Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

State Estimation with Probability Constraints

Mario Rotea and Carlos Lana
Purdue University
West Lafayette, IN 47907, USA

Abstract— **This paper considers a state estimation problem for a discrete-time linear system driven by a Gaussian random process. The second order statistics of the input process and state initial condition are uncertain. However, the probability that the state and input satisfy linear constraints during the estimation interval is known. A minimax estimation problem is formulated to determine an estimator that minimizes the worst-case mean square error criterion, over the uncertain second order statistics, subject to the probability constraints. It is shown that a solution to this constrained state estimation problem is given by a Kalman filter for appropriately chosen input and initial condition models. These models are obtained from a finite dimensional convex optimization problem. The application of this estimator to an aircraft tracking problem quantifies the improvement in estimation accuracy obtained from the inclusion of probability constraints in the minimax formulation.**

I. INTRODUCTION

The Kalman filter has had an enormous technological impact since it was first presented in the early 60's [1][2]. The strengths of the Kalman filter, and other closely related estimators, include the efficiency with which it can be implemented and its optimality with respect to the mean square estimation error. The underlying modeling assumptions in the Kalman filter are a linear system generating the data and known statistics for the initial condition, the process noise, and the measurement noise.

In many practical problems the system or the exogenous input statistics are not known exactly. Often, this uncertainty is neglected and the estimator is designed using nominal values. It is known that this approach may lead to poor estimation performance [3][4][5][6]. This motivates the search for estimators which are robust (less sensitive) to parameter variations and uncertainties. One of the available techniques for robust estimation is the well-known minimax approach. Given a specific performance metric (e.g., mean square error), the minimax approach yields an estimator that minimizes the worst-case performance index over all possible values of the uncertain parameters. An early account of this approach is in [3], while [5] and [6] describe recent work in this area. One of the critiques to the minimax approach is that it may be too conservative in certain problems [7].

In many practical applications the state satisfies known constraints. For example, in most problems one has an idea of upper and lower bounds on state variables. In theory, this prior knowledge should lead to improved estimation accuracy, which in turn may reduce the conservatism of the minimax approach. Unfortunately, incorporating state constraints requires a nonlinear framework for estimation for which there are no known solutions with the simplicity of the Kalman filter.

One approach to incorporate constraints is the use of a receding horizon approximation [8][9]. In this strategy, an optimization problem is solved over a fixed-size data window each time a new measurement becomes available. This optimization problem delivers a state estimate, or prediction, that satisfies the known constraints. The approach has been currently extended to problems with modeling uncertainty [10][11]. Even though this technique is gaining terrain in a number of applications, it requires an optimization problem to be solved with the arrival of each new measurement. From a theoretical point of view, the receding horizon approach is only an approximation to the original optimal estimation problem.

In this paper we show how to construct an estimator that takes advantage of prior knowledge of state and input constraints yet the estimator retains the simplicity of the Kalman filter. *The key is not to enforce hard constraints in the state and input but to replace them with constraints satisfied in probability.* This constraint softening allows the incorporation of prior knowledge into the estimation problem without increasing the on-line computation relative to the one of the conventional Kalman filter.

We consider a minimax state estimation problem for a discrete-time linear system. The uncertainty in the problem is in the covariance matrices of both the exogenous input (e.g., process and measurement noise) and the state initial condition. We also assume that, during the estimation interval, the state and the exogenous input satisfy linear constraints with certain minimal probability. We show that the causal linear estimator that minimizes the worst-case mean square estimation error over all possible values of the uncertainty, subject to the probability constraints, is a conventional Kalman filter whose exogenous input and initial condition covariance matrices can be obtained by solving a convex optimization problem off line. A simple aircraft tracking problem is given to quantify the merit of our approach. In this example we show that the inclusion of probability constraints into the minimax formulation leads to a nontrivial improvement in estimation accuracy.

Recently, related ideas have been proposed to handle linear constraints on the states in model predictive control [12]. We are grateful to the reviewer that pointed out this reference. Our ideas and results have been developed independently of

This work was supported in part by awards ECS-0080832 and CMS-0511913 from National Science Foundation.
Contact author, rotea@purdue.edu

the work in [12], which only gives a suboptimal solution to MPC problems with probability constraints.

