An exact minimum variance filter for a class of discrete time systems with random parameter perturbations

K. Ponomareva [a], P. Date [a]

[a]*The Department of Mathematical Sciences, School of Information Systems, Computing and Mathematics, Brunel University, UK.*

Abstract

An exact, closed-form minimum variance filter is designed for a class of discrete time uncertain systems which allows for both multiplicative and additive noise sources. The multiplicative noise model includes a popular class of models (Cox-Ingersoll-Ross type models) in econometrics. The parameters of the system under consideration which describe the state transition are assumed to be subject to stochastic uncertainties. The problem addressed is the design of a filter that minimizes the trace of the estimation error variance. Sensitivity of the new filter to the size of parameter uncertainty, in terms of the variance of parameter perturbations, is also considered. We refer to the new filter as the 'perturbed Kalman filter' (PKF) since it reduces to the traditional (or unperturbed) Kalman filter as the size of stochastic perturbation approaches zero. We also consider a related approximate filtering heuristic for univariate time series and we refer to filter based on this heuristic as approximate perturbed Kalman filter (APKF). We test the performance of our new filters on three simulated numerical examples and compare the results with unperturbed Kalman filter that ignores the uncertainty in the transition equation. Through numerical examples, PKF and APKF are shown to outperform the traditional (or unperturbed) Kalman filter in terms of the size of the estimation error when stochastic uncertainties are present, even when the size of stochastic uncertainty is inaccurately identified.

Key words: Uncertain discrete-time systems; Robust state estimation; Minimum variance filter; Multiplicative noise.

1 Introduction

The estimation of the state variables given the noisy measurements is one of the fundamental problems in control and signal processing. It is well known that Kalman filter, based on the minimization of the variance of the estimation error, requires

an accurate model with Gaussian noise; see, for example, [1]. However, precise modelling of systems is usually difficult or impossible and system parameters may vary with time or be affected by disturbances. This has motivated studies on robust Kalman filter design that will guarantee an upper bound on the filtering error covariance for any parameter uncertainty. Filtering with guaranteed error for uncertain systems was first considered in [2]. In recent years, several results have been derived on the design of robust estimators that give an upper bound on the error variance for any allowed modelling uncertainty, see [3]-[9] and references therein. The idea of seeking the upper bound on the error variance has been recently applied to discrete-time Markovian jump systems with parameter and noise uncertainty in [10].

A Riccati equation based approach was adopted in [3], [6], [8], [9] and [11] to deal with parameter uncertainty of norm-bounded type. More recently, in [12], this method has been extended to uncertain discrete-time nonlinear systems where nonlinear functions are assumed to be unknown but within a conic region, characterized as a Lipschitz condition on the system state and control signal residuals.

An alternative approach is based on the linear matrix inequality (LMI) method and involves using an interior point algorithm for convex optimization. This interior-point method for convex optimization has also been used to recursively compute the minimal confidence ellipsoid for the state in [14]. The LMI approach has been applied to solve the robust $\mathcal{H}_2$ and $\mathcal{H}_\infty$ filtering for systems with norm-bounded uncertainty, integral quadratic constraints and polytopic uncertainty, see [4], [5], [14], [15] and [16]-[17]. A problem with both deterministic (unknown-but-bounded) and stochastic uncertainties has been considered in [11], [18], [19]-[20] and more recently in [21]. In these papers this stochastic uncertainty is expressed as a multiplicative noise. Unlike the case of the additive noise, the second order statistics of the multiplicative noise is usually unknown, as it depends on the real state of the system.

The approach taken in our paper differs from the contributions mentioned above in several crucial aspects. First, we consider discrete time-varying uncertain systems with both multiplicative and additive noise sources. The system under consideration is assumed to be subject to stochastic uncertainties in the parameters of the transition equation. Our class of systems includes square-root affine systems, *i.e.* Cox-Ingersoll-Ross type models, first introduced in [22], which are frequently used in econometrics and finance literature, and also models that arise in the area of stochastically sampled digital control systems, see [23] and references therein for more details. Secondly, the problem addressed is the design of a filter that analytically minimizes the trace of estimation error covariance matrix. In [23], an optimal linear estimator for linear discrete-time systems with stochastic parameters that are statistically independent in time was derived. In this paper we provide an independent derivation of the results in [23] for a class of perturbed systems. In contrast with the optimization based worst case approaches, we yield an exact, closed form expression for this variance minimizing filter. For univariate time series, we de-

rive some results on the deviation of eigenvalues of the state transition matrix of our new filter from those of the state transition matrix of Kalman estimator for the same system, when the uncertainty is ignored. This analysis of eigenvalues of the perturbed system complements the results on asymptotic stability in [23]. We show that our new perturbed filter is a well-behaved function of the model parameters, in the sense that it converges to the traditional Kalman filter as the variance of stochastic uncertainty tends to zero. We consider approximate filtering heuristic for univariate time series, inspired by the exact method proposed here, which appears to work well for a wider class of nonlinear systems. We also demonstrate through extensive numerical experiments that our new filters perform better than the unperturbed Kalman filter even when the size of uncertainty is poorly identified.

The rest of the paper is organized as follows. In section 2, the minimum variance filtering problem for discrete time-varying systems subject to parameter uncertainty and multiplicative noises is formulated and the equations for the new perturbed Kalman filter are derived for linear as well as certain non-linear cases. Section 3 deals with the sensitivity of the new filter to the parameter perturbation. A new filtering heuristic is proposed in section 4, which is applicable to a wider class of nonlinear systems. Three numerical examples are given in section 5 [1] and some concluding remarks are in section 6. Proofs of the results in sections 3-4 are provided in the Appendix.

2 Derivation of the new perturbed Kalman filter

Consider the following discrete-time state-space system:

$$\boldsymbol{\mathcal{X}}(k+1) = \mathbf{A}\boldsymbol{\mathcal{X}}(k) + \boldsymbol{\Delta A}(k)\boldsymbol{\mathcal{X}}^{\gamma}(k) + \mathbf{B}\mathbf{w}(k+1), \tag{1}$$
$$\boldsymbol{\mathcal{Y}}(k) = \mathbf{C}\boldsymbol{\mathcal{X}}(k) + \mathbf{D}\mathbf{v}(k), \tag{2}$$

where $\gamma \in \{0, 0.5, 1\}$; $\boldsymbol{\mathcal{X}}(k)$ and $\boldsymbol{\mathcal{Y}}(k)$ are the respective state vector and measurement vector at time $t(k)$; $\mathbf{A, B, C}$ and $\mathbf{D}$ are given deterministic matrices; and $\mathbf{v}(k), \mathbf{w}(k)$ are uncorrelated zero mean Gaussian random variables with identity covariance matrix. $\boldsymbol{\mathcal{X}}^{\gamma}(k)$ indicates a vector whose each element is the corresponding element of $\boldsymbol{\mathcal{X}}(k)$ raised to the power γ and a positive value is chosen when $\gamma = 0.5$. We will comment on the choice of γ at the end of this section. The time increment $t(k) - t(k-1)$ is assumed constant for all k. Matrix $\mathbf{A}$ in transition equation is perturbed by a random $n \times n$ matrix $\boldsymbol{\Delta A}(k)$, where the elements $\boldsymbol{\Delta A}_{ij}(k)$ are discrete time processes with zero mean and a constant covariance matrix $\boldsymbol{P_{\Delta A_{ij}}}$. This model structure allows for a large number of uncorrelated noise sources, potentially $O(n^2)$ when $\boldsymbol{\Delta A}$ has independent perturbations, instead of $O(n)$ if the

---

[1] The numerical results in section 5 are also reported in the first author's doctoral thesis [24]

3

noise is purely additive. In practice, $\boldsymbol{\Delta A}(k)$ may be derived from the covariance matrix of parameter estimates if the parameters are obtained from noisy data. Obviously, some of the entries in $\boldsymbol{\Delta A}(k)$ can be identically zero. $\boldsymbol{\Delta A}(k)$ and the noise sources $\mathbf{w}(k)$, $\mathbf{v}(k)$ are assumed to be uncorrelated for all k.

