# Robust Kalman Filters for Linear Time-Varying Systems With Stochastic Parametric Uncertainties

Fan Wang, *Member, IEEE*, and Venkataramanan Balakrishnan, *Member, IEEE*

***Abstract*—We present a robust recursive Kalman filtering algorithm that addresses estimation problems that arise in linear time-varying systems with stochastic parametric uncertainties. The filter has a one-step predictor-corrector structure and minimizes an upper bound of the mean square estimation error at each step, with the minimization reduced to a convex optimization problem based on linear matrix inequalities. The algorithm is shown to converge when the system is mean square stable and the state space matrices are time invariant. A numerical example consisting of equalizer design for a communication channel demonstrates that our algorithm offers considerable improvement in performance when compared with conventional Kalman filtering techniques.**

***Index Terms*—Linear matrix inequality, linear time-varying systems, robust Kalman filters, stochastic parametric uncertainty.**

## I. Introduction

THE NOTATIONS in this paper are fairly standard. $E[\cdot]$ denotes the expectation of a random variable (matrix). Var$(\cdot)$ denotes the variance of a random variable (vector). $P > 0 (P \geq 0)$ means that $P$ is a symmetric and positive definite (positive semi-definite) matrix. $A \gneqq B$ means that $A \geq B$ and $A \neq B$. $\mathbf{Tr}(\cdot)$ is the trace of a matrix. $\mathbf{Co}\{\cdot\}$ denotes a convex hull. $\mathbf{diag}(\cdot)$ defines a (block) diagonal matrix. $\|\cdot\|$ is the matrix norm, that is, the largest singular value of a matrix. $^\dagger$ denotes the Moore–Penrose pseudo inverse of a matrix. $(*)^T$ and $(**)^T$ are used in some places to represent the $(2,1)$ and $(3,1)$ terms of a symmetric matrix when the $(1,2)$ and $(1,3)$ terms are given. $\hat{x}(m\,|\,n)$ is the estimation of $x(m)$ with observations up to time $n$. If $m > n$, $\hat{x}(m\,|\,n)$ is known as a predicted estimation. If $m = n$, $\hat{x}(m\,|\,n)$ is known as a filtered estimation. For discrete-time systems considered in the paper, the state matrix $A$ is said to be stable if all the eigenvalues of $A$ are strictly inside the unit circle. A discrete-time system $x(k+1) = A(k)x(k)$, where $A(k)$ is a random process, is said to be *mean square stable* (see [1]) if for all initial conditions $x(0)$, we have $\lim_{k\to\infty} E[x(k)x(k)^T] = 0$, i.e., $\lim_{k\to\infty} x(k) = 0$ almost surely.

Consider the linear system

$$\begin{aligned} x(k+1) &= \mathcal{A}(k)x(k) + \mathcal{B}(k)u_i(k) \\ y(k) &= \mathcal{C}(k)x(k) + \mathcal{D}(k)u_m(k) \end{aligned} \tag{1}$$

where $k = 0, 1, 2, \ldots, x(k) \in \mathbf{R}^n$ is the state, $u_i(k) \in \mathbf{R}^{n_w}$ is the input noise, $y(k) \in \mathbf{R}^{n_y}$ is the measured output, and $u_m(k) \in \mathbf{R}^{n_u}$ is the measurement noise, with $u_i$ and $u_m$ being independent white noise random processes. Equation (1) models systems that are affected by both actuator and sensor noises ($u_i$ and $u_m$, respectively). A fundamental problem associated with such systems is that of state estimation, i.e., the optimal estimation of the state $x(k)$ from the noisy measurements $\{y(i), i = 0, 1, \ldots, k\}$; the corresponding state estimate is denoted $\hat{x}(k\,|\,k)$. Such estimation problems arise in several applications in signal processing, communications, and control; see, for example, [2], [3], and the references therein.

Recursive minimum mean-square error (MMSE) estimators form an important class of optimal state estimators for system (1) and have many applications in signal processing, communications, and automatic control [3]–[7]. MMSE estimators minimize the expected value of the square of the estimation error, i.e., $E[(x(k) - \hat{x}(k\,|\,k))^T(x(k) - \hat{x}(k\,|\,k))]$, at each $k$. When the random processes $u_i$ and $u_m$ are Gaussian, it turns out that the MMSE estimator is a linear filter whose coefficients can be determined by solving a Riccati difference equation. (This is the celebrated Kalman filter.) When $u_i$ and $u_m$ are not Gaussian, the Kalman filter yields the best linear MMSE estimator. An important (and desirable) property of the Kalman filtering algorithm is that it converges when system (1) is time invariant and detectable as well as stabilizable [8].

The Kalman filter consists of the following two parts:[1]

1) One-step prediction update:

$$\begin{aligned} x_f(k) &= \mathcal{A}(k-1)x_f(k-1) \\ &\quad + K(k-1)(\mathcal{C}(k-1)x_f(k-1) - y(k-1)) \\ \hat{x}(k\,|\,k-1) &= x_f(k). \end{aligned} \tag{2a}$$

2) Filtered estimation update:

$$\hat{x}(k\,|\,k) = \hat{x}(k\,|\,k-1) + F(k)(\mathcal{C}(k)\hat{x}(k\,|\,k-1) - y(k)). \tag{2b}$$

When the matrices $\mathcal{A}(k), \mathcal{B}(k), \mathcal{C}(k)$, and $\mathcal{D}(k)$ in (1) can be measured exactly, computing the Kalman gains $K(k-1)$ and $F(k)$ in (2) is equivalent to a quadratic optimization problem: one that can be solved analytically [8]. However, in many cases, there exist uncertainties in model parameters and/or model structure because of errors from system identification or model reduction routines; see, for example, [4], [7], [10]–[12]. The performance of estimators designed without accounting for these uncertainties can be severely degraded and sometimes even unacceptable [13], [12]. Thus, estimators

Manuscript received February 12, 2001; revised January 2, 2002. This work was supported by the Office of Naval Research under Contract N00014-97-1-0640. The associate editor coordinating the review of this paper and approving it for publication was Dr. Olivier Besson.

F. Wang is with the Advanced Radio Technology, GTSS, Motorola, Inc., Arlington Heights, IL 60004 USA (e-mail: fanw@cig.mot.com).

V. Balakrishnan is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-1285 USA (e-mail: ragu@ecn.purdue.edu).

Publisher Item Identifier S 1053-587X(02)02387-5.

[1] A more complete introduction to Kalman filter can be found in the literature; see, for example, [8] and [9].

TABLE I
COMPARISON OF SEVERAL ESTIMATION METHODS

| | Input signal | Uncertainty | Design criterion | Filter structure |
|---|---|---|---|---|
| Robust Kalman | white noise | white noise | MMSE | recursive |
| Kalman ([8]) | white noise | — | MMSE | recursive |
| Robust $H_2$ ([15]) | white noise | norm bounded | $H_2$ | LTI |
| Robust $H_\infty$ ([14]) | $\ell_2$ signal | norm bounded | $H_\infty$ | LTI |
| Stochastic $H_\infty$ ([17]) | mean energy bounded | white noise | stochastic $H_\infty$ | LTI |
| Game theoretic estimator ([12]) | 2-norm bounded | norm bounded | finite horizon quadratic objective | recursive |

must be designed with graceful performance degradation in the presence of modeling errors. This issue of "robust estimation" has been addressed in a number of recent publications; see, for example, [12], [14]–[16], and the references therein. By assuming the input signal has limited total energy (or $\ell_2$-norm), linear time-invariant (LTI) filters have been designed (see [14] and [17]) to optimize the steady-state performance measured via the $H_\infty$ norm (or $\ell_2$-gain) of the map from the noise input to the estimation error. In another scenario where the input signal is white noise with limited power, linear time-invariant filters have been designed (see [15]) to optimize the steady-state performance measured via the $H_2$ norm of the map from the noise input to the estimation error.

The models considered in robust estimation problems fall under two classes. One class consists of a linear time-invariant system affected by parametric uncertainties that are deterministic and typically known only to lie in some bounded set [12], [14], [18]. The second class consists of linear time-invariant systems affected by stochastic uncertainties, which can also be viewed as a multiplicative noise inputs; see [17], [19], and [20].

Table I summarizes the characteristics of several estimation problems and their solutions. While the above-cited works on robust estimation in the literature provide a fairly complete solution to several steady-state estimation problems, the solutions are linear time-invariant filters, and none of them consider the transient behavior of their estimation algorithms. Indeed, even the conventional Kalman filter is initialized in an *ad hoc* fashion, leaving room for improvement in its transient performance.

In this paper, we consider MMSE estimation problems for linear time-varying systems affected by stochastic uncertainties, with a view toward optimizing the transient performance of the estimation. The stochastic uncertainties that we consider affect the system matrices; in addition, we assume that the correlation of the state initial condition is known only to lie in a polytope.[2] For such systems, starting with the standard one-step predictor-corrector filter structure (2), we develop a recursive estimation algorithm where at each step, an upper bound of the mean square of the estimation error over all possible uncertainties is minimized. The minimization is performed via numerical convex optimization over linear matrix inequalities (LMIs). We will refer to our algorithm as the robust Kalman filtering algorithm. As a by-product of our robust filtering algorithm, we also obtain a technique for optimally initializing a recursive filtering algorithm, for instance, the conventional Kalman filter.