Due to space limitations the proofs of the technical results are not included. The proofs may be obtained from the corresponding author. The notation used in this paper is fairly standard. Time-dependent quantities and their associated discrete-time sequences are denoted with the same symbol, with omission of the time index for the sequences. The Kronecker delta sequence is denoted by δ. The operators $E\{\cdot\}$ and $\mathrm{tr}\{\cdot\}$ denote expected value and trace, respectively. The operator $d_i\{\cdot\}$ denotes the i-th diagonal entry of a square matrix. The identity matrix is denoted by I. Given Hermitian matrices A and B, the inequality $A \geq B$ ($A > B$) means that $A - B$ is positive semidefinite (definite).

II. Problem formulation

Consider the time-varying discrete-time linear system defined by the following equations

$$
\begin{aligned}
x(k+1) &= A(k)x(k) + B(k)w(k), \quad x(0) = x_0 && \text{(1a)}\\
z(k) &= C_1(k)x(k) + D_1(k)w(k) && \text{(1b)}\\
y(k) &= C_2(k)x(k) + D_2(k)w(k) && \text{(1c)}
\end{aligned}
$$

where $x(k)$ is the state vector, $w(k)$ an exogenous input vector, and $z(k)$ and $y(k)$ the system outputs. Only $y(k)$ is measured for $0 \leq k \leq T$, where T is given. The exogenous input w is assumed to be a Gaussian random sequence with the following statistics:

$$
\begin{aligned}
&E\{w(k)\} = 0 && \text{(2a)}\\
&E\{w(k)w^*(j)\} = Q(k)\delta(k-j) && \text{(2b)}
\end{aligned}
$$

for $0 \leq k, j \leq T$. That is, we assume that w is Gaussian white noise with covariance matrix sequence Q. The initial condition x_0 is a Gaussian random vector characterized by

$$
\begin{aligned}
&E\{x_0\} = \bar{x}_0 && \text{(3a)}\\
&E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^*\} = X_0. && \text{(3b)}
\end{aligned}
$$

We also assume that $w(k)$ and x_0 are independent for all $k \in [0, T]$, which implies

$$
E\{w(k)x_0^*\} = 0. \tag{4}
$$

The remaining assumptions are the following:

A1) The system matrices in (1) are known for $k \in [0, T]$.

A2) The mean value $\bar{x}_0$ of the system initial condition is known.

A3) Sets $\mathcal{Q}$ and $\mathcal{X}_0$ are known such that the covariance matrix sequence Q and the covariance matrix X_0 satisfy $Q(k) \in \mathcal{Q}$, for $k \in [0, T]$, and $X_0 \in \mathcal{X}_0$.

A4) For every $k \in [0, T]$ and $Q(k) \in \mathcal{Q}$, we have

$$
\begin{aligned}
&B(k)Q(k)D_2^*(k) = 0 && \text{(5a)}\\
&D_2^*(k)Q(k)D_2(k) > 0. && \text{(5b)}
\end{aligned}
$$

From equations (1), (2), and (5), it follows that the effective process noise Bw and the effective measurement noise Dw are uncorrelated and that no part of the measured output y is noise free.

Given any $k \in [0, T]$, our objective is to estimate the unknown state vector $x(k)$ from the measurements $y(0)$, $y(1)$, …, $y(k)$. More specifically, we seek a linear causal operator $\mathcal{F}$ such that

$$
\hat{x}(k) = (\mathcal{F}(\bar{x}_0; y))(k) \tag{6}
$$

provides an estimate of $x(k)$ for $k \in [0, T]$. Notice that the mean $\bar{x}_0$ is the unbiased estimate of $x(0)$ prior to obtaining any measurement.

The performance of the estimator $\hat{x}$, or $\mathcal{F}$, is evaluated using the estimation error

$$
\epsilon(k) = x(k) - \hat{x}(k). \tag{7}
$$

In this paper we restrict our attention to linear causal operators $\mathcal{F}$ with the following properties:

$$
\begin{aligned}
&\mathcal{F} : \mathcal{R}^{n_x} \times \ell_2^{n_y}[0, T] \mapsto \ell_2^{n_x}[0, T] && \text{(8a)}\\
&\mathcal{F} \text{ is a bounded operator} && \text{(8b)}
\end{aligned}
$$

where n_x and n_y are the dimensions of the state vector $x(k)$ and the measured output $y(k)$, respectively. We shall refer to estimators $\hat{x}$ satisfying (6) and (8) as *admissible estimators*. We denote the set of admissible estimators by $\mathcal{A}$.