Instead of applying the standard linear estimation theory to our systems of equations and writing the optimal estimator using [1], as in [23], our aim is to derive the equations for a recursive minimum variance filter for state $\boldsymbol{\mathcal{X}}(k)$, given the measurements $\boldsymbol{\mathcal{Y}}(k), \boldsymbol{\mathcal{Y}}(k-1), \cdots, \boldsymbol{\mathcal{Y}}(0)$, which will reduce to Kalman filter if $\boldsymbol{P_{\Delta A}}$ is 0. As in the case of any recursive estimation algorithm, we start by assuming that $\hat{\boldsymbol{\mathcal{X}}}(k|k)$ is known. We then write the predicted mean, i.e. $\hat{\boldsymbol{\mathcal{X}}}(k+1|k)$, and the updated mean, $\hat{\boldsymbol{\mathcal{X}}}(k+1|k+1)$, in the same form as the Kalman filter, and work out the covariance at the time step $t(k+1)$ using equation (1). The standard prediction and update equations for a linear filter and the system described by (1) are

$$\hat{\boldsymbol{\mathcal{X}}}(k+1|k) = \mathbf{A}\hat{\boldsymbol{\mathcal{X}}}(k|k), \tag{3}$$

$$\hat{\boldsymbol{\mathcal{X}}}(k+1|k+1) = \hat{\boldsymbol{\mathcal{X}}}(k+1|k) + \bar{\boldsymbol{K}}(k+1)(\boldsymbol{\mathcal{Y}}(k+1) - \mathbf{C}\hat{\boldsymbol{\mathcal{X}}}(k+1|k)). \tag{4}$$

Our aim is to find a filter gain $\bar{\boldsymbol{K}}(k+1)$ that would minimize the state covariance, which we will denote by $\bar{\boldsymbol{P}}(k+1|k+1)$. Combining the equations (3) and (4) and using the fact that $\boldsymbol{\Delta A}(k)$ and $\boldsymbol{\mathcal{X}}(k)$ are uncorrelated, the expression for the covariance can be easily shown to be

$$\bar{\boldsymbol{P}}(k+1|k+1) = \bar{\boldsymbol{K}}(k+1)\mathbf{DD}'\bar{\boldsymbol{K}}(k+1) + (\mathbf{I} - \bar{\boldsymbol{K}}(k+1)\mathbf{C})\bar{\boldsymbol{P}}(k+1|k)(\mathbf{I} - \bar{\boldsymbol{K}}(k+1)\mathbf{C})', \tag{5}$$

where

$$\bar{\boldsymbol{P}}(k+1|k) = \mathbf{A}\bar{\boldsymbol{P}}(k|k)\mathbf{A}' + \mathbf{BB}' + \widetilde{\boldsymbol{P}}(k|k), \tag{6}$$

and $\widetilde{\boldsymbol{P}}(k|k)$ is given by:

$$\widetilde{\boldsymbol{P}}_{\boldsymbol{ij}}(k|k) = \sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{E}(\boldsymbol{\Delta A}_{ik}\boldsymbol{\Delta A}_{jl})\mathbb{E}(\boldsymbol{\mathcal{X}}_k^\gamma \boldsymbol{\mathcal{X}}_l^\gamma). \tag{7}$$

If the elements of $\boldsymbol{\Delta A}(k)$ are uncorrelated at time step $t(k)$ then $\widetilde{\boldsymbol{P}}(k|k)$ is a diagonal matrix and its expression simplifies considerably to:

$$\widetilde{\boldsymbol{P}}_{\boldsymbol{ii}}(k|k) = \sum_{j=1}^{n} \boldsymbol{P_{\Delta A}}_{ij}(\bar{\boldsymbol{P}}_{jj}(k|k) + (\hat{\boldsymbol{\mathcal{X}}}_j(k|k))^2) \text{ if } \gamma = 1,$$

$$= \sum_{j=1}^{n} \boldsymbol{P_{\Delta A}}_{ij}\hat{\boldsymbol{\mathcal{X}}}_j(k|k) \text{ if } \gamma = 0.5,$$

$$= \sum_{j=1}^{n} \boldsymbol{P_{\Delta A}}_{ij} \text{ if } \gamma = 0.$$

Here $\hat{\boldsymbol{\mathcal{X}}}_j(k|k)$ is the jth element of vector $\hat{\boldsymbol{\mathcal{X}}}(k|k)$. We need to find $\bar{\boldsymbol{K}}(k+1)$ that would minimize the trace of the covariance, i.e. tr $\bar{\boldsymbol{P}}(k+1|k+1)$. Now, the partial

4

derivative of the trace with respect to the filter gain matrix is

$$\frac{\partial \operatorname{tr} \bar{\boldsymbol{P}}(k+1|k+1)}{\partial \bar{\boldsymbol{K}}(k+1)} = -2\bar{\boldsymbol{P}}(k+1|k)\mathbf{C}' + 2\bar{\boldsymbol{K}}(k+1)(\mathbf{C}\bar{\boldsymbol{P}}(k+1|k)\mathbf{C}' + \mathbf{DD}'). \quad (8)$$

Setting this partial derivative to zero leads to the following expression for $\bar{\boldsymbol{K}}(k{+}1)$:

$$\bar{\boldsymbol{K}}(k+1) = \bar{\boldsymbol{P}}(k+1|k)\mathbf{C}'[\mathbf{C}\bar{\boldsymbol{P}}(k+1|k)\mathbf{C}' + \mathbf{DD}']^{-1}. \quad (9)$$

It can be verified that this is indeed a minimum by examining the Hessian of the tr $\bar{\boldsymbol{P}}(k{+}1|k{+}1)$, please see [25] for details of differentiation of a scalar function of a matrix. Given $\hat{\boldsymbol{\mathcal{X}}}(k|k), \bar{\boldsymbol{P}}(k|k)$ are specified, the sequence of operations specified by equations (7), (6), (9), (5), (3) and (4) (in this order) completely define our recursive filter, with equation (7) determining a $\gamma-$dependent term. Comparing equation (6) to that of the unperturbed Kalman filter, we note that there is an extra term, namely $\widetilde{\boldsymbol{P}}(k|k)$, in the expression for $\bar{\boldsymbol{P}}(k{+}1|k)$ and this is due to the random parameter perturbation in equation (1). If $\boldsymbol{\Delta A}(k)$ was zero for all k, the expression for the predicted covariance $\bar{\boldsymbol{P}}(k+1|k)$ would be the same as one in the original Kalman filter equations for an unperturbed state space system (with $\boldsymbol{\Delta A}(k) = 0$ in (1)). The same would happen to the equations for the updated covariance $\bar{\boldsymbol{P}}(k+1|k+1)$ and filter gain $\bar{\boldsymbol{K}}(k+1)$.

Equation (7) indicates why we are using three specific values of γ, since any other values of γ would mean having to compute other moments of $\boldsymbol{\mathcal{X}}(k|k)$ recursively. In section 4, we will look at an approximate filtering heuristic, inspired by the exact method for the three values of γ proposed here, which appears to work well for certain other values of γ. The model structure employed here is still quite flexible and includes traditional linear state space systems ($\gamma = 0$), Cox-Ingersoll-Ross type models used in econometrics [22] ($\gamma = 0.5$) and multiplicative noise models ($\gamma = 1$) that are sometimes useful in digital control systems.

3 Stability and sensitivity of filter to parameter perturbations

Let us rewrite equation (4) in terms of $\hat{\boldsymbol{\mathcal{X}}}(k|k)$ for the unperturbed case:

$$\hat{\boldsymbol{\mathcal{X}}}(k+1|k+1) = (\mathbf{I} - \mathbf{K}(k+1)\mathbf{C})\mathbf{A}\hat{\boldsymbol{\mathcal{X}}}(k|k) + \mathbf{K}(k+1)\boldsymbol{\mathcal{Y}}(k+1). \quad (10)$$

We know that the above system is stable in the case of unperturbed Kalman filter ((10), as mentioned before), i.e. all the eigenvalues of $(\boldsymbol{I}-\mathbf{K}(k{+}1)\boldsymbol{C})\boldsymbol{A}$ are inside the unit circle if the pair $(\mathbf{C}, \mathbf{A})$ is observable, see [26] and references therein. As seen in the previous section, the filter gain of the perturbed Kalman filter differs from the unperturbed Kalman filter by an additive factor which depends on the variance of perturbations. It is of interest to see what happens to the stability of this matrix, which maps $\hat{\boldsymbol{\mathcal{X}}}(k|k)$ to $\hat{\boldsymbol{\mathcal{X}}}(k+1|k+1)$, under parameter perturbations.

3.1 Univariate measurement case

In this section, we consider the eigenvalues of the matrix $(\boldsymbol{I} - \mathbf{K}(k+1)\boldsymbol{C})\boldsymbol{A}$, with $\mathbf{K}(k+1)$ from the traditional (or unperturbed) Kalman filter and compare it with the eigenvalues of $(\boldsymbol{I} - \bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A}$, with $\bar{\boldsymbol{K}}(k+1)$ from the perturbed Kalman filter, introduced above. We consider a case when measurement equation (2) is scalar, i.e. $\mathbf{C}$ is a $1 \times n$ matrix and $\mathbf{D}$ is a scalar, whereas $\mathbf{A}$ is $n \times n$ matrix.

To start with, we write down the equations for Kalman gain and the predicted covariance for the unperturbed Kalman filter for the system in (1) with $\boldsymbol{\Delta A} = 0$:

$$\mathbf{K}(k + 1) = \mathbf{P}(k + 1|k)\mathbf{C}'[\mathbf{C}\mathbf{P}(k + 1|k)\mathbf{C}' + \mathbf{DD}']^{-1},$$
$$\mathbf{P}(k + 1|k) = \mathbf{A}\mathbf{P}(k|k)\boldsymbol{A}' + \mathbf{BB}',$$

and similar expressions for the perturbed Kalman filter from the previous section are repeated for convenience:

$$\bar{\boldsymbol{K}}(k + 1) = \bar{\boldsymbol{P}}(k + 1|k)\mathbf{C}'[\mathbf{C}\bar{\boldsymbol{P}}(k + 1|k)\mathbf{C}' + \mathbf{DD}']^{-1}, \tag{11}$$
$$\bar{\boldsymbol{P}}(k + 1|k) = \mathbf{A}\bar{\boldsymbol{P}}(k|k)\boldsymbol{A}' + \mathbf{BB}' + \widetilde{\bar{\boldsymbol{P}}}(k|k). \tag{12}$$

