

Robust guaranteed-cost adaptive quantum phase estimation

Shibdas Roy,^{1,2,*} Dominic W. Berry,^{3,†} Ian R. Petersen,^{4,‡} and Elanor H. Huntington^{4,5,§}

¹*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*

²*Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583*

³*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*

⁴*Research School of Engineering, Australian National University, Canberra, ACT 2601, Australia*

⁵*Centre for Quantum Computation and Communication Technology, Australian Research Council, Australia*

(Received 11 January 2017; published 11 May 2017)

Quantum parameter estimation plays a key role in many fields like quantum computation, communication, and metrology. Optimal estimation allows one to achieve the most precise parameter estimates, but requires accurate knowledge of the model. Any inevitable uncertainty in the model parameters may heavily degrade the quality of the estimate. It is therefore desired to make the estimation process robust to such uncertainties. Robust estimation was previously studied for a varying phase, where the goal was to estimate the phase at some time in the past, using the measurement results from both before and after that time within a fixed time interval up to current time. Here, we consider a robust guaranteed-cost filter yielding robust estimates of a varying phase in real time, where the current phase is estimated using only past measurements. Our filter minimizes the largest (worst-case) variance in the allowable range of the uncertain model parameter(s) and this determines its guaranteed cost. It outperforms in the worst case the optimal Kalman filter designed for the model with no uncertainty, which corresponds to the center of the possible range of the uncertain parameter(s). Moreover, unlike the Kalman filter, our filter in the worst case always performs better than the best achievable variance for heterodyne measurements, which we consider as the tolerable threshold for our system. Furthermore, we consider effective quantum efficiency and effective noise power, and show that our filter provides the best results by these measures in the worst case.

DOI: [10.1103/PhysRevA.95.052322](https://doi.org/10.1103/PhysRevA.95.052322)

I. INTRODUCTION

Quantum parameter estimation plays a central role in many fields, such as quantum computing [1], communication [2–4], and metrology [5]. It involves estimating a classical variable, such as an optical phase shift, of a quantum system [6]. This is required since measurements are corrupted with unavoidable noises in practice and it is important to estimate as accurately as possible the desired variable from such noisy measurements. Also, often the variable of interest is not directly accessible to measurement. As an example, the detection of an optical beam always results in the measurement of the amplitude of the beam. It is, however, an important task to estimate the phase of an optical beam from the available measurement results for the amplitude of the field. In communications, for example, it is more reliable to encode information into the phase rather than the amplitude or intensity of the optical field [7]. Moreover, adaptive measurements (i.e., involving feedback) allow for demonstrably better estimation of phase than nonadaptive estimation. Quantum phase estimation has therefore been extensively studied in this context [8–15].

Optimal estimation strategies have played significant roles in improving the achievable precision in quantum phase estimation [16–18]. For example, adaptive estimation with a Kalman filter has provided mean-square errors less than the standard quantum limit (SQL) [16,19]. The SQL is the minimum phase estimation error that can be obtained with

a coherent beam using a perfect heterodyne technique and sets an important benchmark for the quality of a measurement. However, optimal estimation schemes provide optimal performance only if the parameters underlying the system are accurately known. Any uncertainty in our knowledge of these parameters can heavily degrade the estimation performance under certain circumstances. In the presence of large uncertainty, the worst-case mean-square errors can be higher than the acceptable thresholds in practical engineering systems. Such a threshold often determines the point beyond which a system has a risk of breaking down or becoming unusable. It is therefore important to make the estimation process robust to such uncertainties [20,21]. A robust filter is one that provides acceptable performance for the full possible range of the uncertain system parameter(s), and the worst case is the situation where the parameters result in the largest mean-square error.

For quantum estimation problems that can be approximated as having linearized dynamics, many of the principles of robust classical estimation can be applied. There is a rich set of classical estimation strategies in modern control theory for achieving robustness for systems with explicitly introduced uncertainties in a systematic state-space setting [20,21]. Such classical robust estimation principles as applied to quantum estimation problems are little studied. Some of the authors have earlier studied in Ref. [22] the problem of robustly estimating a continuously varying phase, modeled as a stochastic noisy process, of a squeezed state of light [23,24]. In particular, a robust fixed-interval smoother was constructed that provides guaranteed worst-case performance when compared to an optimal smoother in adaptive quantum-optical phase estimation. A smoother involves deducing an estimate based on both past and future measurements with respect to the time at which the

*roy_shibdas@yahoo.co.in

†dominic.berry@mq.edu.au

‡i.r.petersen@gmail.com

§elanor.huntington@anu.edu.au

phase is to be estimated [25,26]. This noncausal estimator is of relevance to applications such as gravitational wave detection [27,28], where obtaining more precise estimates is of greater interest than real-time estimates.

However, for applications such as quantum computing and communication, it is instead required to obtain real-time estimates as precisely as possible. In such cases, it is filtering, involving only past measurements, that can be used and should be made to yield as accurate real-time estimates as possible. In this paper, we consider a robust guaranteed-cost filter [29] in an adaptive quantum phase estimation process of the form considered in Ref. [16], using a coherent state of light. Such a robust filter helps in achieving real-time guaranteed worst-case performance, in relation to the SQL taken as the tolerable threshold, over a Kalman filter. The phase to be estimated is modeled as an Ornstein-Uhlenbeck (OU) process [Eq. (3.1)] as in Refs. [16,17,22].

The robust smoother illustrated in Ref. [22] was shown to consist of a forward robust filter and a backward robust filter, the estimates of which were combined to yield the desired smoothed estimate. However, the approach of Ref. [22] considered an energy-bounded description of noise [30]. By contrast, we here use the approach of Ref. [29] that considered a white noise description, which is the actual noise involved here [Eqs. (3.4) and (3.5)]. In addition, such a robust filter [Eq. (4.7)], as its name suggests, comes with a guaranteed cost [Eq. (4.6)] that quantifies an upper bound on the mean-square error of estimates obtained from the filter for the uncertain system [Eq. (5.14)]. In other words, the worst-case mean-square error of the robust filter is guaranteed to be bounded by the aforementioned cost.

The robust guaranteed-cost filter considered here has previously been applied by three of the authors to a coherent state of light to yield some preliminary results [19,31]. This paper builds on Ref. [19] and presents more in-depth results, insights, and analysis relevant to the physical problem at hand. We demonstrate that with uncertainty in the system [Eq. (4.2)], the Kalman filter performance [Eq. (5.13)] gets increasingly worse with respect to the SQL [Eq. (B3)], whereas the robust filter [Eq. (5.14)] consistently beats the SQL in the worst-case scenario. We explicitly illustrate that the robust filter is guaranteed to beat the Kalman filter as well as the SQL in the worst-case situation.

Moreover, we define an effective quantum efficiency [Eq. (8.3)] for our filters, such that the best estimation scheme yields the same variance as the suboptimal estimation scheme with no loss [32]. We illustrate that our robust filter always exhibits the maximum achievable effective quantum efficiency in the worst-error case, whereas the Kalman filter suffers heavily in this regard. Furthermore, we define an effective noise power [Eqs. (9.23) and (9.22)] for our filters, such that the mean-square error for the optimal filter with added noise is the same as that for a suboptimal filter. This is inspired by a similar result for a constant phase [33,34]. We see that our robust filter always admits the minimum possible effective noise power in the worst-error case, but the Kalman filter suffers from much higher effective noise owing to the uncertainty in the model.

The structure of the paper is as follows. Section II provides an introduction to the theory of estimation using filters. Section III discusses the optimal adaptive estimation process

using a Kalman filter for an exact system model. Section IV deduces the robust filter for the uncertain system model. Section V shows how to calculate the filter mean-square errors for the uncertain system. Section VI concerns the comparison of the mean-square errors of the Kalman and robust filters for the uncertain system. Section VII illustrates that the worst-case mean-square error of the robust filter always beats that of the Kalman filter as well as the SQL. Section VIII introduces a quantity called the effective quantum efficiency for the Kalman and robust filters and compares the filters in terms of this quantity. Section IX defines another quantity called the effective noise power for the filters, which are then compared in terms of this quantity. Finally, Section X ends the paper with concluding remarks.

II. ESTIMATION USING FILTERS

Here we look at the problem of estimating a set of time-varying parameters describing the state of a system, given noisy measurements (output) of observable quantities along with a model relating the observations to the underlying state, corrupted by noise (input). Such a dynamical system is conveniently modeled in a compact state-space form as a set of input, output, and state variables related by first-order differential equations, where the variables are expressed as vectors. If the system is linear and time invariant, the differential and algebraic equations are written in matrix form as follows [35]:

$$\dot{x}(t) = Ax(t) + Bv(t), \quad (2.1)$$

$$y(t) = Cx(t) + Dw(t), \quad (2.2)$$

$$z(t) = Lx(t). \quad (2.3)$$

Here x is the state vector, y is the output vector, z is the vector of variables to be estimated, v and w are the inputs and are (possibly vector) white noise processes that may be correlated, t is the time variable, and A , B , C , D , and L are suitable matrices of compatible dimensions. In particular, A is referred to as the state matrix.

A suitable estimator for the above system (2.1)–(2.3) is then a filter of the form

$$\dot{\hat{x}}(t) = F\hat{x}(t) + Ky(t), \quad (2.4)$$

$$\hat{z}(t) = L\hat{x}(t), \quad (2.5)$$

where \hat{x} is the estimate of the state x , and \hat{z} is the estimate of the desired quantity z . The input to the filter is the output vector y of the system, and the filter output is the estimate \hat{z} . In the Laplace domain, we get

$$(sI - F)\hat{x}(s) = Ky(s), \quad (2.6)$$

$$\hat{z}(s) = L\hat{x}(s), \quad (2.7)$$

where s is the Laplace variable. The transfer function, relating the filter output to the filter input, is therefore

$$G(s) = \frac{\hat{z}(s)}{y(s)} = L(sI - F)^{-1}K. \quad (2.8)$$

The estimation error is $\varepsilon = z - \hat{z}$. Then, the optimal steady-state filter is the system (2.4), (2.5), and (2.8) with F and K

suitably chosen to minimize the cost,

$$J_c = \lim_{t \rightarrow \infty} \langle \varepsilon^T(t) \varepsilon(t) \rangle, \quad (2.9)$$

where $\langle \cdot \rangle$ denotes the expectation.

The optimal estimator that minimizes the above cost is given by the Kalman filter. It operates recursively on streams of noisy input data to produce a statistically optimal estimate of the underlying system state. The steady-state Kalman filter [Eqs. (2.4) and (2.5)] for the system (2.1)–(2.3) is obtained from the solution to the algebraic Riccati equation [20,22,36],

$$AP + PA^T + BQB^T - PC^T(DRD^T)^{-1}CP = 0, \quad (2.10)$$

which is quadratic in P , the error-covariance matrix for the state. The positive-semidefinite solution (if one exists) for P is the stabilizing solution that is desired. Here, we have assumed the following about the white noises v and w (in particular that they are uncorrelated):

$$\langle v(t)v^T(\tau) \rangle = Q\delta(t - \tau), \quad (2.11)$$

$$\langle w(t)w^T(\tau) \rangle = R\delta(t - \tau), \quad (2.12)$$

$$\langle v(t)w^T(\tau) \rangle = 0, \quad (2.13)$$

where $\delta(\cdot)$ is the Dirac delta function. Then, the Kalman filter cost [Eq. (2.9)] is

$$J_c = LPL^T, \quad (2.14)$$

which is the error covariance for the desired quantity, and the relevant matrices of the filter are

$$F = A - KC, \quad (2.15)$$

$$K = PC^T(DRD^T)^{-1}. \quad (2.16)$$

Here the matrix K is the gain matrix of the Kalman filter, when L is identity, and, therefore, is commonly referred to as the Kalman gain.

Essentially, the Kalman filter recursively calculates a new prediction of the system's state upon averaging it with a new measurement, weighted suitably based on the estimated certainty: values with higher certainty are trusted more. The Kalman filter needs only the last “best guess” instead of the entire history of a system's state to calculate a new one. The Kalman gain determines the relative certainty of the current state and the measurements and, therefore, the filter performance [35].

III. OPTIMAL FILTER

The adaptive estimator used in Ref. [16] involves a feedback filter and an offline smoother, which yields the final optimal estimate with a delay with respect to the estimation time. The feedback filter used in Ref. [16] is suboptimal and would be optimal when using a Kalman filter [19]. Here, we consider the Kalman filter alone, the output of which is the desired optimal real-time estimate, which is also used to adapt the local oscillator phase through feedback in the adaptive estimation process. First, we define the process and measurement models for our system in a state-space setting [22].

A. System model (exact)

An OU noise process modulates the phase $\phi(t)$, to be estimated, of the continuous coherent beam of light [16] and constitutes our process model:

$$\dot{\phi}(t) = -\lambda\phi(t) + \sqrt{\kappa}v(t), \quad (3.1)$$

where $\lambda > 0$ is the inverse correlation time of the phase, $\kappa > 0$ is the magnitude of the phase variation, and $v(t)$ is a zero-mean white Gaussian noise with unity amplitude.