[2]Our framework is perhaps related closest with the one in [10], where the uncertainties affect the noise moments; a game-theoretic argument is used to establish the existence of an optimal recursive scheme.

We will demonstrate through an example that the robust Kalman filter can provide much improved transient performance when compared with the conventional Kalman filter. Perhaps more important, for systems with stochastic parametric uncertainties, the performance of the conventional Kalman filter can be severely degraded, whereas the performance of the robust Kalman filter degrades fairly gracefully.

As with the conventional Kalman filtering algorithm for time-varying systems, the convergence of the robust Kalman filtering algorithm that we present is not guaranteed in general. However, we prove convergence in the special case of the estimation of a system with time-invariant state-space matrices and stochastic parametric uncertainties, provided that the uncertain system is mean-square stable. Moreover, we show that the conventional Kalman filter is a special case of the proposed robust Kalman filtering algorithm for systems with no uncertainties.

The organization of the paper is as follows. In Section II, we discuss the mathematical framework underlying our problem and make some preliminary remarks. In Section III, we describe the derivation of a robust Kalman filtering algorithm that minimizes an upper bound of the mean square of the estimation error at each step. We also present the convergence property of this recursive algorithm and its connection with the conventional Kalman filter. In Section IV, we apply the filtering technique developed in this paper to design equalizers for a communication channel. The proofs are given in Appendix A.

## II. PROBLEM SETUP

We consider the following linear time-varying system:[3]

$$\begin{aligned} x(k+1) &= A_\Delta(k)x(k) + B_\Delta(k)w(k) \\ y(k) &= C_\Delta(k)x(k) + D_\Delta(k)w(k) \\ z(k) &= L(k)x(k) \end{aligned} \tag{3a}$$

where

$$\begin{aligned} A_\Delta(k) &= A(k) + \sum_{i=1}^{m} A_i^s(k)\zeta_i^s(k) \\ B_\Delta(k) &= B(k) + \sum_{i=1}^{m} B_i^s(k)\zeta_i^s(k) \\ C_\Delta(k) &= C(k) + \sum_{i=1}^{m} C_i^s(k)\zeta_i^s(k) \\ D_\Delta(k) &= D(k) + \sum_{i=1}^{m} D_i^s(k)\zeta_i^s(k) \end{aligned} \tag{3b}$$

[3]$w(k)$ is a vector containing both input noise and measurement noise.

$x(k) \in \mathbf{R}^n, y(k) \in \mathbf{R}^{n_y}$, and $w(k) \in \mathbf{R}^{n_w}$. $z(k) \in \mathbf{R}^{n_z}$ is the signal we wish to estimate. $w$ is a zero-mean white noise process and satisfies $E[w(i)w(j)^T] = \delta(i-j)I$, where $\delta(k)$ is the Dirac Delta function. $\zeta_i^s, i = 1, \ldots, m$ are zero-mean random processes with $E[\zeta_i^s(k)\zeta_j^s(l)] = \delta(i-j)\delta(k-l)$. The initial state $x(0)$ of system (3a) is a random vector, with its correlation $X(0) = E[x(0)x(0)^T]$ known only to lie in a polytope $\mathbf{Co}\{X_1(0), \ldots, X_p(0)\}$. The random processes $w$, $\zeta_i^s$, $i = 1, \ldots, m$, and the random vector $x(0)$ are mutually independent.

System (3) is said to be *mean square stable* (see [1]) if, with $w(k) = 0$ for $k = 0, 1, 2, \ldots$, we have

$$\lim_{k \to \infty} E[x(k)x(k)^T] = 0$$

for any initial condition $x(0)$.

Our objective is to design an optimal robust filter of the one-step predictor-corrector form given in (2), where $\hat{z}(k) = L(k)\hat{x}(k \,|\, k)$ is an estimate of $z(k)$ (see Fig. 1). (The case $L(k) = I$ corresponds to state estimation.) Specifically, since the correlation $X(k)$ of the state at each $k$ depends on the correlation $X(0)$ of the initial condition, $X(k)$ is uncertain when $X(0)$ is uncertain. We wish to find the optimal Kalman gains $K(k-1)$ and $F(k)$ to minimize the maximum value of the mean square of the estimation error $E\|z(k) - \hat{z}(k)\|^2$ over all allowable values for $X(k)$. (Here, the expectation is taken with respect to the random initial state, the input and measurement noises, as well as the stochastic parametric uncertainties $\zeta_i^s$).

Compared with the steady state $H_2$ and $H_\infty$ filters in [14], [15], and [17], the filters we design are recursive and optimize the transient performance. In addition, the filters we design are robust to the stochastic parametric uncertainties, in comparison with the conventional (nonrobust) Kalman filters [8], [9], which are designed based on nominal models with no uncertainties.

## III. Robust Kalman Filtering

### A. System Model

We begin by rewriting (3) in an equivalent form

$$\begin{aligned} x(k+1) &= A(k)x(k) + v_x(k) \\ y(k) &= C(k)x(k) + v_y(k) \\ z(k) &= L(k)x(k) \end{aligned}$$

where

$$\begin{aligned} v_x(k) = &\sum_{i=1}^{m} A_i^s(k)\zeta_i^s(k)x(k) \\ &+ \left(B(k) + \sum_{i=1}^{m} B_i^s(k)\zeta_i^s(k)\right) w(k) \end{aligned}$$

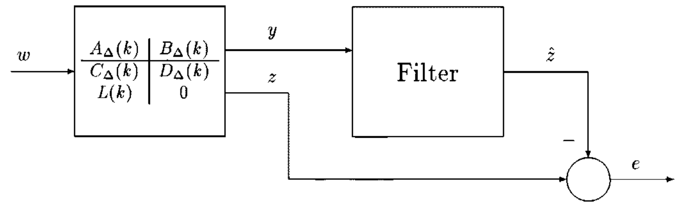



Fig. 1. Estimation of linear time-varying systems with stochastic parametric uncertainties.

and

$$\begin{aligned} v_y(k) = &\sum_{i=1}^{m} C_i^s(k)\zeta_i^s(k)x(k) \\ &+ \left(D(k) + \sum_{i=1}^{m} D_i^s(k)\zeta_i^s(k)\right) w(k). \end{aligned}$$

$v_x$ and $v_y$ are random processes. Since $\zeta_i^s$, $i = 1, \ldots, m$ and $w$ are independent zero-mean white-noise processes, the mean and the variance of $v_x$ and $v_y$ satisfy the following conditions for every $k = 0, 1, 2, \ldots$:

$$\begin{gathered} E[v_x(k)] = 0, \quad E[v_y(k)] = 0 \\ \mathrm{Var}\left(\begin{bmatrix} v_x(k) \\ v_y(k) \end{bmatrix}\right) = \begin{bmatrix} \hat{B}(k)\hat{B}(k)^T & \hat{B}(k)\hat{D}(k)^T \\ \hat{D}(k)\hat{B}(k)^T & \hat{D}(k)\hat{D}(k)^T \end{bmatrix} \end{gathered}$$

where $\hat{B}(k)$ and $\hat{D}(k)$ are defined at the bottom of the page, and $X(k) = E[x(k)x(k)^T]$. We also note that $v_x$ and $v_y$ are white-noise random processes so that when $i \neq j$, $E[v_x(i)v_x(j)^T] = 0$, $E[v_y(i)v_y(j)^T] = 0$, and $E[v_x(i)v_y(j)^T] = 0$. In addition, $E[x(i)v_x(j)^T] = 0$ and $E[x(i)v_y(j)^T] = 0$ for $i \leq j$. Therefore, random processes $[v_x, v_y]$ and $[\hat{B}(k)v(k), \hat{D}(k)v(k)]$ have the same (first- and second-order) statistics, where $v$ is a zero-mean unit-variance white noise random process.

Thus, the model in (3) can be rewritten as

$$\begin{aligned} x(k+1) &= A(k)x(k) + \hat{B}(k)v(k) \\ y(k) &= C(k)x(k) + \hat{D}(k)v(k) \\ z(k) &= L(k)x(k). \end{aligned} \tag{4}$$

If the variance $X(0) = E[x(0)x(0)^T]$ is known, then it is easily verified that $X(k)$ is uniquely determined by the following recursion:

$$\begin{aligned} X(k+1) = h(X(k)) &\triangleq A(k)X(k)A(k)^T + B(k)^T B(k) \\ &+ \sum_{j=1}^{m} \left(A_j^s(k)X(k)A_j^s(k)^T + B_j^s(k)B_j^s(k)^T\right). \end{aligned} \tag{5}$$

However, in our setting, $X(0)$ is only known to lie in a polytope $\mathbf{Co}\{X_1(0), \ldots, X_p(0)\}$. Since in the recursion (5) $X(k+1)$ is

---

$$\begin{aligned} \hat{B}(k) &= \begin{bmatrix} A_1^s(k)X(k)^{1/2} & \cdots & A_m^s(k)X(k)^{1/2} & B(k) & B_1^s(k) & \cdots & B_m^s(k) \end{bmatrix} \\ \hat{D}(k) &= \begin{bmatrix} C_1^s(k)X(k)^{1/2} & \cdots & C_m^s(k)X(k)^{1/2} & D(k) & D_1^s(k) & \cdots & D_m^s(k) \end{bmatrix} \end{aligned}$$

a linear function of $X(k)$, it can be shown recursively that $X(k)$ also lies in a polytope $\mathbf{Co}\{X_1(k), \dots, X_p(k)\}$, where

$$X_i(k+1) = A(k)X_i(k)A(k)^T + B(k)B(k)^T + \sum_{j=1}^{m} \left(A_j^s(k)X_i(k)A_j^s(k)^T + B_j^s(k)B_j^s(k)^T\right). \quad (6)$$

If the system is mean square stable and the state-space matrices are time-invariant, the polytope converges to a fixed point as $k \to \infty$; see Proposition A.3 in Appendix A.