Suppose that both filters start with the same initial mean and covariance, $\hat{\boldsymbol{\mathcal{X}}}(0|0)$ and $\boldsymbol{P}(0|0)$. The idea is to keep track of the differences between $\bar{\boldsymbol{P}}(k + 1|k)$ and $\mathbf{P}(k+1|k)$, so that at each time step $t(k)$ we can express eigenvalues of the perturbed KF in terms of the ones from the unperturbed KF.
Defining

$$\mathbf{S}(k + 1) = \mathbf{C}\mathbf{P}(k + 1|k)\mathbf{C}' + \mathbf{DD}', \tag{13}$$

in case of the unperturbed Kalman filter we have

$$(\boldsymbol{I} - \mathbf{K}(k + 1)\boldsymbol{C})\boldsymbol{A} = (\boldsymbol{I} - \mathbf{P}(k + 1|k)\mathbf{C}'\mathbf{S}(k + 1)^{-1}\mathbf{C})\boldsymbol{A}. \tag{14}$$

The following result gives an exact expression for $\bar{\boldsymbol{P}}(k|k)$ in terms of $\mathbf{P}(k|k)$, which proves to be useful in establishing the necessary relationship in perturbed and unperturbed eigenvalues:

Proposition 1:

$$\bar{\boldsymbol{P}}(k|k) = \mathbf{P}(k|k) + \Delta\bar{\boldsymbol{P}}(k|k), \tag{15}$$

where the recursion for $\Delta\bar{\boldsymbol{P}}(k|k)$ is as follows:

$$\Delta\bar{\boldsymbol{P}}(0|0) = 0,$$
$$\Delta\bar{\boldsymbol{P}}(k|k) = \Delta\bar{\boldsymbol{P}}(k|k - 1) + \mathbf{P}(k|k - 1)\mathbf{C}'\mathbf{S}(k)^{-1}\mathbf{C}\mathbf{P}(k|k - 1)$$
$$- (\mathbf{P}(k|k - 1) + \Delta\bar{\boldsymbol{P}}(k|k - 1))\phi(k)(\mathbf{P}(k|k - 1) + \Delta\bar{\boldsymbol{P}}(k|k - 1)),$$

and where $\bar{\boldsymbol{P}}(k|k - 1) = \mathbf{P}(k|k - 1) + \Delta\bar{\boldsymbol{P}}(k|k - 1)$,

$$\Delta\bar{\boldsymbol{P}}(k|k - 1) = \widetilde{\bar{\boldsymbol{P}}}(k - 1|k - 1) + \mathbf{A}\Delta\bar{\boldsymbol{P}}(k - 1|k - 1)\boldsymbol{A}'.$$

Finally, $\phi(k) = \mathbf{C}'\mathbf{S}(k)^{-1}[\mathbf{I} + \mathbf{C}\Delta\bar{\boldsymbol{P}}(k|k-1)\mathbf{C}'\mathbf{S}(k)^{-1}]^{-1}$.
Proof: See Appendix. ∎

Let us define

$$\boldsymbol{\alpha}(k+1) = C\Delta\bar{\boldsymbol{P}}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}. \tag{16}$$

Given the dimensions of $\mathbf{C}$, $\mathbf{S}(k+1)$ is a positive scalar and hence $\boldsymbol{\alpha}(k+1)$ is a scalar as well. As the covariance of the perturbation matrix $\boldsymbol{\Delta A}$, i.e. $\boldsymbol{P_{\Delta A}} \to 0$ and $\boldsymbol{\alpha}(k+1) \to 0$, $\Delta\bar{\boldsymbol{P}}(k|k-1) \to 0$ and $\Delta\bar{\boldsymbol{P}}(k|k) \to 0$. Hence $\bar{\boldsymbol{P}}(k+1|k) \to \mathbf{P}(k+1|k)$. Using this relationship, we obtain an expression similar to (14) for the perturbed case for a small enough perturbation, i.e. for $0 < \boldsymbol{\alpha}(k+1) < 1$.

Proposition 2:

$$(\boldsymbol{I} - \bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A} = \left(\mathbf{I} - \frac{(\mathbf{P}(k+1|k) + \Delta\bar{\boldsymbol{P}}(k+1|k))}{(1 + \boldsymbol{\alpha}(k+1))}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right)\mathbf{A}. \tag{17}$$

Proof: See Appendix. ∎

It is of interest to compare the eigenvalues of the perturbed KF and the eigenvalues of the traditional KF. It is also desirable to try to express the size of the eigenvalue perturbation as a function of the size of parameter perturbation. Suppose that matrix $\mathbf{A}$ has eigenvalues $\mu_1, \mu_2, ..., \mu_n$; $\lambda_1(k+1), \lambda_2(k+1), ..., \lambda_n(k+1)$ and $\bar{\lambda}_1(k+1), \bar{\lambda}_2(k+1), ..., \bar{\lambda}_n(k+1)$ are eigenvalues of (14) and (17) respectively. We would look into relationship between $\lambda_i(k+1)$ and μ_i first. We know that $\det(\mathbf{A}) = \prod_{i=1}^{n}\mu_i$ and $\text{tr}(\mathbf{A}) = \sum_{i=1}^{n}\mu_i$. Considering the determinant and the trace of (14) gives us the following:

$$\frac{\prod_{i=1}^{n}\lambda_i(k+1)}{\prod_{i=1}^{n}\mu_i} = \det\left(\boldsymbol{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right), \tag{18}$$

$$= \mathbf{D}\mathbf{D}'\mathbf{S}(k+1)^{-1} \text{ if } n = 2,$$

$$\sum_{i=1}^{n}\lambda_i(k+1) = \text{tr}(\mathbf{A}) - \text{tr}\left(\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\mathbf{A}\right). \tag{19}$$

Considering the determinant and the trace of (17), we obtain our main result for the eigenvalues of the perturbed filter.

Proposition 3:

$$(3a)\frac{\prod_{i=1}^{n}\bar{\lambda}_i(k+1)}{\prod_{i=1}^{n}\mu_i} = \det\left(\mathbf{I} - (\mathbf{P}(k+1|k) + \Delta\bar{\boldsymbol{P}}(k+1|k))\boldsymbol{\psi}(k+1)\right), \qquad (20)$$

$$= \frac{\mathbf{DD}'\mathbf{S}(k+1)^{-1}}{(1+\boldsymbol{\alpha}(k+1))} \text{ if } n=2,$$

$$(3b)\sum_{i=1}^{n}\bar{\lambda}_i(k+1) = \text{tr}\,(\mathbf{A}) - \text{tr}\left((\mathbf{P}(k+1|k) + \Delta\bar{\boldsymbol{P}}(k+1|k))\boldsymbol{\psi}(k+1)\mathbf{A}\right), \quad (21)$$

where $\boldsymbol{\psi}(k+1) = \frac{\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}}{1+\boldsymbol{\alpha}(k+1)}$ and $\boldsymbol{\alpha}(k+1)$ is as defined earlier in this section.
Proof: See Appendix. ∎

We look at the perturbation of eigenvalues due to parameter perturbations next. Note that (20) and (21) are equalities and *not* upper or lower bounds. Let us consider (18) and (20):

$$\frac{\prod_{i=1}^{n}\lambda_i(k+1)}{\prod_{i=1}^{n}\mu_i} = \det\left(\mathbf{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right),$$

$$\frac{\prod_{i=1}^{n}\bar{\lambda}_i(k+1)}{\prod_{i=1}^{n}\mu_i} = \det\left(\mathbf{I} - \frac{\mathbf{P}(k+1|k) + \Delta\bar{\boldsymbol{P}}(k+1|k)}{1+\boldsymbol{\alpha}(k+1)}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right).$$

Defining $\widetilde{\boldsymbol{P}}(k+1|k) = \frac{(\mathbf{P}(k+1|k)+\Delta\bar{\boldsymbol{P}}(k+1|k))}{(1+\boldsymbol{\alpha}(k+1))}$ for simplicity, we note as $\Delta\bar{\boldsymbol{P}}(k+1|k) \to 0$, we have $\boldsymbol{\alpha}(k+1) \to 0$, $\widetilde{\boldsymbol{P}}(k+1|k) \to \mathbf{P}(k+1|k)$ and $\text{tr}\,(\widetilde{\boldsymbol{P}}(k+1|k)) \to \text{tr}\,(\mathbf{P}(k+1|k))$. Hence $\prod_{i=1}^{n}\bar{\lambda}_i(k+1)$ is bounded in the same way as $\prod_{i=1}^{n}\lambda_i(k+1)$.

