Robust Filtering for Discrete-Time Systems with Bounded Noise and Parametric Uncertainty

Laurent El Ghaoui and Giuseppe Calafiore

***Abstract*—This note presents a new approach to finite-horizon guaranteed state prediction for discrete-time systems affected by bounded noise and unknown-but-bounded parameter uncertainty. Our framework handles possibly nonlinear dependence of the state-space matrices on the uncertain parameters. The main result is that a minimal confidence ellipsoid for the state, consistent with the measured output and the uncertainty description, may be recursively computed in polynomial time, using interior-point methods for convex optimization. With n states, l uncertain parameters appearing linearly in the state-space matrices, with rank-one matrix coefficients, the worst-case complexity grows as $O(l(n+l)^{3.5})$. With unstructured uncertainty in all system matrices, the worst-case complexity reduces to $O(n^{3.5})$.**

***Index Terms*—Convex optimization, Kalman filtering, LMIs, set-membership filtering, unknown-but-bounded uncertainty.**

I. INTRODUCTION

This note is concerned with the problem of state estimation and filtering for discrete-time systems subject to unknown-but-bounded noise and parameter uncertainty affecting possibly *every* system matrix. The problem of state estimation for systems with uncertainty goes back to the early days of automatic control and signal processing, and several approaches exist in the literature up to this date, e.g., the stochastic approach (Kalman filtering theory), the H_∞ filtering theory, and the deterministic, or set-membership, approach.

It is now well known that the standard Kalman filter [1] requires an accurate model of the process under consideration, and assumes only *additive* uncertainty on the process and measurement equations, in the form of Gaussian noise. If these requirements are not met, the Kalman filter may lead to poor performance, see for instance [26]. This fact motivated further research in the direction of robustness in the stochastic setting, see, e.g., [4], [14], [23], [28].

Robust filtering has also been extensively studied in an H_∞ framework. In this setting, the exogenous input signal is assumed to be energy bounded rather than Gaussian. An H_∞ filter is designed such that the worst-case "gain" of the system is minimized, [15], [19].

The approach taken in this note is derived from the deterministic interpretation of the discrete-time Kalman filter given in [3]. The deterministic filter in [3] was shown to give a state estimate in the form of an ellipsoidal set of all possible states consistent with the given measurements and a deterministic additive description of the noise. The idea of propagating ellipsoids of confidence for systems with ellipsoidal noise goes back a long way; precursors in this field include Kurzhanski [16], Schweppe [25], whose ideas were later developed by Chernousko [6], Maskarov and Norton [18] and Ovseevich [22]. These authors consider the case with additive noise, assuming that the state-space process matrices are exactly known, in parallel to Kalman filtering; see [17] for a study of this parallel.

Manuscript received September 13, 1999; revised June 1, 2000. Recommended by Associate Editor J. Chen.

L. El Ghaoui is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720-1770 USA (e-mail: elghaoui@eecs.berkeley.edu).

G. Calafiore is with the Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy (e-mail: calafiore@polito.it).

Publisher Item Identifier S 0018-9286(01)06616-8.

The main contribution of this note is to extend the above mentioned set-membership approach to the case when structured uncertainty affects the system matrices. A similar approach has been considered in [24], where *unstructured* uncertainty described by a "Sum Quadratic Constraint" is assumed on the system.

The key result presented is that ellipsoids of confidence of minimal "size" (sum of semiaxis lengths or volume) can be recursively computed in polynomial time, via interior-point methods for convex optimization [21]. A similar problem, stated in the context of *static* systems, is explored in [9], while pure state prediction (without measurement information) is studied in [10].

A. Notation

For a square matrix X, $X \succ 0$ (resp. $X \succeq 0$) means X is symmetric, and positive–definite (resp. semidefinite). For a matrix U, $U_\perp$ denotes any orthogonal complement of U, i.e., a matrix of maximal rank such that $UU_\perp = 0$, and $U^\dagger$ denotes the (Moore–Penrose) pseudo-inverse of U.

Ellipsoids will be described as $\mathcal{E}(E, \hat{x}) = \{x: x = \hat{x} + Ez, \|z\| \le 1\}$, where $\hat{x} \in \mathbb{R}^n$ is the center, and $E \in \mathbb{R}^{n,n}$ is the *shape matrix* of the ellipsoid. This representation can handle "flat" ellipsoids, such as points or intervals. An alternative description involves the squared shape matrix $P = EE^T$, $P \succeq 0$: $\mathcal{E} = \{x: P \succeq (x - \hat{x})(x - \hat{x})^T\}$. When $P \succ 0$, the previous expression is also equivalent to $\mathcal{E} = \{x: (x - \hat{x})^T P^{-1}(x - \hat{x}) \le 1\}$.

The "size" of an ellipsoid is a function of the squared shape matrix P, and will be denoted $f(P)$. Throughout this note, $f(P)$ is either $\mathbf{Tr}(P)$, which corresponds to the sum of squares of the semiaxes lengths, or $\log\det(P)$, which is related to the volume.

II. PRELIMINARIES AND SETUP

We consider the following class of uncertain discrete-time systems

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \mathbf{M}(\Delta_k) \begin{bmatrix} x_k \\ w_k \\ v_k \end{bmatrix} \tag{1}$$

where it is assumed that the initial state x_0 belongs to a given ellipsoid $\mathcal{E}(E_0, \hat{x}_0)$, and $w_k \in \mathbb{R}^{n_w}$, $v_k \in \mathbb{R}^{n_v}$ are unknown-but-bounded noise signals, which are assumed to belong to a unit sphere, i.e., $\|w_k\| \le 1$, $\|v_k\| \le 1$, $\forall k$. This formalism allows us to consider the case when independent and norm-bounded signals affect the state dynamics and the sensor equations separately, as in the deterministic version of the classical Kalman filtering setup, see, e.g., [3], [25]. The case of noise signals bounded in ellipsoids is of course a trivial extension of this setup.