### B. Robust Kalman Filter Recursions

While (4) is similar to the setting for the conventional Kalman filtering problem, we note that the system matrices $\hat{B}(k)$ and $\hat{D}(k)$ depend on the second moment of the states and are, thus, uncertain. For this reason, the optimal linear recursive MMSE filter design problem is considerably harder than the conventional Kalman filtering problem, which can be solved analytically [8].

With uncertain system (3), at each $k$, we are given the following:

- N1) $L(k)$, as well as $A(k)$, $B(k)$, $C(k)$, and $D(k)$, which are the measurements of the system matrices $A_\Delta(k)$, $B_\Delta(k)$, $C_\Delta(k)$, and $D_\Delta(k)$. Note that the measurement noise in system matrices is characterized by the terms of $\zeta_i^s$ in (3b).
- N2) $y(k)$, which is the measurement of the noisy output. Note that the measurement noise in output is characterized by the terms of $w$ in (3a).
- N3) From (6), the vertices $X_1(k), \dots, X_p(k)$ of the polytope in which $X(k)$ lies.

Our objective is to design a linear MMSE estimator

$$\begin{aligned} x_f(k) &= A(k-1)x_f(k-1) + K(k-1) \\ &\quad \times (C(k-1)x_f(k-1) - y(k-1)) \\ \hat{x}(k \mid k-1) &= x_f(k) \qquad (7a) \\ \hat{x}(k \mid k) &= \hat{x}(k \mid k-1) + F(k)(C(k)\hat{x}(k \mid k-1) - y(k)) \qquad (7b) \end{aligned}$$

where the Kalman gains $K(k-1)$ and $F(k)$ are obtained from the solution of the following minimax optimization problem:

$$\begin{aligned} &\underset{K(k-1),F(k)}{\texttt{Minimize:}} \quad \max_{X(k)} \quad E[\|z(k) - \hat{z}(k)\|^2] \\ &\texttt{Subject to:} \qquad \text{Data (N1–N3)} \\ &\qquad\qquad\qquad\qquad \text{Equations (7) and (4).} \qquad (8) \end{aligned}$$

The minimax problem (8) is difficult to solve directly. We therefore first define an upper bound on the quantity

$$J(K(k-1), F(k)) \triangleq \max_{X(k)} \quad E[\|z(k) - \hat{z}(k)\|^2]$$

and then determine $K(k-1)$ and $F(k)$ to minimize the upper bound. We will show that this results in a convex optimization problem.

We now proceed with the derivation of an upper bound on $J(K(k-1), F(k))$. The estimation algorithm we will derive is a recursive algorithm. The initialization of the recursions will be discussed in Section III-A. At this stage, we focus our attention on the recursions. We assume that at each $k$, we have available a matrix $P(k-1)$ with

$$P(k-1) \geq E[(x(k-1) - \hat{x}(k-1 \mid k-2)) \times (x(k-1) - \hat{x}(k-1 \mid k-2))^T].$$

Then

$$\begin{aligned} &E[(x(k) - \hat{x}(k \mid k-1))(x(k) - \hat{x}(k \mid k-1))^T] \\ &\quad \leq \sup_{X(k-1)} (A(k-1) + K(k-1)C(k-1)) \\ &\qquad \times P(k-1)(A(k-1) + K(k-1)C(k-1))^T \\ &\qquad + (\hat{B}(k-1) + K(k-1)\hat{D}(k-1)) \\ &\qquad \times (\hat{B}(k-1) + K(k-1)\hat{D}(k-1))^T \qquad (9) \\ &\quad \triangleq f(P(k-1), K(k-1)) \qquad (10) \end{aligned}$$

where $X(k-1)$ lies in $\mathbf{Co}\{X_1(k-1), \dots, X_p(k-1)\}$.

Then, we can obtain $P(k) \geq E[(x(k) - \hat{x}(k \mid k-1))(x(k) - \hat{x}(k \mid k-1))^T]$ simply by requiring that

$$f(P(k-1), K(k-1)) \leq P(k). \qquad (11)$$

Next, we have

$$\begin{aligned} &E[(x(k) - \hat{x}(k \mid k))(x(k) - \hat{x}(k \mid k))^T] \\ &\quad \leq \sup_{X(k)=h(X(k-1))} (I + F(k)C(k))P(k)(I + F(k)C(k))^T \\ &\qquad + (F(k)\hat{D}(k))(F(k)\hat{D}(k))^T \qquad (12) \\ &\quad \triangleq \tau(F(k), P(k)) \qquad (13) \end{aligned}$$

where $X(k-1)$ lies in $\mathbf{Co}\{X_1(k-1), \dots, X_p(k-1)\}$. Then, for any matrix $M(k)$ that satisfies

$$\tau(F(k), P(k)) \leq M(k) \qquad (14)$$

we have

$$E[(z(k) - \hat{z}(k))(z(k) - \hat{z}(k))^T] \leq L(k)M(k)L(k)^T$$

which implies that $\mathbf{Tr}(L(k)M(k)L(k)^T)$ is an upper bound of the objective function in (8).

We can now formulate the problem of determining $K(k-1)$ and $F(k)$ to minimize an upper bound on the the mean square of the estimation error over all possible values for the corresponding state correlation matrix as follows:

$$\begin{aligned} &\underset{P(k),M(k),K(k-1),F(k)}{\texttt{Minimize:}} \quad \mathbf{Tr}(L(k)M(k)L(k)^T) \\ &\texttt{Subject to:} \quad \text{Conditions (11) and (14)} \qquad (15) \end{aligned}$$

where $P(k), M(k), K(k-1)$, and $F(k)$ are optimization variables.

While the optimization problem (15) has no analytical solution in general, we establish via the following theorem that it it can be reformulated as a convex optimization problem: that of minimizing a linear objective subject with linear matrix inequality (LMI) constraints. This problem can be solved numerically very efficiently using standard algorithms [21], [22] so that Theorem 3.1 provides for an efficient and effective numerical solution of problem (15); for details on LMIs, see [1] and the references therein.

*Theorem 3.1:* Consider the optimization problem

$$\underset{Q(k),M(k),Y(k),F(k)}{\texttt{Minimize:}} \quad \mathbf{Tr}(L(k)M(k)L(k)^T) \tag{16a}$$

Subject to:

$$\begin{bmatrix} Q(k) & T_{12}(k-1) & T_{13,i}(k-1) \\ T_{12}(k-1)^T & Q(k-1) & 0 \\ T_{13,i}(k-1)^T & 0 & I \end{bmatrix} \geq 0 \tag{16b}$$

$$\begin{bmatrix} M(k) & I+F(k)C(k) & F(k)\hat{D}_i(k) \\ I+C(k)^TF(k)^T & Q(k) & 0 \\ \hat{D}_i(k)^TF(k)^T & 0 & I \end{bmatrix} \geq 0$$
$$i=1,\ldots,p \tag{16c}$$

where

$$\begin{aligned} T_{12}(k-1) &= Q(k)A(k-1)+Y(k)C(k-1) \\ T_{13,i}(k-1) &= Q(k)\hat{B}_i(k-1)+Y(k)\hat{D}_i(k-1) \\ \hat{B}_i(k) &= \Big[A_1^s(k)X_i(k)^{1/2}\cdots A_m^s(k)X_i(k)^{1/2} \\ &\qquad B(k)\ B_1^s(k)\cdots B_m^s(k)\Big] \\ \hat{D}_i(k) &= \Big[C_1^s(k)X_i(k)^{1/2}\cdots C_m^s(k)X_i(k)^{1/2} \\ &\qquad D(k)\ D_1^s(k)\cdots D_m^s(k)\Big] \end{aligned} \tag{16d}$$

and $X_i(k)$ is defined in (6).

1) Suppose that $P(k), M(k), K(k-1)$, and $F(k)$ satisfy (11) and (14); then, $Q(k)=P(k)^\dagger, M(k), Y(k)=P(k)^\dagger K(k-1)$, and $F(k)$ satisfy (16b) and (16c).
2) Suppose that $Q(k), M(k), Y(k)$, and $F(k)$ satisfy (16b) and (16c); then, $P(k)=Q(k)^\dagger, M(k), K(k-1)=Q(k)^\dagger Y(k)$, and $F(k)$ satisfy (11) and (14).
3) The optimal value of (16) equals the optimal value of (15).