Now let us consider (19) and (21):

$$\sum_{i=1}^{n}\lambda_i(k+1) = \text{tr}\,(\mathbf{A}) - \text{tr}\left(\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{CA}\right),$$

$$\sum_{i=1}^{n}\bar{\lambda}_i(k+1) = \text{tr}\,(\mathbf{A}) - \text{tr}\left(\frac{(\mathbf{P}(k+1|k)+\Delta\bar{\boldsymbol{P}}(k+1|k))}{(1+\boldsymbol{\alpha}(k+1))}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{CA}\right).$$

Observe also that covariances $\mathbf{P}(k+1|k)$ and $\widetilde{\boldsymbol{P}}(k+1|k)$ are both symmetric and positive semidefinite. Hence

$$\gamma_{min}\,\text{tr}\,(\mathbf{P}(k+1|k)) \leq \text{tr}\,(\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{CA}) \leq \gamma_{max}\,\text{tr}\,(\mathbf{P}(k+1|k)),$$

$$\gamma_{min}\,\text{tr}\,(\widetilde{\boldsymbol{P}}(k+1|k)) \leq \text{tr}\,(\widetilde{\boldsymbol{P}}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{CA}) \leq \gamma_{max}\,\text{tr}\,(\widetilde{\boldsymbol{P}}(k+1|k)),$$

where γ_{min} and γ_{max} are respectively the smallest and the largest eigenvalues of the matrix $\frac{1}{2}(\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{CA} + (\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{CA})')$, see [27] for a similar example of bounds on matrix eigenvalues. Thus $\sum_{i=1}^{n}\bar{\lambda}_i(k+1)$ is bounded in the same way as $\sum_{i=1}^{n}\lambda_i(k+1)$. For the case $n=2$, $\bar{\lambda}_i$ are fully defined by their sum and product because of the characteristic equation $\bar{\lambda}^2 - (\bar{\lambda}_1 + \bar{\lambda}_2)\bar{\lambda} + \bar{\lambda}_1\bar{\lambda}_2 = 0$. Therefore, we conclude that for a small enough perturbation, $\bar{\lambda}_i$ won't differ by too much from

λ_i. Hence it is clear that the perturbed Kalman filter will have a transient behaviour very similar to an unperturbed Kalman filter provided the random perturbations are small. Even though some of the equality results are only for scalar systems with two (or less) eigenvalues, it gives a useful qualitative insight into the impact of perturbations on the transient behaviour of the filter.

3.2 *Multivariate measurement case*

Note that *Proposition* 2 and *Proposition* 3 can be both extended to cover the case when $\mathbf{C}$ is a $n \times n$ matrix and $\mathbf{D}$ is a $n \times 1$ vector. In this case $\boldsymbol{\alpha}(k+1)$ given in (16) is a matrix. Provided that $|||\boldsymbol{\alpha}(k+1)||| < 1$, where $||| * |||$ is a matrix norm, we have from [28]:

$$(\boldsymbol{I} + \boldsymbol{\alpha}(k+1))^{-1} = \sum_{l=0}^{\infty}(-\boldsymbol{\alpha}(k+1))^l.$$

In this case *Proposition* 2 becomes:
Proposition 4:

$$(\boldsymbol{I} - \bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A} = \left(\boldsymbol{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C} - \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C}\right)\boldsymbol{A}, \tag{22}$$

where

$$\Delta\bar{\boldsymbol{K}}(k+1) = \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\sum_{l=1}^{\infty}(-\boldsymbol{\alpha}(k+1))^l + \Delta\bar{\boldsymbol{P}}(k+1|k)\boldsymbol{\phi}(k+1)$$

and $\boldsymbol{\phi}(k+1)$ is as in *Proposition* 1.
Proof: See Appendix. ∎
Taking the trace and determinant of (22) will provide us with the sum and product of eigenvalues $\bar{\lambda}_i$ for the general case.
Proposition 5:

$$(5a)\prod_{i=1}^{n}\bar{\lambda}_i(k+1) = \det\left(\mathbf{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C} + \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C}\right)\prod_{i=1}^{n}\mu_i, \tag{23}$$

$$(5b)\sum_{i=1}^{n}\bar{\lambda}_i(k+1) = \operatorname{tr}(\mathbf{A}) - \operatorname{tr}\left((\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C} + \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\mathbf{A}\right). \tag{24}$$

Proof: See Appendix. ∎
We note as $\Delta\bar{\boldsymbol{P}}(k+1|k) \to 0$, we have $\boldsymbol{\alpha}(k+1) \to 0$ and hence $\Delta\bar{\boldsymbol{K}}(k+1) \to 0$. Comparing (23) with (18) and (24) with (19), we observe that $\prod_{i=1}^{n}\bar{\lambda}_i(k+1)$ is bounded in the same way as $\prod_{i=1}^{n}\lambda_i(k+1)$ and $\sum_{i=1}^{n}\bar{\lambda}_i(k+1)$ is bounded in

the same way as $\sum_{i=1}^{n} \lambda_i(k+1)$. Therefore, we conclude that for a small enough perturbation, $\bar{\lambda}_i$ won't differ by too much from λ_i in general case as well.

4 Approximate filtering for $\gamma \geq 1.5$

In this section we would consider a univariate state-space system as in (1) with $\gamma = \frac{l}{2}$ with $l = 3, 4, \ldots$. These values of γ come up in constant elasticity of variance (CEV) type models, see [29] for more details. In this particular case, $\widetilde{\bar{P}}(k-1|k-1) = P_{\Delta A}\mathbb{E}(\mathcal{X}^{2\gamma}(k-1))$, where positive values of $\mathcal{X}$ are chosen whenever there are two roots. Hence in order to find the covariance $\bar{P}(k|k-1)$, optimal gain $\bar{K}(k)$ and the updated state $\hat{\mathcal{X}}(k|k)$ at time step $t(k)$, we would need to know $\mathbb{E}(\mathcal{X}^{2\gamma}(k-1))$ from $t(k-1)$. In general there is no closed form solution for this for $l \geq 3$ as we are only propogating the first two moments of $\mathcal{X}$ throughout the filter recursions. However we can carry out moment matching approximation of $\Delta A(k-1)\mathcal{X}^{\gamma}(k-1)$ by rewriting it as a noise term with the same first two moments as $\mathbf{F}(k-1)\mathbf{u}(k-1)$. Here $\mathbf{u}(k-1)$ is a Gaussian random variable, uncorrelated with the state and noise terms $\mathbf{w}(k)$ and $\mathbf{v}(k)$, and $\mathbf{F}(k-1)^2 = P_{\Delta A}\mathbb{E}\left(\mathcal{X}^{2\gamma}(k-1)\right)$. This way we would match the first two moments of term $\Delta A(k-1)\mathcal{X}^{\gamma}(k-1)$ exactly, and higher order moments would be proportional to those of $\mathbf{u}(k-1)$, in particular, the odd moments would be zero.

Using (1) we are interested in rewriting $\mathcal{X}(k) - \hat{\mathcal{X}}(k|k)$ as an expression depending on the state, Δ_s, and an expression consisting solely of noise terms, Δ_n :

Proposition 6

$$\mathcal{X}(k) - \hat{\mathcal{X}}(k|k) = \Delta_s + \Delta_n,$$

where $\Delta_s = \left(I - \bar{K}(k)C\right)A\left(\mathcal{X}(k-1) - \hat{\mathcal{X}}(k-1|k-1)\right)$ and $\Delta_n = \left(I - \bar{K}(k)C\right)(\mathbf{B}\mathbf{w}(k) + \mathbf{F}(k-1)\mathbf{u}(k-1)) - \bar{K}(k)\mathbf{D}\mathbf{v}(k)$.

Proof: See Appendix. ∎

Since $\mathbf{u}(k), \mathbf{w}(k)$ and $\mathbf{v}(k)$ are zero mean uncorrelated Gaussian random variables, we have $\mathbb{E}(\mathbf{w}^j(k)) = \mathbb{E}(\mathbf{v}^j(k)) = \mathbb{E}(\mathbf{u}^j(k)) = 0$ for any odd integer j. Thus $\mathbb{E}(\Delta_n)^j = 0$. In the case of $l = 3$ and $\gamma = 1.5$, we can raise equation (29) to the power 3, and take expectations of both sides. This would allow us to find the expression for $\mathbb{E}(\mathcal{X}(k))^3$:

Proposition 7

$$\mathbb{E}(\mathcal{X}^3(k)) = 3\mathbf{P}(k|k)\hat{\mathcal{X}}(k|k) + \hat{\mathcal{X}}^3(k|k) + (I - \bar{K}(k)C)^3\mathbf{A}^3*$$
$$* \left(\mathbb{E}(\mathcal{X}^3(k-1)) - 3\mathbf{P}(k-1|k-1)\hat{\mathcal{X}}(k-1|k-1) - \hat{\mathcal{X}}^3(k-1|k-1)\right).$$

10

Proof: See Appendix. ∎

In general we have

$$\mathbb{E}\left(\boldsymbol{\mathcal{X}}(k) - \hat{\boldsymbol{\mathcal{X}}}(k|k)\right)^q = \mathbb{E}(\boldsymbol{\Delta_s} + \boldsymbol{\Delta_n})^q, \tag{25}$$

for some integer power q. Provided at time step $t(k)$ we know $\mathbb{E}(\boldsymbol{\mathcal{X}}(k-1))^i$ for $i = 1, 2, ..., 2\gamma$, we can recursively find $\mathbb{E}(\boldsymbol{\mathcal{X}}(k))^j$ for $j = 3, ..., 2\gamma$ using the following proposition.

Proposition 8

$$\sum_{i=0}^{q} \binom{q}{i} \mathbb{E}(\boldsymbol{\mathcal{X}}(k))^i (\hat{\boldsymbol{\mathcal{X}}}(k|k))^{q-i} = \sum_{j=0,2,...,q} \binom{q}{j} \mathbb{E}(\boldsymbol{\Delta_n})^j \mathbb{E}(\boldsymbol{\Delta_s})^{q-j} \text{ if q is even,}$$

$$= \sum_{j=0,2,...,q-3} \binom{q}{j} \mathbb{E}(\boldsymbol{\Delta_n})^j \mathbb{E}(\boldsymbol{\Delta_s})^{q-j} \text{ if q is odd,}$$

$$\tag{26}$$

for $q = 3, ..., 2\gamma$.

Proof: See Appendix. ∎

Note that minimum variance derivation of the equations for $\bar{\boldsymbol{K}}(k)$ in (9) still holds for these values of γ as well. We would call this filter as approximate perturbed Kalman filter (APKF) as it is a linear filter providing a solution by approximating the perturbation term by a noise term. We look at the performance of the new filters in simulation examples in the next section.