The normalized output photocurrent of the adaptive homodyne detection of the above coherent beam is

$$\begin{aligned} I(t)dt &= 2|\alpha| \sin[\phi(t) - \hat{\phi}(t)]dt + dW(t), \\ &\simeq 2|\alpha|[\phi(t) - \hat{\phi}(t)]dt + dW(t), \end{aligned} \quad (3.2)$$

where $\hat{\phi}(t)$ is the phase estimate output of the (feedback) filter, α is the coherent amplitude of the beam, and $W(t)$ is a Wiener process arising from quantum vacuum fluctuations. Also, here we have taken a linear approximation of the sine function, since the phase estimate would be close to the true phase owing to the feedback. The measurement is appropriately scaled to yield a measurement model as follows:

$$\theta(t) := \frac{I(t)}{2|\alpha|} + \hat{\phi}(t) = \phi(t) + \frac{1}{2|\alpha|}w(t), \quad (3.3)$$

where $w(t) := dW/dt$ is another zero-mean white Gaussian noise with unity amplitude.

Thus we have the following state-space model:

$$\begin{aligned} \dot{\phi}(t) &= -\lambda\phi(t) + \sqrt{\kappa}v(t) \quad (\text{process model}), \\ \theta(t) &= \phi(t) + \frac{1}{2|\alpha|}w(t) \quad (\text{measurement model}), \end{aligned} \quad (3.4)$$

where

$$\begin{aligned} \langle v(t)v(\tau) \rangle &= \delta(t - \tau), \\ \langle w(t)w(\tau) \rangle &= \delta(t - \tau), \\ \langle v(t)w(\tau) \rangle &= 0. \end{aligned} \quad (3.5)$$

B. Kalman filter

The steady-state Kalman filter for the model (3.4) is obtained from Eqs. (2.4), (2.5), and (2.15) as

$$\begin{aligned} \dot{\hat{\phi}}(t) &= -\lambda\hat{\phi}(t) + K(\theta(t) - \hat{\phi}(t)) \\ &= -(\lambda + K)\hat{\phi}(t) + K\phi(t) + \frac{K}{2|\alpha|}w(t). \end{aligned} \quad (3.6)$$

Here, the Kalman gain K is derived from the error covariance $P = \sigma^2$ of the Kalman filter. The algebraic Riccati equation for P from Eq. (2.10) simplifies for our system to

$$-2\lambda P - 4|\alpha|^2 P^2 + \kappa = 0. \quad (3.7)$$

The stabilizing solution P^+ of Eq. (3.7) is found to be

$$P^+ = \frac{\kappa}{\lambda + \sqrt{\lambda^2 + 4\kappa|\alpha|^2}}, \quad (3.8)$$

and the Kalman gain K from Eq. (2.16) is then

$$K = -\lambda + \sqrt{\lambda^2 + 4\kappa|\alpha|^2}. \quad (3.9)$$

The transfer function (2.8) of the Kalman filter here is

$$G_K(s) := \frac{\hat{\phi}(s)}{\theta(s)} = \frac{K}{s + \lambda + K}. \quad (3.10)$$

Equations (3.8), (3.9), and (3.6) are key equations.

IV. ROBUST FILTER

In this section, we construct a robust guaranteed-cost filter of the form defined in Ref. [29], as outlined in Appendix A, corresponding to the optimal filter discussed earlier. First, we define our uncertain model.

A. Uncertain model

We consider uncertainty in the parameter λ as follows:

$$\lambda \rightarrow \lambda + \mu\delta\lambda, \quad (4.1)$$

where $|\delta| \leq 1$ is unknown, and $\mu \in [0, 1)$ is the level of uncertainty in the model.

Then, the equivalent of Eq. (3.4) for the uncertain case is

$$\begin{aligned} \dot{\phi}(t) &= -(\lambda + \mu\delta\lambda)\phi(t) + \sqrt{\kappa}v(t) \quad (\text{process model}), \\ \theta(t) &= \phi(t) + \frac{1}{2|\alpha|}w(t) \quad (\text{measurement model}). \end{aligned} \quad (4.2)$$

B. Guaranteed-cost filter

The steady-state robust filter is then obtained by solving the following Riccati equation for model (4.2):

$$-(4|\alpha|^2 - \epsilon\lambda^2)Q^2 - 2\lambda Q + \left(\frac{\mu^2}{\epsilon} + \kappa\right) = 0, \quad (4.3)$$

which is obtained from Eq. (A4) in Theorem 1 from Appendix A. Here Q is the upper bound for the error covariance of the robust filter. Also, ϵ is a positive constant, with additional conditions on its value given in Theorem 1 in Appendix A.

The stabilizing solution Q^+ of Eq. (4.3) is found to be

$$Q^+ = \frac{\frac{\mu^2}{\epsilon} + \kappa}{\lambda + \sqrt{\lambda^2 + (4|\alpha|^2 - \epsilon\lambda^2)\left(\frac{\mu^2}{\epsilon} + \kappa\right)}}. \quad (4.4)$$

The optimum value of ϵ for which the above bound [Eq. (4.4)] is minimum is obtained as

$$\epsilon_{\text{opt}} = \frac{\lambda\mu(1 - \mu) + \mu\sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}}{\kappa\lambda}. \quad (4.5)$$

Substituting this value of ϵ in Eq. (4.4) yields

$$Q^+ = \frac{\kappa}{\lambda(1 - \mu) + \sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}}. \quad (4.6)$$

The robust filter equation is then given as

$$\begin{aligned} \dot{\hat{\phi}}(t) &= (-\lambda + \epsilon_{\text{opt}}Q^+\lambda^2)\hat{\phi}(t) + 4|\alpha|^2Q^+(\theta(t) - \hat{\phi}(t)) \\ &= (-\lambda(1 - \mu) - 4|\alpha|^2Q^+)\hat{\phi}(t) \\ &\quad + 4|\alpha|^2Q^+\phi(t) + 2|\alpha|Q^+w(t). \end{aligned} \quad (4.7)$$

Note that this does not depend on δ , which is the unknown quantity, so the filter can be used. Equations (4.6) and (4.7)

are two of the main results in this paper. The transfer function of the robust filter is

$$G_R(s) := \frac{\hat{\phi}(s)}{\theta(s)} = \frac{4|\alpha|^2Q^+}{s + \lambda(1 - \mu) + 4|\alpha|^2Q^+}. \quad (4.8)$$

Clearly, when $\mu = 0$, we have $\epsilon_{\text{opt}} = 0$, so that with no uncertainty in the system, Eqs. (4.4) and (4.6) reduce to Eq. (3.8). In other words, Eq. (4.6) is the optimal mean-square error for the case where λ is known and λ is replaced with $\lambda(1 - \mu)$. Likewise, with $\mu = 0$, the robust filter equation (4.7) reduces to the Kalman filter equation (3.6), and the robust filter transfer function (4.8) reduces to the Kalman filter transfer function (3.10).

Note that the robust filter [Eqs. (4.7) and (4.8)] is the same as the Kalman filter with λ replaced by $\lambda(1 - \mu)$. Essentially, a robust filter and a Kalman filter are two different filter design algorithms; i.e., a Kalman filter is designed to be optimal with respect to a certain fixed model and a robust filter is designed to provide acceptable performance for an uncertain system model. However, there is no reason why the Kalman filter and the robust filter design methods should not lead to the same filter in certain situations, even though they are for different models.

V. FILTER MEAN-SQUARE ERRORS FOR UNCERTAIN SYSTEM

In the previous sections we considered the Kalman filter and its error for $\delta = 0$ and the robust filter's maximum error. In this section we consider the mean-square errors for the robust and Kalman filters as functions of the unknown parameter δ . We employ a Lyapunov equation [22] to obtain the desired results.

The uncertain process model from Eq. (4.2) is augmented with the filter equation, i.e., Eq. (3.6) for the Kalman filter or Eq. (4.7) for the robust filter, to yield the following state-space model:

$$\dot{\bar{\phi}}(t) = \bar{A}\bar{\phi}(t) + \bar{B}\bar{w}(t), \quad (5.1)$$

where $\bar{\phi}(t) := \begin{bmatrix} \phi(t) \\ \hat{\phi}(t) \end{bmatrix}$ and $\bar{w}(t) := \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}$. Here,

$$\bar{A} = \begin{bmatrix} -(\lambda + \mu\delta\lambda) & 0 \\ \Omega & -\Lambda \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \sqrt{\kappa} & 0 \\ 0 & \frac{\Omega}{2|\alpha|} \end{bmatrix}. \quad (5.2)$$

In the case of the Kalman filter,

$$\Omega = K = -\lambda + \sqrt{\lambda^2 + 4\kappa|\alpha|^2}, \quad (5.3)$$

$$\Lambda = \lambda + K = \sqrt{\lambda^2 + 4\kappa|\alpha|^2}. \quad (5.4)$$

In the case of the robust filter,

$$\Omega = 4|\alpha|^2Q^+ = -\lambda(1 - \mu) + \sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}, \quad (5.5)$$

$$\Lambda = \lambda(1 - \mu) + 4|\alpha|^2Q^+ = \sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}. \quad (5.6)$$

The steady-state state covariance matrix P_S is obtained by solving a Lyapunov equation as follows:

$$\bar{A}P_S + P_S\bar{A}^T + \bar{B}\bar{B}^T = 0. \quad (5.7)$$

Here, P_S is the symmetric matrix

$$P_S = \langle \bar{\phi}(t) \bar{\phi}^T(t) \rangle = \begin{bmatrix} P_1 & P_2 \\ P_2 & P_3 \end{bmatrix}, \quad (5.8)$$

where $P_1 = \langle \phi^2(t) \rangle$, $P_2 = \langle \phi(t) \hat{\phi}(t) \rangle$, and $P_3 = \langle \hat{\phi}^2(t) \rangle$. Then, the estimation error can be written as

$$\varepsilon(t) = \phi(t) - \hat{\phi}(t) = [1 \quad -1] \bar{\phi}(t), \quad (5.9)$$

which is of mean zero, since all the quantities determining $\varepsilon(t)$ are of mean zero. The error covariance is then given as

$$\sigma^2 = \langle \varepsilon^2(t) \rangle = P_1 - 2P_2 + P_3. \quad (5.10)$$

Upon solving Eq. (5.7), we get the following:

$$\begin{aligned} P_1 &= \frac{\kappa}{2(\lambda + \mu\delta\lambda)}, \\ P_2 &= \frac{\Omega\kappa}{2(\lambda + \mu\delta\lambda)(\Lambda + \lambda + \mu\delta\lambda)}, \\ P_3 &= \frac{\Omega^2}{8|\alpha|^2\Lambda} + \frac{\Omega^2\kappa}{2\Lambda(\lambda + \mu\delta\lambda)(\Lambda + \lambda + \mu\delta\lambda)}. \end{aligned} \quad (5.11)$$

Substituting Eqs. (5.11) in Eq. (5.10) and simplifying, we get

$$\sigma^2 = \kappa \frac{\lambda + \Lambda(1 + \mu\delta)}{2\Lambda(1 + \mu\delta)(\Lambda + \lambda + \mu\delta\lambda)} + \frac{\Omega^2}{8|\alpha|^2\Lambda}. \quad (5.12)$$

Thus, the mean-square error σ_K^2 of the Kalman filter for the uncertain system is

$$\begin{aligned} \sigma_K^2 &= \kappa \frac{\lambda + (K + \lambda)(1 + \mu\delta)}{2(1 + \mu\delta)(K + \lambda)(K + 2\lambda + \mu\delta\lambda)} \\ &\quad + \frac{K^2}{8|\alpha|^2(K + \lambda)}, \end{aligned} \quad (5.13)$$

and the mean-square error σ_R^2 of the robust filter for the uncertain system is

$$\sigma_R^2 = \kappa \frac{\lambda(1 - \mu)^2 + J(1 + \mu\delta)}{2J(1 + \mu\delta)(\lambda + \mu\delta\lambda + J)} + \frac{(4|\alpha|^2 Q^+)^2}{8|\alpha|^2 J}, \quad (5.14)$$

where $J := \lambda(1 - \mu) + 4|\alpha|^2 Q^+$. Equations (5.13) and (5.14) constitute two more main equations in this paper.

VI. COMPARISON OF THE MEAN-SQUARE ERRORS OF THE FILTERS

The mean-square errors of the Kalman and robust filters can be computed using the technique mentioned in the previous section for the uncertain system for $-1 \leq \delta \leq 1$ and plotted on the same graph for comparison. We used the following values for the various parameters for consistency with Ref. [17]: $\lambda = 5.9 \times 10^4$ rad/s, $\kappa = 1.9 \times 10^4$ rad/s, and $|\alpha|^2 = 1 \times 10^6$ s $^{-1}$. Figure 1 shows the comparisons for $\mu = 0.5$ and $\mu = 0.8$. Also shown on the plots are the upper bounds Q^+ of the robust filter mean-square error. Moreover, the *optimal limit* for arbitrary δ (i.e., the minimum achievable mean-square error if the exact value of λ was known at each δ) is found to be [as in Eq. (3.8)]

$$\sigma_{\text{opt}}^2 = \frac{1}{4|\alpha|^2} \left[-(\lambda + \mu\delta\lambda) + \sqrt{(\lambda + \mu\delta\lambda)^2 + 4\kappa|\alpha|^2} \right], \quad (6.1)$$

which is always less than or equal to σ^2 .