Given any admissible estimator $\mathcal{F}$ of the form (6), its performance can be quantified using the time-averaged mean square error criterion (MSE) defined as

$$
J(\mathcal{F}; Q, X_0) = \frac{1}{T+1} \sum_{k=0}^{T} E\{\epsilon^*(k)W(k)\epsilon(k)\} \tag{9}
$$

where W is a given positive definite matrix sequence. Notice that this performance measure is written as an explicit function of the operator $\mathcal{F}$, the matrix sequence Q, and the matrix X_0. In this paper, we are interested in the following minimax estimation problem

$$
\begin{aligned}
&J_{\mathrm{opt}} = \inf_{\mathcal{F}} \sup_{Q,\, X_0} J(\mathcal{F}; Q, X_0) && \text{(10a)}\\
&\quad \text{subject to}\\
&\quad \mathcal{F} \in \mathcal{A},\ Q(k) \in \mathcal{Q},\ X_0 \in \mathcal{X}_0, \text{ and} && \text{(10b)}\\
&\quad \mathrm{Prob}\{z(k) \leq h(k)\} \geq \gamma,\ \forall k \in [0, T] && \text{(10c)}
\end{aligned}
$$

where the inequalities in (10c) are elementwise for given vectors $h(k)$ and γ.

The practical application of this minimax problem formulation is as follows. If the probability constraints (10c) are ignored, then (10) is a standard minimax estimation problem, where one seeks an estimator that minimizes the worst-case MSE over the set of covariance matrix sequences $\mathcal{Q}$ and the set of covariance matrices $\mathcal{X}_0$, respectively. Assume now that more is known about the behavior of the states and the input in the form of linear constraints. For example, we may have high confidence that the state and the input satisfy certain constraints during the estimation interval. In theory, this additional knowledge should lead to improved estimation accuracy. The formulation in (10) achieves that by adding probability constraints that limit the choices of the covariance matrix sequence Q and the covariance matrix

X_0 in the calculation of the worst-case MSE. This is a very natural way of incorporating prior knowledge of constraints into an optimal estimation problem. *As it is shown in this paper, the minimax problem (10) is no more difficult than a Kalman filter problem. In fact, a solution $\mathcal{F}_{\mathrm{opt}}$ is exactly a conventional Kalman filter for suitable chosen covariance matrix X_0 and covariance sequence Q.*

III. Probability Constraints

The next lemma gives precise conditions on the system parameters that are necessary and sufficient for the satisfaction of probability constraints of the form (10c). The result can be used to precisely characterize the exogenous input covariance matrix sequence Q and the initial condition covariance matrix X_0 that are consistent with the probability constraints.

Lemma 1 (Representation of probability constraints): Consider the system defined by equations (1) through (4). Let $X(k)$ and $r(k)$ denote the unique solutions to

$$X(k+1) = A(k)X(k)A^*(k) + B(k)Q(k)B^*(k) \quad (11a)$$
$$r(k+1) = A(k)r(k) \quad (11b)$$

with initial conditions $X(0) = X_0$ and $r(0) = \bar{x}_0$, respectively. Let $c > \frac{1}{2}$ be given and assume that θ_c is the unique solution to

$$c = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta_c} e^{-\frac{1}{2}z^2} dz. \quad (12)$$

Then, given any scalar b and vectors a_1 and a_2, the probability constraint

$$\mathrm{Prob}\left\{a_1^* x(k) + a_2^* w(k) \le b\right\} \ge c \quad (13)$$

is satisfied if and only if

$$a_1^* X(k) a_1 + a_2^* Q(k) a_2 \le \left(\frac{b - a_1^* r(k)}{\theta_c}\right)^2 \quad (14a)$$
$$b - a_1^* r(k) \ge 0. \quad (14b)$$

IV. Main Results

The results in this section show that a solution to the minimax optimization problem (10) can be obtained in two steps:

Step 1. Compute a covariance matrix sequence Q_{opt} and a covariance matrix $X_{0,\mathrm{opt}}$ that maximize a suitably constructed linear cost function over a convex set. The precise formulation is given subsequently in Theorems 1 and 2.

Step 2. Compute a standard Kalman Filter for the system in equation (1), under assumptions (2) through (5), with exogenous input and initial condition models set to $Q = Q_{\mathrm{opt}}$ and $X_0 = X_{0,\mathrm{opt}}$.

We shall make the following additional assumptions on the data of the minimax problem (10):

A5) The entries γ_i of the probability bound vector γ introduced in (10c) are strictly greater than $1/2$.

A6) The mean state vector $r(k)$ introduced in (11b) satisfies the elementwise inequality constraints $C_1(k)r(k) \le h(k)$, for all $k \in [0, T]$.

A7) The sets $\mathcal{Q}$ and $\mathcal{X}_0$ are compact and convex.