5 Numerical Examples

We consider three different cases for three different values of γ to illustrate and contrast the performance of the proposed filter with that of the Kalman filter.

5.1 Case when $\gamma = 1$

In this section we consider a numerical example with the following state-space equations.

$$\boldsymbol{\mathcal{X}}(k+1) = \mathbf{A}\boldsymbol{\mathcal{X}}(k) + \boldsymbol{\Delta A}(k)\boldsymbol{\mathcal{X}}^{\gamma}(k) + \mathbf{B}\mathbf{w}(k+1),$$
$$\boldsymbol{\mathcal{Y}}(k) = \mathbf{C}\boldsymbol{\mathcal{X}}(k) + \mathbf{D}\mathbf{v}(k),$$

where $\gamma = 1$,

$$\mathbf{A} = \begin{pmatrix} 0 & -0.5 \\ 1 & 1 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} -6 \\ 1 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} -100 & 10 \end{pmatrix}$$

and $\mathbf{D} = 1$. $\mathbf{v}(k), \mathbf{w}(k)$ are uncorrelated Gaussian random variables. Perturbation matrix $\mathbf{\Delta A}(k)$ has zero mean, the matrix elements have the following covariance matrix:

$$\boldsymbol{P_{\Delta A}} = \begin{pmatrix} 0.12 & 0.02 \\ 0.15 & 0.1 \end{pmatrix}.$$

Initial conditions are $\boldsymbol{\mathcal{X}}(0) = \begin{pmatrix} 1 & 0 \end{pmatrix}', \hat{\boldsymbol{\mathcal{X}}}(0|0) = \begin{pmatrix} 0 & 0 \end{pmatrix}'$ and

$$\mathbf{P}(0|0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In order to compare the performance of both the perturbed (PFK) and the unperturbed Kalman filters (KF), we consider the average root mean square error(avRMSE) for the state, where RMSE for a sample path l is given by

$$RMSE_l = \sqrt{\frac{1}{2F} \sum_{j=1}^{F} ((\boldsymbol{\mathcal{X}}_1(j) - \widehat{\boldsymbol{\mathcal{X}}}_1(j))^2 + (\boldsymbol{\mathcal{X}}_2(j) - \widehat{\boldsymbol{\mathcal{X}}}_2(j))^2)},$$

and AvRMSE over a given number of sample paths L is defined by $\frac{1}{L} \sum_{l=1}^{L} RMSE_l$. We use $F = 100$ and $L = 100$. Comparison of the variance of RMSE (VAR) for both filters is also shown in tables 1-3 and is calculated as $\frac{1}{L} \sum_{l=1}^{L} (RMSE_l - AvRMSE)^2$. VAR provides a measure of variation in the filter performance across different sample paths. We consider three cases: when the real covariance of perturbation matrix $\mathbf{\Delta A}(k)$ is known exactly and when two different cases where it is assumed to be incorrectly estimated. If the model is calibrated from data, it is likely that the parameters are imperfectly known and even the size of uncertainty (in terms of its covariance matrix) is not known exactly. It is of interest to see whether our filter performs well if this is indeed the case. We use

$$\widetilde{\boldsymbol{P}_{\Delta A1}} = \begin{pmatrix} 0.2 & 0.1 \\ 0.05 & 0.15 \end{pmatrix} \text{ and } \widetilde{\boldsymbol{P}_{\Delta A2}} = \begin{pmatrix} 0.25 & 0.15 \\ 0.05 & 0.2 \end{pmatrix}, \text{ but keep } \boldsymbol{P_{\Delta A}} \text{ in the equa-}$$

tions of PKF (*i.e.* $\boldsymbol{P_{\Delta A}}$ is the *incorrect* covariance matrix used when the correct covariances in the data generating system are $\widetilde{\boldsymbol{P}_{\Delta A1}}, \widetilde{\boldsymbol{P}_{\Delta A2}}$). The results with incorrectly specified uncertainty as above are presented in tables 2 and 3 respectively, while table 1 presents the results when $\boldsymbol{P_{\Delta A}}$ is exactly known. A comparison of a simulated path with the paths generated by PKF and KF for $\boldsymbol{\mathcal{X}}_1$, with $\widetilde{\boldsymbol{P}_{\Delta A1}}$ as the real covariance of the perturbation matrix $\mathbf{\Delta A}$, is shown in Fig.1.

12

Table 1. Comparison of AvRMSE when $P_{\Delta A}$ is estimated correctly

	KF	PKF	Improvement
AvRMSE	14.750	5.927	59.8%
VAR	54.061	2.285	95.8%

Table 2. Comparison of AvRMSE when $P_{\Delta A 1}$ is not estimated correctly

	KF	PKF	Improvement
AvRMSE	18.775	8.102	56.9%
VAR	472.414	22.214	95.3%

Table 3. Comparison of AvRMSE when $P_{\Delta A 2}$ is not estimated correctly

	KF	PKF	Improvement
AvRMSE	37.557	16.088	57.1%
VAR	1465.857	158.557	89.2%

Fig. 1. Simulated paths
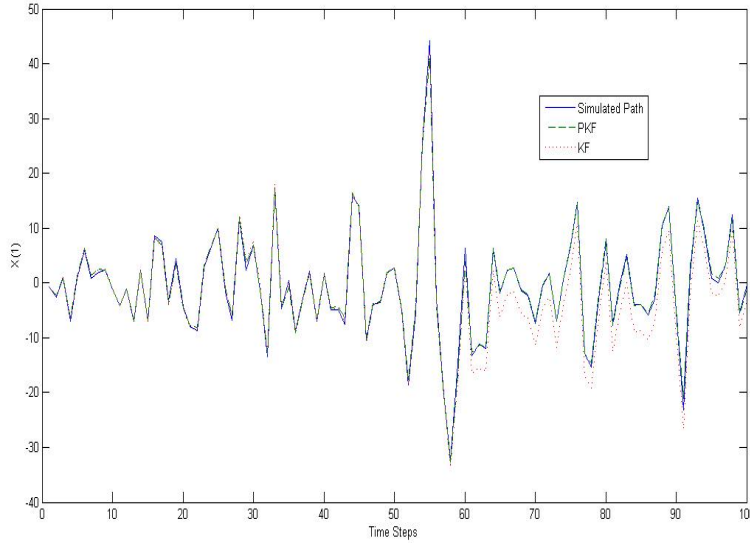
We can see from all the three tables that PKF provides more accurate state estimates in terms of both the measures of error when compared to unperturbed Kalman filter, in both the cases when $P_{\Delta A}$ is estimated correctly and when it is not. Note that acknowledging that the model is not precise and accounting for the random parameter uncertainties, makes the filter more robust even to poor estimates of the parameter uncertainty (in terms of variance of ΔA). This observation is in line with the intuition that unperturbed filter is highly *tuned* to the system parameters and the addition of $\widetilde{\boldsymbol{P}}(k|k)$ provides a *de-tuning* effect, thereby making the filter

13

more robust to uncertainty.

5.2 Case when $\gamma = 0.5$

As another example, we consider a nonlinear system

$$
\begin{aligned}
\boldsymbol{\mathcal{X}}(k+1) &= \mathbf{A}\boldsymbol{\mathcal{X}}(k) + \boldsymbol{\Delta A}(k)\boldsymbol{\mathcal{X}}^{\gamma}(k) + \mathbf{B}\mathbf{w}(k+1), \\
\boldsymbol{\mathcal{Y}}(k) &= \mathbf{C}\boldsymbol{\mathcal{X}}(k) + \mathbf{D}\mathbf{v}(k),
\end{aligned}
$$

with $\gamma = 0.5$, $\mathbf{A} = 0.9$, $\mathbf{B} = 0.1$, $\mathbf{C} = 1$ and $\mathbf{D} = 0.01$. $\mathbf{v}(k), \mathbf{w}(k)$ are uncorrelated Gaussian random variables. Initial conditions are assumed to be $\boldsymbol{\mathcal{X}}(0) = 0.1$, $\hat{\boldsymbol{\mathcal{X}}}(0|0) = 0$ and $\mathbf{P}(0|0) = 1$.
We will compare the performance of the new perturbed Kalman filter(PKF) with unperturbed Kalman filter(KF). Average root mean square error is calculated for both filters as follows:

$$
RMSE = \sqrt{\frac{1}{F}\sum_{j=1}^{F}((\boldsymbol{\mathcal{X}}(j) - \widehat{\boldsymbol{\mathcal{X}}}(j))^2)},
$$

$F = 100$ and we calculate AvRMSE over 100 paths, as before. As in the previous example, these errors are also calculated for the case when covariance $\boldsymbol{P_{\Delta A}}$ is not estimated correctly. Results are summarised in tables 4-5.

Table 4. Comparison of AvRMSE when $\boldsymbol{P_{\Delta A}}$ is estimated correctly

$\boldsymbol{P_{\Delta A}}$	KF	PKF	Improvement
0.2	0.009418	0.009131	3.1%
0.3	0.009422	0.008954	5.0%
0.4	0.009661	0.008921	10.1%

Table 5. Comparison of AvRMSE when $\boldsymbol{P_{\Delta A}}$ is not estimated correctly

True $\boldsymbol{P_{\Delta A}}$	Assumed $\boldsymbol{P_{\Delta A}}$	KF	PKF	Improvement
0.4	0.2	0.009965	0.009219	7.5%
0.3	0.2	0.009491	0.009052	4.6%
0.2	0.3	0.009355	0.009078	3.0%
0.2	0.4	0.009418	0.009129	2.8%

We can see from these tables that perturbed Kalman filter provides better accuracy when compared to unperturbed Kalman filter for different values of the perturbation

matrix variance. The improvement is more pronounced when the variance of the uncertainty is larger, as can be expected.