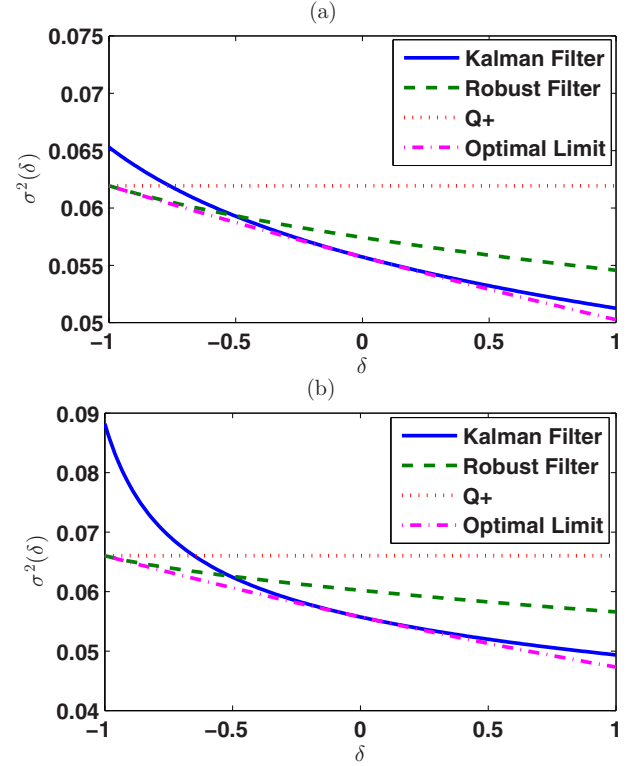


FIG. 1. Comparison of the mean-square errors of the Kalman and robust filters for (a) $\mu = 0.5$ and (b) $\mu = 0.8$.

All the curves (except Q^+) monotonically decrease with increasing δ . This is obvious from the respective expressions (5.13), (5.14), and (6.1). The Kalman filter mean-square error is lower than the robust filter mean-square error at $\delta = 0$, because the Kalman filter is optimized for the parameters of the exact model. Also, note that $K + \lambda = \sqrt{\lambda^2 + 4\kappa|\alpha|^2} \geq \sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2} = J$. Here, $K + \lambda$ is the corner frequency of the Kalman filter, whereas J is that of the robust filter. The corner frequency of the process model from Eq. (4.2) is $\lambda(1 + \mu\delta)$, which rises with increasing δ . Thus, the Kalman filter outperforms the robust filter with δ approaching $+1$, owing to the larger corner frequency of the Kalman filter, resulting in less phase information getting truncated compared to the robust filter.

However, as δ approaches -1 , the robust filter is superior to the Kalman filter. Since the corner frequency of the process model is now less than λ , the lower corner frequency of the robust filter provides an advantage. At $\delta = -1$, the worst-case mean-square error $\sigma_w^2 := \sigma^2(\delta = -1, \mu)$ of the robust filter is quantified by Q^+ , which is much lower than that of the Kalman filter. This advantage is greater with $\mu = 0.8$ than with $\mu = 0.5$. Indeed, as noted earlier, Q^+ from Eq. (4.6) is the optimal mean-square error [Eq. (3.8)] with λ replaced by $\lambda(1 - \mu)$ (i.e., $\lambda + \mu\delta\lambda$, where δ equals -1 and corresponds to the worst-case situation).

The fact that the relative guaranteed worst-case performance of the robust filter with respect to the optimal filter improves with increasing level of uncertainty μ in the model is illustrated in Fig. 2(a). We have also shown the worst-case SQL, which is computed for each μ as in Appendix B. Note

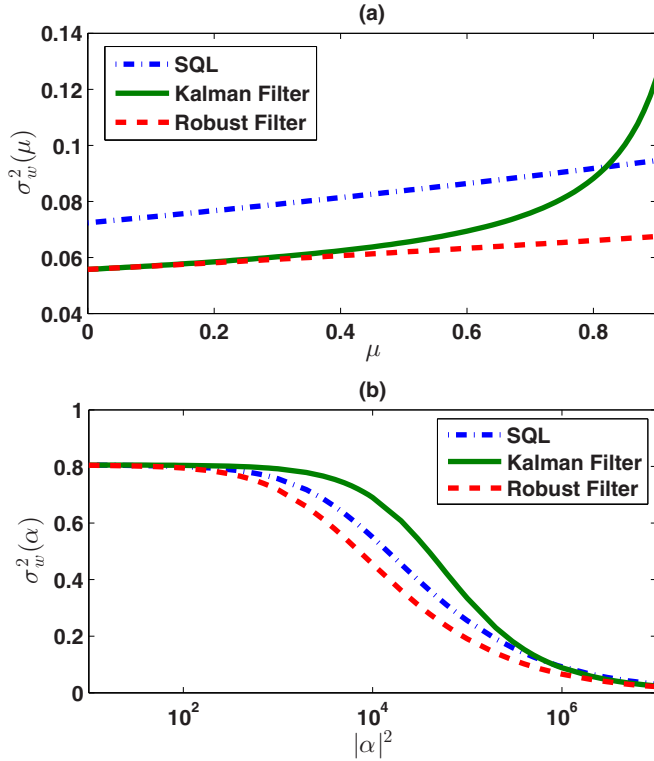


FIG. 2. Comparison of the worst-case mean-square errors of the Kalman and robust filters as a function of (a) μ and (b) $|\alpha|^2$. For (b), we fix $\mu = 0.8$.

from the plot that the worst-case robust filter mean-square error beats the SQL for all values of μ , whereas the Kalman filter worst-case mean-square error increasingly exceeds the SQL at higher values of μ . Indeed, the worst-case corner frequency of the uncertain process is $\lambda(1 - \mu)$, which goes down with increasing value of μ , while the corner frequency of the Kalman filter remains unchanged at $\sqrt{\lambda^2 + 4\kappa|\alpha|^2} \geq \lambda(1 - \mu)$, allowing for less noise getting filtered out with rising μ . Consequently, the mean-square error of the Kalman filter keeps rising with increasing μ to exceed the SQL. On the other hand, the corner frequency $\sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}$ of the robust filter consistently falls with rising μ , allowing for the mean-square error to stay below the SQL throughout.

Figure 2(b) shows a comparison of the worst-case mean-square errors of the two filters as a function of the photon flux $|\alpha|^2$ with the uncertainty level set at $\mu = 0.8$. We also plot the SQL for comparison. We see that our robust filter exhibits an optimal photon number for which its worst-case performance is the best relative to the Kalman filter. This is similar to the observation made in Ref. [22] (that the robust smoother admitted an optimal photon number when compared to the optimal smoother). Also, the worst-case mean-square error of the robust filter is always lower than the SQL, while that of the Kalman filter is not. Indeed, for low values of photon flux, the corner frequency $\sqrt{\lambda^2 + 4\kappa|\alpha|^2}$ of the Kalman filter is well beyond the corner frequency $\lambda(1 - \mu)$ of the uncertain phase, resulting in its mean-square error being quite high. However, with rising $|\alpha|^2$ beyond the point where λ^2 is comparable to $4\kappa|\alpha|^2$, the value of λ (that remains fixed

here) becomes increasingly less dominant and, therefore, the impact of any uncertainty in λ on the estimation precision keeps reducing. The behavior of the robust filter is similar, although its corner frequency $\sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}$ clearly stays much lower than that of the Kalman filter, resulting in lower mean-square errors. Also, the mean-square error of the robust filter starts falling earlier compared to the Kalman filter with rising $|\alpha|^2$ at the point where $\lambda^2(1 - \mu)^2$ is comparable to $4\kappa|\alpha|^2$. A similar observation can be made for the SQL too, the worst-case corner frequency for which is $\lambda(1 - \mu) + K_{\text{SQL}} = \sqrt{\lambda^2(1 - \mu)^2 + 2\kappa|\alpha|^2}$ from Eq. (B4).

All these plots indicate that the worst-case mean-square error, quantified by Q^+ , of our robust filter is guaranteed to beat that of the Kalman filter for all situations. Moreover, the robust filter worst-case mean-square error beats the SQL, even when the Kalman filter worst-case mean-square error does not. So if the SQL is considered as the tolerable threshold of the mean-square error for our system, our robust filter guarantees that the mean-square error never exceeds this threshold. This is exactly why robust design methods are powerful for practical engineering applications (as also noted in Ref. [22]).

VII. ERROR ANALYSIS

In this section, we explicitly prove that the robust filter mean-square error is guaranteed to beat the Kalman filter mean-square error and the SQL in the worst-case situation for all admissible values of λ , κ , $|\alpha|$, and μ .

a. Robust filter vs Kalman filter. Here, we intend to show that the worst-case Kalman filter mean-square error, i.e., Eq. (5.13) with $\delta = -1$, is always beaten by the worst-case robust filter mean-square error, given by Q^+ from Eq. (4.6). As mentioned earlier, it is immediately obvious from Eq. (4.6), when compared to Eq. (3.8), that Q^+ is the optimal mean-square error at $\delta = -1$. Therefore, it can be no larger than the mean-square error at this value of δ of the Kalman filter, which is constructed to be optimal for the exact model having $\delta = 0$ and is suboptimal everywhere else including at $\delta = -1$.

b. Robust filter vs SQL. Here, we intend to show that the worst-case robust filter error covariance, quantified by Q^+ from Eq. (4.6), always beats the SQL, given by Eq. (B3) with $\delta = -1$, for all $\lambda, \kappa, |\alpha| \geq 0$ and $0 \leq \mu < 1$.

Subtracting Eq. (4.6) from Eq. (B3) and simplifying with $\delta = -1$, we get

$$\begin{aligned} P_{\text{SQL}}(\delta = -1) - Q^+ &= \\ &= \frac{-\lambda(1 - \mu) + \sqrt{\lambda^2(1 - \mu)^2 + 2\kappa|\alpha|^2}}{2|\alpha|^2} \\ &\quad - \frac{-\lambda(1 - \mu) + \sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}}{4|\alpha|^2}. \end{aligned} \quad (7.1)$$

One can see that, in order for the above quantity to be greater than or equal to zero, we need to have

$$2\sqrt{\lambda^2(1 - \mu)^2 + 2\kappa|\alpha|^2} - \lambda(1 - \mu) \geq \sqrt{\lambda^2(1 - \mu)^2 + 4\kappa|\alpha|^2}, \quad (7.2)$$

which is proven in Appendix C.

Hence, $P_{\text{SQL}}(\delta = -1) - Q^+ \geq 0$; i.e., the robust filter mean-square error for the uncertain system is always less than

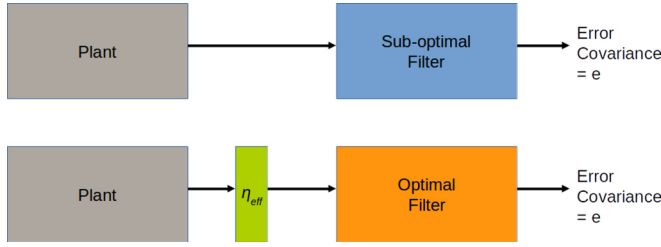


FIG. 3. A suboptimal estimation process may yield the same mean-square error as the optimal estimation process preceded by a loss.

or equal to the SQL in the worst-case situation. This is a key result in our paper.

VIII. EFFECTIVE QUANTUM EFFICIENCY

In all of the above, we have assumed perfect detector efficiency. That is, the photon flux $|\alpha|^2$ equals the square of the absolute amplitude of the input coherent state $\alpha_c e^{i\phi(t)}$. In practice, the photon flux is

$$|\alpha|^2 = \eta_d |\alpha_c|^2, \quad (8.1)$$

where $0 < \eta_d \leq 1$ is the homodyne detector efficiency.

The inefficiency of detectors increases the mean-square error, just as using a suboptimal filter does. This suggests considering an *effective quantum efficiency* for our filters. That is, the effective quantum efficiency is that which would increase the mean-square error as much as using the suboptimal filter (see Fig. 3). The effective quantum efficiency η_{eff} for a suboptimal filter is defined to be that such that the optimal filter with this efficiency gives the same mean-square error as the suboptimal filter with unit efficiency. This effective efficiency is inspired by that in Ref. [32], where a measurement scheme is mathematically identical to a combination of loss followed by a better measurement. The difference is that here we are just matching the mean-square error.

Consider the plant here to be our uncertain model from Eq. (4.2). Also, suppose a suboptimal filter yields mean-square error e as shown in the figure. Then, the mean-square error yielded by the optimal (Kalman) filter preceded by a loss introduced in the form of η_{eff} as in the figure will be equal to e if we have

$$e = \frac{1}{4\eta_{\text{eff}}|\alpha|^2} \left(-\lambda_u + \sqrt{\lambda_u^2 + 4\kappa\eta_{\text{eff}}|\alpha|^2} \right), \quad (8.2)$$

where the right-hand side has been obtained with $\lambda_u := (\lambda + \mu\delta\lambda)$ for the uncertain system (4.2) as Eq. (3.8) was obtained for the exact model (3.4). Upon solving Eq. (8.2) for η_{eff} , we get

$$\eta_{\text{eff}} = \frac{\kappa - 2e(\lambda + \mu\delta\lambda)}{4|\alpha|^2 e^2}. \quad (8.3)$$

This equation is one of the main results in this paper.