The uncertainty on the system matrices is assumed to be represented in linear fractional representation (LFR) form, i.e., for any given Δ

$$\mathbf{M}(\Delta) = M + \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \Delta(I - H\Delta)^{-1} [R_1 \quad R_2 \quad R_3] \tag{2}$$

where

$$M = \begin{bmatrix} A & B & 0 \\ C & 0 & D \end{bmatrix}$$

and $A, B, C, D, L_1, L_2, R_1, R_2, R_3$, and $H \in \mathbb{R}^{n_q, n_p}$ are given matrices. The uncertainty matrix Δ is in general time-varying and structured, and satisfies a given norm bound $\Delta \in \mathbf{\Delta}_1 \doteq \{\Delta \in \mathbf{\Delta}:$

$\|\Delta\| \le 1\}$, where $\mathbf{\Delta}$ is a subspace of $\mathbb{R}^{n_p, n_q}$, called the *structure subspace*. We also introduce the linear subspace $\mathcal{B}(\mathbf{\Delta})$, constructed from the subspace $\mathbf{\Delta}$, and referred to as the *scaling subspace*

$$\mathcal{B}(\mathbf{\Delta}) = \left\{(S, T, G): \forall \Delta \in \mathbf{\Delta},\ S\Delta = \Delta T,\ G\Delta = -\Delta^T G^T\right\}. \tag{3}$$

The above linear fractional representation of the uncertainty has great generality and is widely used in control theory, see for instance [13]. This framework includes the case when parameters perturb each coefficient of the data matrices in a polynomial or rational manner, as seen in the representation lemma given in [8], as well as more classical uncertainty models, such as norm-bounded unstructured uncertainty, and additive perturbations on the state and measurement equations.

Well-Posedness Assumption: We will make the standing assumption that the representation (2) is well-posed over $\mathbf{\Delta}_1$, meaning that $\det(I - H\Delta) \neq 0$ for all $\Delta \in \mathbf{\Delta}_1$. A well-known sufficient condition for well-posedness, which also arises in μ analysis problems [13], is given by

$$\begin{aligned} \exists\, S, T, G:\quad & H^T T H + H^T G + G^T H \prec S, \\ & (S, T, G) \in \mathcal{B}(\mathbf{\Delta}), \qquad S \succeq 0,\ T \succeq 0. \end{aligned} \tag{4}$$

If the system is well posed, we can rewrite the system equations equivalently as

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_k + L_1 p_k, \\ y_k &= Cx_k + Dv_k + L_2 p_k \\ q_k &= R_1 x_k + R_2 w_k + R_3 v_k + H p_k \\ p_k &= \Delta q_k, \qquad \Delta \in \mathbf{\Delta}_1 \end{aligned} \tag{5}$$

where p_k, q_k are perturbation signals.

Quadratic Embedding of LFRs: The main advantage of LFRs is that it enables to approximate an uncertain input–output relation by a set of quadratic constraints. This fact is stated in the following lemma, whose proof is omitted for brevity.

Lemma 1: For arbitrary vectors p, q, the property

$$p = \Delta q, \quad \text{for some } \Delta \in \mathbf{\Delta}_1 \tag{6}$$

implies that the following quadratic inequalities in (p, q) hold: For every $(S, T, G) \in \mathcal{B}(\mathbf{\Delta})$, with $S \succeq 0, T \succeq 0$

$$\begin{bmatrix} q \\ p \end{bmatrix}^T \begin{bmatrix} T & G \\ G^T & -S \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix} \ge 0. \tag{7}$$

Moreover, when $\mathbf{\Delta} = \mathbb{R}^{n_p, n_q}$ (unstructured uncertainty) the above quadratic embedding is nonconservative, meaning that property (7) implies (6). △

Using the above result, we can devise a quadratic *outer* approximation for the system equations (5), valid for every triple $(S, T, G) \in \mathcal{B}(\mathbf{\Delta})$, with $S \succeq 0, T \succeq 0$

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_k + L_1 p_k \\ y_k &= Cx_k + Dv_k + L_2 p_k \\ q_k &= R_1 x_k + R_2 w_k + R_3 v_k + H p_k \\ 0 &\le \begin{bmatrix} q_k \\ p_k \end{bmatrix}^T \begin{bmatrix} T & G \\ G^T & -S \end{bmatrix} \begin{bmatrix} q_k \\ p_k \end{bmatrix}. \end{aligned} \tag{8}$$

The above (outer) quadratic approximations for LFRs, when used in conjunction with the $\mathcal{S}$-procedure (see for instance [5]) is a key element in our approach. General results on the tightness of this embedding are given in [2], [12].

III. Robust Predictive Filter

The aim of the robust predictive filter is to determine a confidence ellipsoid $\mathcal{E}(E_+, \hat{x}_+)$ for the state at the next time instant x_{k+1}, given the measurement information at the time instant k, and given that x_k belongs to a current ellipsoid of confidence $\mathcal{E}(E, \hat{x})$. Therefore, we look for P_+, $\hat{x}_+$ such that

$$(x_{k+1} - \hat{x}_+)^T P_+^{-1} (x_{k+1} - \hat{x}_+) \le 1 \tag{9}$$

whenever a) (1) holds for some $\Delta_k \in \mathbf{\Delta}_1$, b) x_k is in $\mathcal{E}(E, \hat{x})$, and c) the noise terms w_k, v_k are bounded in unit spheres, i.e., $\|w_k\| \le 1$, $\|v_k\| \le 1$.