Notice that no restriction is imposed on the system by assuming that the mean state $r(k)$ satisfies the constraints $C_1(k)r(k) \le h(k)$. If any of these constraints is not met then the associated probability constraint is not feasible and the minimax problem makes no sense.

To state our results we shall make use of the matrix sequence H, with elements defined by

$$H(k) = \left(h(k) - C_1(k)r(k)\right)\left(h(k) - C_1(k)r(k)\right)^*. \quad (15)$$

Theorem 1 (Optimal worst-case MSE): Consider the optimal minimax estimation problem (10), under assumptions A1) through A7). Let θ_{γ_i} denote the solution to the integral equation (12) with $c = \gamma_i$. The optimal worst-case MSE J_{opt} can be obtained by solving the following optimization problem

$$J_{\mathrm{opt}} = \sup_{Q, X_0} \frac{1}{T+1} \sum_{k=0}^{T} \mathrm{tr}\{Y(k)W(k)\} \quad (16a)$$

subject to the following constraints $\forall k \in [0, T]$:

$$Q(k) \in \mathcal{Q},\ X_0 \in \mathcal{X}_0 \quad (16b)$$
$$Y(k) = Y^-(k) - K(k)C_2(k)Y^-(k) \quad (16c)$$
$$K(k) = Y^-(k)C_2^*(k)M^{-1}(k) \quad (16d)$$
$$M(k) = C_2(k)Y^-(k)C_2^*(k) + D_2(k)Q(k)D_2^*(k) \quad (16e)$$
$$Y^-(0) = X(0) = X_0 \quad (16f)$$
$$Y^-(k+1) = A(k)Y(k)A^*(k) + B(k)Q(k)B^*(k) \quad (16g)$$
$$X(k+1) = A(k)X(k)A^*(k) + B(k)Q(k)B^*(k) \quad (16h)$$
$$d_i\left\{C_1(k)X(k)C_1^*(k) + D_1(k)Q(k)D_1^*(k) - \theta_{\gamma_i}^{-2}H(k)\right\} \le 0, \quad \forall i = 1, \ldots, n_z \quad (16i)$$

where H is defined in (15) and n_z is the number of probability constraints in (10c).

The optimization problem in (16) has a linear cost function and constraints given by difference matrix equations. Equations (16c) through (16g) are the well-known equations corresponding to the Kalman filter for the system. The constraint imposed by these equations is, in general, nonlinear in Q and X_0. Theorem 2 below shows that this constraint may be replaced with a linear matrix inequality, which, given the convexity of $\mathcal{Q}$ and $\mathcal{X}_0$, leads to an equivalent convex program for solving (16). Prior to stating this result, we give a formula for an estimator that solves the minimax problem (10).

Corollary 1 (Optimal minimax estimator $\mathcal{F}_{\mathrm{opt}}$): If Q_{opt} and $X_{0,\mathrm{opt}}$ denote a solution to (16), then an admissible estimator $\mathcal{F}_{\mathrm{opt}}$ that solves the minimax problem (10) is of the form

$$\hat{x}(k) = x^-(k) + K_{\mathrm{opt}}(k)(y(k) - C_2(k)x^-(k)) \quad (17a)$$
$$x^-(k+1) = A(k)\hat{x}(k) \quad (17b)$$

where $x^-(0) = \bar{x}_0$ and the gain sequence K_{opt} is from (16) with $Q = Q_{\mathrm{opt}}$ and $X_0 = X_{0,\mathrm{opt}}$.

Notice that the estimator given in this corollary is a Kalman filter with covariance matrices, for the exogenous input and initial condition, set to the ones that yield the worst-case performance J_{opt}. These matrices depend only on the a priori information of the system and can be computed from Theorem 2 before running the estimator. Therefore, the on-line implementation of this estimator reduces to the implementation of a standard Kalman filter.