Now, consider Fig. 1 and note that the filters are suboptimal everywhere except the Kalman filter at $\delta = 0$ and the robust filter at $\delta = -1$. Then, the mean-square errors σ_K^2 from Eq. (5.13) or σ_R^2 from Eq. (5.14) for these filters (that were plotted in Fig. 1) may be obtained by an effective filter,

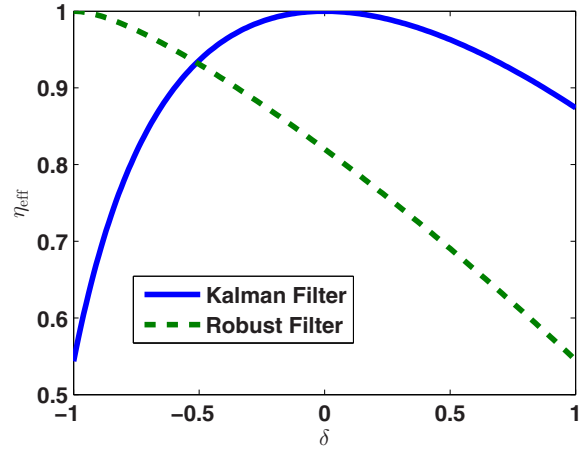


FIG. 4. Effective quantum efficiency: Comparison of the effective quantum efficiencies of the Kalman and robust filters for the uncertain system as a function of δ with $\mu = 0.8$.

involving an optimal Kalman filter preceded by a loss channel with transmissivity η_{eff} given by Eq. (8.3) with $e = \sigma_K^2$ or σ_R^2 , respectively.

At this point, let us ensure that η_{eff} from Eq. (8.3) satisfies the lower and upper bounds of 0 and 1, respectively, as desired. Note from Eq. (8.3) that we have $\eta_{\text{eff}} > 0$ if

$$e < \frac{\kappa}{2(\lambda + \mu\delta\lambda)}. \quad (8.4)$$

The right-hand side above is P_1 from Eq. (5.11). This P_1 is simply the state covariance $\langle \phi^2(t) \rangle$ for the uncertain process model in Eq. (4.2). On the other hand, e is the error covariance $\langle (\phi(t) - \hat{\phi}(t))^2 \rangle$. Note that the right-hand side of Eq. (8.4) equals the error covariance if the phase estimate is zero. In most practical situations, the phase estimate would be better than this, giving a positive efficiency.

Next, note from Eq. (8.3) that we have $\eta_{\text{eff}} \leq 1$ if

$$\begin{aligned} 4|\alpha|^2 e^2 + 2(\lambda + \mu\delta\lambda)e - \kappa &\geq 0 \\ \Leftrightarrow \left(e + \frac{(\lambda + \mu\delta\lambda) - \sqrt{(\lambda + \mu\delta\lambda)^2 + 4\kappa|\alpha|^2}}{4|\alpha|^2} \right) \\ &\times \left(e + \frac{(\lambda + \mu\delta\lambda) + \sqrt{(\lambda + \mu\delta\lambda)^2 + 4\kappa|\alpha|^2}}{4|\alpha|^2} \right) \\ &\geq 0. \end{aligned} \quad (8.5)$$

Clearly, the second term in the product on the left-hand side above is positive. The first term in the product is merely $e - \sigma_{\text{opt}}^2$, which is evidently greater than or equal to zero, since the estimation mean-square error of any filter is greater than or equal to that of the optimal filter. Hence, Eq. (8.5) is satisfied.

Now, it is interesting to see how the effective quantum efficiency η_{eff} from Eq. (8.3) varies with various parameters for our Kalman and robust filters for the uncertain system. Figure 4 shows a comparison of the effective quantum efficiencies of the filters as a function of δ for $\mu = 0.8$. The efficiency η_{eff} for the robust filter is unity in the worst-error case, i.e., at $\delta = -1$, since the robust filter is optimal for this value of δ . On the other hand, η_{eff} for the Kalman filter is far below unity in the

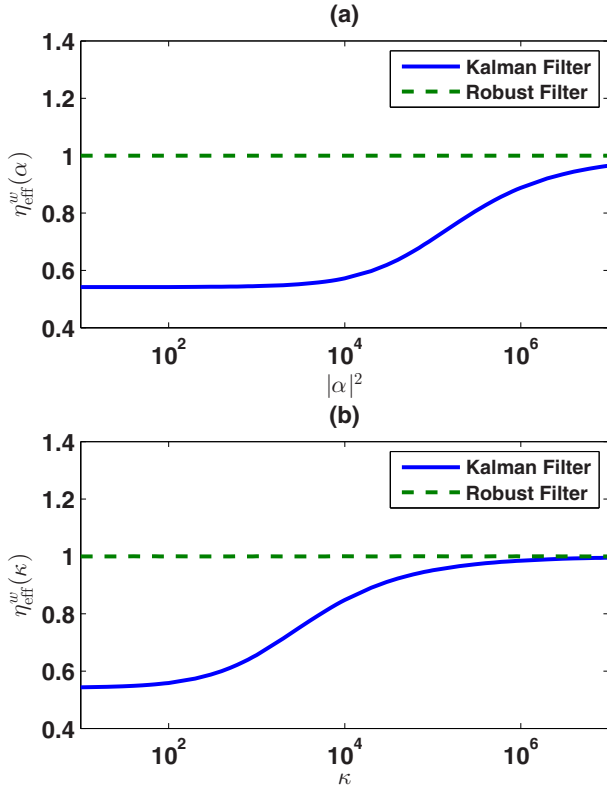


FIG. 5. Effective quantum efficiency: Comparison of the worst-error case effective quantum efficiencies of the Kalman and robust filters as a function of (a) $|\alpha|^2$ and (b) κ . Here, we fix $\mu = 0.5$.

worst-error case. Also, note that $\eta_{\text{eff}} = 1$ for the Kalman filter at $\delta = 0$, since the Kalman filter is optimal for the exact model, which corresponds to $\delta = 0$.

It is further interesting to note our robust filter's advantage over the Kalman filter in terms of the effective quantum efficiency in the worst-error case situation (i.e., at $\delta = -1$), which we denote as η_{eff}^w . Figures 5(a) and 5(b) show comparisons of η_{eff}^w of the filters as functions of $|\alpha|^2$ and κ , respectively, for $\mu = 0.5$. Since the robust filter is always optimal at $\delta = -1$, its effective quantum efficiency always remains unity in the worst-error case, whereas that of the Kalman filter varies significantly and approaches $\eta_{\text{eff}}^w = 1$ of the robust filter for increasingly higher values of κ and $|\alpha|^2$. Indeed, for low values of κ or $|\alpha|^2$, the corner frequency $\sqrt{\lambda^2 + 4\kappa|\alpha|^2}$ of the Kalman filter is well beyond the worst-case corner frequency $\lambda(1 - \mu)$ of the uncertain phase, yielding poor estimates and therefore poor effective quantum efficiency. However, beyond the point at which $4\kappa|\alpha|^2$ becomes comparable to λ^2 , the value of λ (that remains fixed here) becomes increasingly less dominant and therefore the impact of any uncertainty in λ on the quality of the estimate starts diminishing, resulting in improved effective quantum efficiency. Thus, our robust filter always achieves the maximum effective quantum efficiency that is possible in the worst-case situation, unlike the Kalman filter. This is because, while the robust filter is designed to be optimal in the worst-error case, the Kalman filter is optimal for the exact model without uncertainty.

IX. EFFECTIVE NOISE POWER

In Sec. VIII, we defined effective quantum efficiency for a suboptimal measurement with respect to the optimal measurement. However, the effect of the loss is an increase in the error covariance, which can be viewed as resulting from a scaling of the phase (the same model involving the same differential equations with suitably scaled variables) and not the suboptimal filter. We illustrate this in the following.

The measurement model equation is

$$\theta(t) = \phi(t) + \frac{1}{2|\alpha|} w(t). \quad (9.1)$$

With efficiency η we have

$$\theta(t) = \phi(t) + \frac{1}{2\sqrt{\eta}|\alpha|} w(t). \quad (9.2)$$

Now let us scale the time to $q = \sqrt{\eta}t$. Then, in terms of the new time we have a new Gaussian noise $w_1(q) = \eta^{1/4}w(t)$, which gives

$$\theta(q) = \phi(q) + \frac{1}{2\eta^{1/4}|\alpha|} w_1(q) \quad (9.3)$$

or

$$\eta^{1/4}\theta(q) = \eta^{1/4}\phi(q) + \frac{1}{2|\alpha|} w_1(q). \quad (9.4)$$

Let us define $\theta_1(q) := \eta^{1/4}\theta(q)$ and $\phi_1(q) := \eta^{1/4}\phi(q)$. In terms of these new quantities the equation is

$$\theta_1(q) = \phi_1(q) + \frac{1}{2|\alpha|} w_1(q). \quad (9.5)$$

The modified equation for ϕ is

$$\eta^{1/2}\dot{\phi}(q) = -\lambda\phi(q) + \eta^{1/4}\sqrt{\kappa}v_1(q) \quad (9.6)$$

or

$$\dot{\phi}_1(q) = -\eta^{-1/2}\lambda\phi_1(q) + \sqrt{\kappa}v_1(q). \quad (9.7)$$

Defining $\lambda_1 := \eta^{-1/2}\lambda$, we get

$$\dot{\phi}_1(q) = -\lambda_1\phi_1(q) + \sqrt{\kappa}v_1(q). \quad (9.8)$$

That is, the loss on $|\alpha|$ gives the same plant but with a scaled time, a scaled λ , and scaled phases.

Now consider the filter

$$\dot{\hat{\phi}}(t) = -\lambda\hat{\phi}(t) + K(\theta(t) - \hat{\phi}(t)). \quad (9.9)$$

In terms of the scaled time,

$$\eta^{1/2}\dot{\hat{\phi}}(q) = -\lambda\hat{\phi}(q) + K(\theta(q) - \hat{\phi}(q)). \quad (9.10)$$

Considering scaled variables as well and transferring the η to the other side,

$$\dot{\hat{\phi}}_1(q) = -(\eta^{-1/2}\lambda)\hat{\phi}_1(q) + (\eta^{-1/2}K)(\theta_1(q) - \hat{\phi}_1(q)). \quad (9.11)$$

Let us define $K_1 := \eta^{-1/2}K$, so this equation becomes

$$\dot{\hat{\phi}}_1(q) = -\lambda_1\hat{\phi}_1(q) + K_1(\theta_1(q) - \hat{\phi}_1(q)). \quad (9.12)$$

If the Kalman gain was previously calculated for a coherent amplitude of $\eta^{1/2}\alpha$, then $K = -\lambda + \sqrt{\lambda^2 + 4\kappa\eta|\alpha|^2}$. That means that $K_1 := \eta^{-1/2}K = -\lambda_1 + \sqrt{\lambda_1^2 + 4\kappa|\alpha|^2}$. This, in

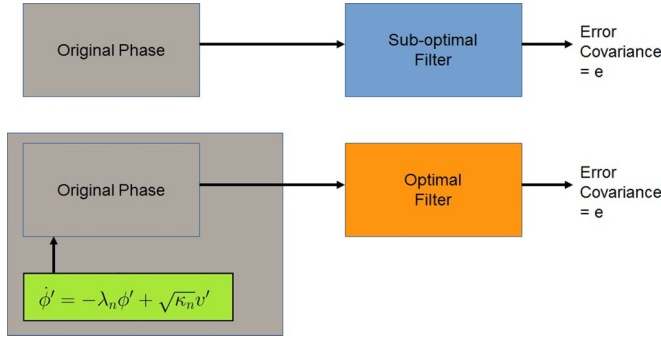


FIG. 6. A suboptimal estimation process of a phase may yield the same mean-square error as the optimal estimation process of the phase plus additional noise.

turn, means that K_1 is the desired Kalman gain for the state with scaled λ without loss. The error covariance is therefore just multiplied by the scaling factor between ϕ_1 and ϕ , and is therefore $\eta^{-1/2}$ times the error covariance without loss. Thus, the effect of the loss can be absorbed into scaling of the various variables to give the same differential equations as before, but still with the optimal Kalman filter (and not the suboptimal filter). In other words, the loss is equivalent to having a phase that varies by a larger amount, which would be like adding some noise and trying to measure the joint phase variation plus noise accurately.

Here, we intend to define an *effective noise power* for the OU phase under measurement, motivated by the result that a suboptimal measurement of a constant phase is equivalent to an optimal measurement on a different state [33,34]. In some cases the different state can correspond to the original state with added phase noise. Here, we instead consider added phase noise with an optimal measurement for the varying phase case. We consider an additional OU noise $\dot{\phi}' := -\lambda_n \phi' + \sqrt{\kappa_n} v'$, added to the original phase ϕ . This is depicted schematically in Fig. 6. We measure $\phi + \phi'$, but estimate only ϕ and not $\phi + \phi'$. This is distinct from trying to estimate $\phi + \phi'$, which is directly equivalent to a scaled scheme of measurement with the optimal filter (similar to the case with effective quantum efficiency discussed above).