The following theorem contains our main result for the computation of the one-step-ahead confidence ellipsoid.

Theorem 1: An ellipsoid of confidence $\mathcal{E}_+ = \mathcal{E}(P_+, \hat{x}_+)$ can be obtained by solving the optimization problem in the variables P_+, x_+, τ_x, τ_w, τ_v, S, G, T

$$\begin{gathered} \text{minimize } f(P_+) \text{ subject to} \\ (S, T, G) \in \mathcal{B}(\mathbf{\Delta}), \quad S \succeq 0, \quad T \succeq 0, \quad \tau_x, \tau_w, \tau_v \ge 0 \\ \left[\begin{array}{c|c} P_+ & \Phi_1(\hat{x}_+)\Psi \\ \hline \Psi^T \Phi_1^T(\hat{x}_+) & \Psi^T(\Upsilon(\tau_x, \tau_w, \tau_v) - \Omega(S, T, G))\Psi \end{array}\right] \succ 0 \end{gathered} \tag{10}$$

where

$$\Phi_1(\hat{x}_+) \doteq [\, A\hat{x} - \hat{x}_+ \quad AE \quad B \quad 0 \quad L_1 \,]$$

and

$$\Psi = [\, C\hat{x} - y_k \quad CE \quad 0 \quad D \quad L_2 \,]_\perp; \tag{11}$$

$$\Upsilon(\tau_x, \tau_w, \tau_v) \doteq \mathbf{diag}(1 - \tau_x - \tau_w - \tau_v,\ \tau_x I,\ \tau_w I,\ \tau_v I,\ 0); \tag{12}$$

$$\Omega(S, T, G) \doteq \Phi^T \begin{bmatrix} T & G \\ G^T & -S \end{bmatrix} \Phi; \tag{13}$$

$$\Phi \doteq \begin{bmatrix} R_1\hat{x} & R_1 E & R_2 & R_3 & H \\ 0 & 0 & 0 & 0 & I \end{bmatrix} \tag{14}$$

and $f(P_+)$ measures the size of the ellipsoid, either $f(P_+) = \mathbf{Tr}\, P_+$, or $f(P_+) = \log\det P_+$. △

Proof: See Appendix A.

When the ellipsoid size is measured by the trace function, the ellipsoid update reduces to a semidefinite programming (SDP) problem. In this case, the update can be performed in polynomial-time using recently developed interior-point algorithms [21], [27] and related software [11]. However, the complexity of the algorithm (using a general-purpose SDP code) is still high, mainly due to the presence of $O(n^2)$ variables appearing in P_+, which makes the complexity of the problem grow as $O(n^{6.5})$, where n is the number of states (see [27] for details on complexity of SDP's). In the case of minimum-volume ellipsoids, the above formulation is not even convex in P_+.

We remark that the previous result provides a set-valued (ellipsoidal) estimate for the state, which could be useful for instance in robust optimization-based control, model validation [24], and robust collision avoidance applications [7]. On the other hand, if a noise-free estimate of the state is desired, then the confidence-set information could be neglected, and the centers of the confidence ellipsoids could be taken as optimal estimates of the system states.

Notice also that, in the case when no uncertainty is present on the system matrices, and only the deterministic disturbances w_k, v_k act on the system, the results given by Theorem 1 coincide with those provided by classical deterministic ellipsoidal bounding algorithms, see for instance [18] and references therein.

We next show how to eliminate the variable P_+, and transform the problem into a convex optimization problem with much better complexity properties. This alternative formulation will handle both the trace and volume as objective functions.

A. Decoupled Filtering Recursions

In this section, we give explicit expressions for the shape and center of $\mathcal{E}_+$, in terms of the optimal values of a certain convex optimization problem. This results in decoupled equations that are similar in spirit to the standard Kalman predictor equations. This new formulation will be used later to obtain an algorithm with better complexity properties than the general problem obtained in Theorem 1. The following technical lemma will be needed in the sequel.

Lemma 2: Let X_{ij}, $1 \le i \le j \le 3$ be matrices of appropriate size, with X_{ii} square and symmetric. The problem (in variables X, Z)

$$\text{minimize } f(X) \text{ subject to } \begin{bmatrix} X & Z & X_{13} \\ Z^T & X_{22} & X_{23} \\ X_{13}^T & X_{23}^T & X_{33} \end{bmatrix} \succeq 0 \tag{15}$$

is feasible if and only if

$$\begin{bmatrix} X_{22} & X_{23} \\ X_{23}^T & X_{33} \end{bmatrix} \succeq 0. \tag{16}$$

In this case, problem (15) is equivalent to the problem (in variable X only)

$$\text{minimize } f(X) \text{ subject to } \begin{bmatrix} X & X_{13} \\ X_{13}^T & X_{33} \end{bmatrix} \succeq 0 \tag{17}$$

and it admits unique optimal variables, given by $X = X_{13}X_{33}^{\dagger}X_{13}^T$, $Z = X_{13}X_{33}^{\dagger}X_{23}^T$. △

Proof: See Appendix B.

