A Framework for State-Space Estimation with Uncertain Models

Ali H. Sayed, *Fellow, IEEE*

***Abstract*—This paper develops a framework for state-space estimation when the parameters of the underlying linear model are subject to uncertainties. Compared with existing robust filters, the proposed filters perform regularization rather than deregularization. It is shown that, under certain stabilizability and detectability conditions, the steady-state filters are stable and that, for quadratically-stable models, the filters guarantee a bounded error variance. Moreover, the resulting filter structures are similar to various (time- and measurement-update, prediction, and information) forms of the Kalman filter, albeit ones that operate on corrected parameters rather than on the given nominal parameters. Simulation results and comparisons with $\mathcal{H}_\infty$, guaranteed-cost, and set-valued state estimation filters are provided.**

***Index Terms*—Estimation, guaranteed-cost design, $\mathcal{H}_\infty$-filtering, Kalman filtering, parametric uncertainty, quadratic stability, regularized least-squares, set-valued estimation, steady-state filter.**

I. Introduction

The Kalman filter is the optimal linear least-mean-squares estimator for systems that are described by linear state-space Markov models (see, e.g., [1], [2]). Since its inception in the early 1960s, it has played a significant role in numerous fields ranging from orbit determination, to finance, to communications, to control, and other fields.

A central premise in the Kalman filter theory is that the underlying state-space model is accurate. When this assumption is violated, the performance of the filter can deteriorate appreciably. (See, e.g., the edited volumes [3], [4], as well as [5], which contain several discussions and articles on practical issues in Kalman filtering design. See also the simulation examples further ahead in Section VII.) This filter sensitivity to modeling errors has led to several works in the literature on the development of robust state-space filters; robust in the sense that they attempt to limit, in certain ways, the effect of model uncertainties on the overall filter performance. Three distinctive approaches to state-space estimation in this regard are $\mathcal{H}_\infty$ filtering, set-valued estimation, and guaranteed-cost designs.

The $\mathcal{H}_\infty$ approach attempts to construct filters that bound the 2-induced norm of the operator that maps the disturbances to the estimation errors (see, e.g., [6]–[14]). One limitation of $\mathcal{H}_\infty$ designs for online (i.e., recursive) filter operation is that they require continuous testing of a certain existence condition. When the condition fails at any particular iteration, the desired $\mathcal{H}_\infty$ performance is lost and the filter can diverge. This is a consequence of the fact that $\mathcal{H}_\infty$ filters minimize certain indefinite quadratic forms and, as a result, they perform de-regularization (see, e.g., [15]). One method to ameliorate this difficulty is to increase the value of the robustness parameter γ; this however comes at the expense of decreasing the robustness of the filter (see again the simulations in Section VII). This issue is of such concern for on-line filtering operations that several tools and tuning methods have been studied in the literature, especially for linear time-invariant models, with the objective of enabling the designer to check *a priori* whether a prescribed robustness level γ can be guaranteed by an $\mathcal{H}_\infty$ filter over intervals of arbitrary lengths (see, e.g., the works [16], [17]).

A second useful approach to robust estimation is the set-valued estimation approach. In this design, one attempts to construct ellipsoids around state estimates that are consistent with the observations and subject to certain norm constraints on the noise disturbances (see, e.g., [18]–[20] and the many references in the latter edited volume). Some extensions to handle model uncertainties are described in [14], [21], [22], with the reference [22] considering a class of model uncertainties that can be described by integral (or sum) quadratic constraints. Here again one is faced with the requirement of checking for certain existence conditions, which can be an impediment to on-line filtering—see the simulation results in Section VII.

A third well-studied approach to robust estimation is the guaranteed-cost paradigm. Here one attempts to construct state-space estimators that guarantee that the steady-state variance of the state estimation error is upper bounded by a certain constant value for all admissible uncertainties in the model (see, e.g., [14], [23]–[25] and also [26], [27]). The solution usually involves some design parameters that need to be selected adequately (or tuned) in order to guarantee the existence of a positive-definite stabilizing solution of a certain discrete algebraic Riccati equation—see Section VI-D. The arguments and the derivations in most cases (see, e.g., [14], [23], [25]) are limited to time-invariant and quadratically-stable nominal models in steady-state operation. Extensions of the results to finite-horizon time-variant models are considered in [28], [29]. The solution in [28], however, leads to a more involved filter structure and suffers from instability problems, as acknowledged by the authors [28, p. 185] and also observed in simulations. The solution in [29] is one that is consistent with the steady-state filters developed in [14], [23], [25]. It again requires testing of certain existence conditions, which can be a limitation for on-line operation. The discussion in [29] further elaborates on sufficient conditions for guaranteed operation over arbitrarily long intervals of time.

Manuscript received January 28, 2000; revised September 15, 2000. Recommended by Associate Editor G. De Nicolao. This work was supported in part by a Grant from the National Science Foundation under Award ECS-9820765.

The author is with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail: sayed@ee.ucla.edu).

Publisher Item Identifier S 0018-9286(01)06620-X.

In this paper, we develop a robust procedure for state-space estimation in the presence of modeling uncertainties. Compared with the standard Kalman filter, which is known to minimize the regularized residual norm at each iteration, the new filters are designed to minimize the worst-possible regularized residual norm over the class of admissible uncertainties at each iteration. In addition, compared with the aforementioned robust formulations, the resulting filters perform data regularization rather than deregularization; a property that circumvents existence conditions and is convenient for online operation. The new filters are also shown to lead to stable steady-state performance and, for quadratically stable models, they are further shown to guarantee bounded error variances. Moreover, the proposed framework applies to a general class of parametric uncertainties, specified through the selection of a modeling function $\phi(x)$—see the remark in the paragraph following (7).

We start our exposition in the next section by formulating a generic regularized least-squares problem for models with data uncertainties. Once this is done, we shall then focus on the state-space estimation problem in some detail.

Notation: For a column vector z and a positive-definite matrix W, we write $||z||^2$ and $||z||_W^2$ to denote the Euclidean norm and its weighted version, namely, $z^T z$ and $z^T W z$, respectively. Also, for brevity, we may sometimes write $A^T W(\cdot)$ instead of $A^T W A$ especially when the factor A admits a long expression (see, e.g., the last expression in the statement of Theorem 3).

II. Regularized Least-Squares with Uncertainties

As is well known (see, e.g., [2]), many estimation techniques rely on solving regularized least-squares problems of the form

$$\min_x \left[x^T Q x + (Ax-b)^T W (Ax-b)\right] \tag{1}$$

where $x^T Q x$ is a regularization term with $Q = Q^T > 0$ and $W = W^T \geq 0$ is a weighting matrix.[1] The unknown vector x is n-dimensional, while A is $N \times n$ and b is $N \times 1$. Both A and b are assumed known, and the solution of (1) is

$$\hat{x} = \left[Q + A^T W A\right]^{-1} A^T W b. \tag{2}$$

In practice, the nominal data $\{A, b\}$ are often subject to disturbances and/or uncertainties. Such errors can degrade the performance of the estimator (2). For example, if the actual data matrix were $(A+\delta A)$, for some unknown perturbation δA, then the estimator (2) that is designed based on A alone, and without accounting for the existence of δA, can perform poorly.

This motivated us to introduce in [30] a robustified version of (1) that can account for a general class of uncertainties in the data $\{A, b\}$. Thus let $J(x, y)$ denote a two-variable cost function of the form

$$J(x, y) = x^T Q x + R(x, y)$$

where the residual $R(x, y)$ is defined by

$$R(x, y) \triangleq (Ax - b + Hy)^T W (Ax - b + Hy).$$

Here, H is an $N \times m$ known matrix and y is an $m \times 1$ unknown perturbation vector. Comparing the expression for $R(x, y)$ with the term $(Ax-b)^T W (Ax-b)$ that appears in (1), we see that we are representing possible sources of uncertainties in A and b by the additional term Hy. The matrix H provides the designer with the freedom of restricting the uncertainty y to certain range spaces. While y itself is not known, we assume that what is known is a bound on its Euclidean norm, say $||y|| \leq \phi(x)$, for some known (linear or nonlinear) nonnegative function $\phi(x)$. Observe that the bound on y is allowed to depend on x.

