and the measurement made on the bath are known) and so θ only parametrizes the Hamiltonian evolution of the system. This is the most interesting case and simplifies the treatment. The extension to the case where the measurement is known but the free system evolution is not unitary but is rather described by a Markovian master equation is also straightforward. Now the measurement operator is constrained by the completeness relation Eq. (24) and this requires that

$$\int d\mu_{\theta,0}(I)(Idt) = 0, \tag{34}$$

$$\int d\mu_{\theta,0}(I)(I^*dt)(Idt) = dt. \tag{35}$$

These moments mean that we may identify Idt as a complex Wiener increment under the measure $d\mu_{\theta,0}(I)$. However, in order to specify this measure completely, we must also specify the remaining second-order moment of the Wiener increment (clearly, this must also be of order dt). We will say that

$$\int d\mu_{\theta,0}(I)(Idt)(Idt) = u\,dt, \tag{36}$$

where we need $|u| \leq 1$. In line with our assumption that the measurement interaction and the measurement on the bath is known, we will require that u is independent of θ. The case $u=0$ corresponds in the quantum optical setting to heterodyne detection, while $|u|=1$ corresponds to homodyne detection with some local oscillator phase. Note that these moments are independent of θ and so we can drop the subscript θ for this measure on I from here on. Since the moments of Idt under $d\mu_0(I)$ indicate that we consider Idt to be a complex Wiener increment, we adopt the Ito rules

$$(Idt)^2 = u\,dt, \qquad (I^*dt)(Idt) = dt. \tag{37}$$

Now we would like to know the observed statistics of I under the physical measure $d\mu(I)$. There are two kinds of conditioned expectation values for operators $\hat{a}$ in this problem. Expectation values conditioned on a particular value of the unknown parameter will be denoted $\overline{a}_\theta = \mathrm{Tr}[\hat{a}\rho_\theta]$. On the other hand, expectation values conditioned only on the history of measurement results will be denoted $\overline{a} = \mathrm{Tr}[\hat{a}\rho]$. Now we know from the preceding discussion that

$$d\mu(I) = \int_\theta d\mu(\theta)d\mu_0(I)\mathrm{Tr}[\Omega^\dagger_{\theta,I}\Omega_{\theta,I}\rho_\theta] \tag{38}$$

$$= d\mu_0(I)\int_\theta d\mu(\theta)\mathrm{Tr}[(1 + I^*\hat{c}dt + I\hat{c}^\dagger dt)\rho_\theta] \tag{39}$$

$$= d\mu_0(I)(1 + I^*dt\overline{c} + Idt\overline{c^\dagger}). \tag{40}$$

Hence, the expected value of I is

$$\langle I\rangle=\int d\mu(I)I=u\overline{c^{\dagger}}+\overline{c}. \tag{41}$$

From Eq. (40) we can see that the second-order moments of Idt are independent of the state and of θ and are equal to the second-order moments under $d\mu_0(I)$. Thus, the transformation from the measure $d\mu_0(I)$ to $d\mu(I)$ is a transformation of drift similar to a Girsanov transformation [17] and we can identify Idt with

$$Idt=u\overline{c^{\dagger}}+\overline{c}dt+dW, \tag{42}$$

where dW is a complex Wiener increment under the measure $d\mu(I)$ obeying $dW^2=udt,dW^*dW=dt$.

On the other hand, the probability measure for the measurement trajectories conditioned on a given value of θ is

$$d\mu_{\theta}(I)=d\mu_0(I)\mathrm{Tr}[\Omega^{\dagger}_{\theta,I}\Omega_{\theta,I}\rho_{\theta}] \tag{43}$$

$$=d\mu_0(I)(1+I^*dt\overline{c}_{\theta}+Idt\overline{c^{\dagger}_{\theta}}). \tag{44}$$

Using Eq. (32), it is now straightforward (keeping terms up to second order in Idt) to update the probability for θ conditioned on I

$$d\mu_{\mathbf{I}_{[0,t+dt)}}=[1+(\overline{c}_{\theta}-\overline{c})(I^*dt-u^*\overline{c}dt-\overline{c^{\dagger}}dt)+(\overline{c^{\dagger}_{\theta}}-\overline{c^{\dagger}})\times(Idt-\overline{c}dt-u\overline{c^{\dagger}}dt)]d\mu_{\mathbf{I}_{[0,t)}}(\theta). \tag{45}$$