Now, we notice that one can always choose Ψ in such a way that its first row is the (transpose of the) first unit vector. A suitable matrix Ψ is, therefore, of the form

$$\Psi = \left[\begin{array}{c|c} 1 & 0 \\ \hline \psi_1 & \Psi_2 \end{array}\right]. \tag{18}$$

We introduce the following notation:

$$\begin{aligned} &Q(\tau_x, \tau_w, \tau_v, S, T, G) \\ &\quad \doteq \Psi^T(\Upsilon(\tau_x, \tau_w, \tau_v) - \Omega(S, T, G))\Psi \\ &\quad \doteq \begin{bmatrix} q_{11}(\tau_x,\tau_w,\tau_v,S,T,G) & q_{12}^T(\tau_x,\tau_w,\tau_v,S,T,G) \\ q_{12}(\tau_x,\tau_w,\tau_v,S,T,G) & Q_{22}(\tau_x,\tau_w,\tau_v,S,T,G) \end{bmatrix} \\ &f_1 \doteq A\hat{x} + [AE \;\; B \;\; 0 \;\; L_1]\psi_1. \end{aligned} \tag{19}$$

The decoupled robust filtering equations are then given in the following theorem.

Theorem 2: Consider the convex optimization problem in the variables $\tau_x, \tau_w, \tau_v, S, G, T$

$$\begin{aligned} &\inf f\left(KQ_{22}^{-1}(\tau_x, \tau_w, \tau_v, S, T, G)K^T\right) \text{ subject to} \\ &\quad Q(\tau_x, \tau_w, \tau_v, S, T, G) \succ 0, \\ &\quad (S, T, G) \in \mathcal{B}(\mathbf{\Delta}), \quad S \succeq 0, \quad T \succeq 0, \tau_x, \tau_w, \tau_v \ge 0, \end{aligned} \tag{20}$$

where $K = [AE \quad B \quad 0 \quad L_1]\Psi_2$. If the above problem is feasible, then the optimal ellipsoid is unique. At the optimum, the optimal shape matrix satisfies

$$P_+ = KQ_{22}^{\dagger}(\tau_x, \tau_w, \tau_v, S, T, G)K^T \tag{21}$$

while the optimal center of the ellipsoid is given by

$$\begin{aligned} \hat{x}_+ = f_1 - KQ_{22}^{\dagger}(\tau_x, \tau_w, \tau_v, S, T, G) \\ \cdot q_{12}(\tau_x, \tau_w, \tau_v, S, T, G). \end{aligned} \tag{22}$$

Proof: In view of the structure (18) of Ψ, we can rewrite the main LMI in (10) as

$$\left[\begin{array}{c|cc} P_+ & f_1 - \hat{x}_+ & K \\ \hline (f_1 - \hat{x}_+)^T & q_{11} & q_{12}^T \\ K^T & q_{12} & Q_{22} \end{array}\right] \succ 0 \tag{23}$$

where q_{11}, q_{12}, Q_{22}, and f_1 are defined in (19), and

$$K \doteq [AE \;\; B \;\; 0 \;\; L_1]\Psi_2.$$

The statements of the theorem then easily follow applying Lemma 2 to the LMI (23), with $Z = f_1 - \hat{x}_+$, and the other matrices defined appropriately. □

We remark that the classical well-posedness condition recalled in (4) implies that the ellipsoid of confidence computed by means of Theorem 2 is bounded at each step. Moreover, it is easily shown that the well-posedness condition (4) holds if and only if problem (20) is strictly feasible. Well-posedness therefore insures boundedness of the optimal ellipsoid at each step.

B. Summary: Filter Recursion

The robust predictive filter can be implemented recursively as follows:

1) select a time horizon T_h. Form an LFR of the system, and find a basis of the scaling subspace $\mathcal{B}(\mathbf{\Delta})$;
2) start with an initial ellipsoid of confidence $\mathcal{E}_0 = \mathcal{E}(\hat{x}_0, E_0)$. Set $k = 0$, $E = E_0$, $\hat{x} = \hat{x}_0$;
3) given E, $\hat{x}$, and current measurement y, solve the convex optimization problem (20), and find associated optimal scaling variables S, T, G;
4) form the matrix P_+ and center x_+ as given by (21) and (22);
5) find (using Cholesky factorization) a matrix E_+ such that $P_+ = E_+E_+^T$;
6) set $\hat{x} = \hat{x}_+$, $E = E_+$. If $k \ge T_h$, exit. Otherwise, set $k = k+1$ and go to Step 3.

C. Complexity Analysis

In this section, we outline how the interior-point methods described in [21] can be used to solve the optimization problem (20). We here stress the fact that the result of Theorem 2 dramatically improves the complexity of the SDP formulation obtained in Theorem 1. We begin by assuming that the size function is given by the trace $f(P) = \mathbf{Tr}(P)$.

A General Problem: Problem (20) can be expressed as

$$\begin{aligned} &\inf \alpha \text{ subject to } \alpha \ge \mathbf{Tr}\left(KR(s)^{-1}K^T\right), \\ &Q(s) := \begin{bmatrix} R(s) & r(s) \\ r(s)^T & q(s) \end{bmatrix} \succ 0, \quad S(s) \succeq 0 \end{aligned} \tag{24}$$

where vector s contains the free variables, and $Q(s)$, $S(s)$ are symmetric matrices affine in s. Here, $q(s)$ is the scalar, lower-right block in $Q(s)$. The constraint $S(s) \succ 0$ reflects the original constraints on the scaling variables S, T, and τ_x, τ_w, τ_v. The matrix $S(s)$ is a block diagonal matrix, with k diagonal blocks of size $\mu_i \times \mu_i$ each, where $\mu = [\mu_1, \ldots, \mu_k]$ is a vector describing the uncertainty structure. We first discuss in general terms the complexity of this problem, as a function of the size of $Q(s)$, N; the number of free variables, N_s; and the size and structure of the matrix scalings, which is described by μ.