The model can then be given in state-space form as

$$\begin{aligned} \dot{\bar{\phi}} &= A\bar{\phi} + B\bar{v}, \\ \theta' &= C\bar{\phi} + Dw, \\ z &= L\bar{\phi}, \end{aligned} \quad (9.13)$$

where z determines the parameter to be estimated, and

$$\begin{aligned} \bar{\phi} &:= \begin{bmatrix} \phi \\ \phi' \end{bmatrix}, \quad \bar{v} := \begin{bmatrix} v \\ v' \end{bmatrix}, \\ A &= \begin{bmatrix} -\lambda_u & 0 \\ 0 & -\lambda_n \end{bmatrix}, \quad B = \begin{bmatrix} \sqrt{\kappa} & 0 \\ 0 & \sqrt{\kappa_n} \end{bmatrix}, \\ C &= [1 \quad 1], \quad D = \frac{1}{2|\alpha|}, \quad L = [1 \quad 0]. \end{aligned} \quad (9.14)$$

Here, $\lambda_u = \lambda + \mu\delta\lambda$ and v, v' , and w are mutually uncorrelated Gaussian white noises, each with unity amplitude. Also, we take $\lambda_n = \lambda_u$ for simplicity. Note that while the new

measurement θ' considers the measurement of the combination $\phi + \phi'$ (through C), the variable to be estimated, z , is just the original phase ϕ (through L). Then, if the optimal steady-state estimate is given by \hat{z} , the optimal mean-square estimation error $\langle (z - \hat{z})^2 \rangle$ is obtained by solving the following Kalman filter algebraic Riccati equation:

$$AP + PA^T - PC^T(DD^T)^{-1}CP + BB^T = 0, \quad (9.15)$$

where P is the symmetric matrix

$$P := \begin{bmatrix} P_1 & P_2 \\ P_2 & P_3 \end{bmatrix}, \quad (9.16)$$

such that the desired mean-square estimation error is

$$\langle (z - \hat{z})^2 \rangle = LPL^T = P_1. \quad (9.17)$$

Now, expanding Eq. (9.15), we get the following equations:

$$-2\lambda_u P_1 + \kappa - 4|\alpha|^2(P_1 + P_2)^2 = 0, \quad (9.18)$$

$$-2\lambda_u P_2 - 4|\alpha|^2(P_1 + P_2)(P_2 + P_3) = 0, \quad (9.19)$$

$$-2\lambda_u P_3 + \kappa_n - 4|\alpha|^2(P_2 + P_3)^2 = 0. \quad (9.20)$$

Upon solving the system of equations (9.18)–(9.20) for P_1, P_2 , and P_3 , we get

$$\begin{aligned} P_1 &= \frac{\kappa(2|\alpha|^2\kappa_n(\kappa + \kappa_n) - \kappa\lambda_u^2 + \kappa\lambda_u\beta)}{4|\alpha|^2(\kappa + \kappa_n)^2\lambda_u}, \\ P_2 &= \frac{-2|\alpha|^2\kappa\kappa_n(\kappa + \kappa_n) - \kappa\kappa_n\lambda_u^2 + \kappa\kappa_n\lambda_u\beta}{4|\alpha|^2(\kappa + \kappa_n)^2\lambda_u}, \\ P_3 &= \frac{\kappa_n(2|\alpha|^2\kappa(\kappa + \kappa_n) - \kappa_n\lambda_u^2 + \kappa_n\lambda_u\beta)}{4|\alpha|^2(\kappa + \kappa_n)^2\lambda_u}, \\ \beta &:= \sqrt{4|\alpha|^2(\kappa + \kappa_n) + \lambda_u^2}. \end{aligned} \quad (9.21)$$

Next, we deduce $\kappa_n \geq 0$, such that the optimal filter for Eq. (9.13) yields a variance (9.17) equal to the estimation mean-square error e of a given arbitrary filter for the original model [Eq. (4.2)]. Thus, we should have $P_1 = e$ and κ_n may be derived to be

$$\kappa_n = \frac{\kappa\lambda_u(2|\alpha|e - \sqrt{\kappa - 2\lambda_u e})}{|\alpha|(\kappa - 2\lambda_u e)}. \quad (9.22)$$

The *effective noise power* of the phase $\phi + \phi'$ being measured is then given by

$$\kappa_{\text{eff}} := [\sqrt{\kappa} \quad \sqrt{\kappa_n}] \begin{bmatrix} \sqrt{\kappa} \\ \sqrt{\kappa_n} \end{bmatrix} = \kappa + \kappa_n. \quad (9.23)$$

This is because one may verify that the phase being measured (not estimated) is effectively another OU process as below obtained from Eq. (9.13):

$$\dot{\phi} + \dot{\phi}' = -\lambda_u(\phi + \phi') + [\sqrt{\kappa} \quad \sqrt{\kappa_n}] \begin{bmatrix} v \\ v' \end{bmatrix}. \quad (9.24)$$

Equations (9.23) and (9.22) are two more main results.

Note that in measurement theory for a constant phase, modifying the state suitably and using an optimal measurement gives the same result as a suboptimal measurement [33,34]. As a result, it yields the same overall probability distribution for the estimation errors, not just the mean-square error. This means that the measurement can be regarded as being the

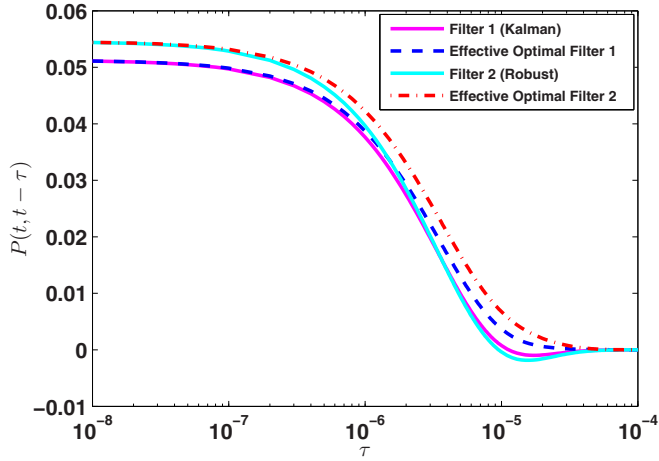


FIG. 7. Comparison between the two-time error correlations of various filters as a function of the time difference τ for $\mu = 0.5$ and $\delta = 1$. The various traces in the plot are as follows: (i) the solid pink line is the two-time error correlation at each τ of the Kalman filter constructed for the exact model and applied to the uncertain system; (ii) the solid cyan line is the two-time error correlation at each τ of the robust filter constructed for the uncertain model and applied to the uncertain system; (iii) the dotted blue line is the two-time error correlation at each τ of the optimal filter constructed for the uncertain model plus an added noise corresponding to the Kalman filter of trace (i); and (iv) the dotted red line is the two-time error correlation at each τ of the optimal filter constructed for the uncertain model plus an added noise corresponding to the robust filter of trace (ii).

same in Fig. 6 in a deeper sense than considered so far here by demonstrating that it will reproduce the same probability distribution for the measurement errors. This can be tested for a varying phase, as considered here, by demonstrating the equivalence not just in terms of mean-square errors but also in terms of the two-time error correlations.

Figure 7 shows the comparison of the two-time error correlations of the various filters for $\delta = 1$, $\mu = 0.5$, $\lambda_n = \lambda_u$, and κ_n as derived in Eq. (9.22) and, as used before, $\lambda = 5.9 \times 10^4$ rad/s, $\kappa = 1.9 \times 10^4$ rad/s, and $|\alpha|^2 = 1 \times 10^6$ s $^{-1}$. The two-time error correlations are computed as in Appendix D. Clearly, the effective optimal filter we obtained for the Kalman filter is not equivalent to the Kalman filter at this value of δ in terms of the two-time error correlations; likewise for the robust filter. Thus, we observe that with an added OU noise as considered here, the effective optimal filter is equivalent to the suboptimal filter (our Kalman filter at $\delta \neq 0$ or our robust filter at $\delta \neq -1$) in terms of mean-square estimation errors (this is expected since κ_n of the added OU noise is derived such that the mean-square errors match up) only, and not two-time error correlations as well. There are analytic reasons, as presented in Appendix E, that suggest that the two-time error correlations cannot possibly be made to match up. However, for the purposes of our work here, where our focus is to define and compare the effective noise powers for our different estimators, it suffices to use only mean-square estimation errors.

We now show in Fig. 8 a comparison of the effective noise powers of the filters as a function of δ for $\mu = 0.5$. The effective noise power κ_{eff} for the robust filter is at a minimum

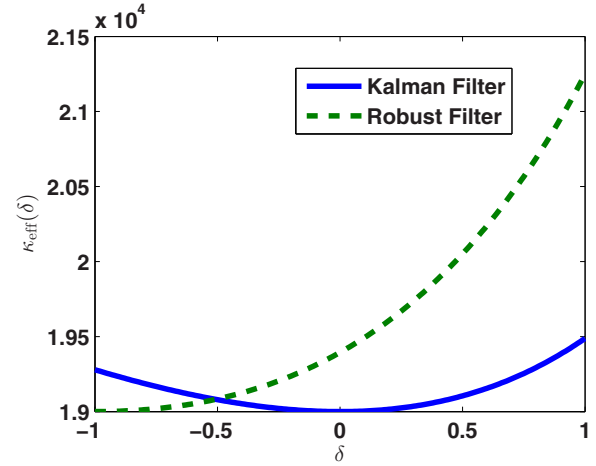


FIG. 8. Effective noise power: Comparison of the effective noise powers of the Kalman and robust filters for the uncertain system as a function of δ with $\mu = 0.5$.

and equal to κ in the worst-error case, i.e., at $\delta = -1$, since the robust filter is optimal for this value of δ . By contrast, κ_{eff} for the Kalman filter is more than that of the robust filter in the worst-error case. Also, note that κ_{eff} is at a minimum (i.e., equal to κ) for the Kalman filter at $\delta = 0$, since the Kalman filter is optimal for the exact model, which corresponds to $\delta = 0$.

It is interesting to note our robust filter's advantage over the Kalman filter in terms of the effective noise power in the

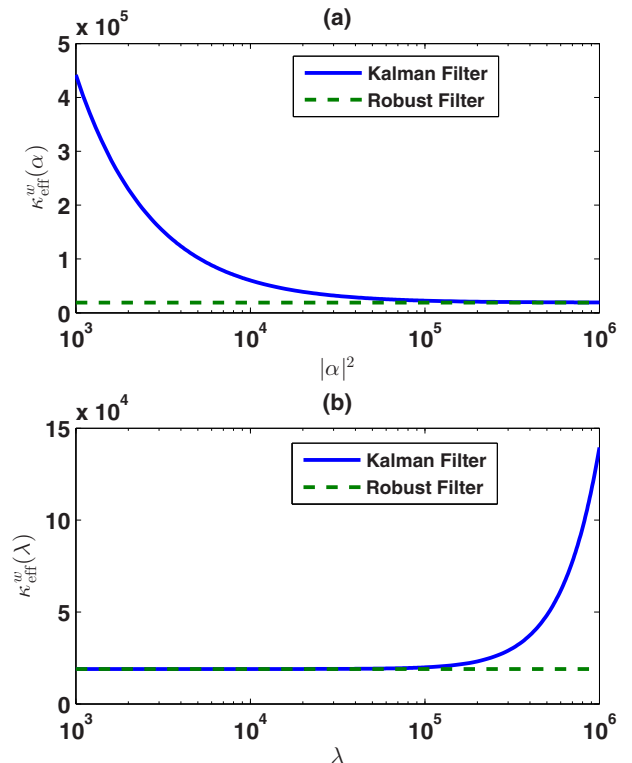


FIG. 9. Effective noise power: Comparison of the worst-error case effective noise powers of the Kalman and robust filters as a function of (a) $|\alpha|^2$ and (b) λ . Here, we fix $\mu = 0.5$.

worst-error case situation (i.e., at $\delta = -1$), that we denote as κ_{eff}^w . Figures 9(a) and 9(b) show comparisons of κ_{eff}^w of the filters as functions of $|\alpha|^2$ and λ , respectively, with $\mu = 0.5$. We note that the worst-error case effective noise power of the robust filter always remains constant at κ , while that of the Kalman filter approaches $\kappa_{\text{eff}}^w = \kappa$ of the robust filter for higher values of $|\alpha|^2$ and lower values of λ . Intuitively, since the worst-case mean-square error rises for the Kalman filter with decreasing $|\alpha|^2$, the corresponding effective optimal filter incurs higher noise power. Moreover, with rising λ , the size of the uncertainty in the model increases, resulting in higher effective noise power for the Kalman filter.

All these plots show that when considering the worst-case scenario, our robust filter always incurs the least effective noise possible, whereas the Kalman filter often effectively suffers from much higher noise due to the uncertainty in the model. This is again because of the fact that the robust filter design involves optimizing the error in the worst-case situation, while that for the Kalman filter considers zero uncertainty in the model.