Theorem 2 (Equivalent convex problem): Under the assumptions of Theorem 1, the optimal worst-case MSE J_{opt}, and a solution Q_{opt} and $X_{0,\mathrm{opt}}$ to (16), can be obtained by solving the following convex problem:

$$J_{\mathrm{opt}} = \sup_{Q,X_0,Z} \frac{1}{T+1}\sum_{k=0}^{T} \mathrm{tr}\{Z(k)W(k)\} \quad (18a)$$

subject to the following constraints $\forall k \in [0,T]$:

$$Q(k) \in \mathcal{Q},\ X_0 \in \mathcal{X}_0,\ Z(k) > 0 \quad (18b)$$

$$\begin{bmatrix} Z(0) & 0 \\ 0 & -D_2(0)Q(0)D_2^*(0) \end{bmatrix} \leq \begin{bmatrix} I \\ C_2(0) \end{bmatrix} X_0 \begin{bmatrix} I \\ C_2(0) \end{bmatrix}^* \quad (18c)$$

$$\begin{bmatrix} Z(k+1) & 0 \\ 0 & -D_2(k+1)Q(k+1)D_2^*(k+1) \end{bmatrix} \leq$$
$$\begin{bmatrix} I \\ C_2(k+1) \end{bmatrix} \begin{bmatrix} A^*(k) \\ B^*(k) \end{bmatrix}^* \begin{bmatrix} Z(k) & 0 \\ 0 & Q(k) \end{bmatrix} \begin{bmatrix} A^*(k) \\ B^*(k) \end{bmatrix} \begin{bmatrix} I \\ C_2(k+1) \end{bmatrix}^*, \quad (18d)$$

and equations (16h), (16i), and $X(0) = X_0$, (18e)

where (18d) is evaluated from $k = 0$ to $k = T-1$ only. Moreover, if Q_{opt} and $X_{0,\mathrm{opt}}$ solve (18), then these matrices also solve (16).

We conclude this section with an analysis result to calculate the worst-case MSE criterion for a particular class of estimators with observer-like structure. This result is used in the subsequent example section to compare the performance of the optimal minimax estimator $\mathcal{F}_{\mathrm{opt}}$ given in Corollary 1 with the performance of a conventional minimax estimator that does not take into account probability constraints on the state and input sequences.

Let $\mathcal{F}_L$ denote a state estimator of the form

$$\hat{x}(k) = x^-(k) + L(k)(y(k) - C_2(k)x^-(k)) \quad (19a)$$
$$x^-(k+1) = A(k)\hat{x}(k) \quad (19b)$$

where $x^-(0) = \bar{x}_0$ and L is a matrix sequence of appropriate dimensions. It is easy to show that $\mathcal{F}_L \in \mathcal{A}$. We are interested in evaluating the worst-case MSE criterion of $\mathcal{F}_L$ under the probability constraints of the minimax problem (10). That is, we want to calculate

$$J_{\mathrm{worst}}(\mathcal{F}_L) = \sup_{Q,X_0} J(\mathcal{F}_L; Q, X_0) \quad (20a)$$

subject to

$$Q(k) \in \mathcal{Q},\ X_0 \in \mathcal{X}_0, \text{ and} \quad (20b)$$
$$\mathrm{Prob}\{z(k) \leq h(k)\} \geq \gamma,\ \forall k \in [0,T] \quad (20c)$$

where the cost function $J(\mathcal{F}_L; Q, X_0)$ is the MSE criterion defined in (9). The next lemma shows how to calculate the worst-case MSE criterion $J_{\mathrm{worst}}(\mathcal{F}_L)$.

Lemma 2 (Worst-case MSE): Consider the system defined in (1) and the estimator $\mathcal{F}_L$ defined in (19). Suppose that assumptions A1) through A7) hold. The worst-case MSE criterion $J_{\mathrm{worst}}(\mathcal{F}_L)$ can be computed by solving the following convex program

$$J_{\mathrm{worst}} = \sup_{Q,\,X_0} \frac{1}{T+1}\sum_{k=0}^{T} \mathrm{tr}\{Y(k)W(k)\} \quad (21a)$$

subject to the following constraints $\forall k \in [0,T]$:

$$Q(k) \in \mathcal{Q},\ X_0 \in \mathcal{X}_0 \quad (21b)$$
$$Y(k) = (I - L(k)C_2(k))\,Y^-(k)\,(I - L(k)C_2(k))^* + L(k)D_2(k)Q(k)(L(k)D_2(k))^* \quad (21c)$$
$$Y^-(0) = X(0) = X_0 \quad (21d)$$
$$Y^-(k+1) = A(k)Y(k)A^*(k) + B(k)Q(k)B^*(k), \quad (21e)$$

and equations (16h) and (16i). (21f)

V. Numerical Example

Consider the aircraft tracking problem in the two dimensional space z_1-z_2 depicted in Figure 1. The aircraft is assumed to move in a nearly constant velocity mode with expected trajectory defined by the segment going from waypoint A to waypoint B [13]. The goal is to estimate the actual position and velocity of the aircraft from noisy position measurements provided by a radar. The estimation is carried out during a fix time window determined by the expected travel time from A to B.