*Proof:* Let $X(k-1)=\sum_{i=1}^p \lambda_i X_i(k-1)$, where $\lambda_i \in [0,1]$ and $\sum_{i=1}^p \lambda_i = 1$. Since $X(k)$ [see (5)] is a linear function in $X(k-1)$, we have $X(k)=\sum_{i=1}^p \lambda_i X_i(k)$ for the same set of $\lambda_i,\ i=1,\ldots,p$. By Schur's complements lemma, (11) is equivalent to

$$\sum_{i=1}^p \lambda_i \begin{bmatrix} P(k) & \hat{T}_{12}(k) & \hat{T}_{13}(k)\mathcal{X}_i(k-1) \\ (*)^T & P(k-1) & 0 \\ (**)^T & 0 & \mathcal{X}_i^2(k-1) \end{bmatrix} \geq 0 \tag{17}$$

where $(*)$ is the (1, 2) term of the symmetric matrix, and $(**)$ is the (1, 3) term of the symmetric matrix

$$\begin{aligned} \hat{T}_{12}(k) &= (A(k-1)+K(k-1)C(k-1))P(k-1) \\ \hat{T}_{13}(k) &= \hat{B}_i(k-1)+K(k-1)\hat{D}_i(k-1). \end{aligned}$$

Similarly, (14) is equivalent to

$$\sum_{i=1}^p \lambda_i \begin{bmatrix} M(k) & \tilde{T}_{12}(k) & \tilde{T}_{13}(k)\mathcal{X}_i(k) \\ (*)^T & P(k) & 0 \\ (**)^T & 0 & \mathcal{X}_i^2(k) \end{bmatrix} \geq 0 \tag{18}$$

where $\mathcal{X}_i(k) = \mathbf{diag}(X_i(k)^{1/2},\ldots,X_i(k)^{1/2},I,\ldots,I)$, and

$$\tilde{T}_{12}(k) = (I+F(k)C(k))P(k), \quad \tilde{T}_{13}(k) = F(k)\hat{D}_i(k).$$

Since (17) and (18) hold for all $\lambda_i \in [0,1]$ such that $\sum_{i=1}^p \lambda_i = 1$, they are equivalent, respectively, to

$$\begin{bmatrix} P(k) & \hat{T}_{12}(k) & \hat{T}_{13}(k) \\ (*)^T & P(k-1) & 0 \\ (**)^T & 0 & I \end{bmatrix} \geq 0 \tag{19}$$

and

$$\begin{bmatrix} M(k) & (I+F(k)C(k))P(k) & F(k)\hat{D}_i(k) \\ (*)^T & P(k) & 0 \\ (**)^T & 0 & I \end{bmatrix} \geq 0. \tag{20}$$

1) Suppose that $P(k), M(k), K(k-1)$, and $F(k)$ satisfy (11) and (14) and thus satisfy LMIs (19) and (20). Then, multiplying LMI (19) on the left and right by $\mathbf{diag}(P(k)^\dagger, P(k-1)^\dagger, I)$ yields LMI (16b); multiplying LMI (20) on the left and right by $\mathbf{diag}(I, P(k)^\dagger, I)$ yields LMI (16c), where $Q(k)=P(k)^\dagger$ and $Y(k)=P(k)^\dagger K(k-1)$. In addition, $P(k)\geq 0$ implies $Q(k)\geq 0$.
2) Following the same line as in item 1), multiplying LMI (16b) on the left and right by $\mathbf{diag}(Q(k)^\dagger, Q(k-1)^\dagger, I)$ yields LMI (19); multiplying LMI (16c) on the left and right by $\mathbf{diag}(I, Q(k)^\dagger, I)$ yields LMI (20), where $P(k)=Q(k)^\dagger$ and $K(k-1)=Q(k)^\dagger Y(k)$. Therefore, the result of item 2) is established.
3) Item 1) implies that the optimal value of (16) is less than or equal to the optimal value of (15). Item 2) implies that the optimal value of (15) is less than or equal to the optimal value of (16). Therefore, we established that the optimal values of (16) and (15) are equal. ■

## *C. Initialization of the Robust Kalman Filter*

Theorem 3.1 paves the way for a robust Kalman filtering algorithm. To start the algorithm, we need to initialize $Q(0)$, which is the process that we describe next. Let $x_f(0)=0$.[4] Then, we have $Q(0)=E[x(0)x(0)^T]^\dagger = X(0)^\dagger$, where $X(0)\in \mathbf{Co}\{X_1(0),\ldots,X_p(0)\}$. Using Theorem 3.1, $Q(1)$ can be computed as an optimizer to the following problem:

$$\underset{Q(1),M(1),Y(1),F(1)}{\texttt{Minimize:}} \quad \mathbf{Tr}(L(1)M(1)L(1)^T)$$

Subject to:

$$\begin{bmatrix} Q(1) & T_{12}(0) & T_{13,i}(0) \\ T_{12}(0)^T & X(0)^\dagger & 0 \\ T_{13,i}(0)^T & 0 & I \end{bmatrix} \geq 0$$

$$\begin{bmatrix} M(1) & I+F(1)C(1) & F(1)\hat{D}_i(1) \\ I+C(1)^TF(1)^T & Q(1) & 0 \\ \hat{D}_i(1)^TF(1)^T & 0 & I \end{bmatrix} \geq 0$$

$$X(0)\in \mathbf{Co}\{X_1(0),\ldots,X_p(0)\}, \quad i=1,\ldots,p \tag{21}$$

where $M(1)=M(1)^T, Q(1)=Q(1)^T$. We then have $P(1)=Q(1)^\dagger$ and $K(0)=Q(1)^\dagger Y(1)$.

[4]If $E[x(0)]=a$ is known and $a\neq 0$, we may define $x_f(0)=a$, and $Q(0)$ becomes $(X(0)-aa^T)^\dagger$. The initialization process (22) can still be applied by replacing $X(0)$ by $X(0)-aa^T$ and $X_i(0)$ by $X_i(0)-aa^T$, respectively.

The matrix inequalities in (21) are not linear in all the variables because the first LMI constraint has the term $X(0)^\dagger$. However, following the same line as in the proof of Theorem 3.1, it is easily shown that (21) is equivalent to the LMI

$$\underset{Q(1),M(1),Y(1),F(1)}{\texttt{Minimize:}} \ \mathbf{Tr}(L(1)M(1)L(1)^T)$$

Subject to:

$$\begin{bmatrix} Q(1) & T_{12}(0)X_i(0) & T_{13,i}(0) \\ X_i(0)T_{12}^T(0) & X_i(0) & 0 \\ T_{13,i}(0)^T & 0 & I \end{bmatrix} \geq 0, \ i=1,\ldots,p$$
$$\begin{bmatrix} M(1) & I+F(1)C(1) & F(1)\hat{D}_i(1) \\ I+C(1)^TF(1)^T & Q(1) & 0 \\ \hat{D}_i(1)^TF(1)^T & 0 & I \end{bmatrix} \geq 0 \tag{22}$$

where $T_{12}(0), T_{13,i}(0), \hat{D}_i(1)$ are defined in (16d), and $K(0) = Q(1)^\dagger Y(1)$.

*Remark:* The solution to the optimization problem (21) can also be applied toward optimally initializing other recursive algorithms, for instance, the conventional Kalman filtering algorithms, with an attendant improvement in transient performance.

### D. Robust Kalman Filtering Algorithm

We now summarize the various steps in our robust Kalman filtering algorithm.

*Robust Kalman Filtering Algorithm:*

Step 1) Solve (22) to initialize $Q(1), M(1), K(0)$, and $F(1)$. Let $k = 1$.

Step 2) At time $k+1$, let

$$X_i(k+1) = A(k)X_i(k)A(k)^T + B(k)B(k)^T + \sum_{j=1}^{m}\left(A_j^s(k)X_i(k)A_j^s(k)^T + B_j^s(k)B_j^s(k)^T\right) \quad i=1,\ldots,p. \tag{23}$$

Step 3) Solve the optimization problem (16) for $Q(k), M(k)$ and $K(k-1), F(k)$.

Step 4) Repeat Steps 2 and 3.

We note that if we have $x_f(0) = E[x(0)]$ in the estimator (7), then the robust Kalman filtering algorithm is unbiased.

In the following, we will discuss the connection of the robust Kalman filter with the conventional Kalman filter. We will also provide the condition for the convergence of the recursions in the robust Kalman filtering algorithm.

### E. Connection With the Conventional Kalman Filtering Algorithm

If there is no stochastic parametric uncertainty in system (3), i.e., $\zeta_i^s(k) = 0$, it can be shown that the robust Kalman filter reduces to the conventional Kalman filter. Therefore, the robust Kalman filter can be viewed as an extension of the conventional Kalman filter to systems with stochastic parametric uncertainties.