A basic idea for solving a problem such as (24) is to associate a *barrier* for the feasible set, and solve a sequence of unconstrained minimization problems, involving a weighted combination of the barrier and the (linear) objective. The complexity of a path-following interior-point method as described in [21, p. 93] depends on our ability of finding a "self-concordant barrier" associated with the constraints. When such a barrier is known, the number of iterations grows as $O(\theta^{1/2})$, where θ is the "parameter of the barrier." The cost of each iteration is proportional to that of computing the gradient g and Hessian H of the barrier, and solving the linear system $Hd = g$, where the unknown d is the search direction. We note that in practice, the number of iterations is almost independent of problem size.

We can associate to problem (24) a self-concordant barrier, and find its parameter. Indeed, a direct consequence of the result [21, Prop. 5.1.8] is that the function

$$\begin{aligned} F(\alpha,\, s) =& -\log\left(\alpha - \mathbf{Tr}\left(KR(s)^{-1}K^T\right)\right) \\ &- \log\det Q(s) - \log\det S(s) \end{aligned} \tag{25}$$

is a self-concordant barrier for problem (24), with parameter $\theta = N + 1 + \sum_k \mu_k$. A tedious but straightforward calculation shows that the gradient and Hessian of the barrier can be computed in time $O(\nu)$, where

$$\nu = N_s^3 + N_s^2\left(N^2 + \sum_{i=1}^{k}\mu_i^2\right) + N_s\left(N^3 + \sum_{i=1}^{k}\mu_i^3\right). \tag{26}$$

Complexity of Robust Filtering: Let us specialize the above results to two specific instances of robust filtering. Assume first that the uncertainty matrix Δ comprises l uncertain scalar parameters, each appearing r times on the diagonal of Δ (r is related to the degree to which each parameter appears in the state-space representation of the system). We will express the complexity of the algorithm in terms of n (the order of the system), l (the number of uncertain scalar parameters), and r (which measures the degree of nonlinearity).

Thus, in our notation, we have $n_p = n_q = lr$. Also, $S = T$ is a symmetric, block-diagonal matrix, with l blocks, each of size $r \times r$, while G is a skew-symmetric matrix with the same structure. Therefore, $\mu_k = r$, $k = 1, \ldots, l$, and problem (20) involves a total of $N_s = O(lr^2)$ variables. The matrix $Q(s)$ is at most of row size $N := n + n_w + n_v + n_p - 1 = O(n + lr)$, the precise number depending on the rank of the matrix appearing in the right hand side of (18). The cost of each iteration is therefore given by (26), with

$$\begin{aligned} \nu &= \left(lr^2\right)^3 + \left(lr^2\right)^2\left((n+lr)^2 + lr^2\right) + lr^2\left((n+lr)^3 + lr^3\right) \\ &= O\left(lr^2\left(n + lr^2\right)(n+lr)^3\right). \end{aligned}$$

Since the parameter of the barrier (25) is $\theta = O(n + lr)$, the total complexity estimate is $O((n+lr)^{0.5}\nu)$.

Assuming $r = 1$ (e.g., parameters appear linearly in the state-space matrices, with rank-one matrix coefficients) results in a total complexity of $O(l(n+l)^{3.5})$. We note that, for *fixed* number of uncertain parameters (precisely, for fixed l and r), the complexity estimate is $O(n^{3.5})$, which is comparable to the case of standard Kalman filtering.

When *unstructured*, additive uncertainty is present on $A,\, B,\, C,\, D$, then $\mu = [1,\, 1,\, 1,\, 1]$, and $\theta = O(n)$, from which it can be easily verified that the total complexity in the unstructured case grows as $O(n^{3.5})$. As noted above, the number of iterations is almost constant in practice, so the *practical* complexity is $O(n^3)$.

Minimum-Volume Ellipsoids: The above results can be extended to the case when a minimum-volume ellipsoid is sought. Indeed, when $f(P) = \log\det P$, we simply minimize the objective $\log\det(KR(s)^{-1}K^T)$ under the constraints of problem (20), which can be done using path-following interior-point methods for self-concordant functions, as proved in [20]. Complexity estimates are similar to the trace case.

IV. Example

To illustrate the results, we consider a simple numerical example which has been used as a benchmark in [4], [14], [28], and is therefore useful for comparison purposes. The numerical results were implemented using the SDP formulation of Theorem 1, with a general-purpose SDP code [11]

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 0 & -0.5 \\ 1 & 1 + 0.3\delta_k \end{bmatrix} x_k + 0.02\begin{bmatrix} -6 \\ 1 \end{bmatrix} w_k \\ y_k &= [-100 \quad 10]x_k + 0.02v_k, \end{aligned}$$

with $|\delta_k| \le 1$, $\|w_k\| \le 1$, $\|v_k\| \le 1$, and assuming the initial state belongs to the ellipsoid $\mathcal{E}(E_0,\, \hat{x}_0)$, with $E_0 = 3I$, $\hat{x}_0 = 0$. The signal to be estimated is $z(k) = [1 \quad 0]x(k)$. The LFR uncertainty representation specializes to $H = 0$, $L_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $L_2 = 0$, $R_1 = [0 \quad 0.3]$, $R_2 = R_3 = 0$. The scaling subspace is in this case described by $S = T = \lambda$ (a scalar), $G = 0$. The system was simulated using deterministic, boundary-visiting sequences for the noise and the uncertainty. The results obtained with the robust filter, using $f(P) = \mathbf{Tr}(P)$, are shown in Fig. 1(a). The bounds on the signal $z(k)$ are obtained projecting the state ellipsoid along the output direction.