5.3 Case when $\gamma = 1.5$

As an example for our APKF, we consider a nonlinear system

$$
\begin{aligned}
\mathcal{X}(k+1) &= \mathbf{A}\mathcal{X}(k) + \mathbf{\Delta A}(k)\mathcal{X}^{\gamma}(k) + \mathbf{B}\mathbf{w}(k+1), \\
\mathcal{Y}(k) &= \mathbf{C}\mathcal{X}(k) + \mathbf{D}\mathbf{v}(k),
\end{aligned}
$$

with $\gamma = 1.5$, $\mathbf{A} = 0.9$, $\mathbf{B} = 0.1$, $\mathbf{C} = 1$ and $\mathbf{D} = 0.01$. $\mathbf{v}(k), \mathbf{w}(k)$ are uncorrelated Gaussian random variables. Initial conditions are assumed to be $\mathcal{X}(0) = 0.1$, $\hat{\mathcal{X}}(0|0) = 0$ and $\mathbf{P}(0|0) = 1$, as per example in the previous subsection.
We will compare the performance of the new approximate perturbed Kalman filter (APKF) with unperturbed Kalman filter (KF). Average root mean square error is calculated for both filters as in the previous subsection. As in the previous example, these errors are also calculated for the case when covariance $\boldsymbol{P_{\Delta A}}$ is not estimated correctly. Results are summarized in tables 6-7.
We can see from these tables that approximate perturbed Kalman filter provides better accuracy when compared to unperturbed Kalman filter for different values of the perturbation matrix variance. The improvement is more pronounced when the variance of the uncertainty is larger, as can be expected.

Table 6. Comparison of AvRMSE when $\boldsymbol{P_{\Delta A}}$ is estimated correctly

$\boldsymbol{P_{\Delta A}}$	KF	APKF	Improvement
0.2	0.009542	0.009204	3.5%
0.3	0.010312	0.009219	10.6%
0.4	0.012105	0.009235	23.7%

Table 7. Comparison of AvRMSE when $\boldsymbol{P_{\Delta A}}$ is not estimated correctly

True $\boldsymbol{P_{\Delta A}}$	Assumed $\boldsymbol{P_{\Delta A}}$	KF	APKF	Improvement
0.4	0.2	0.012087	0.009232	23.6%
0.3	0.2	0.011253	0.009329	17.1%
0.2	0.3	0.009527	0.009214	3.2%
0.2	0.4	0.009583	0.009297	3.0%

6 Summary

The contributions of this paper are as follows.

Firstly, this paper provides an independent derivation of an exact recursive linear minimum variance filter for a class of uncertain discrete time systems. The class of systems under consideration include linear models with state multiplicative uncertainty as well as square root affine models as special cases.

Secondly, we have analyzed the sensitivity of the new filter to the size of the parameter perturbation. We have also provided results for the product and sum of the eigenvalues of the new perturbed filter in terms of the perturbation parameter. We have deduced that the perturbed Kalman filter will have a transient behaviour very similar to an unperturbed Kalman filter, provided the random perturbations are small. Our analysis of the eigenvalues of the new perturbed filter complements the asymptotic results for similar systems in the literature.

Thirdly, in this paper we have also provided a new filtering heuristic for a wider class of nonlinear systems for the special univariate case of $\gamma = \frac{l}{2}$, for any positive integer $l \geq 3$.

Finally, we have demonstrated the utility of the new methods through comprehensive numerical examples. These examples illustrate the improved accuracy achieved by the new filters when compared to traditional (or unperturbed) Kalman filter. Critically, the examples indicate that the perturbed filters perform better than the unperturbed Kalman filter even when the size of uncertainty is poorly identified. This has important implications in cases where the model is calibrated from data.

Current research is focussed on linking the results in this paper with likelihood-based calibration of latent state models and extending the results to the parameter uncertainty in the measurement equation.

Appendix

Proof of Proposition 1

We will prove this using a simple mathematical induction argument. Suppose that both filters have the same initial covariance $\mathbf{P}(0|0) = \bar{\boldsymbol{P}}(0|0)$. Also let $\Delta\bar{\boldsymbol{P}}(0|0) = 0$, so that the proposition holds for $k = 0$. Let's assume that (15) holds for some $k = m$ with $m \geq 0$, i.e. we have

$$\bar{\boldsymbol{P}}(m|m) = \mathbf{P}(m|m) + \Delta\bar{\boldsymbol{P}}(m|m),$$

where $\Delta\bar{\mathbf{P}}(m|m) = \Delta\bar{\mathbf{P}}(m|m-1) + \mathbf{P}(m|m-1)\mathbf{C}'\mathbf{S}(m)^{-1}\mathbf{C}\mathbf{P}(m|m-1) - (\mathbf{P}(m|m-1)+\Delta\bar{\mathbf{P}}(m|m-1))\boldsymbol{\phi}(m)(\mathbf{P}(m|m-1)+\Delta\bar{\mathbf{P}}(m|m-1))$. Considering expression for $\bar{\mathbf{P}}(m+1|m)$ first

$$\begin{aligned}
\bar{\mathbf{P}}(m+1|m) &= \mathbf{A}\bar{\mathbf{P}}(m|m)\mathbf{A}' + \mathbf{BB}' + \widetilde{\bar{\mathbf{P}}}(m|m) \\
&= \mathbf{A}\mathbf{P}(m|m)\mathbf{A}' + \mathbf{BB}' + \mathbf{A}\Delta\bar{\mathbf{P}}(m|m)\mathbf{A}' + \widetilde{\bar{\mathbf{P}}}(m|m) \\
&= \mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m).
\end{aligned}$$

Then expression for $\bar{\mathbf{P}}(m+1|m+1)$ becomes

$$\bar{\mathbf{P}}(m+1|m+1) = \bar{\mathbf{P}}(m+1|m) - \bar{\mathbf{P}}(m+1|m)\mathbf{C}'(\mathbf{C}\bar{\mathbf{P}}(m+1|m)\mathbf{C}' + \mathbf{DD}')^{-1}\mathbf{C}\bar{\mathbf{P}}(m+1|m)$$
$$= \mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)$$
$$- \left(\mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)\right)\boldsymbol{\phi}(m+1)\mathbf{C}\left(\mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)\right).$$

Rearranging the terms as before, we obtain

$$\begin{aligned}
\bar{\mathbf{P}}(m+1|m+1) &= \mathbf{P}(m+1|m) - \mathbf{P}(m+1|m)\mathbf{C}'\mathbf{S}(m+1)^{-1}\mathbf{C}\mathbf{P}(m+1|m) \\
&+ \Delta\bar{\mathbf{P}}(m+1|m) + \mathbf{P}(m+1|m)\mathbf{C}'\mathbf{S}(m+1)^{-1}\mathbf{C}\mathbf{P}(m+1|m) \\
&- \left(\mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)\right)\boldsymbol{\phi}(m+1)\left(\mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)\right) \\
&= \mathbf{P}(m+1|m+1) + \Delta\bar{\mathbf{P}}(m+1|m+1),
\end{aligned}$$

where $\Delta\bar{\mathbf{P}}(m+1|m+1) = \Delta\bar{\mathbf{P}}(m+1|m) + \mathbf{P}(m+1|m)\mathbf{C}'\mathbf{S}(m+1)^{-1}\mathbf{C}\mathbf{P}(m+1|m) - \left(\mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)\right)\boldsymbol{\phi}(m+1)\left(\mathbf{P}(m+1|m) + \Delta\bar{\mathbf{P}}(m+1|m)\right)$ as required. Hence by induction we have proven *Proposition* 1 holds for all $k \geq 0$.