Consider the following optimization problem of the conventional Kalman filtering algorithm (see [8] for details):

$$P^*(k+1) = \min_{K(k)} f(P^*(k), K(k)),^5 \quad \text{and}$$
$$K^*(k) = \text{argmin}_{K(k)} f(P^*(k), K(k)) \tag{24a}$$

where

$$f(P(k),K(k)) = (A+K(k)C)P(k)(A+K(k)C)^T + (K(k)D+B)(K(k)D+B)^T \tag{24b}$$

and where $(A,B,C,D)$ are the state space matrices of the system. The optimal solution to (24a) is

$$K^*(k) = -(AP^*(k)C^T + BD^T)(CP^*(k)C^T + DD^T)^{-1}$$
$$P^*(k+1) = (A+K^*(k)C)P^*(k)(A+K^*(k)C)^T + (K^*(k)D+B)(K^*(k)D+B)^T. \tag{24c}$$

Here, we assume $DD^T$ is nonsingular, which is a standard assumption in Kalman filtering; see, for example, [2], [3], and [8]. The well-known condition for the convergence of the recursions in (24) is given by Proposition A.2 in the Appendix.

The recursions in the robust Kalman filter are based on solving an optimization problem (16) or equivalently (15) at each step. In order to show the equivalence between the robust and the conventional Kalman filtering algorithms under the condition $\zeta_i^s(k) = 0$, it suffices to show that the optimal solution to (24) is the same as the optimal solution to (15), in which $\hat{B}(k) = B, \hat{D}(k) = D$.

*Theorem 3.2:* Let $P(0) \geq 0$ be given. Let $\{P^*(k), K^*(k)\}$ be the sequence consisting of the solutions to the optimization problem (24). Let $\{\tilde{P}^*(k), \tilde{M}^*(k), \tilde{K}^*(k), \tilde{F}^*(k)\}$ be the sequence consisting of the solutions to the optimization problem

$$\min_{P(k),M(k),K(k-1),F(k)} \mathbf{Tr}(M(k)), \quad k=1,2\ldots \tag{25a}$$

where

$$f(P(k-1),K(k-1)) = (A+K(k-1)C)P(k-1)(A+K(k-1)C)^T + (K(k-1)D+B)(K(k-1)D+B)^T \tag{25b}$$
$$P(k) = f(P(k-1),K(k-1)) \tag{25c}$$
$$M(k) = (I+F(k)C)P(k)(I+F(k)C)^T + F(k)DD^TF(k)^T \tag{25d}$$

and $DD^T$ is assumed to be positive definite.

Then

$$P^*(k) = \tilde{P}^*(k), \quad \text{and} \quad K^*(k) = \tilde{K}^*(k), \quad k=1,2,\ldots.$$

We note that (25) is a special case of (15) with $\hat{B}(k) = B, \hat{D}(k) = D$. We also note that if the system is time varying, the equivalence between the robust and the conventional Kalman filters still holds by replacing $(A,B,C,D)$ with $(A(k),B(k),C(k),D(k))$ in Theorem 3.2. The proof of Theorem 3.2 is given in Appendix A.

[5]Although the set of positive semidefinite matrices is only a partially ordered set, it is well known that the problem $\min_{K(k)} f(P^*(k), K(k))$ is well defined; see [8].

As a simple corollary of Theorem 3.2, we have the following conclusion: If the correlation matrix $X(0)$ is exactly known in advance, the polytope covering $X(k)$ is a fixed point at each $k$; see the recursion (5). The optimization problem (16) turns out to be equivalent to the optimization problem encountered in the conventional Kalman filtering problem, with $B$ replaced by $\hat{B}(k)$ and $D$ replaced by $\hat{D}(k)$ in (24). In this case, the MMSE estimation problem can be solved analytically, and the solution satisfies the Riccati difference equation in (24).

### *F. Convergence of the Robust Kalman Filter*

If the system state-space matrices are time invariant, it is well known that the recursion in the conventional Kalman filtering algorithm converges to a steady-state estimator under the conditions listed in Proposition A.2. If system (3) is mean square stable, the robust Kalman filtering algorithm has a similar convergence property.

*Theorem 3.3:* Consider the state estimation problem $(L = I)$ with system (3), with all the state space matrices being time invariant and $DD^T$ being positive definite.[5]

1) If there exists $Q = Q^T$ with $Q > 0$ such that

$$A^T Q A - Q + \sum_{j=1}^{m} \left(A_j^s\right)^T Q \left(A_j^s\right) < 0 \tag{26}$$

then the system (3) is mean square stable, and $\lim_{k\to\infty} X(k) = X$, where $X$ is unique and independent of the initial condition $X(0) \in \mathbf{Co}\{X_1(0), \ldots, X_p(0)\}$. Moreover, if the steady-state matrices $(A, \bar{B}, C, \bar{D})$, where

$$\begin{aligned} \bar{B} &= [A_1^s X \quad \ldots \quad A_m^s X \quad B \quad B_1^s \quad \ldots, B_m^s] \\ \bar{D} &= [C_1^s X \quad \ldots \quad C_m^s X \quad D \quad D_1^s \quad \ldots \quad D_m^s] \end{aligned} \tag{27}$$

satisfy the following condition:

2) $(A - \bar{B}\bar{D}^T(\bar{D}\bar{D}^T)^{-1}C, \bar{B}(I - \bar{D}^T(\bar{D}\bar{D}^T)^{-1}\bar{D}))$ is stabilizable, or

$$\begin{bmatrix} A - e^{j\omega}I & \bar{B} \\ C & \bar{D} \end{bmatrix} \text{ has full row rank for all } \omega \in [0, 2\pi)$$

then the robust Kalman filter converges to a steady-state LTI estimator, i.e., for any $P(0) \geq 0$, we have $\lim_{k\to\infty} P^*(k) = \bar{P}^* \geq 0, \lim_{k\to\infty} K^*(k) = \bar{K}^*, \lim_{k\to\infty} F^*(k) = \bar{F}^*$ and $(A + \bar{K}^* C)$ is stable, where $\{P^*(k), K^*(k), F^*(k)\}$ is the sequence consisting of the solutions to the robust Kalman filter.

The proof of Theorem 3.3 is given in Appendix A.

## IV. Numerical Example: Equalizer for Communication Channels

We present an application of the robust Kalman filtering techniques proposed in this paper toward the design of equalizers for a communication channel. Consider the following system:

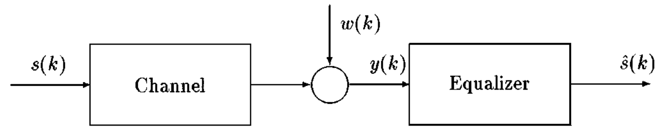



Fig. 2. Communication channel with an equalizer.

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} &= \begin{bmatrix} 0.9 & 0.5 \\ 0 & 0.9 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} s(k) \\ y(k) &= [1+\zeta(k) \quad 1+\zeta(k)] \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \\ &\quad + (5+\zeta(k))s(k) + w(k). \end{aligned} \tag{28}$$

$s$ is a signal that has a power of 0 dB and is transmitted through the channel. $w$ is a white noise that corrupts the received signal $y$, with a power of $-$ 10 dB. The channel model (28) is affected by time-varying uncertainties $\zeta$ that are a combination of both deterministic and stochastic parametric uncertainties (see for example, [4]). The initial conditions of the state vectors $x_1(0)$ and $x_2(0)$ are random variables and satisfy the second moment conditions

$$\begin{aligned} E[x_1(0)x_1(0)] \leq 1, \quad E[x_2(0)x_2(0)] \leq 1 \\ E[x_1(0)x_2(0)] = 0, \quad E[x_2(0)x_1(0)] = 0. \end{aligned}$$

For the channel (28), we design an equalizer to estimate the input signal $s(k)$ (Fig. 2).

We first add one more state variable in (28) so that the new model of the channel is now

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} &= \begin{bmatrix} 0.9 & 0.5 & 1 \\ 0 & 0.9 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} s(k+1) \\ y(k) &= [1+\zeta(k) \quad 1+\zeta(k) \quad 5+\zeta(k)] \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + w(k) \\ z(k) &= [0 \quad 0 \quad 1] \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}. \end{aligned} \tag{29}$$

Assuming a zero mean white noise model for the input signal $s$, we design an equalizer using the robust Kalman filtering techniques developed in this paper. The first case considered is when the channel is on its nominal condition and does not have any uncertainty, i.e., $\zeta(k) = 0$; this is an ideal channel, and the corresponding equalizer is a conventional Kalman filter. In Fig. 3, we compare a Kalman filter initialized optimally by solving (22), with Kalman filters initialized using an *ad hoc* scheme [2] ($P(0) = \eta I$, where $\eta = 5, 10$, and 20). The mean square error (MSE) estimates are obtained by averaging 200 runs. Under the condition of $\zeta(k) = 0$ in channel (29), it can be easily checked that the Kalman filter for this system converges (see the convergence condition in Proposition A.2). Therefore, after ten steps,

[5]This condition can be relaxed to $DD^T + \sum_{i=1}^m (C_i^s X (C_i^s)^T + D_i^s (D_i^s)^T) > 0$ where $\lim_{k\to\infty} X(k) = X$.