For illustration purposes, we also estimated the signal $z(k)$ using a standard Kalman filter, assuming a process noise variance $\sigma_w = 0.333$, measurement noise variance $\sigma_v = 0.333$, and initial state covariance equal to the identity. The results obtained with the Kalman filter are shown in Fig. 1(b), where the bounds indicate 3σ confidence regions.

This example clearly illustrates that the Kalman filter (which neglects the uncertainty on the system matrices) may provide central estimates that are completely erroneous (bias). Also, the (stochastic) confidence intervals provided by the Kalman filter are indeed tighter than their deterministic counterparts computed via the robust filter, but they do not guarantee the containment of the true signal $z(k)$.

V. Conclusion

The main contribution of this note is a technique that is able to handle 1) *uncertainty* in all the system matrices, and 2) *structure information* about the uncertainty, in filtering problems for uncertain discrete-time systems. The estimates and their (deterministic) ellipsoids of confidence are computed in polynomial-time using convex optimization, for both the minimum-volume and minimum-trace cases. The numerical complexity of the proposed filtering algorithms is comparable to that of the standard Kalman filter. The presented results are valid over a finite time horizon; infinite horizon and convergence issues are subject of ongoing research.

The presented method seems to be mostly suitable to applications with nonstationary processes or signals. It is expected that this technique, and the related approaches explored in [9], [10], should be applicable in a variety of contexts, ranging from robust failure detection to localization problems, and identification of systems with structured uncertainty.

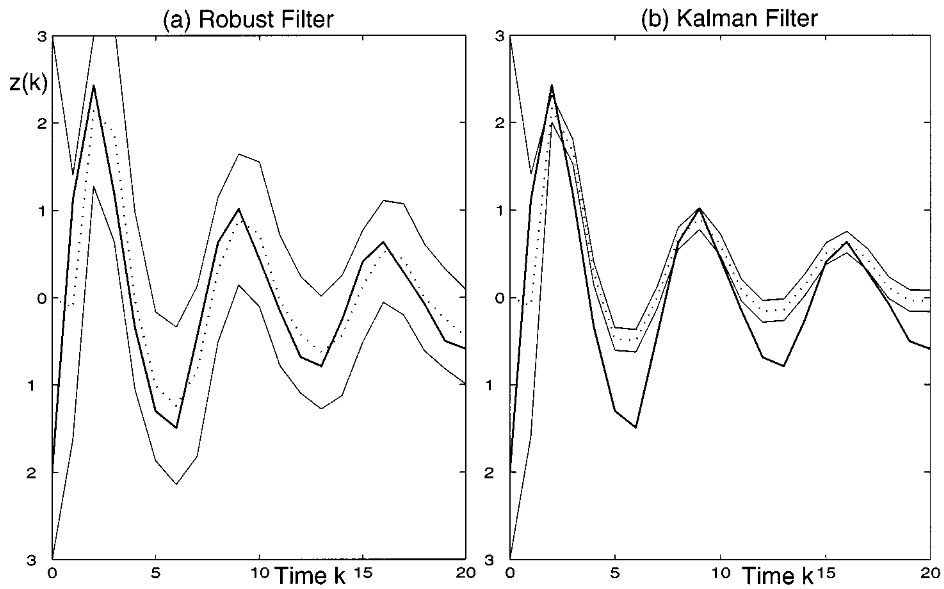

Fig. 1. Estimation of $z(k)$ using (a) the robust deterministic filter and (b) a standard Kalman filter. The thick lines represent $z(k)$, the dotted lines represent the central estimates, the solid lines represent the bounds on the estimates [ellipsoidal projections for (a), and 3σ confidence regions, for (b)].

APPENDIX A
PROOF OF THEOREM 1

Applying the quadratic embedding lemma, condition (9) is satisfied whenever conditions a), b), and c) below it are satisfied, if there exist $(S, T, G) \in \mathcal{B}(\mathbf{\Delta})$, with $S \succeq 0$, $T \succeq 0$, such that $(x_{k+1} - \hat{x}_+)^T P_+^{-1}(x_{k+1} - \hat{x}_+) \leq 1$ whenever (8) holds, and $x_k = \hat{x} + Ez_k$, $\|z_k\| \leq 1$, $\|w_k\| \leq 1$, $\|v_k\| \leq 1$.

Eliminating the equality constraints for q_k, x_{k+1}, x_k, the above conditions may be rewritten via a set of quadratic inequalities in the vector $\xi^T = [1 \; z_k^T \; w_k^T \; v_k^T \; p_k^T]$, namely

$$\xi^T \Phi_1(\hat{x}_+)^T P_+^{-1} \Phi_1(\hat{x}_+)\xi \leq 1$$

whenever

$$\begin{aligned} \xi^T \Phi_2^T \Phi_2 \xi &\leq 0 \\ \xi^T \Omega(S, T, G)\xi &\geq 0 \\ \xi^T \, \mathbf{diag}(-1, I, 0, 0, 0)\xi &\leq 0 \\ \xi^T \, \mathbf{diag}(-1, 0, I, 0, 0)\xi &\leq 0 \\ \xi^T \, \mathbf{diag}(-1, 0, 0, I, 0)\xi &\leq 0. \end{aligned}$$