This allows us to write down a stochastic Fokker-Planck equation for the probability distribution of θ

$$dp(\theta|\mathbf{I}_{[0,t+dt)})=[(\overline{c}_{\theta}-\overline{c})(I^*dt-u^*\overline{c}dt-\overline{c^{\dagger}}dt)+\mathrm{H.c.}]\times p(\theta|\mathbf{I}_{[0,t)}). \tag{46}$$

Note that under $d\mu(I)$, the innovation $Idt-\overline{c}dt-u\overline{c^{\dagger}}dt$ is a Wiener increment, and thus, has mean zero and is not correlated with either the quantum state or $p(\theta)$. This equation is very similar in form to the Kushner-Stratonovich equation that arises in classical state estimation problems [19]. In order to be able to propagate this equation for the probability distribution of θ we must also be able to update the conditioned state ρ_{θ}, and hence, the expectation values $\overline{c}_{\theta}$. From Eq. (26), we can show that ρ_{θ} obeys the stochastic master equation (SME)

$$d\rho_{\theta}=-i[H(\theta),\rho_{\theta}]dt+\mathcal{D}[\hat{c}]\rho_{\theta}dt+\mathcal{H}[\hat{c}(I^*dt-\overline{c^{\dagger}}dt-u^*\overline{c}_{\theta}dt)]\rho_{\theta}. \tag{47}$$

Equation (46) and the family of stochastic master equations (47) describe the quantum parameter estimation problem for measurements with continuous measurement currents such as optical homodyne detection. As we indicated at the start of this section, and as in the algorithm discussed in [3,8], a family of quantum states conditioned on the measurement record and on different values of θ is propagated using appropriate SME's while the conditioned probability distribution for θ is propagated using a stochastic Fokker-Planck equation of the kind that arises in classical estimation problems. As we shall see below, it is possible to solve these equations for certain linear models, such as force estimation, due to position measurement on a free particle or oscillator. In general, it will be necessary to integrate these equations numerically after first discretizing θ. In principle, this is straightforward although the discretization must be sufficiently fine that a good approximation for the mean $\overline{c}+u\overline{c^{\dagger}}$ is maintained at all times and this will usually involve a prohibitive computational cost. One way of avoiding this is to consider a linear variant of this update equation that is, in fact, more closely allied to the algorithm in [3]. This variant is an analogue both of the linear version of the stochastic master equation [18] and of the Zakai equation that is the linear counterpart to the Kushner-Stratonovich equation [19] in classical state estimation. This linear variant does not preserve the normalization of $p(\theta|I)$ but does not depend on $u\overline{c}+\overline{c^{\dagger}}$ and yet still propagates the relative probabilities of different values of θ.

The basic observation is that in the Bayes' rule Eq. (32), the measure $d\mu(r)$ is independent of θ and only ensures the normalization of $d\mu_r(\theta)$. If we are only interested in the relative likelihood of different values of θ, we may consider unnormalized measures $d\overline{\mu_r}(\theta)$ on the space of possible θ and replace $d\mu(r)$ by any measure on r independent of θ. In particular, for our example of continuous measurements we may choose

$$d\overline{\mu}_{\mathbf{I}_{[0,t+dt)}}(\theta)d\mu_0(I)=d\mu_{\theta}(I)d\overline{\mu}\mathbf{I}_{[0,t)}(\theta). \tag{48}$$

Substituting from Eq. (44), we get

$$d\tilde{p}(\theta|\mathbf{I}_{[0,t+dt)})=(\overline{c}_{\theta}I^*dt+\overline{c^{\dagger}_{\theta}}Idt)\tilde{p}(\theta|\mathbf{I}_{[0,t+dt)}). \tag{49}$$

Under this linear propagation equation, the dynamics of the unnormalized distribution $\tilde{p}(\theta|\mathbf{I}_{[0,t)})$ may be calculated for each value of θ independently. This will make it possible to calculate relative probabilities of a discrete set of possible values of θ given a particular sequence of measurement results with no constraints on the discretization of θ.