Proof of Proposition 2

In section 3 we defined the following relationship between $\bar{\mathbf{P}}(k+1|k)$ and $\mathbf{P}(k+1|k)$:

$$\bar{\mathbf{P}}(k+1|k) = \mathbf{P}(k+1|k) + \Delta\bar{\mathbf{P}}(k+1|k).$$

Here $\Delta\bar{\mathbf{P}}(k+1|k) = \widetilde{\bar{\mathbf{P}}}(k|k) + \mathbf{A}\Delta\bar{\mathbf{P}}(k|k)\mathbf{A}'$ and the recursion for $\bar{\mathbf{P}}(k|k)$ is as in *Proposition* 1. Rewriting equation (11) and using definition of $\mathbf{S}(k+1)$ from (13),

$$\begin{aligned}
\bar{\mathbf{K}}(k+1) &= (\mathbf{P}(k+1|k) + \Delta\bar{\mathbf{P}}(k+1|k))\mathbf{C}'(\mathbf{S}(k+1) + \mathbf{C}\Delta\bar{\mathbf{P}}(k+1|k)\mathbf{C}')^{-1} \\
&= (\mathbf{P}(k+1|k) + \Delta\bar{\mathbf{P}}(k+1|k))\boldsymbol{\phi}(k+1) \\
&= \frac{(\mathbf{P}(k+1|k) + \Delta\bar{\mathbf{P}}(k+1|k))\mathbf{C}'\mathbf{S}(k+1)^{-1}}{1 + \boldsymbol{\alpha}(k+1)}.
\end{aligned}$$

For small enough perturbation, i.e. $0 < \alpha(k+1) < 1$, we can expand $(1 + \alpha(k+1))^{-1}$ as a power series and then write:

$$\bar{K}(k+1) = \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\left(1 - \frac{\alpha(k+1)}{1+\alpha(k+1)}\right) + \Delta\bar{P}(k+1|k)\phi(k+1)$$

$$= \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1} + \Delta\bar{P}(k+1|k)\phi(k+1) - \alpha(k+1)\mathbf{P}(k+1|k)\phi(k+1)$$
$$\tag{27}$$

$$= \mathbf{K}(k+1) + \Delta\bar{K}(k+1),$$

where the perturbation in the Kalman gain, to account for the random perturbation in the model parameters, is given by

$$\Delta\bar{K}(k+1) = \frac{\Delta\bar{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}}{1+\alpha(k+1)} - \alpha(k+1)\frac{\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}}{1+\alpha(k+1)}.$$

Rearranging (27) we get the desired result:

$$(\boldsymbol{I} - \bar{K}(k+1)C)\boldsymbol{A} = (\boldsymbol{I} - (\mathbf{K}(k+1) + \Delta\bar{K}(k+1))C)\boldsymbol{A}$$
$$= \left(\boldsymbol{I} - \frac{(\mathbf{P}(k+1|k) + \Delta\bar{P}(k+1|k))}{(1+\alpha(k+1))}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right)\boldsymbol{A}.$$

Proof of Proposition 3

To prove equation (3a), we take determinant of (17)

$$\det{(\boldsymbol{I} - \bar{K}(k+1)C)\boldsymbol{A}} = \det{(A)}\det{\left(\boldsymbol{I} - \frac{(\mathbf{P}(k+1|k) + \Delta\bar{P}(k+1|k))}{(1+\alpha(k+1))}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right)}.$$

When $n = 2$, we can use the fact that for any matrix $\mathbf{B}$, $\det(\boldsymbol{I} - \mathbf{B}) = 1 - \text{tr}(\mathbf{B}) + \det(\mathbf{B})$. Remember that $\alpha(k+1) = \mathbf{C}\Delta\bar{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}$. Also note that $\det{(\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C})} = 0$ and from definition of $\mathbf{S}(k+1)$ in (13) we have $\text{tr}{(\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C})} = 1 - \mathbf{D}\mathbf{D}'\mathbf{S}(k+1)^{-1}$. This leads to

$$\det{(\boldsymbol{I} - \bar{K}(k+1)C)\boldsymbol{A}} = \det{(A)}\left(1 - \text{tr}{\left(\frac{(\mathbf{P}(k+1|k) + \Delta\bar{P}(k+1|k))}{(1+\alpha(k+1))}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right)}\right),$$

$$= \det{(\mathbf{A})}\left(1 - \frac{\text{tr}{(\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C})}}{1+\alpha(k+1)}\right) - \det{(\mathbf{A})}\left(\frac{\text{tr}{(\mathbf{C}\Delta\bar{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1})}}{1+\alpha(k+1)}\right),$$

$$= \det{(\mathbf{A})}\left(1 - \frac{1 - \mathbf{D}\mathbf{D}'\mathbf{S}(k+1)^{-1}}{1+\alpha(k+1)} - \frac{\alpha(k+1)}{1+\alpha(k+1)}\right),$$

$$= \det{(\mathbf{A})}\frac{\mathbf{D}\mathbf{D}'\mathbf{S}(k+1)^{-1}}{(1+\alpha(k+1))}.$$

To prove equation (3b), we take trace of (17)

$$\text{tr}{\left((\boldsymbol{I} - \bar{K}(k+1)C)\boldsymbol{A}\right)} = \text{tr}(\mathbf{A}) - \text{tr}{\left(\mathbf{I} - \frac{(\mathbf{P}(k+1|k) + \Delta\bar{P}(k+1|k))}{(1+\alpha(k+1))}\mathbf{C}'\mathbf{S}(k+1)^{-1}\mathbf{C}\right)},$$

which leads to the required result.

Proof of Proposition 4

Using power series for $\boldsymbol{\alpha}(k+1)$ we can rewrite $\bar{\boldsymbol{K}}(k+1)$ as:

$$\bar{\boldsymbol{K}}(k+1) = (\mathbf{P}(k+1|k) + \Delta\bar{\boldsymbol{P}}(k+1|k))\boldsymbol{\phi}(k+1)$$
$$= \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\left(\boldsymbol{I} + \sum_{l=1}^{\infty}(-\boldsymbol{\alpha}(k+1))^l\right) + \Delta\bar{\boldsymbol{P}}(k+1|k)\boldsymbol{\phi}(k+1)$$
$$= \mathbf{K}(k+1) + \Delta\bar{\boldsymbol{K}}(k+1),$$

where

$$\mathbf{K}(k+1) = \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}$$

and

$$\Delta\bar{\boldsymbol{K}}(k+1) = \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\sum_{l=1}^{\infty}(-\boldsymbol{\alpha}(k+1))^l + \Delta\bar{\boldsymbol{P}}(k+1|k)\boldsymbol{\phi}(k+1).$$
$$(28)$$

Then using $\Delta\bar{\boldsymbol{K}}(k+1)$ we obtain

$$(\boldsymbol{I} - \bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A} = \left(\boldsymbol{I} - (\mathbf{K}(k+1) + \Delta\bar{\boldsymbol{K}}(k+1))\boldsymbol{C}\right)\boldsymbol{A},$$
$$= \left(\boldsymbol{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\boldsymbol{C} - \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C}\right)\boldsymbol{A},$$

as required.

Proof of Proposition 5

In order to proof (5a) we start by taking the determinant of (22) and use expression in *Proposition* 4:

$$\prod_{i=1}^{n}\bar{\lambda}_i(k+1) = \det\left((\mathbf{I} - \bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A}\right)$$
$$= \det\left(\mathbf{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\boldsymbol{C} + \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C}\right)\det(\mathbf{A})$$
$$= \det\left(\mathbf{I} - \mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\boldsymbol{C} + \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C}\right)\prod_{i=1}^{n}\mu_i.$$

For the proof of (5b) we take the trace of (22) to obtain the required result:

$$\sum_{i=1}^{n}\bar{\lambda}_i(k+1) = \text{tr}\left((I - \bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A}\right)$$
$$= \text{tr}(\mathbf{A}) - \text{tr}\left((\mathbf{P}(k+1|k)\mathbf{C}'\mathbf{S}(k+1)^{-1}\boldsymbol{C} + \Delta\bar{\boldsymbol{K}}(k+1)\boldsymbol{C})\boldsymbol{A}\right).$$

Proof of Proposition 6

In order to prove *Proposition 6* we start by using (1) while rewriting $\mathbf{\Delta A}(k-1)\boldsymbol{\mathcal{X}}^\gamma(k-1)$ as $\mathbf{F}(k-1)\mathbf{u}(k-1)$. We expand $\hat{\boldsymbol{\mathcal{X}}}(k|k)$ using (10) and rearrange the terms:

$$
\begin{aligned}
\boldsymbol{\mathcal{X}}(k) - \hat{\boldsymbol{\mathcal{X}}}(k|k) &= \mathbf{A}\boldsymbol{\mathcal{X}}(k-1) + \mathbf{F}(k-1)\mathbf{u}(k-1) + \mathbf{B}w(k) \\
&\quad - \left(\mathbf{A}\hat{\boldsymbol{\mathcal{X}}}(k-1|k-1) + \bar{K}(k)(\boldsymbol{\mathcal{Y}}(k) - \mathbf{CA}\hat{\boldsymbol{\mathcal{X}}}(k-1|k-1))\right) \\
&= \left(\mathbf{I} - \bar{K}(k)\mathbf{C}\right)\mathbf{A}\left(\boldsymbol{\mathcal{X}}(k-1) - \hat{\boldsymbol{\mathcal{X}}}(k-1|k-1)\right) \\
&\quad + \left(\mathbf{I} - \bar{K}(k)\mathbf{C}\right)\left(\mathbf{B}w(k) + \mathbf{F}(k-1)\mathbf{u}(k-1)\right) - \bar{K}(k)\mathbf{D}v(k).
\end{aligned}
\tag{29}
$$

Proof of Proposition 7

We take expectations of both sides of (29) raised to the power 3:

$$
\begin{aligned}
\mathbb{E}\left(\boldsymbol{\mathcal{X}}(k) - \hat{\boldsymbol{\mathcal{X}}}(k|k)\right)^3 &= \mathbb{E}(\mathbf{\Delta_s} + \mathbf{\Delta_n})^3 \\
&= \mathbb{E}(\mathbf{\Delta_s^3}) + 3\mathbb{E}(\mathbf{\Delta_s^2\Delta_n}) + 3\mathbb{E}(\mathbf{\Delta_s\Delta_n^2}) + \mathbb{E}(\mathbf{\Delta_n^3}) \\
&= \mathbb{E}(\mathbf{\Delta_s^3}) \\
&= \mathbb{E}\left((\mathbf{I} - \bar{K}(k)\mathbf{C})\mathbf{A}\left(\boldsymbol{\mathcal{X}}(k-1) - \hat{\boldsymbol{\mathcal{X}}}(k-1|k-1)\right)\right)^3 \\
&= (\mathbf{I} - \bar{K}(k)\mathbf{C})^3\mathbf{A}^3\mathbb{E}\left(\boldsymbol{\mathcal{X}}(k-1) - \hat{\boldsymbol{\mathcal{X}}}(k-1|k-1)\right)^3.
\end{aligned}
\tag{30}
$$