Here, Ω is defined in (13), and

$$\begin{aligned} \Phi_1(\hat{x}_+) &\doteq [\, A\hat{x} - \hat{x}_+ \quad AE \quad B \quad 0 \quad L_1 \,] \\ \Phi_2 &\doteq [\, C\hat{x} - y_k \quad CE \quad 0 \quad D \quad L_2 \,]. \end{aligned}$$

A sufficient condition for the previous conditions to hold is given by the $\mathcal{S}$-procedure (see, e.g., [5]): there exist nonnegative scalars τ_x, τ_y, τ_w, τ_v such that

$$\begin{aligned} &\xi^T \Phi_1^T(\hat{x}_+) P_+^{-1} \Phi_1(\hat{x}_+)\xi - \tau_y \xi^T \Phi_2^T \Phi_2 \xi \\ &\quad - \xi^T \Upsilon(\tau_x, \tau_w, \tau_v)\xi + \xi^T \Omega(S, T, G)\xi < 0 \end{aligned} \tag{27}$$

where Υ is defined in (12). A necessary and sufficient condition for (27) to hold for all ξ is

$$\begin{aligned} &\Phi_1^T(\hat{x}_+) P_+^{-1} \Phi_1(\hat{x}_+) - \tau_y \Phi_2^T \Phi_2 - \Upsilon(\tau_x, \tau_w, \tau_v) \\ &\quad + \Omega(S, T, G) \prec 0. \end{aligned}$$

Let now Ψ be an orthogonal complement of Φ_2, i.e., a matrix of full rank such that $\Phi_2 \Psi = 0$. Then, using the elimination lemma (see [5]) we have that the above matrix inequality is satisfied for some value of τ_y, if and only if the following inequality (where τ_y does not appear) is satisfied:

$$\begin{aligned} &\Psi^T \Phi_1^T(\hat{x}_+) P_+^{-1} \Phi_1(\hat{x}_+)\Psi - \Psi^T(\Upsilon(\tau_x, \tau_w, \tau_v) \\ &\quad - \Omega(S, T, G))\Psi \prec 0. \end{aligned}$$

Using Schur complements, the previous condition is rewritten in the form

$$\left[\begin{array}{c|c} P_+ & \Phi_1(\hat{x}_+)\Psi \\ \hline \Psi^T \Phi_1^T(\hat{x}_+) & \Psi^T(\Upsilon(\tau_x, \tau_w, \tau_v) - \Omega(S, T, G))\Psi \end{array}\right] \succ 0 \tag{28}$$

which is an LMI condition in the problem variables P_+, $\hat{x}_+$, τ_x, τ_w, τ_v, S, G, T. The optimal ellipsoid of confidence based on the above sufficient condition is then determined minimizing $f(P_+)$, which results in the optimization problem presented in Theorem 1. □

APPENDIX B
PROOF OF LEMMA 2

By the Schur complement rule, the LMI constraint in (15) holds if and only if

$$\begin{aligned} &\begin{bmatrix} X - X_{\rm opt} & Z - Z_{\rm opt} \\ (Z - Z_{\rm opt})^T & \tilde{X}_{22} \end{bmatrix} \succeq 0 \\ &\begin{bmatrix} X_{13} \\ X_{23} \end{bmatrix} (I - X_{33} X_{33}^\dagger) = 0 \end{aligned} \tag{29}$$

where

$$\begin{aligned} X_{\rm opt} &= X_{13} X_{33}^\dagger X_{13}^T \\ Z_{\rm opt} &= X_{13} X_{33}^\dagger X_{23}^T \\ \tilde{X}_{22} &= X_{22} - X_{23} X_{33}^\dagger X_{23}^T. \end{aligned}$$

Problem (15) is thus equivalent to the problem of minimizing $f(X)$ subject to the above constraints. The equality in (29) is automatically enforced when (16) holds, and problem (17) is feasible. When this is the case, problem (15) is equivalent to problem (17). We further note that the inequality in (29) is equivalent to

$$X \succeq X_{\mathrm{opt}} + (Z - Z_{\mathrm{opt}})\tilde{X}_{22}^{\dagger}(Z - Z_{\mathrm{opt}})^T,$$
$$(Z - Z_{\mathrm{opt}})\left(I - \tilde{X}_{22}\tilde{X}_{22}^{\dagger}\right) = 0.$$

Both in the case of trace and log-determinant, the function $f(X)$ is concave on the cone of positive–definite matrices. This implies that the optimal value of X, Z are $X = X_{\mathrm{opt}}$, $Z = Z_{\mathrm{opt}}$, as claimed. □

ACKNOWLEDGMENT

This note has benefited from interesting discussions and valuable input from several people, including R. Balakrishnan, S. Boyd, E. Féron, A. Kurzhanski, and R. Tempo. The authors particularly thank A. Nemirovski for his help regarding Section III-C. Useful comments from the reviewers and the Associate Editor are also gratefully acknowledged.

REFERENCES

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

[2] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, "Robust semidefinite programming," in *Semidefinite Programming and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds. Norwell, MA: Kluwer, Feb. 2000.

[3] D. P. Bertsekas and I. B. Rhodes, "Recursive state estimation for a set-membership description of uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 117–128, Feb. 1971.

[4] P. Bolzern, P. Colaneri, and G. De Nicolao, "Optimal design of robust predictors for linear discrete-time systems," *Syst. Control Lett.*, vol. 26, pp. 25–31, 1995.

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

[6] F. L. Chernousko, *State Estimation of Dynamic Systems*. Boca Raton, FL: CRC Press, 1994.

[7] L. El Ghaoui and F. Seigneuret, "Robust optimization methodologies for the free route concept," in *Proc. 1998 Amer. Control Conf.*, vol. 3, USA, 1998, pp. 1797–1799.