Consider now the problem of solving

$$\hat{x} = \arg\min_x \max_{||y|| \leq \phi(x)} J(x, y). \tag{3}$$

Problem (3) can be interpreted as a constrained two-player game problem, with the designer trying to pick an estimate $\hat{x}$ that minimizes the cost while the opponent $\{y\}$ tries to maximize the cost. The game problem is constrained since it imposes a bound [through $\phi(x)$] on how large (or how damaging) the opponent can be. Observe further that the strength of the opponent can vary with the choice of x. We shall assume in the sequel that H and $\phi(x)$ are not identically zero

$$H \neq 0 \quad \text{and} \quad \phi(\cdot) \neq 0 \tag{4}$$

since if either is zero, the game problem (3) trivializes to (1).

A special case of (3) was studied in [31]–[34] with the choices $W = I$, $Q = 0$, $H = I$, and $\phi(x) = \eta ||x||$. It turns out, however, that for treating the state-space estimation problem of this paper one has to allow for nontrivial choices of $\{W, Q, H\}$, as well as for more general choices of $\phi(x)$. The problem in this general case is richer in structure and its solution requires some care to avoid the introduction of multiple regularization parameters, as was shown in [30].

In this paper, we focus on the following specialization of (3)

$$\min_x \max_{\substack{\delta A \\ \delta b}} \left[||x||_Q^2 + ||(A+\delta A)x - (b+\delta b)||_W^2\right] \tag{5}$$

where

- $\{\delta A\}$ — $N \times n$ perturbation matrix to the nominal matrix A,
- δb — denotes an $N \times 1$ perturbation vector to the nominal vector b;
- $\{\delta A, \delta b\}$ — assumed to satisfy a model of the form

$$[\,\delta A \quad \delta b\,] = H\Delta[\,E_a \quad E_b\,] \tag{6}$$

where Δ is an arbitrary contraction, $||\Delta|| \leq 1$, and $\{H, E_a, E_b\}$ are known quantities of appropriate dimensions (E_b is a column vector). Perturbation models of the form (6) are common in robust filtering and control and can arise from tolerance specifications on physical parameters (see [35] for an example).

[1]When Q is sign-indefinite, as is common in $\mathcal{H}_\infty$ formulations (e.g., [11], [15]), we say that the solution performs deregularization. In such situations, the least-squares problem need not always have a minimum; the Hessian matrix with respect to x needs to be positive–definite.

In order to see how (5) is a special case of (3), we rewrite the cost in (5) as

$$x^T Q x + [Ax - b + (\delta A x - \delta b)]^T W [Ax - b + (\delta A x - \delta b)]$$

so that with Hy defined as

$$Hy \triangleq \delta A x - \delta b = H\Delta(E_a x - E_b)$$

and y as $y = \Delta(E_a x - E_b)$, problem (5) can be verified to be equivalent to the following:

$$\min_x \max_{\|y\| \le \|E_a x - E_b\|} \left[\|x\|_Q^2 + \|Ax - b + Hy\|_W^2\right]$$

which is a special case of (3) for the particular choice $\phi(x) = \|E_a x - E_b\|$. One can also handle the case in which the uncertainties $\{\delta A, \delta b\}$ in (5) are unrelated yet bounded, say

$$\|\delta A\| \le \eta, \qquad \|\delta b\| \le \eta_b \tag{7}$$

for some nonnegative scalars $\{\eta, \eta_b\}$, instead of (6). In this case, problem (5) reduces to a special case of (3) for the choice $\phi(x) = \eta\|x\| + \eta_b$—see [36].

The formulation (3) is more general than (6) and (7) in that it allows for other classes of perturbations through the choice of the function $\phi(x)$. The solution of (3) in this general case is discussed in [30], [37]. In this paper, we focus on perturbations of the form (6). Our arguments are such that they can be extended to other classes of perturbations. The following result is proven in [30] and [37].

Theorem 1 (Solution): The problem (5) and (6) has a unique solution $\hat{x}$ that is given by [compare with (2)]

$$\hat{x} = \left[\hat{Q} + A^T \hat{W} A\right]^{-1} \left[A^T \hat{W} b + \hat{\lambda} E_a^T E_b\right] \tag{8}$$

where the modified weighting matrices $\{\hat{Q}, \hat{W}\}$ are obtained from $\{Q, W\}$ via

$$\hat{Q} \triangleq Q + \hat{\lambda} E_a^T E_a \tag{9}$$

$$\hat{W} \triangleq W + WH\left(\hat{\lambda} I - H^T W H\right)^{\dagger} H^T W \tag{10}$$

and the nonnegative scalar parameter $\hat{\lambda}$ is determined from the optimization

$$\hat{\lambda} = \arg \min_{\lambda \ge \|H^T W H\|} G(\lambda) \tag{11}$$

where the function $G(\lambda)$ is defined as

$$G(\lambda) = \|x(\lambda)\|_Q^2 + \lambda\|E_a x(\lambda) - E_b\|^2 + \|Ax(\lambda) - b\|_{W(\lambda)}^2. \tag{12}$$

Here

$$W(\lambda) \triangleq W + WH\left(\lambda I - H^T W H\right)^{\dagger} H^T W \tag{13}$$

$$Q(\lambda) \triangleq Q + \lambda E_a^T E_a \tag{14}$$

and

$$x(\lambda) \triangleq \left[Q(\lambda) + A^T W(\lambda) A\right]^{-1} \left[A^T W(\lambda) b + \lambda E_a^T E_b\right]. \tag{15}$$

[The notation $X^{\dagger}$ denotes the pseudoinverse of X.]

Moreover, the value of the resulting optimal cost of (5) is equal to $G(\hat{\lambda})$, and is given by

$$G\left(\hat{\lambda}\right) = [1 \quad \hat{x}^T] \begin{bmatrix} \hat{\lambda}\|E_b\|^2 & -\hat{\lambda} E_b^T E_a \\ -\hat{\lambda} E_a^T E_b & \hat{Q} \end{bmatrix} \begin{bmatrix} 1 \\ \hat{x} \end{bmatrix} + (A\hat{x} - b)^T \hat{W} (A\hat{x} - b).$$

$\diamondsuit$

A. Structure of the Solution

We thus see that the solution of (5) and (6) requires that we first determine an optimal nonnegative scalar parameter, $\hat{\lambda}$, which corresponds to the minimizing argument of the function $G(\lambda)$ over the semi-open interval $[\|H^T W H\|, \infty)$. For convenience of notation, we shall denote the lower bound on λ by λ_l, i.e.,

$$\lambda_l \triangleq \left\|H^T W H\right\|. \tag{16}$$

Compared with the solution (2) of the standard regularized least-squares problem (1), we see that the expression for $\hat{x}$ in (8) is distinct in two important ways.

1) First, the weighting matrices $\{Q, W\}$ need to be replaced by corrected versions $\{\hat{Q}, \hat{W}\}$. These corrections are defined in terms of the optimal parameter $\hat{\lambda}$ and they are also dependent on the uncertainty model.
2) Second, the right-hand side of (8) contains an additional term that is equal to $\hat{\lambda} E_a^T E_b$. This means that, with $\hat{\lambda}$ given, the $\hat{x}$ in (8) can be interpreted as the solution to a regularized least-squares problem of the form

$$\min_x \left([1 \quad x^T] \begin{bmatrix} \hat{\lambda}\|E_b\|^2 & -\hat{\lambda} E_b^T E_a \\ -\hat{\lambda} E_a^T E_b & \hat{Q} \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} + (Ax - b)^T \hat{W} (Ax - b)\right)$$

with a *cross-coupling* term between x and unity. Observe further from (8) that $\hat{x}$ satisfies an orthogonality relation of the form

$$-\hat{\lambda} E_a^T E_b + \hat{Q}\hat{x} + A^T \hat{W}(A\hat{x} - b) = 0 \tag{17}$$

which is typical of regularized least-squares problems with coupling.

B. The Minimization of $G(\lambda)$

For any value of λ in the interval $[\lambda_l, \infty)$, the matrix $W(\lambda)$ in (13) is nonnegative-definite so that $G(\lambda) \ge 0$ for $\lambda \ge \lambda_l$ (it may become negative for $\lambda < \lambda_l$). In addition, it can be proven that the function $G(\lambda)$ has a *unique* minimum in the interval $[\lambda_l, \infty)$; and hence that it has a unique global minimum and no local minima [37]. This indicates that the determination of $\hat{\lambda}$ can be sought via optimization routines that need not be concerned with the possibilities of local minima.