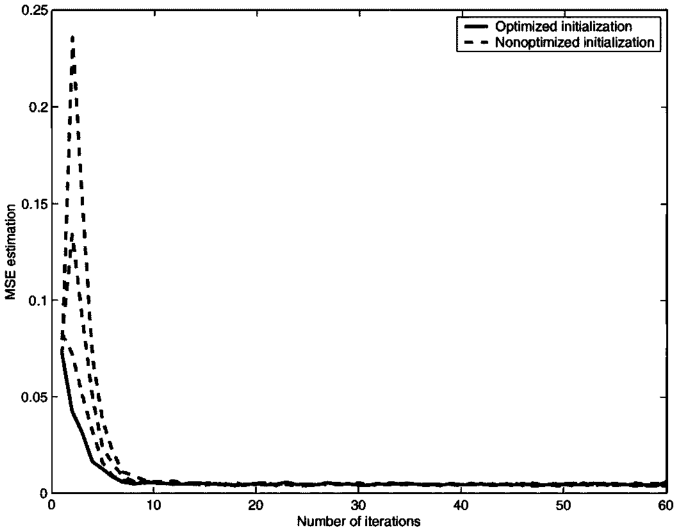



Fig. 3. Illustration of the performance improvement of the conventional Kalman filter with optimal initialization.

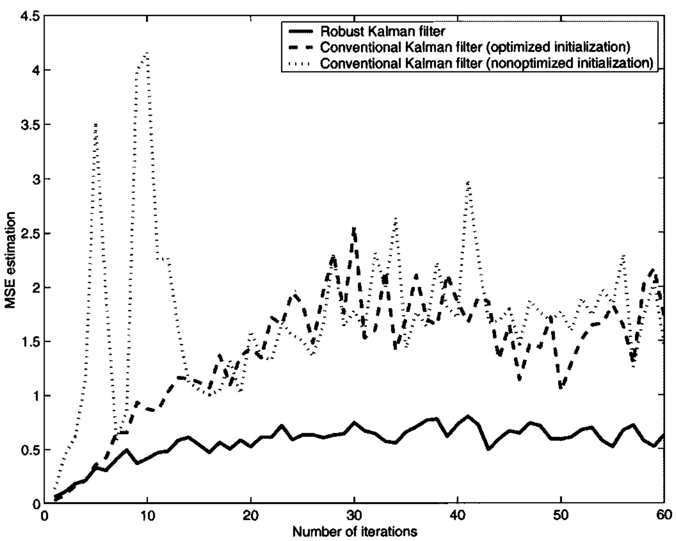



Fig. 4. Performance comparison between the robust and conventional Kalman filters.

we can see that the MSE of Kalman filters with different initializations are close to each other and converge to a constant. However for the transient performance (less than 10 steps), the *ad hoc* initialized Kalman filter is significantly worse than that of the optimally initialized Kalman filter. This example indicates that the optimal initialization by solving (22) can be superior to an *ad hoc* initialization and can improve the transient performance of conventional Kalman filters.

In the second example, we compare the performance of the conventional Kalman filtering algorithm with the robust Kalman filtering algorithm in the presence of uncertainties. Specifically, the time-varying uncertainties satisfy $\zeta(k) = 0.1\zeta_d(k) + 0.5\zeta_s(k)$, where $\zeta_d(k)$ is deterministic, and $|\zeta_d(k)| < 1$ for all $k$ and can be measured in real time; $\zeta_s$ is a zero-mean white noise process with a power of 0 dB and is independent of $w$ and $s$. Fig. 4 shows a comparison of the experimentally obtained mean square error values, averaged over 200 runs, obtained with the conventional Kalman filters and the robust Kalman filter. For a fair comparison, we include the simulation results of both the conventional Kalman filter initialized using an *ad hoc* scheme and the conventional Kalman filter initialized optimally by solving (22). For the nonoptimally initialized conventional Kalman filter, the performance is significantly worse than that of the other two filtering algorithms, which are initialized optimally. For the optimally initialized conventional Kalman filter, since it has a similar initialization as the robust Kalman filter, its transient performance at the beginning of the recursion (less than ten steps) is similar to that of the robust Kalman filter. However, since the conventional Kalman filter does not consider the uncertainties in system matrices, its performance degrades significantly thereafter. With the robust Kalman filtering algorithm, an upper bound of the mean square estimation error over all possible uncertainties is minimized recursively, and therefore, the performance is much improved. It can be seen from Fig. 4 that after ten steps of recursions, the robust Kalman filter yields a much lower mean-squared error than that with the conventional Kalman filters.

## V. Conclusion

We have developed a robust Kalman filtering algorithm for linear time-varying systems with stochastic parametric uncertainties. We have shown that for systems without uncertainties, the robust Kalman filter reduces to the conventional Kalman filter; for systems with stochastic parametric uncertainties, it offers significant improvement in performance. If the system is mean square stable and the state-space matrices are time invariant, the robust Kalman filter converges to a steady-state estimator. Our filtering algorithm is formulated as a convex optimization problem with linear matrix inequality constraints and can be implemented numerically efficiently. We have established that the techniques presented in the paper can be used to optimally initialize other recursive algorithms, including the conventional Kalman filtering algorithm, with an attendant improvement in transient performance. Finally, we have shown via a numerical example that the techniques developed in this paper can be applied toward the design of equalizers for communication channels with much improved results.

## Appendix A

*Lemma A.1 ([1, p. 131]):* The system

$$x(k+1) = A_\Delta(k)x(k) \tag{30}$$

where $x(k) \in \mathbf{R}^n$

$$A_\Delta(k) = A + \sum_{i=1}^{m} A_i^s \zeta_i^s(k)$$

and $\zeta_i^s, i = 1, \ldots, m$ are zero-mean random processes with $E[\zeta_i^s(k)\zeta_j^s(l)] = \delta(i-j)\delta(k-l)$, is mean square stable if and only if there exists a matrix $Q = Q^T \in \mathbf{R}^{n\times n}$ with $Q > 0$ such that

$$A^TQA - Q + \sum_{i=1}^{m} (A_i^s)^T Q (A_i^s) < 0. \qquad (31)$$

Moreover, the quadratic Lyapunov function $V(x(k)) = E[x(k)^TQx(k)]$ is a monotonically decreasing function of $k$, with $\lim_{k\to\infty} V(x(k)) = 0$.

Note that the correlation of the state in (30) satisfies the recursion

$$X(k+1) = AX(k)A^T + \sum_{j=1}^{m} A_j^s X(k) \left(A_j^s\right)^T$$

where $X(k) = E[x(k)x(k)^T]$, and $X(0)$ is the initial condition. Lemma A.1 implies that (31) is sufficient and necessary to have $\lim_{k\to\infty} X(k) = 0$.

The following result about the convergence of the conventional Kalman filtering algorithm is well known. More details about Proposition A.2 can be found in the literature; see, for example [8].

*Proposition A.2:* Let $A \in \mathbf{R}^{n\times n}, B \in \mathbf{R}^{n\times n_w}, C \in \mathbf{R}^{n_y\times n}$, and $D \in \mathbf{R}^{n_y\times n_w}$. Suppose $DD^T > 0$. If one of the following conditions holds.

1) $(A, C)$ is detectable, and $(A - BD^T(DD^T)^{-1}C, B(I - D^T(DD^T)^{-1}D))$ is stabilizable;
2) $(A, C)$ is detectable, and

$$\begin{bmatrix} A - e^{j\omega}I & B \\ C & D \end{bmatrix}$$

has full row rank for all $\omega \in [0, 2\pi)$;

then for any initial condition $P(0) \geq 0$, the sequence of solutions $\{P^*(k), K^*(k)\}$ to the recursion (24) converges to a unique steady-state solution

$$\lim_{k\to\infty} P^*(k) = P^* \quad \text{and} \quad \lim_{k\to\infty} K^*(k) = K^*.$$

Furthermore, $P^* \geq 0, K^* = -(AP^*C^T + BD^T)(CP^*C^T + DD^T)^{-1}$, and $(A + K^*C)$ is stable.

*Proposition A.3:* Let $\{X(k)\}$ be the sequence consisting of the solutions to the recursions

$$X(k+1) = AX(k)A^T + BB^T + \sum_{j=1}^{m} \left(A_j^s X(k) \left(A_j^s\right)^T + B_j^s \left(B_j^s\right)^T\right), \quad k = 0, 1, \ldots \qquad (32)$$

with initial condition $X(0) \geq 0$, where $A, B, A_j^s, B_j^s$ are constant matrices. The sequence $\{X(k)\}$ is convergent, and the limit is independent of $X(0)$ if and only if there exists some matrix $Q > 0$ such that

$$A^TQA - Q + \sum_{j=1}^{m} \left(A_j^s\right)^T Q \left(A_j^s\right) < 0 \qquad (33)$$

i.e., the system is mean square stable.

*Proof:* We first show that the sequence $\{X(k)\}$ obtained from (32) is convergent, and the limit is independent of $X(0)$ if and only if the sequence $\{\hat{X}(k)\}$ obtained from

$$\hat{X}(k+1) = A\hat{X}(k)A^T + \sum_{j=1}^{m} A_j^s \hat{X}(k) \left(A_j^s\right)^T \qquad (34)$$

is convergent, and $\lim_{k\to\infty} \hat{X}(k) = 0$.

Suppose the limit of the sequence $\{\hat{X}(k)\}$ exists and that $\lim_{k\to\infty} \hat{X}(k) = 0$. For any $X(0)$, define $\hat{X}(k) = X(k+1) - X(k)$. $\hat{X}(k)$ satisfies the recursion (34) and $\lim_{k\to\infty} \hat{X}(k) = 0$. Then, it can be shown that $\{X(k)\}$ is convergent.