This formalism for the estimation of a classical parameter in quantum dynamics may readily be generalized to the case where there is more than one unknown parameter or where the parameter undergoes some known time dependence as in the previous section. Another interesting situation that may be treated straightforwardly in this formalism is correlating the measurement results from two quantum measurements, both of which depend on θ. Here, we have assumed that apart from the measurement the dynamics of the quantum system is unitary. If this is not true (as is the case for less than perfectly efficient detection, for example) then it is straightforward to show that the first term of Eq. (47) is simply replaced by a Liouvillian term describing the noisy dynamics of the system, thus

$$d\rho_{\theta}=\mathcal{L}(\theta)\rho_{\theta}dt+\mathcal{D}[\hat{c}]\rho_{\theta}dt+\mathcal{H}[\hat{c}(I^*dt-\overline{c^{\dagger}_{\theta}}dt-u^*\overline{c}_{\theta}dt)]\rho_{\theta}. \tag{50}$$

In the next section, we will return the problem of force estimation through continuous position measurement of an oscillator. We will be most interested in finding the optimum (possibly time-dependent) sensitivity of the measurement.

B. Force estimation through continuous position measurement

The general formalism of this section may be reduced to the parameter estimation problem we considered at the start of the paper in the important case of force estimation though continuous position measurement of an oscillator ($\hat{c}=\sqrt{2k}\hat{x},\ u=1,\ H(\theta)=\hat{p}^2/2m+m\omega^2\hat{x}^2/2+\theta\hat{x}$). In this case, it is possible to solve the system of equations (47) and (46) explicitly. We have the system of equations

$$d\rho_\theta=-i[\hat{p}^2/2m+m\omega^2\hat{x}^2/2+\theta\hat{x},\rho_\theta]dt+2k\mathcal{D}[\hat{x}]\rho_\theta dt + \sqrt{2k}\mathcal{H}[\hat{x}]\rho_\theta(Idt-2\sqrt{2k}\overline{x_\theta}dt), \tag{51}$$

$$dp(\theta|\mathbf{I}_{[0,t+dt)})=2\sqrt{2k}(\overline{x_\theta}-\overline{x})(Idt-2\sqrt{2k}\overline{x}dt)\times p(\theta|\mathbf{I}_{[0,t+dt)}). \tag{52}$$

This linear system preserves Gaussian quantum states of the oscillator and Gaussian probability distributions for θ. As a result, we only need to find stochastic equations for the first- and second-order moments of the ρ_θ and $P(\theta|I)$. The procedure is to apply standard master equation techniques [20] combined with the Ito rules for stochastic differential equations to find equations for the moments of $\hat{x}$ and $\hat{p}$, conditioned on a particular value of θ, from Eq. (51) as was done in [5]. The unconditioned moments result from averaging over $p(\theta|I)$:

$$\overline{\Delta x^2}=\int d\theta p(\theta)\mathrm{Tr}[(\hat{x}-\overline{x}_\theta)^2\rho_\theta], \tag{53}$$

$$\overline{\Delta x\Delta p}=\int d\theta p(\theta)(\mathrm{Tr}[(\hat{x}\hat{p}+\hat{p}\hat{x})\rho_\theta]/2-\overline{x}_\theta\overline{p}_\theta), \tag{54}$$

$$\overline{\Delta p^2}=\int d\theta p(\theta)\mathrm{Tr}[(\hat{p}-\overline{p}_\theta)^2\rho_\theta], \tag{55}$$

$$\overline{\Delta x\Delta\theta}=\left(\int d\theta p(\theta)\theta\overline{x}_\theta\right)-\overline{\theta x}, \tag{56}$$

$$\overline{\Delta p\Delta\theta}=\left(\int d\theta p(\theta)\theta\overline{p}_\theta\right)-\overline{\theta p}, \tag{57}$$

$$\overline{\Delta\theta^2}=\int d\theta p(\theta)(\theta-\overline{\theta})^2. \tag{58}$$

These moments form the covariance matrix P and it is a straightforward though tedious exercise to show that it obeys the matrix Riccati equation (20) we derived in the first section. Similarly, the first-order moments obey the equations (19) of the Kalman estimator.

III. STANDARD QUANTUM LIMITS

The preceding sections dealt with the problem of optimal estimation of parameters of the Hamiltonian given a system that is continuously observed. In this section, we will derive the explicit equations of the variances on these estimates.