In addition, a useful observation from many simulations that we have performed is that the function $G(\lambda)$ tends to reach amplitudes close to its minimum value at arguments λ that are generally close to the lower bound λ_l (see, e.g., [38]). We shall use this observation in a future section to suggest a practical approximation for the optimal $\hat{\lambda}$ without the need to explicitly minimize $G(\lambda)$. In particular, we shall later set $\hat{\lambda} = (1+\alpha)\lambda_l$, for some $\alpha > 0$, i.e., we shall set $\hat{\lambda}$ at a multiple of the lower bound. More on this issue later. For now, we continue with the optimal choice for λ.

C. Invertibility Condition

In the state-space context that we study in Section IV, the matrix W will be positive–definite (and, hence, invertible) so that $W(\lambda)$ itself will always be positive-definite (and, hence, also invertible). Therefore, if we restrict the minimization in (11) to the open interval (λ_l, ∞) (i.e., if we only exclude the boundary point λ_l), then the pseudoinverse operation in (13) can be replaced by normal matrix inversion, so that by using the matrix inversion lemma we arrive at the compact expression

$$W^{-1}(\lambda) = W^{-1} - \lambda^{-1} H H^T. \tag{18}$$

This expression will be useful in the sequel; we shall henceforth assume that the boundary point λ_l is excluded and use this simpler expression.

III. Standard State-Space Estimation

We now discuss how to incorporate the earlier discussions into a state-space context. We first review the standard Kalman filtering solution for accurate state-space models.

A. The Kalman Filter

Thus consider a state-space description of the form

$$x_{i+1} = F_i x_i + G_i u_i, \qquad i \geq 0, \tag{19}$$
$$y_i = H_i x_i + v_i \tag{20}$$

where $\{x_0, u_i, v_i\}$ are uncorrelated zero-mean random variables with variances

$$E\left(\begin{bmatrix} x_0 \\ u_i \\ v_i \end{bmatrix}\begin{bmatrix} x_0 \\ u_j \\ v_j \end{bmatrix}^T\right) = \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & Q_i\delta_{ij} & 0 \\ 0 & 0 & R_i\delta_{ij} \end{bmatrix} \tag{21}$$

that satisfy

$$\Pi_0 > 0, \quad R_i > 0, \quad Q_i > 0. \tag{22}$$

Here, δ_{ij} is the Kronecker delta function that is equal to unity when $i = j$ and zero otherwise.

The well-known Kalman filter provides the optimal linear least-mean-squares (l.l.m.s.) estimate of the state variable given prior observations. More specifically, introduce the following predicted and filtered estimates:

$$\hat{x}_i \triangleq \text{l.l.m.s. estimate of } x_i \text{ given } \{y_0, y_1, \ldots, y_{i-1}\}$$
$$\hat{x}_{i|i} \triangleq \text{l.l.m.s. estimate of } x_i \text{ given } \{y_0, y_1, \ldots, y_{i-1}, y_i\}$$

and the corresponding error variances

$$P_i \triangleq E(x_i - \hat{x}_i)(x_i - \hat{x}_i)^T$$
$$P_{i|i} \triangleq E(x_i - \hat{x}_{i|i})(x_i - \hat{x}_{i|i})^T.$$

Then the $\{\hat{x}_i, \hat{x}_{i|i}\}$ can be constructed recursively as follows (see, e.g., [2]):

$$\hat{x}_{i+1} = F_i \hat{x}_{i|i}, \qquad i \geq 0 \tag{23}$$
$$\hat{x}_{i+1|i+1} = \hat{x}_{i+1} + P_{i+1|i+1} H_{i+1}^T R_{i+1}^{-1} e_{i+1} \tag{24}$$
$$e_{i+1} = y_{i+1} - H_{i+1}\hat{x}_{i+1} \tag{25}$$

where

$$P_{i+1} = F_i P_{i|i} F_i^T + G_i Q_i G_i^T \tag{26}$$
$$P_{i+1|i+1} = P_{i+1} - P_{i+1} H_{i+1}^T R_{e,i+1}^{-1} H_{i+1} P_{i+1} \tag{27}$$
$$R_{e,i+1} = R_{i+1} + H_{i+1} P_{i+1} H_{i+1}^T \tag{28}$$

and with initial conditions

$$\hat{x}_{0|0} = P_{0|0}^{-1} H_0^T R_0^{-1} y_0 \tag{29}$$
$$P_{0|0} = \left(\Pi_0^{-1} + H_0^T R_0^{-1} H_0\right)^{-1}. \tag{30}$$

It also holds, when the required inverses exist, that

$$P_{i+1|i+1}^{-1} = P_{i+1}^{-1} + H_{i+1}^T R_{i+1}^{-1} H_{i+1}.$$

Equations (23)–(28) are known collectively as the time- and measurement-update form of the Kalman filter. It can be further seen from these equations that the following prediction forms of the Kalman filter also hold:

$$\hat{x}_{i+1} = F_i \hat{x}_i + F_i P_{i|i} H_i^T R_i^{-1} e_i \tag{31}$$
$$= F_i \hat{x}_i + F_i P_i H_i^T R_{e,i}^{-1} e_i \tag{32}$$

where P_i satisfies the Riccati recursion

$$P_{i+1} = F_i P_i F_i^T + G_i Q_i G_i^T - K_i R_{e,i}^{-1} K_i^T \tag{33}$$
$$K_i = F_i P_i H_i^T \tag{34}$$

with initial conditions $\hat{x}_0 = 0$ and $P_0 = \Pi_0$.

B. A Deterministic Interpretation

Each step (23)–(28) of the time- and measurement-update form of the Kalman filter admits a deterministic interpretation as the solution to a regularized least-squares problem, as we shall now explain (see also [39, pp. 390–391]). We will use this interpretation to motivate a procedure for robust state-space estimation that is based on the result of Theorem 1.

Thus, fix a time instant i and assume that the filtered estimate $\hat{x}_{i|i}$ has already been computed with the corresponding error variance matrix $P_{i|i}$ (assumed positive–definite).[2] Given a new

[2]The final filter recursions are independent of $P_{i|i}^{-1}$ so that the requirement of an invertible $P_{i|i}$ can be relaxed.

measurement y_{i+1}, we then pose the problem of estimating x_i again, along with u_i, by solving

$$\min_{x_i,\, u_i} \left[\|x_i - \hat{x}_{i|i}\|^2_{P^{-1}_{i|i}} + \|u_i\|^2_{Q_i^{-1}} + \|y_{i+1} - H_{i+1}x_{i+1}\|^2_{R^{-1}_{i+1}} \right]. \tag{35}$$

This problem can be interpreted as follows. Given an initial estimate $\hat{x}_{i|i}$ for x_i, one seeks to improve upon it by incorporating the additional information that is provided by the new measurement y_{i+1}. The design criterion is one that minimizes the (regularized) squared residual norm. Now the solution of (35) is known to lead to the filter equations (23)–(28). Indeed, we can rewrite (35) more compactly in the regularized least-squares form (1) with the identifications $W \longleftarrow R^{-1}_{i+1}$

$$x \longleftarrow \text{col}\{x_i - \hat{x}_{i|i},\, u_i\}, \quad b \longleftarrow y_{i+1} - H_{i+1}F_i\hat{x}_{i|i}$$
$$A \longleftarrow H_{i+1}[F_i \quad G_i], \quad Q \longleftarrow \left(P^{-1}_{i|i} \oplus Q_i^{-1}\right).$$

Here, the notation $\text{col}\{a, b\}$ denotes a column vector with entries a and b, and $(a \oplus b)$ denotes a block diagonal matrix with entries a and b.

We shall denote the minimizing arguments of (35) by $\hat{x}_{i|i+1}$ and $\hat{u}_{i|i+1}$. From the solution (2) of any such regularized least-squares problem, we find that

$$\hat{x}_{i|i+1} = \hat{x}_{i|i} + P_{i|i}F_i^T H^T_{i+1} R^{-1}_{i+1}(y_{i+1} - H_{i+1}\hat{x}_{i+1|i+1}) \tag{36}$$

$$\hat{u}_{i|i+1} = Q_i G_i^T H^T_{i+1} R^{-1}_{i+1}(y_{i+1} - H_{i+1}\hat{x}_{i+1|i+1}) \tag{37}$$

where we introduced the quantity [in agreement with the state-space constraint (19)]:

$$\hat{x}_{i+1|i+1} \triangleq F_i\hat{x}_{i|i+1} + G_i\hat{u}_{i|i+1}. \tag{38}$$

If we further introduce

$$\hat{x}_{i+1} \triangleq F_i\hat{x}_{i|i}, \quad P_{i+1} \triangleq F_iP_{i|i}F_i^T + G_iQ_iG_i^T$$

and substitute (36) and (37) for $\{\hat{x}_{i|i+1}, \hat{u}_{i|i+1}\}$ into the definition (38), we re-establish after some straightforward algebra the time- and measurement-update form (23)–(28) of the Kalman filter.