X. CONCLUSION

In this paper, we studied guaranteed-cost robust adaptive quantum phase estimation for a coherent state with parametric uncertainty explicitly introduced in the model in a systematic state-space setting. Our robust filter comes with a guaranteed cost that quantifies the worst-case mean-square estimation error, which is guaranteed to beat that of the Kalman filter and the SQL for the uncertain system. We showed that if the SQL is taken as the tolerable threshold, our robust filter mean-square error always remains under this threshold unlike the Kalman filter. We illustrated that this guaranteed cost allows for our robust filter to always achieve the maximum effective quantum efficiency and the minimum possible effective noise power in the worst-case scenario. In contrast to the robust fixed-interval smoother from Ref. [22], our robust filter considered here helps overcome practical challenges facing quantum parameter estimation in applications such as quantum computing, where real-time estimates are of greater interest than more precise estimates.

Moreover, in deriving an effective noise power for our robust and Kalman filters, we matched only the mean-square errors of these filters for the uncertain system with the optimal filter with an added noise. We saw that it does not seem possible to also match up the two-time error correlations. Thus, another key result here is that it appears that the results from Refs. [33,34] for a constant phase do not generalize to the case of a varying phase, such that the overall probability distributions of the estimation errors match up.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council (ARC) under Grants No. FL110100020 (I.R.P.) and No. CE110001027 (E.H.H.). S.R. was funded by the Singapore National Research Foundation Grant No. NRF-NRFF2011-07 and the Singapore Ministry of Education Academic Research Fund Tier 1 Project No. R-263-000-C06-112, and is currently funded by the UK National Quantum Technologies Pro-

gramme (Grants No. EP/M01326X/1 and No. EP/M013243/1). D.W.B. is funded by an ARC Future Fellowship (Grant No. FT100100761) and an ARC Discovery Project (Project No. DP160102426). This work was also supported by the U.S. Air Force Office of Scientific Research (AFOSR). This material is based on research sponsored by the Air Force Research Laboratory, under Agreements No. FA2386-12-1-4075 and No. FA2386-16-1-4065 (I.R.P.). The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government. Moreover, S.R. thanks Howard Wiseman, Hongbin Song, Aleksandar Davidovic, Obaid Ur Rehman, Trevor Wheatley, Hidehiro Yonezawa, Mohamed Mabrok, and Katanya Kuntz for discussions in relation to this work.

APPENDIX A: ROBUST GUARANTEED-COST FILTER

We consider here the theory underlying a steady-state robust estimator for a class of uncertain linear systems with norm-bounded uncertainty [29]. This extends the steady-state Kalman filter to the case in which the underlying system is also uncertain.

The class of uncertain systems considered is described by the state equations

$$\begin{aligned} \dot{x}(t) &= [A + D_1 \Delta(t) E_1] x(t) + w_1(t), & x(t_0) &= x_0, \\ y(t) &= [C + D_2 \Delta(t) E_1] x(t) + w_2(t), \end{aligned} \quad (\text{A1})$$

where $x(t) \in \mathbb{R}^n$ is the state, $y(t) \in \mathbb{R}^l$ is the measured output, x_0 is the initial condition which is assumed to be a zero-mean Gaussian random vector, $\Delta(t)$ is a time-varying matrix of uncertain parameters satisfying the bound $\Delta^T(t) \Delta(t) \leq I$, and $w_1(t)$ and $w_2(t)$ are zero-mean white Gaussian noise processes with joint covariance matrix

$$\left\langle \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \begin{bmatrix} w_1^T & w_2^T \end{bmatrix} \right\rangle = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} > 0. \quad (\text{A2})$$

Since we are dealing with the steady-state filtering problem, we assume that the initial time $t_0 \rightarrow -\infty$.

Theorem 1 [29]. Consider the uncertain system (A1), which is assumed to be quadratically stable. Then, there exists a constant $\epsilon^* > 0$, such that for all $\epsilon \in (0, \epsilon^*)$, the Riccati equation

$$AS + SA^T + \epsilon S E_1^T E_1 S + \frac{1}{\epsilon} D_1 D_1^T + V_1 = 0 \quad (\text{A3})$$

has a stabilizing solution $S^+ > 0$. For any such ϵ , the Riccati equation

$$\begin{aligned} & (A - D_1 D_2^T (\epsilon V_2 + D_2 D_2^T)^{-1} C) Q \\ & + Q (A - D_1 D_2^T (\epsilon V_2 + D_2 D_2^T)^{-1} C)^T \\ & + \epsilon Q E_1^T E_1 Q - \epsilon Q C^T (\epsilon V_2 + D_2 D_2^T)^{-1} C Q \\ & + \frac{1}{\epsilon} D_1 [I - D_2^T (\epsilon V_2 + D_2 D_2^T)^{-1} D_2] D_1^T \\ & + V_1 = 0 \end{aligned} \quad (\text{A4})$$

has a stabilizing solution $Q^+ > 0$, such that $Q^+ \leq S^+$. Also, the state estimator

$$\begin{aligned} \dot{\hat{x}}(t) = & (A + \epsilon Q^+ E_1^T E_1) \hat{x}(t) + (\epsilon Q^+ C^T + D_1 D_2^T) \\ & \times (\epsilon V_2 + D_2 D_2^T)^{-1} [y(t) - C \hat{x}(t)] \end{aligned} \quad (\text{A5})$$

has the following property: Given any $\delta > 0$, there exists a matrix $\tilde{Q} > 0$ such that $Q^+ \leq \tilde{Q} < Q^+ + \delta I$ and Eq. (A5) is a quadratic guaranteed-cost state estimator for the system (A1) with cost matrix \tilde{Q} .

Conversely, given any quadratic guaranteed-cost state estimator for the system (A1) with cost matrix \tilde{Q} , there exists a constant $\epsilon > 0$, such that Riccati equations (A3) and (A4) have stabilizing solutions $S^+ > 0$ and $Q^+ > 0$, respectively, and $Q^+ < \tilde{Q}$.

Moreover, the quadratic guaranteed-cost state estimator defined by Eqs. (A4) and (A5) has the following property: The steady-state error covariance matrix at time t satisfies the bound $Q_\Delta(t) \leq Q^+$ for all admissible uncertainties $\Delta(t)$ [29].

Furthermore, the optimal guaranteed-cost state estimator for the uncertain system (A1) can be obtained by choosing $\epsilon > 0$ to minimize $\text{Tr}(Q^+)$, where $Q^+ > 0$ is the stabilizing solution to Riccati equation (A4). This minimization is also subject to the constraint that Eq. (A3) has a stabilizing solution $S^+ > 0$ [29].

APPENDIX B: STANDARD QUANTUM LIMIT

The SQL is set by the minimum error in phase estimation that can be obtained using a perfect heterodyne technique or, in other words, a nonadaptive filtering scheme [22,31,37]. We deduce the minimum error covariance for the case of an OU noise process from Eq. (4.2) using the standard Kalman filtering approach [37]. The underlying principle used in the analysis is that the heterodyne measurement scheme is equivalent to, and incurs the same noise penalty as, the *dual-homodyne* scheme [16].

The process and measurement equations for the steady-state Kalman filter, yielding the SQL, are [37]

$$\begin{aligned} \dot{\phi} = & -(\lambda + \mu\delta\lambda)\phi + \sqrt{\kappa}v \quad (\text{process model}), \\ \vartheta = & \phi + \frac{1}{2|\alpha|}n_1 + \frac{1}{2|\alpha|}n_2 \quad (\text{measurement model}), \end{aligned} \quad (\text{B1})$$

where n_1 is the measurement noise of one of the homodyne detectors HD1 of the dual-homodyne measurement, n_2 is the noise arising from the vacuum entering the empty port of the input beam splitter corresponding to the arm having HD1 (refer to Fig. C1 from Ref. [22]), and ϑ determines the net measurement current obtained from the dual-homodyne scheme.

Then, the algebraic Riccati equation required to be solved is

$$-2(\lambda + \mu\delta\lambda)P_{\text{SQL}} - 2|\alpha|^2 P_{\text{SQL}}^2 + \kappa = 0, \quad (\text{B2})$$

where P_{SQL} is the desired error covariance (SQL).

The stabilizing solution of the above equation is

$$P_{\text{SQL}} = \frac{-(\lambda + \mu\delta\lambda) + \sqrt{(\lambda + \mu\delta\lambda)^2 + 2\kappa|\alpha|^2}}{2|\alpha|^2}. \quad (\text{B3})$$

This determines the desired SQL, which we have included in our plots in Sec. VI. In addition, note that the transfer function of the Kalman filter here is

$$G_{\text{SQL}}(s) := \frac{K_{\text{SQL}}}{s + \lambda_u + K_{\text{SQL}}}, \quad (\text{B4})$$

where $\lambda_u = \lambda + \mu\delta\lambda$ and K_{SQL} is the Kalman gain, found to be $K_{\text{SQL}} = -\lambda_u + \sqrt{\lambda_u^2 + 2\kappa|\alpha|^2}$.

APPENDIX C: PROOF FOR ROBUST FILTER VS SQL

We start with the obvious inequality:

$$\kappa^2|\alpha|^4 \geq 0 \quad (\text{C1})$$

$$\begin{aligned} \Rightarrow & \kappa^2|\alpha|^4 + 2\kappa|\alpha|^2\lambda^2(1-\mu)^2 + \lambda^4(1-\mu)^4 \\ \geq & 2\kappa|\alpha|^2\lambda^2(1-\mu)^2 + \lambda^4(1-\mu)^4 \end{aligned} \quad (\text{C2})$$

$$\begin{aligned} \Rightarrow & (\kappa|\alpha|^2 + \lambda^2(1-\mu)^2)^2 \\ \geq & \lambda^2(1-\mu)^2(\lambda^2(1-\mu)^2 + 2\kappa|\alpha|^2) \end{aligned} \quad (\text{C3})$$

$$\begin{aligned} \Rightarrow & \kappa|\alpha|^2 + \lambda^2(1-\mu)^2 \\ \geq & \lambda(1-\mu)\sqrt{\lambda^2(1-\mu)^2 + 2\kappa|\alpha|^2} \end{aligned} \quad (\text{C4})$$

$$\begin{aligned} \Rightarrow & 2\kappa|\alpha|^2 + \lambda^2(1-\mu)^2 - \lambda(1-\mu) \\ \times & \sqrt{\lambda^2(1-\mu)^2 + 2\kappa|\alpha|^2} \geq \kappa|\alpha|^2 \end{aligned} \quad (\text{C5})$$

$$\begin{aligned} \Rightarrow & 4(2\kappa|\alpha|^2 + \lambda^2(1-\mu)^2) + \lambda^2(1-\mu)^2 \\ - & 4\lambda(1-\mu)\sqrt{\lambda^2(1-\mu)^2 + 2\kappa|\alpha|^2} \\ \geq & 4\kappa|\alpha|^2 + \lambda^2(1-\mu)^2 \end{aligned} \quad (\text{C6})$$

$$\begin{aligned} \Rightarrow & (2\sqrt{\lambda^2(1-\mu)^2 + 2\kappa|\alpha|^2} - \lambda(1-\mu))^2 \\ \geq & 4\kappa|\alpha|^2 + \lambda^2(1-\mu)^2 \end{aligned} \quad (\text{C7})$$

$$\begin{aligned} \Rightarrow & 2\sqrt{\lambda^2(1-\mu)^2 + 2\kappa|\alpha|^2} - \lambda(1-\mu) \\ \geq & \sqrt{\lambda^2(1-\mu)^2 + 4\kappa|\alpha|^2}. \end{aligned} \quad (\text{C8})$$

APPENDIX D: TWO-TIME ERROR CORRELATIONS

Here, we devise methods to compute the two-time error correlations of an optimal Kalman filter and also an arbitrary (suboptimal) filter for a given system.