A. Detection of stationary signals

Let us first introduce the idea of the standard quantum limit in the context of von Neumann measurements. The idea is that a particle is prepared in some optimal way at time 0, such that at time τ, a projective measurement is performed to determine the displacement associated with the force. The optimal preparation is crucial as it has to balance the position and the momentum uncertainty. The optimal preparation leads to the expression of the standard quantum limit. Consider a free particle with a Gaussian wave function $\langle x|\psi\rangle$ and initial parameters $\tilde{x}(0),\tilde{\sigma}(0)$ [see Eq. (2)] and subject to an unknown force θ. The integrated equations of motion (3) are given by

$$\tilde{x}(t)=\tilde{x}_0+\theta[t\tilde{\sigma}(0)/i\hbar+t^2/2m],\quad \tilde{\sigma}(t)=\tilde{\sigma}(0)+i\frac{\hbar}{m}t.$$

Suppose that at time τ we perform a von Neumann measurement of the position. The probability distribution associated with this measurement is given by

$$p(x|\theta)\sim\exp\left(-\frac{\left(x-\frac{\theta t^2}{2m}\right)^2}{|\tilde{\sigma}|^2/\tilde{\sigma}_r}\right). \tag{59}$$

Using Bayes' rule with a flat prior distribution for θ, the variance on the estimate of θ given the measurement result x can easily be derived:

$$\sigma_\theta=\frac{2m^2|\tilde{\sigma}(t)|^2}{\tilde{\sigma}_r(t)t^4}=\frac{2m^2\left[\tilde{\sigma}_r^2(0)+\left(\tilde{\sigma}_i(0)+\frac{\hbar t}{m}\right)^2\right]}{\tilde{\sigma}_r(0)t^4}. \tag{60}$$

This function is heavily dependent on the initial conditions of the wave function of the particle. The standard quantum limit can now be derived by choosing the initial conditions such that σ_θ is minimized. This variance can, in principle, go to zero if we allow $\langle\Delta x\Delta p\rangle$ to be negative, but we will not consider such "contractive" states [21,22] here. We therefore impose the condition $\tilde{\sigma}_i(0)\geqslant 0$ in order to focus our attention on the specific issue of sensitivity optimization. The optimal $\tilde{\sigma}(0)$ is then given by $\tilde{\sigma}(0)=\hbar t/m$, and this leads to the expression of the standard quantum limit:

$$\sigma_\theta=\frac{4\hbar m}{t^3}. \tag{61}$$

It is clear that the square of the amplitude of a detectable force has to be bigger than the variance on its estimation to be detectable. Therefore, the previous formula is the expression of the minimal force that can be detected by a free particle of mass m over a time t. Note that the derived formula exceeds the normal equation for the SQL [14] by a factor 8 as the standard equation is not derived in the context of parameter estimation.

We will now apply an analogous reasoning to a quantum particle subject to continuous measurement. The explicit expression of the variance on the estimated force was given by Eq. (16). As noted at the end of the first section, the resulting variance will be given by the standard quantum limit multiplied by a certain factor. From here on, we will therefore work in the dimensionless picture as defined in Eq. (22). In general, it is very hard to find the explicit expression for the autocorrelation function $g(t,t')$ in Eq. (16). Things get much more feasible if we do not vary the sensitivity during the measurement as the system then becomes stationary. It follows that we can assume that the values of the variances reached their steady-state values given by Eq. (11). After some straightforward linear algebra, the explicit expression for $g(t,t')$ in the case of steady state is given by

$$g(t,t')=\frac{1}{b}\exp[-a(t-t')]\sin[b(t-t')], \tag{62}$$

$$a=\omega\tau\sqrt{\frac{1}{2}\left(-1+\sqrt{1+\frac{(2k)^2}{(\omega\tau)^4}}\right)}, \tag{63}$$

$$b=\omega\tau\sqrt{\frac{1}{2}\left(1+\sqrt{1+\frac{(2k)^2}{(\omega\tau)^4}}\right)}. \tag{64}$$

Due to the stationarity of the variances, the autocorrelation function $g(t,t')$ is indeed only dependent on $(t-t')$, and from here on we will therefore use the notation $g(t,t') = g(t-t')$. The full expression of the variance on our estimate now becomes

$$\frac{1}{\sigma_\theta}=2k\int_0^1 dt\left(\int_0^t dt'\,g(t-t')f(t')\right)^2. \tag{65}$$