IV. Robust State-Space Estimation

Referring again to the state-space model (19) and (20), we shall first study the case of uncertainties in the matrices $\{F_i, G_i\}$ alone. The matrix H_i will be initially assumed known exactly. Later we shall address the general case (see Section IV-F). Thus, consider the uncertain model

$$x_{i+1} = (F_i + \delta F_i)x_i + (G_i + \delta G_i)u_i, \qquad i \geq 0, \tag{39}$$

$$y_i = H_ix_i + v_i \tag{40}$$

where the perturbations in $\{F_i, G_i\}$ are modeled as

$$[\delta F_i \quad \delta G_i] = M_i\Delta_i[E_{f,i} \quad E_{g,i}] \tag{41}$$

for some known matrices $\{M_i, E_{f,i}, E_{g,i}\}$ and for an arbitrary contraction Δ_i, $\|\Delta_i\| \leq 1$. Observe that for generality we are allowing the quantities $\{M_i, E_{f,i}, E_{g,i}\}$ to vary with time. The case of uncertainties in F_i only can be handled by setting $E_{g,i} = 0$. Likewise, the case of uncertainties in G_i only can be handled by setting $E_{f,i} = 0$. Finally, the case of accurate models is obtained by setting $M_i = 0$, $E_{f,i} = 0$, and $E_{g,i} = 0$.

Now assume that at step i we are given an *a priori* estimate for the state x_i. We shall denote this initial estimate by $\hat{x}_{i|i}$. Assume further that we are also given a positive-definite weighting matrix $P_{i|i}$, along with the observation at time $(i+1)$, i.e., y_{i+1}. Using this initial information, one can seek to update the estimate of x_i from $\hat{x}_{i|i}$ to $\hat{x}_{i|i+1}$ by solving

$$\min_{\{x_i, u_i\}} \max_{\delta F_i, \delta G_i} \|x_i - \hat{x}_{i|i}\|^2_{P^{-1}_{i|i}} + \|u_i\|^2_{Q_i^{-1}} + \|y_{i+1} - H_{i+1}x_{i+1}\|^2_{R^{-1}_{i+1}} \tag{42}$$

subject to (39), (40), and (41). Here, the weighting matrices $\{Q_i, R_{i+1}\}$ can be regarded as covariance matrices [as in (22)] with $\{u_i, v_i, x_0\}$ modeled as random variables, or simply as weighting matrices in a purely deterministic context.

Problem (42) can be seen to be the robust version of (35) in the same way that (5) and (6) is the robust version of (1). It can be interpreted as attempting to improve the estimate $\hat{x}_{i|i}$ of x_i by incorporating the additional information that is provided by y_{i+1} and by minimizing the worst-possible (regularized) squared residual norm.

Once the solutions $\{\hat{x}_{i|i+1}, \hat{u}_{i|i+1}\}$ have been found, we can use them to construct an estimate for the future state x_{i+1} as in (38). We shall also use the nominal data $\{F_i, G_i, H_i\}$, the uncertainty model parameters $\{M_i, E_{g,i}, E_{f,i}\}$, and the weighting matrices $\{Q_i, R_i, P_{i|i}\}$ to determine a weighting matrix $P_{i+1|i+1}$ for the next step. With $\{\hat{x}_{i+1|i+1}, P_{i+1|i+1}\}$ so determined we proceed to solve a similar problem at the next iteration and determine $\{\hat{x}_{i+2|i+2}, P_{i+2|i+2}\}$, and so on.

Let us now exhibit the recursions that characterize this construction. To begin with, we note that problem (42) can be written more compactly in the form (5) and (6) with the identifications

$$x \longleftarrow \text{col}\{x_i - \hat{x}_{i|i}\} \tag{43}$$
$$b \longleftarrow y_{i+1} - H_{i+1}F_i\hat{x}_{i|i} \tag{44}$$
$$A \longleftarrow H_{i+1}[F_i \quad G_i] \tag{45}$$
$$\delta A \longleftarrow H_{i+1}M_i\Delta_i[E_{f,i} \quad E_{g,i}] \tag{46}$$
$$\delta b \longleftarrow -H_{i+1}M_i\Delta_iE_{f,i}\hat{x}_{i|i} \tag{47}$$
$$Q \longleftarrow \left(P^{-1}_{i|i} \oplus Q_i^{-1}\right) \tag{48}$$
$$W \longleftarrow R^{-1}_{i+1} \tag{49}$$
$$H \longleftarrow H_{i+1}M_i \tag{50}$$
$$E_a \longleftarrow [E_{f,i} \quad E_{g,i}] \tag{51}$$
$$E_b \longleftarrow -E_{f,i}\hat{x}_{i|i} \tag{52}$$
$$\Delta \longleftarrow \Delta_i. \tag{53}$$

According to Theorem 1, the solution $\{\hat{x}_{i|i+1}, \hat{u}_{i|i+1}\}$ is then found by solving the system of equations

$$\left(\hat{Q} + A^T\hat{W}A\right)\hat{x} = \left(A^T\hat{W}b + \hat{\lambda}_iE_a^TE_b\right) \tag{54}$$

where

$$\hat{Q} = \begin{bmatrix} P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^T E_{f,i} & \hat{\lambda}_i E_{f,i}^T E_{g,i} \\ \hat{\lambda}_i E_{g,i}^T E_{f,i} & Q_i^{-1} + \hat{\lambda} E_{g,i}^T E_{g,i} \end{bmatrix} \quad (55)$$

and

$$\hat{W} \triangleq \hat{R}_{i+1}^{-1} = \left(R_{i+1} - \hat{\lambda}_i^{-1} H_{i+1} M_i M_i^T H_{i+1}^T\right)^{-1}. \quad (56)$$

Moreover, $\hat{\lambda}_i$ is the minimizing argument in the interval

$$\hat{\lambda}_i > \left\| M_i^T H_{i+1}^T R_{i+1}^{-1} H_{i+1} M_i \right\| \triangleq \lambda_{l,i} \quad (57)$$

of the corresponding scalar-valued function $G(\lambda)$ in (12) constructed with the identifications (43)–(53). The expression for $G(\lambda)$ is of course time-dependent and $G(\lambda)$ should be minimized over $\lambda \in (\lambda_{l,i}, \infty)$ at each iteration i. The time-dependency exists even for models with constant parameters $\{F, G, H, M, E_f, E_g\}$. This is because $G(\lambda)$ depends on both $P_{i|i}$ and $\hat{x}_{i|i}$. However, for simplicity of notation here, we have not indicated this time-dependence explicitly.

A. Time and Measurement-Update Form

Now substituting (55) and (56) into (54), we can solve for $\{\hat{x}_{i|i+1}, \hat{u}_{i|i+1}\}$. Substituting the resulting expressions into (38) we can establish, after some considerable algebra, the time- and measurement-update robust algorithm listed in Table I.

The major step in the algorithm is step 3, which consists of recursions that are very similar in nature to the time- and measurement-update form of the Kalman filter [cf. (23)–(28)]. The main difference is that the new recursions operate on modified parameters rather than on the given nominal values. This is where steps 1 and 2 in the algorithm are needed: they correct the parameters to the values necessary for the robust estimation step.[3] In the absence of any modeling uncertainties (i.e., with $M_i = 0$, $E_{f,i} = 0$, and $E_{g,i} = 0$), the recursions can be seen to collapse to those of the Kalman filter, as expected. Note further from the listing in Table I that $\hat{Q}_i^{-1} \geq Q_i^{-1}$ and $\hat{R}_{i+1} \leq R_{i+1}$.