The aircraft state at time t_k is represented by the vector of positions and velocities

$$x(k) = [z_1(t_k)\ \dot{z}_1(t_k)\ z_2(t_k)\ \dot{z}_2(t_k)]^*. \quad (22)$$

The evolution of $x(k)$ is modeled assuming decoupled motion between the z_1 and z_2 coordinates, and constant sampling time $T_s = t_{k+1} - t_k$ [13]. The resulting model is

$$x(k+1) = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} \frac{1}{2}T_s^2 & 0 \\ T_s & 0 \\ 0 & \frac{1}{2}T_s^2 \\ 0 & T_s \end{bmatrix} u(k) \quad (23a)$$
$$y(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x(k) + v(k) \quad (23b)$$

where the acceleration input u and the measurement noise v are zero mean Gaussian white noises with covariances $U(k)$ and $V(k)$, respectively.

The model (23) can be written in the form (1) by defining the random sequence w and its covariance matrix sequence Q as follows

$$w(k) = \begin{bmatrix} u(k) \\ v(k) \end{bmatrix} \qquad Q(k) = \begin{bmatrix} U(k) & 0 \\ 0 & V(k) \end{bmatrix}. \quad (24)$$

The mean $\bar{x}_0$ of the system initial condition is a known vector whose position components correspond to the waypoint A and whose velocity components define a vector with direction A-B. The value of $\bar{x}_0$ is

$$\bar{x}_0 = \frac{1}{\sqrt{2}}[35000m\ \ -200m/s\ \ 35000m\ \ -200m/s]^*. \quad (25)$$

The actual values of the covariance matrices X_0 and $U(k)$ are not available; however, they are known to satisfy

$$X_0 \geq \begin{bmatrix} 10m & 0 & 0 & 0 \\ 0 & 0.6m/s & 0 & 0 \\ 0 & 0 & 10m & 0 \\ 0 & 0 & 0 & 0.6m/s \end{bmatrix}^2 \tag{26a}$$

$$X_0 \leq \begin{bmatrix} 300m & 0 & 0 & 0 \\ 0 & 6m/s & 0 & 0 \\ 0 & 0 & 300m & 0 \\ 0 & 0 & 0 & 6m/s \end{bmatrix}^2 \tag{26b}$$

$$U(k) \geq (0.02m/s^2)^2 I \tag{26c}$$

$$U(k) \leq (0.4m/s^2)^2 I \tag{26d}$$

The measurement noise covariance matrix is known and equal to

$$V(k) = \sigma_v^2 I \tag{27}$$

where σ_v is a known parameter taking one of the following values: $85m$, $120m$, or $147m$. The expected arrival time to waypoint B is 175 seconds. We take this time as the length of the estimation window. Thus, with the sampling time $T_s = 5$ seconds, we get an estimation window length $T = 36$ samples.

The shaded regions in Figure 1 denote the regions where the aircraft is most likely to be as it travels from waypoint A to waypoint B. The first region, which is valid for all time, is defined as a band centered in the expected trajectory. The second region, valid only for the expected arrival time $k = T$, is defined as a square centered at waypoint B. In this example, we assume that the aircraft satisfies the constraints corresponding to the boundaries of the shaded regions in Figure 1 with 80% probability. That is, we assume the following probability constraints, $\forall k \in [0, T]$,

$$\text{Prob}\left\{-z_1(k) + z_2(k) \leq \frac{1340}{\sqrt{2}}m\right\} \geq 0.8 \tag{28a}$$

$$\text{Prob}\left\{z_1(k) - z_2(k) \leq \frac{1340}{\sqrt{2}}m\right\} \geq 0.8 \tag{28b}$$

For $k = T$, we assume

$$\text{Prob}\left\{z_1(T) + z_2(T) \leq \frac{1340}{\sqrt{2}}m\right\} \geq 0.8. \tag{28c}$$

Notice that the constraints (28) can be written as shown in (10c) by defining

$$z(k) = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ \delta(k-T) & 0 & \delta(k-T) & 0 \end{bmatrix} x(k) \tag{29a}$$

$$h(k) = \begin{bmatrix} \frac{1340}{\sqrt{2}}m & \frac{1340}{\sqrt{2}}m & \frac{1340}{\sqrt{2}}m \end{bmatrix}^* \tag{29b}$$