[8] S. Dussy and L. El Ghaoui, "Measurement-scheduled control for the RTAC problem: An LMI approach," *Int. J. Robust Nonlinear Control*, vol. 8, no. 4–5, pp. 377–400, 1998.

[9] L. El Ghaoui and G. Calafiore, "Confidence ellipsoids for uncertain linear equations with structure," in *38th Conf. Decision Control*, vol. 2, Phoenix, AZ, Dec. 1999, pp. 1922–1991.

[10] ——, "Worst-case simulation of uncertain systems," in *Robustness in Identification and Control*. ser. Lecture Notes in Control and Information Sciences, A. Garulli, A. Tesi, and A. Vicino, Eds. London, U.K.: Springer-Verlag, June 1999, vol. 245.

[11] L. El Ghaoui and J.-L. Commeau. (1999, Jan.) `lmitool` *version* 2.0. [Online]. Available: via http://www.ensta.fr/~gropco

[12] L. El Ghaoui, F. Oustry, and H. Lebret, "Robust solutions to uncertain semidefinite programs," *SIAM J. Optim.*, vol. 9, no. 1, pp. 33–52, 1998.

[13] M. K. H. Fan, A. L. Tits, and J. C. Doyle, "Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 25–38, Jan. 1991.

[14] J. C. Geromel, "Optimal linear filtering under parameter uncertainty," *IEEE Trans. Signal Processing*, vol. 47, pp. 168–175, Jan. 1999.

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

[16] A. Kurzhanski and I. Vályi, *Ellipsoidal Calculus for Estimation and Control*. Boston, MA: Birkhäuser, 1997.

[17] A. Kurzhanski, "On the approximation of the solutions of estimation problems for uncertain systems by stochastic filtering equations," *Stochastics*, vol. 23, pp. 104–130, 1988.

[18] D. G. Maskarov and J. P. Norton, "State bounding with ellipsoidal set description of the uncertainty," *Int. J. Control*, vol. 65, no. 5, pp. 847–866, 1996.

[19] K. M. Nagpal and P. P. Khargonekar, "Filtering and smoothing in an H_∞ setting," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 152–166, Feb. 1991.

[20] A. Nemirovski, "Proof of self-concordance of a barrier function," unpublished, Jan. 2001, private communication.

[21] Y. Nesterov and A. Nemirovski, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*. Philadelphia, PA: SIAM, 1994.

[22] A. I. Ovseevich, "On equations of ellipsoids approximating attainable sets," *J. Optim. Theory Appl.*, vol. 95, no. 3, pp. 659–676, 1997.

[23] I. R. Petersen and A. Savkin, *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*. Boston, MA: Birkhäuser, 1999.

[24] A. V. Savkin and I. R. Petersen, "Robust state estimation and model validation for discrete-time systems with a deterministic description of noise and uncertainty," *Automatica*, vol. 34, no. 2, pp. 271–274, 1998.

[25] F. C. Schweppe, *Uncertain Dynamic Systems*. Englewood Cliffs: Prentice-Hall, 1973.

[26] Y. Theodor, U. Shaked, and C. E. de Souza, "A game theory approach to robust discrete-time H_∞ estimation," *IEEE Trans. Signal Processing*, vol. 42, pp. 1486–1495, June 1994.

[27] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, Mar. 1996.

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

On Kalman–Yakubovich–Popov Lemma for Stabilizable Systems

Joaquín Collado, Rogelio Lozano, and Rolf Johansson

***Abstract*—The Kalman–Yakubovich–Popov (KYP) Lemma has been a cornerstone in System Theory and Network Analysis and Synthesis. It relates an analytic property of a square transfer matrix in the frequency domain to a set of algebraic equations involving parameters of a minimal realization in time domain. This note proves that the KYP lemma is also valid for realizations which are stabilizable and observable.**

***Index Terms*—Nonminimal realization, positive-real functions.**

I. INTRODUCTION

Given a square transfer matrix $Z(s)$, the Kalman–Yakubovich–Popov (KYP) Lemma relates an analytic property of a square transfer matrix in the frequency domain to a set of algebraic equations involving parameters of a minimal realization in time domain. See the original references [7], [18], and [13], [20]. Further important developments were given in [3], [12]. The lemma was generalized to the multivariable case in [2]. Extensions and clarifications appeared on [5], [16], and [10]. Clear presentations and

Manuscript received March 11, 1999; revised April 6, 2000 and March 28, 2000. Recommended by Associate Editor G. Tao. The work of the first author was supported in part by Conacyt. This work was elaborated while J. Collado was in a sabbatical year at HEUDIASYC Laboratory, Compiegne, France.

J. Collado is with the Universidad Autónoma de Nuevo León, Facultad de Ingeniería Mecánica y Eléctrica, 66451 San Nicolás, N. L. Mexico, and also with Heudiasyc UMR 6599 CNRS - UTC, Centre de Recherche de Royallieu, BP 20.529 - 60200 Compiegne, France (e-mail: jcollado@hds.utc.fr).

R. Lozano is with Heudiasyc UMR 6599 CNRS - UTC, Centre de Recherche de Royallieu, BP 20.529 - 60200 Compiegne, France (e-mail: rlozano@hds.univ-compiegne.fr).

R. Johansson is with the Lund Institute of Technology, Lund University, Department of Automatic Control, SE 221 00 Lund, Sweden (e-mail: Rolf.Johansson@control.lth.se).

Publisher Item Identifier S 0018-9286(01)06615-6.