TABLE I
LISTING OF THE ROBUST FILTERING ALGORITHM IN TIME- AND MEASUREMENT-UPDATE FORM

<u>Assumed uncertain model</u>: Eqs. (39)–(41). Also, $\Pi_0 > 0$, $R_i > 0$, $Q_i > 0$ are given weighting matrices.

<u>Initial conditions</u>: Set $\hat{x}_{0|0} = P_{0|0} H_0^T R_0^{-1} y_0$ and $P_{0|0} = \left(\Pi_0^{-1} + H_0^T R_0^{-1} H_0\right)^{-1}$

<u>Step 1.</u> If $H_{i+1} M_i = 0$, then set $\hat{\lambda}_i = 0$. Otherwise, construct $G(\lambda)$ of (12) with the identifications (43)–(53) and determine $\hat{\lambda}_i$ by minimizing $G(\lambda)$ over the interval

$$\hat{\lambda}_i > \lambda_{l,i} \triangleq \|M_i^T H_{i+1}^T R_{i+1}^{-1} H_{i+1} M_i\|.$$

<u>Step 2.</u> Replace $\{Q_i, R_{i+1}, P_{i|i}, G_i, F_i\}$ by:

$$\begin{aligned}
\hat{Q}_i^{-1} &= Q_i^{-1} + \hat{\lambda}_i E_{g,i}^T \left[I + \hat{\lambda}_i E_{f,i} P_{i|i} E_{f,i}^T\right]^{-1} E_{g,i} \\
\hat{R}_{i+1} &= R_{i+1} - \hat{\lambda}_i^{-1} H_{i+1} M_i M_i^T H_{i+1}^T \\
\hat{P}_{i|i} &= \left(P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^T E_{f,i}\right)^{-1} \\
&= P_{i|i} - P_{i|i} E_{f,i}^T (\hat{\lambda}_i^{-1} I + E_{f,i} P_{i|i} E_{f,i}^T)^{-1} E_{f,i} P_{i|i} \\
\hat{G}_i &= G_i - \hat{\lambda}_i F_i \hat{P}_{i|i} E_{f,i}^T E_{g,i} \\
\hat{F}_i &= (F_i - \hat{\lambda}_i \hat{G}_i \hat{Q}_i E_{g,i}^T E_{f,i})(I - \hat{\lambda}_i \hat{P}_{i|i} E_{f,i}^T E_{f,i})
\end{aligned}$$

If $\hat{\lambda}_i = 0$, then simply set $\hat{Q}_i = Q_i$, $\hat{R}_{i+1} = R_{i+1}$, $\hat{P}_{i|i} = P_{i|i}$, $\hat{G}_i = G_i$, and $\hat{F}_i = F_i$.

<u>Step 3.</u> Update $\{\hat{x}_{i|i}, P_{i|i}\}$ as follows:

$$\begin{aligned}
\hat{x}_{i+1} &= \hat{F}_i \hat{x}_{i|i} \\
\hat{x}_{i+1|i+1} &= \hat{x}_{i+1} + P_{i+1|i+1} H_{i+1}^T \hat{R}_{i+1}^{-1} e_{i+1} \\
e_{i+1} &= y_{i+1} - H_{i+1} \hat{x}_{i+1} \\
P_{i+1} &= F_i \hat{P}_{i|i} F_i^T + \hat{G}_i \hat{Q}_i \hat{G}_i^T \\
P_{i+1|i+1} &= P_{i+1} - P_{i+1} H_{i+1}^T R_{e,i+1}^{-1} H_{i+1} P_{i+1} \\
R_{e,i+1} &= \hat{R}_{i+1} + H_{i+1} P_{i+1} H_{i+1}^T
\end{aligned}$$

B. The Prediction Form

The recursions of Table I can be manipulated into an alternative so-called prediction form, which propagates the quantities $\{\hat{x}_i, P_i\}$ directly. Thus note that by combining the equations for $\hat{x}_{i+1}$ and $\hat{x}_{i|i}$ from Table I we obtain

$$\hat{x}_{i+1} = \hat{F}_i \hat{x}_i + \hat{F}_i P_{i|i} H_i^T \hat{R}_i^{-1} [y_i - H_i \hat{x}_i], \qquad \hat{x}_0 = 0.$$

Now, a simple algebra will show that $P_{i|i} H_i^T \hat{R}_i^{-1} = P_i H_i^T R_{e,i}^{-1}$ so that we also have

$$\hat{x}_{i+1} = \hat{F}_i \hat{x}_i + \hat{F}_i P_i H_i^T R_{e,i}^{-1} [y_i - H_i \hat{x}_i].$$

The recursion for P_i is obtained as follows. Using the equation for P_{i+1} from Table I we get

$$P_{i+1} = F_i \left[P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^T E_{f,i}\right]^{-1} F_i^T + \hat{G}_i \hat{Q}_i \hat{G}_i^T$$

which in view of the equality

$$P_{i|i}^{-1} = P_i^{-1} + H_i^T \hat{R}_i^{-1} H_i$$

becomes

$$P_{i+1} = F_i \left[P_i^{-1} + H_i^T \hat{R}_i^{-1} H_i + \hat{\lambda}_i E_{f,i}^T E_{f,i}\right]^{-1} F_i^T + \hat{G}_i \hat{Q}_i \hat{G}_i^T.$$

Define the extended matrix

$$\overline{H}_i \triangleq \begin{bmatrix} \hat{R}_i^{-1/2} H_i \\ \sqrt{\hat{\lambda}_i} E_{f,i} \end{bmatrix}$$

where the notation $A^{1/2}$ denotes a square-root factor of a positive-definite matrix A. Then by applying the matrix inversion lemma to the term $P_i^{-1} + H_i^T \hat{R}_i^{-1} H_i + \hat{\lambda}_i E_{f,i}^T E_{f,i}$ we arrive at the following recursion for P_{i+1}:

$$P_{i+1} = F_i P_i F_i^T - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^T + \hat{G}_i \hat{Q}_i \hat{G}_i^T \quad (58)$$

where

$$\overline{K}_i = F_i P_i \overline{H}_i^T, \qquad \overline{R}_{e,i} = I + \overline{H}_i P_i \overline{H}_i^T,$$

and with initial conditions $P_0 = \Pi_0$ and $\hat{R}_0 = R_0$.

[3] Recall from (4) that we are assuming $H_{i+1} M_i \neq 0$ since otherwise problem (42) reduces to the standard Kalman filtering step (35). If at any particular iteration we encounter $H_{i+1} M_i = 0$, then we simply set $\hat{\lambda}_i = 0$ and replace $\hat{\lambda}_i^{-1}$ by $\hat{\lambda}_i^{\dagger}$ (and hence also by zero), so that all the terms involving $\hat{\lambda}_i$ or its inverse disappear from the recursions of Table I and they reduce, at that iteration, to the Kalman filter recursions.

TABLE II
LISTING OF THE ROBUST FILTERING ALGORITHM IN PREDICTION FORM

Assumed uncertain model. Same as in Table 1.

Initial conditions: $\hat{x}_0 = 0$, $P_0 = \Pi_0$, and $\widehat{R}_0 = R_0$.

Step 1a. Using $\{\widehat{R}_i, H_i, P_i\}$ compute $P_{i|i}$:

$$\begin{aligned} P_{i|i} &= (P_i^{-1} + H_i^T \widehat{R}_i^{-1} H_i)^{-1} \\ &= P_i - P_i H_i^T (\widehat{R}_i + H_i P_i H_i^T)^{-1} H_i P_i \end{aligned}$$

Step 1b. Same as step 1 of Table 1.

Step 2. Determine $\{\widehat{Q}_i, \widehat{R}_i, \widehat{F}_i, \widehat{G}_i\}$ as in step 2 of Table 1.

Step 3. Now update $\{\hat{x}_i, P_i\}$ to $\{\hat{x}_{i+1}, P_{i+1}\}$ as follows:

$$\begin{aligned} \hat{x}_{i+1} &= \widehat{F}_i \hat{x}_i + \widehat{F}_i P_i H_i^T R_{e,i}^{-1} e_i \\ e_i &= y_i - H_i \hat{x}_i \\ P_{i+1} &= F_i P_i F_i^T - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^T + \widehat{G}_i \widehat{Q}_i \widehat{G}_i^T \\ \overline{K}_i &= F_i P_i \overline{H}_i^T, \quad \overline{R}_{e,i} = I + \overline{H}_i P_i \overline{H}_i^T \\ \overline{H}_i^T &= \left[\begin{array}{cc} H_i^T \widehat{R}_i^{-T/2} & \sqrt{\hat{\lambda}_i} E_{f,i}^T \end{array} \right] \end{aligned}$$

We should remark that (58) is not a standard Riccati recursion since the product $\hat{G}_i\hat{Q}_i\hat{G}_i^T$ is dependent on P_i (through its dependence on $\hat{P}_{i|i}$). In Section IV-E, we shall see that (58) collapses to the standard form of a discrete-time Riccati recursion in some useful special cases. Table II summarizes the prediction form of the robust algorithm.