The force that acted on the system was assumed to be of the form $\theta(t)=\theta f(t)$ with $f(t)$ a known function. Note that this expression is dimensionless and has to be multiplied by $(2\tau^3)/(\hbar m)$. We next introduce $F(\omega)$ and $G(\omega)$ the Fourier transforms of the functions $f(t)\cdot u_{[0,1]}(t)$ and $g(t)\cdot u_{[0,1]}(t)$, where $u_{[0,1]}(t)$ is the window function over the interval $[0,1]$. The damping effect due to the back-action noise is responsible for broadening the spectrum of the harmonic oscillator with a width of approximately $k/(\omega\tau)$. Basic properties of Fourier transformations lead to the expression:

$$\frac{1}{\sigma_\theta}=\frac{2k}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}d\omega_1 d\omega_2\exp\left(i\frac{\omega_1-\omega_2}{2}\right)\times\frac{\sin[(\omega_1-\omega_2)/2]}{(\omega_1-\omega_2)/2}G(\omega_1)G^*(\omega_2)F(\omega_1)F^*(\omega_2). \tag{66}$$

This formula clearly shows that only the frequencies of the signal $F(\beta)$ near the natural frequencies of the oscillator $G(\beta)$ will be detectable.

Now we shall explicitly calculate the value of σ_θ in some different cases. Let us first of all assume that the spectrum $F(\beta)$ is almost constant for all values where $G(\beta)$ is substantially different from 0, i.e., around $\beta\simeq(\omega\tau)$. This is realistic in some scenarios of interest for the detection of gravitational waves [14]. Let us furthermore assume that $\omega\tau\gg1$, which means that the period of the oscillator is much smaller then the observation time. Next we observe that we are allowed to approximate the $\mathrm{sinc}[(\omega_1-\omega_2)/2]$ function by a delta-Dirac function if the width of the spectrum $G(\beta)$, determined by the number $k/(\omega\tau)$, is much bigger than one. This leads to the expression

$$\frac{1}{\sigma_\theta}\simeq\frac{k|F(\omega\tau)|^2}{2\pi}\int_{-\infty}^{\infty}d\omega|G(\omega)|^2, \tag{67}$$

$$=\frac{k|F(\omega\tau)|^2}{2\pi}\int_{-\infty}^{\infty}d\omega\frac{1}{(a^2+b^2-\omega^2)^2+4a^2b^2}, \tag{68}$$

$$=\frac{|F(b)|^2}{4\omega\tau}\chi\left(\frac{2k/(\omega\tau)^2}{\sqrt{1+[2k/(\omega\tau)^2]^2}}\right), \tag{69}$$

$$\chi(x)=(1-x^2)^{1/4}\sqrt{\frac{1+\sqrt{1+x^2}}{2(1+x^2)}}. \tag{70}$$

The function introduced in the last line is only dependent on $2k/(\omega\tau)^2$, which can be tuned freely by changing the value of our sensitivity. The function $\chi(x)$ reaches its maximum value 1 for small values of x, meaning that optimal detection requires $k\ll(\omega\tau)^2$. The derivation, however, required that $1\ll k/(\omega\tau)$. Therefore, the optimal choice of the sensitivity will be given by a value $(\omega\tau)\ll k\ll(\omega\tau)^2$, leading to the variance on the estimate:

$$\sigma_\theta^2\simeq\frac{4\omega\tau}{|F(b)|^2}\frac{\hbar m}{2\tau^3}=\frac{1}{|F(b)|^2}\frac{2\hbar m\omega}{\tau^2}. \tag{71}$$

This corresponds exactly to the expression of the standard quantum limit for an oscillator [14]. A similar expression can be obtained by explicitly integrating Eq. (65) with $f(t) = \delta(t)$. The conditions under which this SQL can be reached are (1) the total duration of the measurement is much bigger then the period of the oscillator; (2) the spectrum of the signal to be detected is flat around the natural frequencies of the observed oscillator.

We will now investigate what happens if this second condition is not fulfilled. In the extreme case, the force to be detected is constant, corresponding to a δ Dirac function in the frequency domain. Again, under the condition that $1 \ll\omega\tau\ll k/(\omega\tau)$, a good approximation of Eq. (66) becomes:

$$\frac{1}{\sigma_\theta} \simeq k|G(0)|^2 = \frac{1}{(\omega\tau)^2}\frac{2k/(\omega\tau)^2}{1+(2k/(\omega\tau)^2)^2}. \tag{72}$$

The optimal sensitivity is now given by $2k=(\omega\tau)^2$, indicating that one has to choose a much higher sensitivity to detect constant forces than resonant oscillating forces. The expression for the SQL for detecting constant forces with a harmonic oscillator therefore becomes

$$\sigma_\theta \simeq 2(\omega\tau)^2\frac{\hbar m}{2\tau^3} = \frac{m\hbar\omega^2}{\tau}. \tag{73}$$