1. Kalman filter

Given the process and measurement models as follows,

$$\begin{aligned} \dot{x}(t) = & Ax(t) + Bv(t), \quad x(0) = x_0 \quad (\text{process model}), \\ y(t) = & Cx(t) + Dw(t) \quad (\text{measurement model}), \end{aligned} \quad (\text{D1})$$

where

$$\begin{aligned}\langle v(t)v^T(\tau) \rangle &= Q\delta(t-\tau), \\ \langle w(t)w^T(\tau) \rangle &= R\delta(t-\tau), \\ \langle v(t)w^T(\tau) \rangle &= 0, \\ \langle v(t)x^T(0) \rangle &= \langle w(t)x^T(0) \rangle = 0,\end{aligned}\quad (\text{D2})$$

the optimal steady-state Kalman filter is given by

$$\dot{\hat{x}}(t) = A\hat{x}(t) + K[y(t) - C\hat{x}(t)], \quad \hat{x}(0) = \hat{x}_0. \quad (\text{D3})$$

Here, K is the Kalman gain, $K = PC^T(DRD^T)^{-1}$, where P is the error covariance (mean-square error), obtained by solving the algebraic Riccati equation:

$$AP + PA^T - PC^T(DRD^T)^{-1}CP + BQB^T = 0. \quad (\text{D4})$$

Then, the estimation error is

$$\varepsilon(t) = x(t) - \hat{x}(t). \quad (\text{D5})$$

Subtracting the filter equation from the process model,

$$\begin{aligned}\dot{x}(t) - \dot{\hat{x}}(t) &= (A - KC)[x(t) - \hat{x}(t)] + Bv(t) - KDw(t), \\ \dot{\varepsilon}(t) &= (A - KC)\varepsilon(t) + Bv(t) - KDw(t).\end{aligned}\quad (\text{D6})$$

If $\Phi(t,0)$ is the state transition matrix for $(A - KC)$, then the solution of the above equation is

$$\begin{aligned}\varepsilon(t) &= \Phi(t,0)\varepsilon(0) + \int_0^t \Phi(t,s)Bv(s) ds \\ &\quad - \int_0^t \Phi(t,s)KDw(s) ds.\end{aligned}\quad (\text{D7})$$

Post-multiplying Eq. (D6) by $\varepsilon^T(t-\tau)$, where $\tau \geq 0$ and taking the expectation, we get

$$\begin{aligned}\langle \dot{\varepsilon}(t)\varepsilon^T(t-\tau) \rangle &= (A - KC)\langle \varepsilon(t)\varepsilon^T(t-\tau) \rangle \\ &\quad + B\langle v(t)\varepsilon^T(t-\tau) \rangle - KD\langle w(t)\varepsilon^T(t-\tau) \rangle.\end{aligned}\quad (\text{D8})$$

Similarly, we obtain

$$\begin{aligned}\langle \varepsilon(t)\varepsilon^T(t-\tau) \rangle &= \langle \varepsilon(t)\varepsilon^T(t-\tau) \rangle(A - KC)^T \\ &\quad + \langle \varepsilon(t)v^T(t-\tau) \rangle B^T \\ &\quad - \langle \varepsilon(t)w^T(t-\tau) \rangle D^T K^T.\end{aligned}\quad (\text{D9})$$

Define $P(t,t-\tau) := \langle \varepsilon(t)\varepsilon^T(t-\tau) \rangle$. Then, upon adding Eqs. (D8) and (D9), we get

$$\begin{aligned}\dot{P}(t,t-\tau) &= (A - KC)P(t,t-\tau) + P(t,t-\tau)(A - KC)^T \\ &\quad + B\langle v(t)\varepsilon^T(t-\tau) \rangle + \langle \varepsilon(t)v^T(t-\tau) \rangle B^T \\ &\quad - KD\langle w(t)\varepsilon^T(t-\tau) \rangle - \langle \varepsilon(t)w^T(t-\tau) \rangle D^T K^T.\end{aligned}\quad (\text{D10})$$

From Eq. (D7), we get

$$\begin{aligned}\langle \varepsilon(t)v^T(t-\tau) \rangle &= \int_0^t \Phi(t,s)B\langle v(s)v^T(t-\tau) \rangle ds \\ &= \int_0^t \Phi(t,s)BQ\delta(s-t+\tau) ds \\ &= \begin{cases} 0, & \tau > t > 0 \\ \frac{1}{2}BQ, & \tau \in \{0,t\} \\ \Phi(t,t-\tau)BQ, & t > \tau > 0, \end{cases}\end{aligned}\quad (\text{D11})$$

using the following three properties of the Dirac delta function:

$$\begin{aligned}\int_0^q \delta(p) dp &= \int_{-q}^0 \delta(p) dp = \frac{1}{2}, \\ \int_{-q}^q \delta(p) dp &= 1, \\ \int_{-q}^q f(p)\delta(p) dp &= f(0),\end{aligned}\quad (\text{D12})$$

where $q > 0$ and $f(\cdot)$ is some function. Similarly, we get

$$\begin{aligned}\langle v(t)\varepsilon^T(t-\tau) \rangle &= \begin{cases} QB^T\Phi^T(t-\tau,t), & \tau > t > 0 \\ \frac{1}{2}QB^T, & \tau \in \{0,t\} \\ 0, & t > \tau > 0, \end{cases} \\ \langle \varepsilon(t)w^T(t-\tau) \rangle &= \begin{cases} 0, & \tau > t > 0 \\ -\frac{1}{2}KDR, & \tau \in \{0,t\} \\ -\Phi(t,t-\tau)KDR, & t > \tau > 0, \end{cases}\end{aligned}\quad (\text{D14})$$

$$\langle w(t)\varepsilon^T(t-\tau) \rangle = \begin{cases} -RD^T K^T \Phi^T(t-\tau,t), & \tau > t > 0 \\ -\frac{1}{2}RD^T K^T, & \tau \in \{0,t\} \\ 0, & t > \tau > 0. \end{cases}\quad (\text{D15})$$

Then, Eq. (D10) can be expressed as

$$\dot{P}(t,t-\tau) = \begin{cases} (A - KC)P(t,t-\tau) + P(t,t-\tau)(A - KC)^T + \Phi(t,t-\tau)BQB^T + \Phi(t,t-\tau)KDRD^T K^T, & t > \tau > 0 \\ (A - KC)P(t,t-\tau) + P(t,t-\tau)(A - KC)^T + BQB^T + KDRD^T K^T, & \tau \in \{0,t\} \\ (A - KC)P(t,t-\tau) + P(t,t-\tau)(A - KC)^T + BQB^T\Phi^T(t-\tau,t) + KDRD^T K^T\Phi^T(t-\tau,t), & t > \tau > 0. \end{cases}\quad (\text{D16})$$

It can be verified that the above equation for $\tau = 0$ reduces to the standard Kalman filter matrix differential Riccati equation.

At steady state, the left-hand side above is zero, yielding Eq. (D4) at $\tau = 0$. Also, the state transition matrix may be computed as follows:

$$\Phi(t,t-\tau) = e^{(A-KC)\tau} = \Phi^{-1}(t-\tau,t). \quad (\text{D17})$$

Usually, by the notation $P(t, t - \tau)$, we would imply $t \geq \tau \geq 0$, hence the third case in Eqs. (D16) will be ignored since the state transition matrix in this case represents backward-time state transition. However, given $t \geq \tau \geq 0$, the third case allows for verifying that $P(t - \tau, t) = P^T(t, t - \tau)$, as expected. Then, the steady-state matrix equation to be solved to obtain $P(t, t - \tau)$ is

$$0 = (A - KC)P(t, t - \tau) + P(t, t - \tau)(A - KC)^T + e^{(A-KC)\tau} B Q B^T + e^{(A-KC)\tau} K D R D^T K^T. \quad (\text{D18})$$

2. Arbitrary filter

Here, we extend Sec. V to deduce the form of Lyapunov equation required to be solved to obtain the two-time state covariance matrix and therefore the two-time error correlations of an arbitrary filter for a given system state-space model.

Let the system process and measurement models and an arbitrary filter for the system be determined by the following differential equations:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bv(t), \quad x(0) = x_0 \quad (\text{process model}), \\ y(t) &= Cx(t) + Dw(t) \quad (\text{measurement model}), \\ \dot{\hat{x}}(t) &= A_e \hat{x}(t) + B_e y(t) \\ &= A_e \hat{x}(t) + B_e Cx(t) + B_e Dw(t) \quad (\text{filter equation}), \end{aligned} \quad (\text{D19})$$

where again

$$\begin{aligned} \langle v(t)v^T(\tau) \rangle &= Q\delta(t - \tau), \\ \langle w(t)w^T(\tau) \rangle &= R\delta(t - \tau), \\ \langle v(t)w^T(\tau) \rangle &= 0, \\ \langle v(t)x^T(0) \rangle &= \langle w(t)x^T(0) \rangle = 0. \end{aligned} \quad (\text{D20})$$

We obtain an augmented system as follows from the process model and the filter equation:

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{B}\bar{w}(t), \quad (\text{D21})$$

where

$$\begin{aligned} \bar{x}(t) &= \begin{bmatrix} x(t) \\ \hat{x}(t) \end{bmatrix}, \quad \bar{w}(t) = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}, \\ \bar{A} &= \begin{bmatrix} A & 0 \\ B_e C & A_e \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} B & 0 \\ 0 & B_e D \end{bmatrix}. \end{aligned} \quad (\text{D22})$$

Then, we get for $\tau \geq 0$,

$$\langle \dot{\bar{x}}(t)\bar{x}^T(t - \tau) \rangle = \bar{A}\langle \bar{x}(t)\bar{x}^T(t - \tau) \rangle + \bar{B}\langle \bar{w}(t)\bar{w}^T(t - \tau) \rangle. \quad (\text{D23})$$

Similarly,

$$\langle \bar{x}(t)\dot{\bar{x}}^T(t - \tau) \rangle = \langle \bar{x}(t)\bar{x}^T(t - \tau) \rangle \bar{A}^T + \langle \bar{x}(t)\bar{w}^T(t - \tau) \rangle \bar{B}^T. \quad (\text{D24})$$

The steady-state state covariance matrix of the augmented system is given by $P_S(t, t - \tau) = \langle \bar{x}(t)\bar{x}^T(t - \tau) \rangle$. Then,

$$\begin{aligned} \dot{P}_S(t, t - \tau) &= \langle \dot{\bar{x}}(t)\bar{x}^T(t - \tau) \rangle + \langle \bar{x}(t)\dot{\bar{x}}^T(t - \tau) \rangle \\ &= \bar{A}P_S(t, t - \tau) + P_S(t, t - \tau)\bar{A}^T \\ &\quad + \bar{B}\langle \bar{w}(t)\bar{w}^T(t - \tau) \rangle + \langle \bar{x}(t)\bar{w}^T(t - \tau) \rangle \bar{B}^T. \end{aligned} \quad (\text{D25})$$

Note

$$\begin{aligned} \langle \bar{w}(t)\bar{w}^T(t - \tau) \rangle &= \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \delta(s - t + \tau) \\ &=: \bar{Q}\delta(s - t + \tau). \end{aligned} \quad (\text{D26})$$

Now, let $\bar{\Phi}(t, 0)$ be the state transition matrix for \bar{A} . Then, we have

$$\bar{x}(t) = \bar{\Phi}(t, 0)\bar{x}(0) + \int_0^t \bar{\Phi}(t, s)\bar{B}\bar{w}(s) ds. \quad (\text{D27})$$

Then,

$$\begin{aligned} \langle \bar{x}(t)\bar{w}^T(t - \tau) \rangle &= \int_0^t \bar{\Phi}(t, s)\bar{B}\langle \bar{w}(s)\bar{w}^T(t - \tau) \rangle ds \\ &= \int_0^t \bar{\Phi}(t, s)\bar{B}\bar{Q}\delta(s - t + \tau) ds \\ &= \begin{cases} 0, & \tau > t > 0 \\ \frac{1}{2}\bar{B}\bar{Q}, & \tau \in \{0, t\} \\ \bar{\Phi}(t, t - \tau)\bar{B}\bar{Q}, & t > \tau > 0. \end{cases} \end{aligned} \quad (\text{D28})$$

Similarly,

$$\langle \bar{w}(t)\bar{x}^T(t - \tau) \rangle = \begin{cases} 0, & t > \tau > 0 \\ \frac{1}{2}\bar{Q}\bar{B}^T, & \tau \in \{0, t\} \\ \bar{Q}\bar{B}^T\bar{\Phi}^T(t - \tau, t), & \tau > t > 0. \end{cases} \quad (\text{D29})$$

Thus, we get

$$\dot{P}_S(t, t - \tau) = \begin{cases} \bar{A}P_S(t, t - \tau) + P_S(t, t - \tau)\bar{A}^T + \bar{B}\bar{Q}\bar{B}^T\bar{\Phi}^T(t - \tau, t), & \tau > t > 0 \\ \bar{A}P_S(t, t - \tau) + P_S(t, t - \tau)\bar{A}^T + \bar{B}\bar{Q}\bar{B}^T, & \tau \in \{0, t\} \\ \bar{A}P_S(t, t - \tau) + P_S(t, t - \tau)\bar{A}^T + \bar{\Phi}(t, t - \tau)\bar{B}\bar{Q}\bar{B}^T, & t > \tau > 0, \end{cases} \quad (\text{D30})$$

where

$$\bar{\Phi}(t, t - \tau) = e^{\bar{A}\tau} = \bar{\Phi}^{-1}(t - \tau, t). \quad (\text{D31})$$

As before, by the notation $P_S(t, t - \tau)$, we would imply $t \geq \tau \geq 0$, and so, the first case above is ignored. Then, at steady state, $\dot{P}_S(t, t - \tau) = 0$, and the form of Lyapunov

equation required to be solved to obtain the desired two-time state covariance matrix is

$$\bar{A}P_S(t, t - \tau) + P_S(t, t - \tau)\bar{A}^T + e^{\bar{A}\tau}\bar{B}\bar{Q}\bar{B}^T = 0. \quad (\text{D32})$$

Then, noting that the filter estimation error is $\varepsilon(t) = [1 - 1]\bar{x}(t)$, the two-time error correlation for the filter is obtained as follows:

$$\begin{aligned} P(t, t - \tau) &= \langle \varepsilon(t)\varepsilon^T(t - \tau) \rangle \\ &= [1 \ -1]P_S(t, t - \tau) \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \end{aligned} \quad (\text{D33})$$

APPENDIX E: MATCHING TWO-TIME CORRELATIONS

Here, we strive to arrive at a condition under which the two-time error correlations of a suboptimal filter for the uncertain OU phase [Eq. (4.2)] possibly match up with those of the optimal filter for the uncertain OU phase with added extra OU noise [Eq. (9.13)].