$$\gamma = [0.8 \ 0.8 \ 0.8]^* \tag{29c}$$

The estimation objective is to estimate position and velocity from noisy position measurements. The weight sequence W in the MSE criterion (9) is taken to be a fixed diagonal matrix, whose main diagonal is given by the reciprocal of the squares of the entries of the mean initial condition $\bar{x}_0$ defined in (25). In this example, we computed two distinct estimators. In both cases, the covariance matrices $Q(k)$ and X_0 are required to satisfy the constraints defined by (24), (26), and (27). The first estimator is computed enforcing the probability constraints (28). This estimator is obtained from Corollary 1 with covariance sequence Q_{opt} and covariance matrix $X_{0,\text{opt}}$ calculated from Theorem 2. The second estimator is a conventional minimax estimator obtained without explicit use of the probability constraints (28). This conventional minimax estimator is also obtained from Corollary 1 and Theorem 2 but removing the constraints (16h) and (16i), which are associated with the probability constraints (28). Each estimator is computed for the following three levels of measurement noise: σ_v = 85m, 120m, and 147m. Thus, a total of six estimators are obtained.

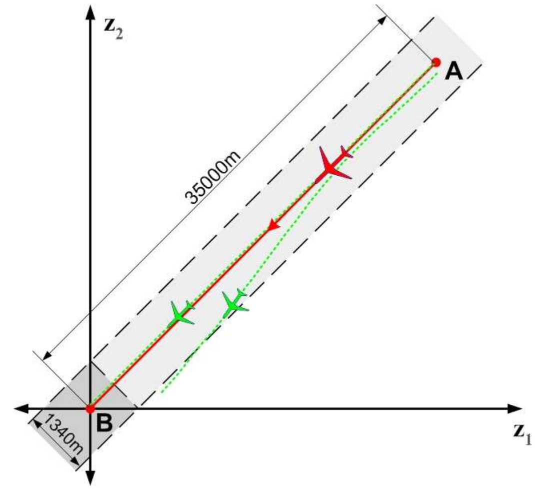

Fig. 1. Schematic of aircraft tracking problem. Solid line: expected aircraft trajectory; dashed line: constraints on aircraft trajectory; dotted line: representative aircraft trajectories.

Table I shows the worst-case MSE J_{worst} for all six estimators. The value J_{worst} was obtained from Lemma 2. Notice that for this analysis the constraints (28) are included so that all six estimators are evaluated using the same prior information; i.e., the probability constraints and the bounds on the covariance matrices. For the minimax estimators designed taking into account the probability constraints, the worst-case MSE from Lemma 2 satisfies $J_{\text{worst}} = J_{\text{opt}}$, where J_{opt} is the optimal minimax MSE from Theorem 2. The last column in the table shows the increase in MSE that results when the probability constraints are not used for estimator design. In this example, the penalty in accuracy for not using this prior information is greater than 20% in all cases.

To complement the worst-case analysis, we now present a statistical analysis of estimation accuracy. For each estimator, we calculated the distribution of the MSE criterion (9) over 3000 randomly generated covariance matrix pairs (U, X_0). These matrix pairs were generated to satisfy the bounds (26) and the probability constraints (28). For each estimator, and randomly generated covariance matrices $Q(k)$ and X_0, the MSE criterion was obtained from (21) without taking the supremum and instead evaluating the matrix sequence Y for the given input and initial state covariances, which is a well-known approach to calculating the MSE criterion [14]. Proceeding in this way, we generated 3000 MSE samples for each estimator. This distribution is representative of the estimators performance when evaluated with input and initial condition statistics that are consistent with the prior knowledge of the system. Table II shows the median of the MSE criterion distributions for each estimator. Once again, the loss of accuracy that results when the probability constraints are ignored is apparent.

Lastly, Figure 2 shows the time evolution of the pointwise mean square error $E\{\epsilon^*(k)W\epsilon(k)\}$ for the two estimators that correspond to the measurement noise level σ_v =120m. This pointwise error is computed as $E\{\epsilon^*(k)W\epsilon(k)\} = \mathrm{tr}\{Y_{\mathrm{worst}}(k)W\}$, where Y_{worst} is a matrix sequence that solves (21). The performance of the optimal minimax estimator, designed with the probability constraints enforced, is shown in the solid line, while the dashed line is used for the conventional minimax design. In this problem, the estimation accuracy improvements occur in the first half of the estimation interval.

TABLE I

COMPARISON OF WORST-CASE MSE CRITERION J_{worst}.

σ_v	Worst-case MSE		
	designs with constraints (28)	designs without constraints (28)	MSE increase
85m	1.026e-3	1.246e-3	21%
120m	1.267e-3	1.552e-3	22%
147m	1.555e-3	1.919e-3	23%

TABLE II

COMPARISON OF MEDIAN VALUES OF MSE CRITERION.