C. The Information Form

The algorithm can also be rewritten in an alternative so-called information form that propagates the inverses of the matrices $P_{i|i}$ rather than the matrices themselves. This form requires the invertibility of F_i. Thus, note that it follows from the time and measurement-update form of Table I that

$$\hat{x}_{i+1|i+1} = \hat{F}_i \hat{x}_{i|i} + P_{i+1|i+1} H_{i+1}^T \hat{R}_{i+1}^{-1} \left[y_{i+1} - H_{i+1} \hat{F}_i \hat{x}_{i|i} \right]$$

so that if we multiply both sides of the above equation by $P_{i+1|i+1}^{-1}$ from the left we obtain the first recursion in Step 3 of Table III, which propagates the normalized state estimate $P_{i|i}^{-1}\hat{x}_{i|i}$. In addition, by inverting the equations for $P_{i+1|i+1}$ and P_{i+1} from Table I we obtain the algorithm listed in Table III.

D. Special Cases

The recursions in all forms can be simplified in some useful special cases: $E_{g,i} = 0$ (i.e., no uncertainty in G_i), $E_{f,i}^T E_{g,i} = 0$ (i.e., the uncertainty in G_i is orthogonal to that in F_i), and $E_{f,i} = 0$ (i.e., no uncertainty in F_i).

In the first case ($E_{g,i} = 0$), it is easy to see that the expressions for $\{\hat{Q}_i, \hat{G}_i, \hat{F}_i\}$ simplify to the following

$$\hat{Q}_i = Q_i, \quad \hat{G}_i = G_i, \quad \hat{F}_i = F_i \left(I - \hat{\lambda}_i \hat{P}_{i|i} E_{f,i}^T E_{f,i} \right).$$

That is, only F_i and R_{i+1} are corrected to $\hat{F}_i$ and $\hat{R}_{i+1}$, respectively.

TABLE III
LISTING OF THE ROBUST FILTERING ALGORITHM IN INFORMATION FORM

Assumed uncertain model. Same as in Table 1 with the additional assumption that F_i is invertible.

Initial conditions: $P_{0|0}^{-1}\hat{x}_{0|0} = H_0^T R_0^{-1} y_0$, $P_{0|0}^{-1} = \Pi_0^{-1} + H_0^T R_0^{-1} H_0$.

Steps 1 and 2. Same as in Table 1.

Step 3. Update $\{P_{i|i}^{-1}\hat{x}_{i|i}, P_{i|i}\}$ as:

$$P_{i+1|i+1}^{-1}\hat{x}_{i+1|i+1} = H_{i+1}^T \widehat{R}_{i+1}^{-1} y_{i+1} + $$

$$+ \left[\left(P_{i+1|i+1}^{-1} - H_{i+1}^T \widehat{R}_{i+1}^{-1} H_{i+1} \right) \widehat{F}_i P_{i|i} \right] P_{i|i}^{-1} \hat{x}_{i|i}.$$

where

$$\begin{aligned} P_{i+1|i+1}^{-1} &= F_i^{-T} \widehat{P}_{i|i}^{-1} F_i^{-1} - K_{\nu,i} R_{\nu,i}^{-1} K_{\nu,i}^T + \\ &\quad + H_{i+1}^T \widehat{R}_{i+1}^{-1} H_{i+1} \\ K_{\nu,i} &= F_i^{-T} \widehat{P}_{i|i}^{-1} F_i^{-1} \widehat{G}_i \\ R_{\nu,i} &= \widehat{Q}_i^{-1} + \widehat{G}_i^T F_i^{-T} \widehat{P}_{i|i}^{-1} F_i^{-1} \widehat{G}_i \end{aligned}$$

In the second case ($E_{f,i}^T E_{g,i} = 0$), it also follows that the expressions for $\{\hat{G}_i, \hat{F}_i\}$ simplify to the same values as above while the expression for $\hat{Q}_i$ can be seen to become

$$\hat{Q}_i = \left(Q_i^{-1} + \hat{\lambda}_i E_{g,i}^T E_{g,i} \right)^{-1}. \tag{59}$$

In the third case ($E_{f,i} = 0$), it follows that $\hat{G}_i = G_i$, $\hat{F}_i = F_i$, and $\hat{Q}_i$ as in (59).

E. The Correction Parameter $\hat{\lambda}_i$

The algorithms of Tables I–III require, at each iteration i, the minimization of a cost function $G(\lambda)$ over the interval $\lambda \in (\lambda_{l,i}, \infty)$, with the lower bound as defined in Table I. As remarked earlier in Section II-B, a reasonable approximation for $\hat{\lambda}_i$ is to set it equal to a multiple of the lower bound, say

$$\hat{\lambda}_i = \begin{cases} (1+\alpha)\lambda_{l,i} & \text{if } \lambda_{l,i} \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{60}$$

for some tuning parameter α (that could be chosen to be time-variant as well); see Fig. 4 for an example comparing a filter implementation that is based on the optimal $\hat{\lambda}_i$ with an implementation that is based on the above approximation for the choice $\alpha = 0.5$.

Using the approximation (60), it is easy to see that the recursion (58) now becomes a standard Riccati recursion in the three special cases considered in the previous section (*viz.*, $E_{g,i} = 0$, $E_{f,i} = 0$, or $E_{f,i}^T E_{g,i} = 0$).

F. Uncertainties in the Measurement Equation

The discussions in the earlier subsections can be extended to the case of uncertainties in all matrices $\{F_i, G_i, H_i\}$, i.e.,

$$\begin{aligned} x_{i+1} &= (F_i + \delta F_i)x_i + (G_i + \delta G_i)u_i \\ y_i &= (H_i + \delta H_i)x_i + v_i. \end{aligned}$$

An examination of the derivation following (42) will reveal that it is the uncertainties in the products $H_{i+1}F_i$ and $H_{i+1}G_i$ that are of interest rather than the individual uncertainties in $\{H_{i+1}, F_i, G_i\}$. Hence, we shall assume that the uncertainties $\{\delta H_i, \delta F_i, \delta G_i\}$ are such that the resulting uncertainties in the products $H_{i+1}F_i$ and $H_{i+1}G_i$ can be modeled as

$$[\,\delta(H_{i+1}F_i) \quad \delta(H_{i+1}G_i)\,] = M_i\Delta_i[\,E_{f,i} \quad E_{g,i}\,]$$

for some known $\{M_i, E_{f,i}, E_{g,i}\}$ and for an arbitrary contraction Δ_i. The identifications (43)–(53) remain valid except for the following three entries:

$$\begin{aligned} \delta A &\longleftarrow M_i\Delta_i[\,E_{f,i} \quad E_{g,i}\,] \\ \delta b &\longleftarrow -M_i\Delta_i E_{f,i}\hat{x}_{i|i} \\ H &\longleftarrow M_i. \end{aligned}$$

In other words, whenever the product $H_{i+1}M_i$ appears in any of the previous recursions it should be replaced by M_i. This affects only two entries in any of the listings of Tables I–III. First, the lower bound on $\hat{\lambda}_i$ now becomes

$$\hat{\lambda}_i > \left\|M_i^T R_i^{-1} M_i\right\| = \lambda_{l,i}.$$

And, second, the definition of $\hat{R}_{i+1}$ is modified to

$$\hat{R}_{i+1} = R_{i+1} - \hat{\lambda}_i^{-1} M_i M_i^T.$$

All other recursions remain unchanged. We shall therefore ignore this extension in the sequel due to its similarity with the cases studied so far.

V. Steady-State Results

We now examine the steady-state performance of the proposed filters when the model parameters are constant, say $\{F, G, H, M, E_f, E_g, Q, R\}$. Only the contraction Δ_i is allowed to vary with time. In particular we shall establish that under certain detectability and stabilizability assumptions the steady-state filters are stable and that, in addition, for quadratically stable models they guarantee bounded error variances.