With reference to Appendix D 2, in our case we have

$$\bar{A} = \begin{bmatrix} -\lambda_u & 0 \\ B_e & A_e \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \sqrt{\kappa} & 0 \\ 0 & \frac{B_e}{2|\alpha|} \end{bmatrix}, \quad \bar{Q} = I, \quad (\text{E1})$$

where $A_e = -(\lambda + K)$ and $B_e = K$ for the Kalman filter from Eq. (3.6), and $A_e = -\lambda - 4|\alpha|^2 Q^+ + \epsilon_{\text{opt}} \lambda^2 Q^+$ and $B_e = 4|\alpha|^2 Q^+$ for the robust filter from Eq. (4.7). Then, we have

$$e^{\bar{A}\tau} = \begin{bmatrix} e^{-\lambda_u \tau} & 0 \\ \frac{B_e(e^{A_e \tau} - e^{-\lambda_u \tau})}{A_e + \lambda_u} & e^{A_e \tau} \end{bmatrix}. \quad (\text{E2})$$

Let us denote

$$P_S := \begin{bmatrix} P_{S1} & P_{S2} \\ P_{S3} & P_{S4} \end{bmatrix}. \quad (\text{E3})$$

Then, we get from Eq. (D32) the following:

$$-2\lambda_u P_{S1} + e^{-\lambda_u \tau} \kappa = 0, \quad (\text{E4})$$

$$(-\lambda_u + A_e)P_{S2} + B_e P_{S1} = 0, \quad (\text{E5})$$

$$(-\lambda_u + A_e)P_{S3} + B_e P_{S1} + \frac{B_e \kappa (e^{A_e \tau} - e^{-\lambda_u \tau})}{A_e + \lambda_u} = 0, \quad (\text{E6})$$

$$B_e (P_{S2} + P_{S3}) + 2A_e P_{S4} + e^{A_e \tau} \frac{B_e^2}{4|\alpha|^2} = 0. \quad (\text{E7})$$

Upon solving the above equations, we get

$$P_{S1} = \frac{e^{-\lambda_u \tau} \kappa}{2\lambda_u}, \quad (\text{E8})$$

$$P_{S2} = -\frac{e^{-\lambda_u \tau} \kappa B_e}{2\lambda_u (A_e - \lambda_u)}, \quad (\text{E9})$$

$$P_{S3} = -\frac{B_e \kappa e^{-\lambda_u \tau} (A_e - \lambda_u) + 2B_e \kappa \lambda_u e^{A_e \tau}}{2\lambda_u (A_e^2 - \lambda_u^2)}, \quad (\text{E10})$$

$$P_{S4} = \frac{\kappa B_e^2 e^{-\lambda_u \tau}}{2\lambda_u (A_e^2 - \lambda_u^2)} + \frac{B_e^2 e^{A_e \tau} (4|\alpha|^2 \kappa + \lambda_u^2 - A_e^2)}{8|\alpha|^2 A_e (A_e^2 - \lambda_u^2)}. \quad (\text{E11})$$

Then, from Eq. (D33), we get $P_\tau := P(t, t - \tau)$:

$$\begin{aligned} P_\tau &= P_{S1} - P_{S2} - P_{S3} + P_{S4} \\ &= e^{-\lambda_u \tau} \kappa \frac{[(A_e + B_e)^2 - \lambda_u^2]}{2\lambda_u (A_e^2 - \lambda_u^2)} \\ &\quad - e^{A_e \tau} \frac{B_e^2}{8|\alpha|^2 A_e} + e^{A_e \tau} \frac{B_e \kappa (2A_e + B_e)}{2A_e (A_e^2 - \lambda_u^2)}. \end{aligned} \quad (\text{E12})$$

Next, with reference to Appendix D 1, we rewrite Eq. (D18) with tildes to avoid ambiguity with notation here:

$$\begin{aligned} 0 &= (A - KC)\tilde{P}_\tau + \tilde{P}_\tau(A - KC)^T \\ &\quad + e^{(A-KC)\tau} BQB^T + e^{(A-KC)\tau} KDRD^T K^T. \end{aligned} \quad (\text{E13})$$

Here, the matrices are as defined in Eq. (9.14), $K = PC^T(DRD^T)^{-1}$, where P is as obtained from Eq. (9.15), and $Q = I$ and $R = 1$. Note that $P = \tilde{P}_{\tau=0} = \tilde{P}(t, t)$. For simplicity, let $K := [K_1 \ K_2]^T$. Clearly, from Eq. (9.21), we have

$$\begin{aligned} K_1 &= 4|\alpha|^2 (P_1 + P_2) = \frac{\kappa(-\lambda_u + \beta)}{\kappa + \kappa_n}, \\ K_2 &= 4|\alpha|^2 (P_2 + P_3) = \frac{\kappa_n(-\lambda_u + \beta)}{\kappa + \kappa_n}. \end{aligned} \quad (\text{E14})$$

Then,

$$A - KC = \begin{bmatrix} -\lambda_u - K_1 & -K_1 \\ -K_2 & -\lambda_u - K_2 \end{bmatrix}, \quad (\text{E15})$$

and the matrix exponential is obtained as

$$e^{(A-KC)\tau} = \begin{bmatrix} \frac{e^{-A_n \tau} K_1 + e^{-\lambda_u \tau} K_2}{K_1 + K_2} & \frac{(e^{-A_n \tau} - e^{-\lambda_u \tau}) K_1}{K_1 + K_2} \\ \frac{(e^{-A_n \tau} - e^{-\lambda_u \tau}) K_2}{K_1 + K_2} & \frac{e^{-\lambda_u \tau} K_1 + e^{-A_n \tau} K_2}{K_1 + K_2} \end{bmatrix}, \quad (\text{E16})$$

where $A_n = \lambda_u + K_1 + K_2$.

Let

$$\tilde{P}_\tau := \begin{bmatrix} \tilde{P}_1 & \tilde{P}_2 \\ \tilde{P}_3 & \tilde{P}_4 \end{bmatrix}. \quad (\text{E17})$$

We then wish to find an added noise such that the two-time error correlations described by \tilde{P}_1 from the above equation match those for the suboptimal filter given by P_τ from Eq. (E12). The only parameter available to vary for the noise is κ_n , which is independent of τ .

Expanding Eq. (E13) and solving for \tilde{P}_1 , the desired two-time error correlation for the optimal filter, we get

$$\tilde{P}_1 = e^{-\lambda_u \tau} \frac{\kappa \kappa_n}{2\lambda_u (\kappa + \kappa_n)} + e^{-\beta \tau} \frac{\kappa^2 (-\lambda_u + \beta)}{4|\alpha|^2 (\kappa + \kappa_n)^2}. \quad (\text{E18})$$

Comparing Eqs. (E12) and (E18), we see that we need the following to hold:

$$A_e = -\beta = -\sqrt{4|\alpha|^2 (\kappa + \kappa_n) + \lambda_u^2}, \quad (\text{E19})$$

$$\frac{(A_e + B_e)^2 - \lambda_u^2}{A_e^2 - \lambda_u^2} = \frac{\kappa_n}{\kappa + \kappa_n}, \quad (\text{E20})$$

$$\frac{B_e \kappa (2A_e + B_e)}{2A_e (A_e^2 - \lambda_u^2)} - \frac{B_e^2}{8|\alpha|^2 A_e} = \frac{\kappa^2 (-\lambda_u + \beta)}{4|\alpha|^2 (\kappa + \kappa_n)^2}. \quad (\text{E21})$$

It turns out that it is not possible to choose a suitable value for only κ_n that will allow for all the three independent conditions above [i.e., Eqs. (E19), (E20), and (E21)] to be satisfied simultaneously. Furthermore, if we consider there is another parameter λ_n available to vary

for the added noise, in addition to κ_n , Eq. (E18) will have three (instead of two) time constants, required to be matched with two time constants from Eq. (E12). Thus, it does not seem possible to match up the two-time error correlations.

-
- [1] M. Hofheinz, H. Wang, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, J. M. Martinis, and A. N. Cleland, *Nature (London)* **459**, 546 (2009).
- [2] R. Slavik, F. Parmigiani, J. Kakande, C. Lundstrom, M. Sjodin, P. A. Andrekson, R. Weerasuriya, S. Sygletos, A. D. Ellis, L. Gruner-Nielsen, D. Jakobsen, S. Herstrom, R. Phelan, J. O'Gorman, A. Bogris, D. Syvridis, S. Dasgupta, P. Petropoulos, and D. J. Richardson, *Nat. Photonics* **4**, 690 (2010).
- [3] K. Inoue, E. Waks, and Y. Yamamoto, *Phys. Rev. Lett.* **89**, 037902 (2002).
- [4] J. Chen, J. L. Habif, Z. Dutton, R. Lazarus, and S. Guha, *Nat. Photonics* **6**, 374 (2012).
- [5] V. Giovannetti, S. Lloyd, and L. Maccone, *Nat. Photonics* **5**, 222 (2011).
- [6] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, UK, 2010).
- [7] E. H. Huntington, Ph.D. dissertation, The Australian National University, Canberra, Australia, 1999.
- [8] H. M. Wiseman, *Phys. Rev. Lett.* **75**, 4587 (1995).
- [9] D. W. Berry and H. M. Wiseman, *Phys. Rev. A* **63**, 013813 (2000).
- [10] D. W. Berry and H. M. Wiseman, *Phys. Rev. A* **65**, 043803 (2002).
- [11] M. Tsang, J. H. Shapiro, and S. Lloyd, *Phys. Rev. A* **79**, 053843 (2009).
- [12] D. T. Pope, H. M. Wiseman, and N. K. Langford, *Phys. Rev. A* **70**, 043812 (2004).
- [13] H. M. Wiseman and R. B. Killip, *Phys. Rev. A* **56**, 944 (1997).
- [14] H. M. Wiseman and R. B. Killip, *Phys. Rev. A* **57**, 2169 (1998).
- [15] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, *Phys. Rev. Lett.* **89**, 133602 (2002).
- [16] T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arao, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington, *Phys. Rev. Lett.* **104**, 093601 (2010).
- [17] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington, and A. Furusawa, *Science* **337**, 1514 (2012).
- [18] K. Iwasawa, K. Makino, H. Yonezawa, M. Tsang, A. Davidovic, E. Huntington, and A. Furusawa, *Phys. Rev. Lett.* **111**, 163602 (2013).
- [19] S. Roy, I. R. Petersen, and E. H. Huntington, in *Proceedings of the Australian Control Conference* (IEEE, 2012), pp. 454–458.
- [20] F. L. Lewis, L. Xie, and D. Popa, *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*, 2nd ed. (CRC Press, Boca Raton, FL, 2008).
- [21] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control* (Prentice-Hall, Englewood Cliffs, NJ, 1996).
- [22] S. Roy, I. R. Petersen, and E. H. Huntington, *New J. Phys.* **17**, 063020 (2015).
- [23] S. Roy, I. R. Petersen, and E. H. Huntington, in *Proceedings of the Conference on Lasers and Electro-Optics* (Optical Society of America, 2013), paper JTh2A.88.
- [24] S. Roy, O. U. Rehman, I. R. Petersen, and E. H. Huntington, in *Proceedings of the European Control Conference* (IEEE, 2014), pp. 896–901.
- [25] M. Tsang, *Phys. Rev. Lett.* **102**, 250403 (2009).
- [26] M. Tsang, *Phys. Rev. A* **80**, 033840 (2009).
- [27] K. Goda, O. Miyakawa, E. E. Mikhailov, S. Saraf, R. Adhikari, K. McKenzie, R. Ward, S. Vass, A. J. Weinstein, and N. Mavalvala, *Nat. Phys.* **4**, 472 (2008).
- [28] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [29] I. R. Petersen and D. C. McFarlane, *IEEE Trans. Autom. Control* **39**, 1971 (1994).
- [30] S. O. R. Moheimani, A. V. Savkin, and I. R. Petersen, *IEEE Trans. Circuits Syst. I: Fundam. Theor. Appl.* **45**, 446 (1998).
- [31] S. Roy, I. R. Petersen, and E. H. Huntington, in *Proceedings of the IEEE International Conference on Control Applications* (IEEE, 2013), pp. 551–555.
- [32] D. Hogg, D. W. Berry, and A. I. Lvovsky, *Phys. Rev. A* **90**, 053846 (2014).
- [33] M. J. W. Hall, *J. Mod. Opt.* **40**, 809 (1993).
- [34] M. J. W. Hall, *J. Phys. A: Math. Theor.* **41**, 255301 (2008).
- [35] S. Roy, Ph.D. dissertation, University of New South Wales, Canberra, Australia, 2015.
- [36] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 3rd ed. (Wiley, New York, 1997).
- [37] S. Roy, I. R. Petersen, and E. H. Huntington, in *Proceedings of the American Control Conference* (IEEE, 2013), pp. 3129–3134.