To show that the limit of $\{X(k)\}$ is unique, suppose that $\{X_1(k)\}$ and $\{X_2(k)\}$ are two sequences of solutions to (32) corresponding to two different initial conditions $X_1(0)$ and $X_2(0)$, with $\lim_{k\to\infty} X_1(k) = X_1$ and $\lim_{k\to\infty} X_2(k) = X_2$. Then, it is easily verified that with $\hat{X}(k) = X_1(k) - X_2(k)$, the sequence $\{\hat{X}(k)\}$ consists of solutions of the recursive equation (34); then, $\lim_{k\to\infty} \hat{X}(k) = 0$ implies $X_1 = X_2$.

Now, suppose the limit of the sequence $\{X(k)\}$ exists and is independent of $X(0)$. We show that $\lim_{k\to\infty} \hat{X}(k) = 0$.

For any $\hat{X}(0)$, let $X_1(0) = \hat{X}(0), X_2(0) = 0$ and $\{X_i(k)\}, i = 1, 2$ be the corresponding sequences from recursion (32). By recursion (34), we have $\hat{X}(k) = X_1(k) - X_2(k)$. Since $\lim_{k\to\infty} X_1(k) = \lim_{k\to\infty} X_2(k)$, we get $\lim_{k\to\infty} \hat{X}(k) = 0$.

Finally by Lemma A.1, (33) is necessary and sufficient for $\{\hat{X}(k)\}$ to converge to zero. This completes the proof. □

*Proof of Theorem 3.2:* Starting with the sequence $\{P^*(k), K^*(k)\}$ consisting of the solutions to the optimization problem (24), it is clear that

$$\tilde{P}^*(k) = P^*(k), \quad \text{and} \quad \tilde{K}^*(k) = K^*(k), \quad k = 1, 2, \ldots$$

must be (possibly nonunique) optimizers for (25) [this follows immediately from (25d)]. Thus, it remains to be shown that $\tilde{P}^*(k) = P^*(k)$ and $\tilde{K}^*(k) = K^*(k)$ are the only candidates as optimizers for (25). In other words, it suffices to show that for a given $\tilde{P}^*(k-1)$, the optimal solutions $\tilde{P}^*(k)$ and $\tilde{K}^*(k)$ in (25) are unique. However, note that given $\tilde{P}^*(k-1)$ and $\tilde{K}^*(k-1)$, $\tilde{P}^*(k)$ is given simply as $f(\tilde{P}^*(k-1), \tilde{K}^*(k-1))$, and thus, we only need to show that given $\tilde{P}^*(k-1)$, $\tilde{K}^*(k-1)$, which solves (25), is unique.

Now, let $\{\tilde{P}^*(k), \tilde{K}^*(k), \tilde{F}^*(k)\}$ be the sequence consisting of the solutions to the optimization problem (25). Assuming an optimal value for $\tilde{P}(k-1)$, we consider the optimization problem (25). For convenience, we introduce new notation for $M(k)$:

$$g(K(k-1), F(k)) = (I + F(k)C)f(P(k-1), K(k-1)) \times (I + F(k)C)^T + F(k)DD^TF(k)^T.$$

We also denote

$$F_{\text{opt}}(K(k-1)) = \underset{F(k)}{\operatorname{argmin}} \mathbf{Tr}(g(K(k-1), F(k))).$$

Note that $\tilde{F}^*(k) = F_{\text{opt}}(\tilde{K}^*(k-1))$.

Now, for any $P(k) = f(P(k-1), K(k-1)) \geq 0$ since $(CP(k)C^T + DD^T) > 0$, it can be easily verified that if $F(k) \neq F_{\text{opt}}(K(k-1))$, then

$$g(K(k-1), F(k)) \gneqq g(K(k-1), F_{\text{opt}}(K(k-1))).$$

Therefore, $\mathbf{Tr}(g(K(k-1), F(k))) > \mathbf{Tr}(g(K(k-1), F_{\text{opt}}(K(k-1))))$. Thus, $F_{\text{opt}}(K(k-1))$ is uniquely determined from $K(k-1)$.

To show that $\tilde{K}^*(k-1)$ is unique, it suffices to show that for any $K(k-1) \neq \tilde{K}^*(k-1), \mathbf{Tr}(g(K(k-1), F_{\text{opt}}(K(k-1)))) > \mathbf{Tr}(g(\tilde{K}^*(k-1), F_{\text{opt}}(\tilde{K}^*(k-1))))$. Let $P_1(k) = f(\tilde{P}(k-1), K(k-1))$ and $P_2(k) = f(\tilde{P}(k-1), \tilde{K}^*(k-1))$. Then, we have $P_1(k) \gneqq P_2(k)$. We now consider two cases.

1) If $F_{\text{opt}}(K(k-1))) \neq F_{\text{opt}}(\tilde{K}^*(k-1))$, then

$$g(K(k-1), F_{\text{opt}}(K(k-1))) \gneqq g(\tilde{K}^*(k-1), F_{\text{opt}}(\tilde{K}^*(k-1))).$$

Therefore, $\mathbf{Tr}(g(K(k-1), F_{\text{opt}}(K(k-1)))) > \mathbf{Tr}(g(\tilde{K}^*(k-1), F_{\text{opt}}(\tilde{K}^*(k-1))))$.

2) If $F_{\text{opt}}(K(k-1)) = F_{\text{opt}}(\tilde{K}^*(k-1))$, then

$$P_1(k)C^T(CP_1(k)C^T + DD^T)^{-1} = P_2(k)C^T(CP_2(k)C^T + DD^T)^{-1}.$$

After simple manipulations, we get

$$DD^T(CP_1(k)C^T + DD^T)^{-1} = DD^T(CP_2(k)C^T + DD^T)^{-1}.$$

Since $DD^T > 0$ and $P_1(K) \geq P_2(K)$, we have $C(P_1(k) - P_2(k)) = 0$. By (25d), this implies that $\mathbf{Tr}(g(K(k-1), F_{\text{opt}}(K(k-1)))) > \mathbf{Tr}(g(\tilde{K}^*(k-1), F_{\text{opt}}(\tilde{K}^*(k-1))))$.

This completes the proof. □

Now, we are ready to prove Theorem 3.3, wherein the robust Kalman filtering algorithm is convergent.

*Proof of Theorem 3.3:* With all the state space matrices in system (3) being time invariant, we have $A(k) = A, B(k) = B$ and $\hat{B}(k), \hat{D}(k)$, defined at the bottom of the page. The correlation of the state $X(k) = E[x(k)x(k)^T]$ satisfies the recursion

$$X(k+1) = AX(k)A^T + BB^T + \sum_{j=1}^{m} \left( A_j^s X(k) \left(A_j^s\right)^T + B_j^s \left(B_j^s\right)^T \right), \quad k = 0, 1, \ldots.$$

The mean square stability of system (3) follows directly from Lemma A.1.

By Proposition A.3, we have $\lim_{k\to\infty} X(k) = X$, where $X$ is unique and independent of the initial condition $X(0) \in \mathbf{Co}\{X_1(0), \ldots, X_p(0)\}$.

With $\lim_{k\to\infty} \hat{B}(k) = \bar{B}$ and $\lim_{k\to\infty} \hat{D}(k) = \bar{D}$, where $\bar{B}$ and $\bar{D}$ are defined in (27), we now define another set of recursions (35a)–(35d), shown at the bottom of the page.

Let $\bar{K}^*(k)$ and $\bar{F}^*(k)$ denote the optimal Kalman gains of (35). Let $K^*(k)$ and $F^*(k)$ denote the optimal Kalman gains of the robust Kalman filter (15). From Theorem 3.2, (35) has the same unique optimal solution of $\bar{P}^*(k)$ and $\bar{K}^*(k)$ as that of (24), with $B$ replaced by $\bar{B}$ and $D$ replaced by $\bar{D}$. From Proposition A.2, we have $\lim_{k\to\infty} \bar{K}^*(k) = \bar{K}^*$, and $(A + \bar{K}^* C)$ is stable. Thus

$$\left\| \sum_{j=1}^{i} \prod_{k=j}^{i} (A + \bar{K}^*(k)C)(A + \bar{K}^*(k)C)^T \right\| \leq T_1(\bar{P}(0))$$

where $T_1(\bar{P}(0))$ is a uniform bound over $i > 0$ and depends only on $\bar{P}(0)$.

Let $P^*(k) = \tilde{f}(P^*(k-1), K^*(k-1))$ denote the optimal solution to (15). From (10) and (35b), it follows that for any small number $\epsilon$, there exists $N$ such that whenever $k \geq N$, we have

$$\|\tilde{f}(P(k), K^*(k)) - \bar{f}(P(k), \hat{K}^*(k))\| \leq T_2(P(k))\epsilon$$

where $\hat{K}^*(k) = \text{argmin}_{K(k)} \bar{f}(P(k), K(k)), T_2(P(k))$ is a constant that depends on $P(k)$. $T_2(P(k))$ is finite if $P(k)$ is bounded.