A. Stability

We consider first the case that involves uncertainties in the system dynamics only. That is, we consider an uncertain model of the form

$$x_{i+1} = (F + \delta F_i)x_i + Gu_i, \qquad i \geq 0 \tag{61}$$

$$y_i = Hx_i + v_i \tag{62}$$

$$\delta F_i = M\Delta_i E_f \tag{63}$$

where only Δ_i (and, hence, δF_i) is allowed to change with time. This model is often studied in the literature of robust filtering.

We further assume that the correction parameter $\hat{\lambda}_i$ is set to a constant value that is equal to a multiple of the admissible lower bound, i.e.,

$$\hat{\lambda}_i = (1+\alpha)\lambda_l = (1+\alpha)\left\|M^T H^T R^{-1} H M\right\| \triangleq \hat{\lambda} \tag{64}$$

for some $\alpha > 0$ chosen by the designer, and for all i.

The prediction form of the robust filter in this case becomes (cf. Table II):

$$\begin{aligned} \hat{x}_{i+1} &= \hat{F}_i\hat{x}_i + \hat{F}_i P_i H^T R_{e,i}^{-1}[y_i - H\hat{x}_i] \\ &= \hat{F}_i\left[I - P_i H^T R_{e,i}^{-1} H\right]\hat{x}_i + \hat{F}_i P_i H^T R_{e,i}^{-1} y_i \end{aligned}$$

where

$$\begin{aligned} \hat{R} &= R - \hat{\lambda}^{-1} HMM^T H^T \\ R_{e,i} &= \hat{R} + HP_iH^T \\ P_{i+1} &= FP_iF^T - \overline{K}_i\overline{R}_{e,i}^{-1}\overline{K}_i^T + GQG^T \\ \overline{K}_i &= FP_i\overline{H}^T \\ \overline{R}_{e,i} &= I + \overline{H}P_i\overline{H}^T \\ \overline{H}^T &= \left[\,H^T\hat{R}^{-T/2} \quad \sqrt{\hat{\lambda}}E_f^T\,\right] \end{aligned}$$

and

$$\begin{aligned} \hat{F}_i &= F\left[I - \hat{\lambda}\left(P_i^{-1} + H^T\hat{R}^{-1}H + \hat{\lambda}E_f^TE_f\right)^{-1}E_f^TE_f\right] \\ &= F\left[I - \hat{\lambda}\left(P_i^{-1} + \overline{H}^T\overline{H}\right)^{-1}E_f^TE_f\right] \\ &= F\left[I - \hat{\lambda}\left(I + P_i\overline{H}^T\overline{H}\right)^{-1}P_iE_f^TE_f\right] \end{aligned}$$

where the last form is independent of P_i^{-1}.

Note that even though the coefficient matrix F is constant, the matrix that appears multiplying $\hat{x}_i$ is time-variant and equal to $\hat{F}_i$. This is in contrast to a Kalman filtering implementation.

Lemma 1 (Two Useful Identities): Let

$$\begin{aligned} F_{c,i} &\triangleq F\left[I - P_i\overline{H}^T\overline{R}_{e,i}^{-1}\overline{H}\right] \\ F_{p,i} &\triangleq \hat{F}_i\left[I - P_iH^TR_{e,i}^{-1}H\right]. \end{aligned}$$

Then $F_{c,i} = F_{p,i}$ and $\hat{F}_iP_iH^TR_{e,i}^{-1} = F_{p,i}P_iH^T\hat{R}^{-1}$.

Proof: The algebra is omitted. ◇

Using the second identity in the lemma, we can rewrite the recursion for the state estimate as

$$\hat{x}_{i+1} = F_{p,i}\hat{x}_i + F_{p,i}P_iH^T\hat{R}^{-1}y_i. \tag{65}$$

We are now in a position to establish the main result of this section concerning the convergence of the robust filter to a stable steady-state filter.

Theorem 2 (Stable Steady-State Filter): Consider the uncertain state-space model (61)–(63) with the corresponding robust filter (65). Assume further that $\{F, \overline{H}\}$ is detectable and $\{F, GQ^{1/2}\}$ is stabilizable. Then, for any initial condition $P_0 = \Pi_0 > 0$ and for any $\alpha > 0$ in (64), the Riccati variable P_i tends to the unique stabilizing and positive semi-definite solution P of the DARE

$$P = FPF^T - FP\overline{H}^T\left(I + \overline{H}P\overline{H}^T\right)^{-1}\overline{H}PF^T + GQG^T.$$

The solution P is stabilizing in the sense that the steady-state closed-loop matrix

$$F_p \triangleq \hat{F}\left[I - PH^TR_e^{-1}H\right]$$

is stable, where

$$\begin{aligned} \hat{F} &= F\left[I - \hat{\lambda}\left(P - P\overline{H}^T\overline{R}_e^{-1}\overline{H}P\right)E_f^TE_f\right] \\ R_e &= \hat{R} + HPH^T, \quad \overline{R}_e = I + \overline{H}P\overline{H}^T. \end{aligned}$$

Proof: The condition $\alpha > 0$ guarantees a positive-definite matrix $\hat{R}$ so that its Cholesky factor, and hence $\overline{H}$, are well defined. Now the detectability of $\{F, \overline{H}\}$ and the stabilizability of $\{F, GQ^{1/2}\}$ are known to guarantee the convergence of P_i to the unique positive semi-definite solution P of the DARE that stabilizes the following matrix (see, e.g., [2])

$$F_c \triangleq F\left[I - P\overline{H}^T\left(I + \overline{H}P\overline{H}^T\right)^{-1}\overline{H}\right].$$

However, we know from the first identity in the previous lemma that this matrix coincides with F_p, and the result is therefore established. $\diamond$

A similar conclusion can be obtained for the uncertain model

$$x_{i+1} = (F + \delta F_i)x_i + (G + \delta G_i)u_i \quad (66)$$
$$y_i = Hx_i + v_i \quad (67)$$
$$[\,\delta F_i \quad \delta G_i\,] = M\Delta_i[\,E_f \quad E_g\,] \quad (68)$$
$$E_f^T E_g = 0 \quad (69)$$

with uncertainties in both F and G that satisfy $E_f^T E_g = 0$. In this case, the same recursions as above will hold with the only exception that the term GQG^T in the recursion for P_{i+1} should be replaced by $G\hat{Q}G^T$ where

$$\hat{Q} = \left(Q^{-1} + \hat{\lambda}E_g E_g^T\right)^{-1}.$$

The identities of Lemma 1 will continue to hold, as well as the conclusion of Theorem 2.

B. Bounded Steady-State Error Variances

We continue with the model (61)–(63) and further assume that it is quadratically stable, i.e., that there exists a positive–definite matrix V such that

$$V - [F + M\Delta E_f]^T V [F + M\Delta E_f] > 0$$

for all contractions Δ. By the small gain theorem of [44] and [45], the quadratic stability requirement is equivalent to the combined conditions of a stable F *and* a bounded norm $\|E_f(zI - F)^{-1}M\|_\infty < 1$.[4]

For such systems we can show that the steady-state robust filter of the previous section guarantees a bounded error variance. To see this, we introduce the estimation error $\tilde{x}_i = x_i - \hat{x}_i$. Then, subtracting the equations

$$x_{i+1} = (F + M\Delta_i E_f)x_i + Gu_i$$
$$\hat{x}_{i+1} = F_p\hat{x}_i + F_p PH^T\hat{R}^{-1}[Hx_i + v_i]$$

we arrive at the following extended state equation:

$$\begin{bmatrix} \tilde{x}_{i+1} \\ \hat{x}_{i+1} \end{bmatrix} = (\mathcal{F} + \delta\mathcal{F}_i)\begin{bmatrix} \tilde{x}_i \\ \hat{x}_i \end{bmatrix} + \mathcal{G}\begin{bmatrix} u_i \\ v_i \end{bmatrix} \quad (70)$$

where

$$\mathcal{F} + \delta\mathcal{F}_i = \begin{bmatrix} F - F_pPH^T\hat{R}^{-1}H & F - F_p - F_pPH^T\hat{R}^{-1}H \\ F_pPH^T\hat{R}^{-1}H & F_p + F_pPH^T\hat{R}^{-1}H \end{bmatrix} + \begin{bmatrix} M\Delta_i E_f & M\Delta_i E_f \\ 0 & 0 \end{bmatrix}$$

and

$$\mathcal{G} = \begin{bmatrix} G & -F_pPH^T\hat{R}^{-1} \\ 0 & F_pPH^T\hat{R}^{-1} \end{bmatrix}.$$

Lemma 2 (Stability of Extended System): The extended model (70) is quadratically stable.