It is now natural to look at what happens in the limit of $\omega\to 0$, that is, if the observed particle is free and only subject to a constant force. In that case the explicit integration of Eq. (64) becomes possible, as a and b both become equal to the sensitivity $\sqrt{k}$. Straightforward but long integrations lead to

$$\sigma_\theta = \frac{8k^{3/2}}{4\sqrt{k}-5+8\exp(-\sqrt{k})\cos(\sqrt{k})-\exp(-2\sqrt{k})[2+\cos(2\sqrt{k})+\sin(2\sqrt{k})]}. \tag{74}$$

Minimization over the sensitivity leads to an expression for the SQL for the detection of a constant force with a free particle subject to continuous observation

$$\sigma_\theta \simeq 3\frac{4\hbar m}{\tau^3}. \tag{75}$$

Note that this expression differs from the corresponding one derived in [4], where calculations were done without properly accounting for the damping effect of measurement back action. Comparing this result with Eq. (61), the variance of our estimate obtained by continuous measurement is three times bigger than if we were doing projective measurements. This is caused by two factors. First, at the end of the continuous measurement, there is still a lot of information encoded about the force in the wave function as the variance on the position at time τ is not at all equal to ∞. Second, the previous result was obtained by assuming that the variances of our Gaussian wave function were in steady state, and this is not necessarily the optimal initial condition. Indeed, it turns out that the optimal initial state (not considering contractive states) of the continuously observed particle is a Gaussian state with well-defined momentum ($\langle\Delta p^2\rangle \ll 1$) and therefore, undefined position $\langle\Delta x^2\rangle \gg 1$. This makes sense as the force to be detected can only be seen because it manifests itself through the momentum. The fact that the position uncertainty is very large is not so bad as the position is continuously observed such that it becomes well defined very quickly. The expression for the variance on the force estimate using this optimally prepared initial state can now be calculated exactly by explicitly solving the Riccati equation (20):

$$\sigma_\theta = \frac{2k^{3/2}[\sinh(2\sqrt{k})+\sin(2\sqrt{k})]}{k[\sinh(2\sqrt{k})+\sin(2\sqrt{k})]-[\cosh(2\sqrt{k})-\cos(2\sqrt{k})]}. \tag{76}$$

Optimization over the sensitivity leads to an enhancement of 2/3 in comparison with the steady state case. An even bigger gain would have been obtained if a projective measurement at the end of the continuous observation were allowed. A realistic way to implement this would be to make the sensitivity very large at the end of the measurement. If the matrix $P(1)$ is the solution of the Riccati equation (20) at time $t=1$, some straightforward calculations show that a projective position measurement reduces the estimator variance by $P_{(3,1)}^2/P_{(1,1)}$. The optimal initial conditions are still given by $\langle\Delta p^2\rangle \ll 1$ and $\langle\Delta x^2\rangle \gg 1$. The exact expression of the variance on the estimate as a function of the sensitivity k is then given by

$$\sigma_\theta = \frac{4k^{3/4}[\cosh(2\sqrt{k})+\cos(2\sqrt{k})]}{k[\cosh(2\sqrt{k})+\cos(2\sqrt{k})]-[\sinh(2\sqrt{k})+\sin(2\sqrt{k})]}. \tag{77}$$

Minimization over the sensitivity leads to the equation

$$\sigma_\theta \simeq 0.752\frac{4\hbar m}{t^3}. \tag{78}$$

Therefore, we have modestly beaten the usual standard quantum limit by optimally preparing the Gaussian wave packet and doing a von Neumann measurement at the end of the continuous measurement. This shows that a continuous measurement together with a projective measurement at the end on an optimally prepared state can reveal more information than only projective measurements. In other words, the balance information gain versus disturbance is a little bit in favor of continuous measurement. Although noise is continuously fed into the system by the sensor, we can extract more information about the classical force.

An even better performance can be obtained if we vary the sensitivity continuously during the measurement (*sensitivity scheduling*). It is indeed the case that back-action noise introduced in the beginning of the measurement does more harm than back-action noise at the end of the measurement, as the random momentum kicks delivered at any given time corrupt all subsequent position readouts. In terms of systems theory, the optimal sensitivity as a function of time is simply an optimal control problem associated with Eq. (20). In this optimal control problem, the cost function is simply the value of the force estimator variance $P_{(3,3)}$ at the final time.