By setting $P(N) = \bar{P}(N)$, we get $\|P^*(N+1) - \bar{P}^*(N+1)\| \leq T_2(P(N))\epsilon$, where $P^*(N+1)$ and $\bar{P}^*(N+1)$ are

---

$$\hat{B}(k) = \left[ A_1^s X(k)^{1/2} \quad \cdots \quad A_m^s X(k)^{1/2} \quad B \quad B_1^s \quad \cdots \quad B_m^s \right]$$
$$\hat{D}(k) = \left[ C_1^s X(k)^{1/2} \quad \cdots \quad C_m^s X(k)^{1/2} \quad D \quad D_1^s \quad \cdots \quad D_m^s \right].$$

---

$$\min_{\bar{P}(k), \bar{M}(k), \bar{K}(k-1), \bar{F}(k)} \mathbf{Tr} \bar{M}(k) \tag{35a}$$

$$\bar{f}(\bar{P}(k-1), \bar{K}(k-1)) \triangleq (A + \bar{K}(k-1)C)\bar{P}(k-1)(A + \bar{K}(k-1)C)^T + (\bar{K}(k-1)\bar{D} + \bar{B})(\bar{K}(k-1)\bar{D} + \bar{B})^T \tag{35b}$$

$$\bar{P}(k) = \bar{f}(\bar{P}(k-1), \bar{K}(k-1)) \tag{35c}$$

$$\bar{M}(k) = (I + \bar{F}(k)C)\bar{P}(k)(I + \bar{F}(k)C)^T + \bar{F}(k)\bar{D}\bar{D}^T\bar{F}(k)^T. \tag{35d}$$

optimal solutions to recursions (15) and (35) respectively. Next, we have

$$\begin{aligned}&\|P^*(N+2)-\bar{P}^*(N+2)\|\\&\quad=\|\tilde{f}(P^*(N+1),K^*(N+1))\\&\qquad-\bar{f}(\bar{P}^*(N+1),\bar{K}^*(N+1))\|\\&\quad\le(T_2(P(N))\|(A+\bar{K}^*(N+1)C)\\&\qquad\times(A+\bar{K}^*(N+1)C)^T\|+T_2(P^*(N+1)))\epsilon.\end{aligned}$$

Recursively, for $i=2,3,\ldots$, we get

$$\begin{aligned}&\|P^*(N+i)-\bar{P}^*(N+i)\|\\&\quad=\|\tilde{f}(P^*(N+i-1),K^*(N+i-1))\\&\qquad-\bar{f}(\bar{P}^*(N+i-1),\bar{K}^*(N+i-1))\|\\&\quad\le\left(T_2(P^*(N+i-1))+\left\|\sum_{j=1}^{i-1}T_2(P^*(N+j-1))\right.\right.\\&\qquad\left.\left.\times\prod_{k=N+j}^{N+i-1}(A+\bar{K}^*(k)C)(A+\bar{K}^*(k)C)^T\right\|\right)\epsilon\\&\quad\le(T_2(P^*(N+i-1))\\&\qquad+\max_{j=0,\ldots,i-2}T_2(P^*(N+j))T_1(P(N)))\epsilon.\end{aligned}$$

Using standard arguments, we obtain that $\{P^*(k)\}$ is also bounded, with the bound depending only on $P(0)$. Therefore, there exist finite constants $T_3(P(0))$ and $T_4(P(0))$ that depend on $P(0)$ such that for any $k\ge 0$, $T_1(P^*(k))\le T_3(P(0))$ and $T_2(P^*(k))\le T_4(P(0))$. Thus, we conclude that $\|P^*(N+i)-\bar{P}^*(N+i)\|$ is bounded and that

$$\|P^*(N+i)-\bar{P}^*(N+i)\|\le[T_4(P(0))T_3(P(0))+T_4(P(0))]\epsilon.$$

Since $\epsilon$ can be made arbitrarily small by choosing $N$ to be large enough, we obtain

$$\lim_{k\to\infty}P^*(k)=\lim_{k\to\infty}\bar{P}^*(k)=\bar{P}^*.$$

Similar arguments establish that $\lim_{k\to\infty}K^*(k)=\lim_{k\to\infty}\bar{K}^*(k)=\bar{K}^*$ and that $\lim_{k\to\infty}F^*(k)=\lim_{k\to\infty}\bar{F}^*(k)=\bar{F}^*$. □

## REFERENCES


[1] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory, Vol. 15 of Studies in Applied Mathematics*. Philadelphia, PA: SIAM, June 1994.

[2] J. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[3] A. Gelb, J. F. Kasper, R. A. Nash, C. F. Price, and A. A. Sutherland, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.

[4] B. Mulgrew and C. F. N. Cowan, *Adaptive Filters and Equalizers*. Boston, MA: Kluwer, 1988.

[5] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[6] S. Marcos, "A network of adaptive Kalman filters for data channel equalization," *IEEE Trans. Signal Processing*, vol. 48, pp. 2620–2627, Sept. 2000.

[7] J. M. Omidi, S. Pasupathy, and P. G. Gulak, "Joint data and Kalman estimation of fading channel using a generalized Viterbi algorithm," *Proc. IEEE Int. Conf. Commun.*, vol. 2, 1996.

[8] B. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[9] G. Goodwin and K. Sin, *Adaptive Filtering, Prediction and Control*. Upper Saddle River, NJ: Prentice-Hall, 1984.

[10] Y. Chen and B. Chen, "Minimax robust deconvolution filters under stochastic parametric and noise uncertainties," *IEEE Trans. Signal Processing*, vol. 42, pp. 32–45, Jan. 1994.

[11] *The Digital Signal Processing Handbook*, V. K. Madisetti and D. B. Williams, Eds., CRC, Boca Raton, FL, 1998, pp. 20.1–20.19.

[12] R. S. Mangoubi, *Robust Estimation and Failure Detection: A Concise Treatment*, London, U.K.: Springer, 1998.

[13] S. Haykin, *Modern Filters*. New York: MacMillan, 1989.

[14] H. Li and M. Fu, "A linear matrix inequality approach to robust $H_\infty$ filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 2338–2350, Sept. 1997.

[15] L. Xie, Y. Soh, and C. de Souza, "Robust Kalman filtering for uncertain discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1310–1314, June 1994.

[16] P. Khargonekar, M. Rotea, and E. Baeyens, "Mixed $H_2/H_\infty$ filtering," *Int. J. Robust Nonlinear Contr.*, vol. 6, pp. 313–330, 1996.

[17] F. Wang and V. Balakrishnan, "Robust estimators for systems with both deterministic and stochastic uncertainties," *Proc. IEEE Conf. Decision Contr.*, pp. 1946–1951, Dec. 1999.

[18] A. Erdogan, B. Hassibi, and T. Kailath, "On linear $H^\infty$ equalization of communication channels," *IEEE Trans. Signal Processing*, vol. 48, pp. 3227–3231, Nov. 2000.

[19] L. E. Ghaoui, "State-feedback control of systems with multiplicative noise via linear matrix inequalities," *Syst. Contr. Lett.*, vol. 24, pp. 223–228, 1995.

[20] D. Hinrichsen and A. J. Pritchard, "Stochastic $H^\infty$," *SIAM J. Contr. Optim.*, vol. 36, no. 5, pp. 1604–1638, 1998.

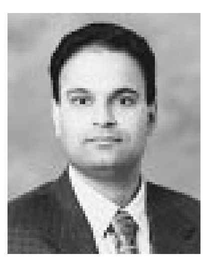
[21] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.

[22] P. Gahinet and A. Nemirovskii, *LMI Control Toolbox: The LMI Lab*. Natick, MA: MathWorks, Inc., 1995.


**Fan Wang** (S'96–M'00) received the B.S. and the M.S. degrees from the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, in 1992 and 1995 respectively, and the Ph.D. degree from the school of electrical and computer engineering, Purdue University, West Lafayette, IN, in 1999.

Since 2000, he has been with the Advanced Radio Technology Group of GTSS, Motorola, Inc., Arlington Heights, IL. His current research interests are in robust control, signal processing, and wireless communications, with an emphasis on the application of convex optimization to problems from these areas.

Dr. Wang is currently a member of the IEEE Control Systems Society Conference Editorial Board. He was a finalist for the 1999 American Control Conference Best Student Paper Award.



**Venkataramanan Balakrishnan** (M'94) received the B.Tech. degree in electronics and communications from the Indian Istitutue of Technology, Madras, in 1985. He then received the M.S. degree in statistics and the Ph.D. degree in electrical engineering in 1992 from Stanford University, Stanford, CA.

Dr. Balakrishnan is an Associate Professor of electrical and computer engineering at Purdue University, West Lafayette, IN. He joined Purdue in 1994 as an Assistant Professor, following post-doctoral stints at Stanford University, California Institute of Technology (Caltech), Pasadena, and the Institute for Systems Research, University of Maryland, College Park. He is a co-author of the monograph *Linear Matrix Inequalities in System and Control Theory* (Philadelphia, PA: SIAM, 1994). His primary research interests are in applying convex optimization to problems in systems and control.

Dr. Balakrishnan received the President of India Gold Medal from the Indian Institute of Technology, Madras, in 1985, the Young Investigator Award from the Office of Naval Research in 1997, the Ruth and Joel Spira Outstanding Teacher Award from the School of Electrical and Computer Engineering at Purdue University in 1998, and the Honeywell Award for excellence in teaching from the School of Electrical and Computer Engineering at Purdue University in 2001.