On expanding both sides of (30) we get

$$
\mathbb{E}\left(\boldsymbol{\mathcal{X}}^3(k) - 3\boldsymbol{\mathcal{X}}^2(k)\hat{\boldsymbol{\mathcal{X}}}(k|k) + 3\boldsymbol{\mathcal{X}}(k)\hat{\boldsymbol{\mathcal{X}}}^2(k|k) - \hat{\boldsymbol{\mathcal{X}}}^3(k|k)\right)
$$
$$
= \mathbb{E}\left(\boldsymbol{\mathcal{X}}^3(k-1) - 3\boldsymbol{\mathcal{X}}^2(k-1)\hat{\boldsymbol{\mathcal{X}}}(k-1|k-1) + 3\boldsymbol{\mathcal{X}}(k-1)\hat{\boldsymbol{\mathcal{X}}}^2(k-1|k-1) - \hat{\boldsymbol{\mathcal{X}}}^3(k-1|k-1)\right) *
$$
$$
* (\mathbf{I} - \bar{K}(k)\mathbf{C})^3\mathbf{A}^3.
\tag{31}
$$

Rearranging equation (31) we get the required result for $\mathbb{E}(\boldsymbol{\mathcal{X}}(k))^3$.

Proof of Proposition 8

Left hand side of equation (25) simplifies to

$$
\mathbb{E}\left(\boldsymbol{\mathcal{X}}(k) - \hat{\boldsymbol{\mathcal{X}}}(k|k)\right)^q = \sum_{i=0}^{q}\binom{q}{i}\mathbb{E}(\boldsymbol{\mathcal{X}}(k))^i(\hat{\boldsymbol{\mathcal{X}}}(k|k))^{q-i}.
\tag{32}
$$

When $i \in \{0, 1, 2\}$, expectations $\mathbb{E}(\boldsymbol{\mathcal{X}}(k))^i$ in (32) are given by filter update equations in section 3. However the problem arises for $i \geq 3$. We have shown in *Proposition* 7 how to find the expectation for $i = 3$. Similarly to find these expectations for $i \geq 3$, we need to raise $\mathbb{E}\left(\boldsymbol{\mathcal{X}}(k) - \hat{\boldsymbol{\mathcal{X}}}(k|k)\right)$ to the power q, where $q = 3, 4, ..., 2\gamma$. Right hand side of equation (25) can be expressed as

$$\mathbb{E}(\boldsymbol{\Delta_s} + \boldsymbol{\Delta_n})^q = \sum_{j=0}^{q} \begin{pmatrix} q \\ j \end{pmatrix} \mathbb{E}(\boldsymbol{\Delta_n})^j \mathbb{E}(\boldsymbol{\Delta_s})^{q-j}. \tag{33}$$

We know that for odd values of j we have $\mathbb{E}(\boldsymbol{\Delta_n})^j = 0$. Also $\mathbb{E}(\boldsymbol{\Delta_s}) = 0$ from the definition. Hence equation (33) reduces to

$$\mathbb{E}(\boldsymbol{\Delta_s} + \boldsymbol{\Delta_n})^q = \sum_{j=0,2,...,q} \begin{pmatrix} q \\ j \end{pmatrix} \mathbb{E}(\boldsymbol{\Delta_n})^j \mathbb{E}(\boldsymbol{\Delta_s})^{q-j} \text{ if q is even,}$$

$$= \sum_{j=0,2,...,q-3} \begin{pmatrix} q \\ j \end{pmatrix} \mathbb{E}(\boldsymbol{\Delta_n})^j \mathbb{E}(\boldsymbol{\Delta_s})^{q-j} \text{ if q is odd.} \tag{34}$$

Putting equations (32) and (34) together will give us the required result.

References

[1] B.D.O. ANDERSON AND J.B. MOORE (1979) Optimal filtering, *Prentice Hall, Englewood Cliffs*.

[2] B.N. JAIN (1975) Guaranteed error estimation in uncertain systems, *IEEE Transactions on Signal Processing*, **20**, 230–232.

[3] I.R. PETERSEN AND D.C. MCFARLANE (1996) Optimal guaranteed cost filtering for uncertain discrete-time linear systems, *International Journal on Robust and Nonlinear Control*, 267–280.

[4] K.A. BARBOSA, C.E. DE SOUZA AND A. TROFINO (2002) Robust H_2 filtering for discrete-time uncertain linear systems using parameter-dependent Lyapunov functions, *Proceedings of American Control Conference, Anchorage, Alaska, USA*, 3224–3229.

[5] J.C. GEROMEL (1999) Optimal linear filtering under parameter uncertainty, *IEEE Transactions on Signal Processing*, **47**, 168–175.

[6] X. ZHU, Y.C. SOH AND L. XIE (2002) Design and analysis of discrete-time robust Kalman filters, *Automatica*, **38**, 1069–1077.

[7] I.R. PETERSEN AND A.V. SAVKIN (1999) Robust Kalman filtering for signal and systems with large uncertainties, *Birkhauser, Boston*.

[8] L. XIE AND Y.C. SOH (1994) Robust Kalman filtering for uncertain systems, *Systems and Control Letters*, **22**, 123–129.

[9] U. SHAKED AND C.E. DE SOUZA (1995) Robust minimum variance filtering, *IEEE Transactions on Signal Processing*, **43**, 2474–2483.

[10] J. ZHU, J. PARK AND K.S. LEE(2007) Robust Kalman filter of discrete-time Markovian jump system with parameter and noise uncertainty, *Proceedings of the 7th WSEAS International Conference on Simulation, Modelling and Optimization, Beijing, China*.

[11] F. YANG, Z. WANG AND Y.S. HUNG (2002) Robust Kalman filtering for discrete time-varying uncertain systems with multiplicative noises, *IEEE Transactions on Automatic Control*, **47**, 1179–1183.

[12] R.F. SUOTO, J.Y. ISHIHARA AND G.A. BORGES (2009) A robust extended Kalman filter for discrete-time systems with uncertain dynamics, measurements and correlated Noise, *Proceedings of American Control Conference, St. Louis, Missouri, USA*.

[13] L. EL GHAOUI AND G. CALAFIORE (2001) Robust filtering for discrete-time systems with bounded noise and parametric uncertainty, *IEEE Transactions on Automatic Control*, **46**, 1084–1089.

[14] C.E. DE SOUZA, M. FU AND K.A. BARBOSA (2008) Robust filtering for uncertain linear discrete-time descriptor systems, *Automatica*, **44**, 792–798.

[15] L. XIE, L. LU, D. ZHANG AND H. ZHANG (2003) Robust filtering for uncertain discrete-time systems: an improved LMI approach, *Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Hawaii*.

[16] K. RAMAKRISHNAN AND G. RAY (2012) Robust stability criteria for a class of uncertain discrete-time systems with time-varying delay, *Applied Mathematical Modelling*, dx.doi.org/10.1016/j.apm.2012.03.045.

[17] Z. WANG, L. HUANG AND X. YANG (2011) H_∞ performance for a class of uncertain stochastic nonlinear Markovian jump systems with time-varying delay via adaptive control method, *Applied Mathematical Modelling*, **35**, 1983–1993.

[18] I. PETERSEN AND M.R. JAMES (1996) Performance analysisand controller synthesis for nonlinear systems with stochastic uncertainty constraints, *Automatica*, **32**, 959–972.

[19] E. GERSHON, U. SHAKED, I. YAESH (2001) $\mathcal{H}_\infty$ control and filltering of

22

discrete-time stochastic systems with multiplicative noise, *Automatica* , **37**, 409–417.

[20] L. EL GHAOUI (1995) State-feedback control of systems with multiplicative noise via linear matrix inequalities, *Systems and Control Letters*, **24**, 223–228.

[21] Z. WANG, F. YANG AND X. LIU (2006) Robust filtering for systems with stochastic non-linearities and deterministic uncertainties, *Journal of Systems and Control Engineering*, **220**, 171–182.

[22] J.C. COX, J.E. INGERSOLL AND S.A.ROSS (1985) A theory of the term structure of interest rates, *Econometrica*, **53**, 385–407.

[23] W.L. DE KONING (1984) Optimal estimation of linear discrete-time systems with stochastic parameters, *Automatica*, **20**, 113–115.

[24] K. PONOMAREVA (2012) Latent state estimation in a class of nonlinear systems, PhD thesis, *Brunel University, School of Information Systems, Computing and Mathematics*.

[25] A. GELB (1986) Applied optimal estimation, *M.I.T. Press, Cambridge, M.A.*.

[26] D. SIMON (2006) Optimal state estimation: Kalman, $\mathcal{H}_\infty$ and nonlinear approaches, *Wiley-Blackwell, Hoboken, New Jersey*.

[27] T. MORI (1988) Comments on "A matrix inequality associated with bounds on solutions of algebraic Riccati and Lyapunov equation", *IEEE Transactions on Automatic Control*, **33**, 1088.

[28] R.A. HORN AND C.R. JOHNSON (1985) Matrix Analysis, Cambridge University Press.

[29] J.C. COX (1996) The constant elasticity of variance option pricing model, *Journal of Portfolio Management*, **22**, 16–17.