This is to be minimized by an appropriate choice of the time variation of the sensitivity. The optimal control can be determined by solving a Bellman equation using techniques of dynamic programming [19]. Due to the nonlinearity of the Riccati equation, this cannot be done analytically. The optimal sensitivity at time τ, however, can easily be obtained: it tends to a Dirac-δ function so as to mimic a projective position measurement. The variance on the estimator after such a projective measurement is reduced by $P^2_{(3,1)}/P_{(1,1)}$. In order to obtain a numerically tractable problem, we define the cost-function $K=P_{(3,3)}(\tau)-P^2_{(3,1)}(\tau)/P_{(1,1)}(\tau)$, the optimal control problem no longer contains a singularity and can be solved numerically. In this second problem, it is assumed that it is possible to make a projective measurement at the final time and the aim is to choose the sensitivity as a function of time such that the information gained during the continuous measurement and due to the projective measurement is maximized. Another way to regularize this problem would be to specify a maximum allowed sensitivity. We discretize the total time in, for example, 50 intervals, and in each interval we assume the sensitivity has a constant value k_j. The solution can then be found by applying some kind of steepest descent algorithm over these 50 variables $\{k_j\}$. It turns out that the optimal $k(t)$ in the case of a free particle ($\omega=0$) is a smooth monotonously but slowly increasing function of time. In this free particle case, the optimal time-varying sensitivity only leads to a marginal gain: the numerical optimization shows that the variance of the estimate becomes very nearly equal to a factor 3/4 of the usual standard quantum limit (61). Nevertheless, we can present this result as a generalization of the usual SQL to include strategies with sensitivity scheduling:

$$\sigma_\theta \simeq \frac{3.000\hbar m}{t^3}. \tag{79}$$

Much greater improvements can be expected from the application of sensitivity scheduling to the case of a continuously observed harmonic oscillator. Indeed, the variance on the position of such a particle is small in the middle of the well and at the borders, while it is big elsewhere. Therefore, the sensitivity should be varied in a sinusoidal manner, so as to measure more precisely at the positions where the variance is small. The optimal variation of sensitivity in time could be determined by solving a similar optimal control problem to the one explained in the previous paragraph. In the limit where projective measurements are allowed, one expects that the optimal variation of sensitivity should correspond to stroboscopic measurement [14], which is indeed well known to beat the usual standard quantum limit.

B. Detection of nonstationary signals

The techniques introduced in the preceding sections can also be used for the estimation of nonstationary signals, as one would have for example in the problem of gravitational wave detection when the arrival time of the signal is unknown. Suppose, for example, that we know that the signal to detect is of the form $\theta(t-t_1)=\theta_0 f(t-t_1)$ with $f(\tau)$ known but amplitude θ_0 and arrival time t_1 unknown. Deciding whether or not the signal has arrived effectively involves hypothesis testing. In a Bayesian framework such as ours we wish to compare the likelihood for the observed measurement results assuming there is no force with the likelihood assuming that the force is acting. For our simple system the easiest way to do this is to assess how closely the residuals of the Kalman filter correspond to a white noise process. For the purposes of this discussion, we will consider a simple example and we will not concern ourselves with finding the optimal protocol.

An effective nonstationary measurement strategy can be implemented by constructing a Kalman filter for system (12) assuming that $\theta=0$ [assuming $f(\tau)=0$ for $\tau<0$]. At times $t<t_1$, the quantity $d\Xi-\overline{x}(t)dt$ is by construction white noise with variance $dt/2k(t)$. From time $t\geqslant t_1$ on however, the force will bias this white noise by an amount $\int_{t_1}^t dt' g(t,t')\theta(t')$ as the $\theta=0$ Kalman filter models the wrong system. This bias will be detectable once it transcends the white noise at time $t_1+\Delta t$:

$$\int_{t_1}^{t_1+\Delta t} dt \int_{t_1}^{t} dt' g(t,t')\,\theta_0 f(t'-t_1) \geqslant \sqrt{\int_{t_1}^{t_1+\Delta t}\frac{dt}{2k(t)}}. \tag{80}$$