σ_v	Median MSE		
	designs with constraints (28)	designs without constraints (28)	MSE increase
85m	0.803e-3	0.896e-3	12%
120m	0.993e-3	1.104e-3	11%
147m	1.216e-3	1.352e-3	11%

VI. CONCLUSIONS

This paper considered the problem of optimal minimax state estimation in a linear time-varying system, driven by a Gaussian random process, subject to pointwise probability constraints on the driving process and the state. This problem is motivated by the possibility of improving the accuracy of minimax estimators by incorporating a priori information in the form of probability constraints on the system variables. The optimality criterion is the worst-case MSE criterion over a given set of initial condition and driving process covariance models, which are consistent with the probability constraints. The optimal minimax state estimator is a Kalman filter with suitably chosen driving input and initial condition covariance matrices. It is shown that these covariances may be obtained from the solution of a convex optimization problem. A numerical aircraft tracking example quantifies the benefits of adding probabilistic constraints in minimax state estimation.

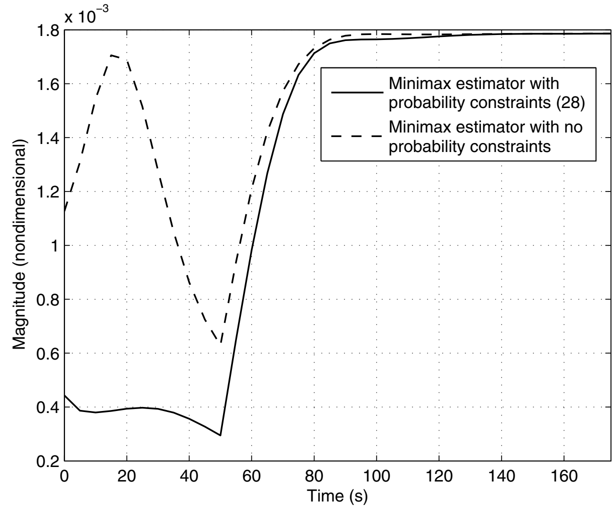

Fig. 2. Variation of pointwise worst-case mean square estimation error with time.

REFERENCES

[1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[2] R. Kalman and R. Bucy, “New results in filtering and prediction theory,” *J. Basic Eng.*, vol. 83, pp. 95 – 108, 1961.

[3] S. A. Kassam and H. V. Poor, “Robust techniques for signal processing: a survey,” *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433 – 481, 1985.

[4] P. Khargonekar, M. Rotea, and E. Baeyens, “Mixed H_2/H_∞ filtering,” *International Journal of Robust and Nonlinear Control*, vol. 6, no. 4, pp. 313–330, 1996.

[5] I. R. Petersen and A. V. Savkin, *Robust Kalman filtering for signals and systems with large uncertainties*. Boston: Birkhauser, 1999.

[6] F. Wang and V. Balakrishnan, “Robust kalman filters for linear time-varying systems with stochastic parametric uncertainties,” *IEEE Transactions on Signal Processing*, vol. 50, no. 4, pp. 803 – 813, 2002.

[7] S. Verdu and H. V. Poor, “Minimax linear observers and regulators for stochastic systems with uncertain second-order statistics.” *IEEE Transactions on Automatic Control*, vol. AC-29, no. 6, pp. 499 – 511, 1984.

[8] C. Rao, J. Rawlings, and J. Lee, “Constrained linear state estimation - a moving horizon approach,” *Automatica*, vol. 37, no. 10, pp. 1619 – 1628, 2001.

[9] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, “Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations,” *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246 – 258, 2003.

[10] Z. Wang, Z. Liu, R. Pei, and X. Ban, “A moving horizon estimation approach to constrained linear system with uncertain model,” *Proceedings of the American Control Conference*, vol. 3, pp. 2726 – 2731, 2003.

[11] A. Alessandri, M. Baglietto, and G. Battistelli, “A minimax receding-horizon estimator for uncertain discrete-time linear systems,” *Proceedings of the American Control Conference*, vol. 1, pp. 205 – 210, 2004.

[12] J. Yan and R. R. Bitmead, “Incorporating state estimation into model predictive control and its application to network traffic control,” *Automatica*, vol. 41, no. 4, pp. 595–604, 2005.

[13] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation -theory algorithms and software*. Wiley-Interscience, 2001.

[14] E. W. Kamen and J. Su, *Introduction to Optimal Estimation*. Springer-Verlag, 1999.