An accurate determination of t_1 will result if we make this Δt as small as possible. The previous equation can again be solved analytically if one has a constant sensitivity and steady-state conditions. To make things easier we assume that the observed particle is free ($\omega=0$), although all calculations can be performed in the more general case too. Let us first assume that the signal to detect is a kick at time t_1: $f(t-t_1)\simeq \delta(t-t_1)\tau$ with τ some measure of the duration of the kick [14]. Introducing the dimensionless parameter $\kappa=\Delta t\sqrt{\hbar k/2m}$, the previous inequality becomes

$$\theta_0 \geqslant \frac{1}{\tau}\sqrt{\frac{\hbar m}{\Delta t}}\frac{\kappa}{1-\exp(-\kappa)[\cos(\kappa)+\sin(\kappa)]}, \tag{81}$$

$$\geqslant \frac{2}{\tau}\sqrt{\frac{\hbar m}{\Delta t}}. \tag{82}$$

In the last step, the optimal κ, related to the optimal sensitivity k, was chosen. The meaning of this equation is clear: a kick with an amplitude θ_0 will only be observed after a time span $\Delta t = 4\hbar m/\tau^2\theta_0^2$. Moreover, the sensitivity has to scale inversely with the square root of Δt.

An analogous treatment applies to the case of a constant force $f(t-t_1)=u_{[0,\infty]}(t-t_1)$. In this case inequality (80) becomes

$$\theta_0 \geqslant \sqrt{\frac{\hbar m}{\Delta t^3}}\frac{\kappa^2}{\exp(-\kappa)\cos(\kappa)+\kappa-1}, \tag{83}$$

$$\geqslant 4.25\sqrt{\frac{\hbar m}{\Delta t^3}}. \tag{84}$$

As expected, the variance has the form of the well-known standard quantum limit, but now in a different setup.

The previous arguments can be refined by using techniques of classical detection theory such as the concept of the matched filter. The results will however be qualitatively similar to the previous ones.

More advanced detection schemes can also be constructed by *adaptively* changing the sensitivity as a real-time function of the measurement record [1]. A possible application of this is a scheme for the detection of a signal with unknown arrival time: first, one chooses the optimal sensitivity for estimating the arrival time, and from the moment the signal is detected, the sensitivity is brought to its optimal value for detecting the amplitude of the signal. More sophisticated versions of this adaptive measurement could be very useful in realistic stroboscopic measurements where the initial phase of the harmonic oscillator is unknown, as the measurement sensitivity could be made a real-time function of the estimated particle position.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation, under Contract Nos. PHY-9987541 and EIA-0086038 the A. P. Sloan Foundation, and the K. U. Leuven. F.V. is grateful to Bart De Moor.

[1] H.M. Wiseman, Phys. Rev. Lett. **75**, 4587 (1995).
[2] G.J. Milburn, Quantum Semiclassic. Opt. **8**, 269 (1996).
[3] H. Mabuchi, Quantum Semiclassic. Opt. **8**, 1103 (1996).
[4] H. Mabuchi, Phys. Rev. A **58**, 123 (1998).
[5] A.C. Doherty, S.M. Tan, A.S. Parkins, and D.F. Walls, Phys. Rev. A **60**, 2380 (1999).
[6] A.C. Doherty and K. Jacobs, Phys. Rev. A **60**, 2700 (1999).
[7] V.P. Belavkin, Rep. Math. Phys. **43**, 405 (1999); V. P. Belavkin, in *Modelling and Control of Systems*, edited by A. Blaquiere (Springer-Verlag, Berlin, 1989).
[8] J. Gambetta and H.M. Wiseman, e-print quant-ph/0103032.
[9] C.M. Caves and G.J. Milburn, Phys. Rev. A **36(12)**, 5543 (1987).
[10] A. Barchielli, L. Lanz, and G.M. Prosperi, Nuovo Cimento Soc. Ital. Fis., B **72**, 79 (1982).
[11] R.L. Hudson and K.R. Parthasarathy, Commun. Math. Phys. **93**, 301 (1984).
[12] V.P. Belavkin and P. Staszewski, Phys. Rev. A **45**, 1347 (1992).
[13] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer-Verlag, Berlin, 2000), 2nd ed.
[14] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, 1992).
[15] H.M. Wiseman and G.J. Milburn, Phys. Rev. A **47**(1), 642 (1993).
[16] H.M. Wiseman and L. Dióśi, eprint quant-ph/0012016.
[17] B. K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer, New York, 1998).
[18] P. Goetsch and R. Graham, Phys. Rev. A **50**(6), 5242 (1994).
[19] P. S. Maybeck, *Stochastic Models, Estimation and Control* (Academic Press, New York, 1982).
[20] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer-Verlag, Berlin, 1994).
[21] H.P. Yuen, Phys. Rev. Lett. **51**, 719 (1983).
[22] M. Ozawa, Phys. Rev. Lett. **60**, 385 (1988).