Proof: Introduce the similarity transformation

$$\mathcal{T} = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}, \quad \mathcal{T}^{-1} = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}.$$

[4]Here, $\|\cdot\|_\infty$ denotes the peak singular value of its argument over values of z on the unit circle.

The system matrices $\{\mathcal{F} + \delta\mathcal{F}_i,\, \mathcal{G}\}$ then reduce to

$$\mathcal{T}(\mathcal{F} + \delta\mathcal{F}_i)\mathcal{T}^{-1} = \begin{bmatrix} F & 0 \\ F_pPH^T\hat{R}^{-1}H & F_p \end{bmatrix} + \begin{bmatrix} M\Delta_i E_f & 0 \\ 0 & 0 \end{bmatrix}$$
$$\mathcal{T}\mathcal{G} = \begin{bmatrix} G & 0 \\ 0 & F_pPH^T\hat{R}^{-1} \end{bmatrix}.$$

The stability of F and F_p now guarantees that the nominal matrix

$$\begin{bmatrix} F & 0 \\ F_pPH^T\hat{R}^{-1}H & F_p \end{bmatrix}$$

is stable. Moreover, the equality

$$E_f(zI - F)^{-1}M = [\,E_f \quad 0\,]\begin{bmatrix} zI - F & 0 \\ -F_pPH^T\hat{R}^{-1}H & zI - F_p \end{bmatrix}^{-1}\begin{bmatrix} M \\ 0 \end{bmatrix}$$

shows that the matrix function on the left-hand side has $\mathcal{H}_\infty$-norm strictly less than one. We thus conclude that the extended system (70) is quadratically stable. $\diamond$

By the result of the above lemma, we conclude that there exists a positive–definite matrix $\mathcal{V}$ such that

$$\mathcal{V} - (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{V}(\mathcal{F} + \delta\mathcal{F}_i)^T > 0$$

for any Δ_i. Now let $\mathcal{M}_i$ denote the covariance matrix of the extended vector $\text{col}\{\tilde{x}_i,\, \hat{x}_i\}$

$$\mathcal{M}_i \triangleq E\begin{bmatrix} \tilde{x}_i \\ \hat{x}_i \end{bmatrix}\begin{bmatrix} \tilde{x}_i \\ \hat{x}_i \end{bmatrix}^T.$$

It follows from (70) that $\mathcal{M}_i$ satisfies the Lyapunov recursion

$$\mathcal{M}_{i+1} = (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{M}_i(\mathcal{F} + \delta\mathcal{F}_i)^T + \mathcal{G}\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\mathcal{G}^T.$$

Using arguments that are common in guaranteed-cost designs (e.g., as in [14], [24]), it is now immediate to establish the following conclusion.

Theorem 3 (Bounded Error-Variance): Under the conditions of Theorem 2, and for a quadratically stable model (61)–(63), the variance of the estimation error of the steady-state robust filter satisfies

$$\lim_{i\to\infty} E\tilde{x}_i\tilde{x}_i^T \le \mathcal{P}_{11}$$

where $\mathcal{P}_{11}$ is the (1, 1)block entry with the smallest trace among all (1, 1) block entries of positive–definite matrices $\mathcal{P}$ that satisfy the inequality

$$\mathcal{P} - \left(\mathcal{F} + \begin{bmatrix} M \\ 0 \end{bmatrix}\Delta[\,E_f \quad E_f\,]\right)\mathcal{P}(\cdot)^T - \mathcal{G}\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\mathcal{G}^T \ge 0$$

for all contractive matrices Δ.

Proof: We use an argument similar to the one in [5, pp. 39, 40]. The existence of $\mathcal{V} > 0$ guarantees the existence of a positive scaling parameter μ such that

$$\mu\mathcal{V} - (\mathcal{F} + \delta\mathcal{F}_i)\mu\mathcal{V}(\cdot)^T - \mathcal{G}\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\mathcal{G}^T > 0$$

so that a $\mathcal{P}$ exists ($\mathcal{P} = \mu\mathcal{V}$) satisfying

$$\mathcal{P} \ge (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{P}(\mathcal{F} + \delta\mathcal{F}_i)^T + \mathcal{G}\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\mathcal{G}^T.$$

Subtracting this inequality from the recursion for $\mathcal{M}_i$ we get

$$\mathcal{P} - \mathcal{M}_{i+1} \ge (\mathcal{F} + \delta\mathcal{F}_i)(\mathcal{P} - \mathcal{M}_i)(\mathcal{F} + \delta\mathcal{F}_i)^T$$

or, equivalently

$$\mathcal{P}-\mathcal{M}_{i+1}=(\mathcal{F}+\delta\mathcal{F}_i)(\mathcal{P}-\mathcal{M}_i)(\mathcal{F}+\delta\mathcal{F}_i)^T+\mathcal{Q}_i$$

for some $\mathcal{Q}_i \geq 0$. The quadratic stability of $\mathcal{F}+\delta\mathcal{F}_i$ then implies that in the limit, as $i\to\infty$, $\mathcal{P}-\mathcal{M}_{i+1}\geq 0$ or $\mathcal{M}_{i+1}\leq\mathcal{P}$. $\diamondsuit$

VI. Comparisons with Other Robust Filters

In this section, we compare the proposed recursions with other robust filters and in the next section we provide some simulation results. We focus on the following uncertain state-space model

$$x_{i+1}=(F_i+\delta F_i)x_i+G_iu_i \tag{71}$$

$$y_i=H_ix_i+v_i \tag{72}$$

$$\delta F_i=M_i\Delta_iE_{f,i} \tag{73}$$

since it is often studied in the literature on robust filtering.

A. Kalman Filtering

We already explained before the differences and similarities between the proposed recursions and those of the Kalman filter. For instance, while both solutions do not require existence conditions, the new recursions operate on modified parameters that take into account the model uncertainties.

B. $\mathcal{H}_\infty$ Filtering

The *a-priori* central $\mathcal{H}_\infty$ filter of level $\gamma>0$ for estimating the combination $s_i=L_ix_i$ of the state vector is given by (see, e.g., [8], [11])

$$\begin{aligned}
\hat{x}_{i+1}&=F_i\hat{x}_i+F_i\tilde{P}_iH_i^T\left(I+H_i\tilde{P}_iH_i^T\right)^{-1}[y_i-H_i\hat{x}_i]\\
\tilde{P}_i^{-1}&=P_i^{-1}-\gamma^{-2}L_i^TL_i\\
P_{i+1}&=F_iP_iF_i^T+G_iQ_iG_i^T-K_iR_{e,i}^{-1}K_i^T\\
R_{e,i}&=\begin{bmatrix}I&0\\0&-\gamma^2I\end{bmatrix}+\begin{bmatrix}H_i\\L_i\end{bmatrix}P_i\,[\,H_i^T\quad L_i^T\,]\\
K_i&=F_iP_i\,[\,H_i^T\quad L_i^T\,]\\
\hat{s}_i&=L_i\hat{x}_i.
\end{aligned}$$

This filter guarantees the following bound:

$$\sup_{\{x_0,u_j,v_j\}}\frac{\sum_{i=0}^{N}|\hat{s}_i-L_ix_i|^2}{x_0^T\Pi_0^{-1}x_0+\sum_{i=0}^{N-1}u_i^TQ_i^{-1}u_i+\sum_{i=0}^{N-1}v_i^Tv_i}\leq\gamma^2$$

for all $0\leq i\leq N$, if and only if, the following conditions are satisfied:

$$P_i^{-1}-\gamma^{-2}L_i^TL_i>0\quad\text{for } 0\leq i\leq N. \tag{74}$$

In general, there is no guarantee beforehand that these conditions will be satisfied for all iterations. In addition, even if the conditions are satisfied over a horizon of length N, there is no guarantee that they will be satisfied over a horizon of longer duration. This is an inconvenience for online operation since when condition (74) fails at any particular iteration, the $\mathcal{H}_\infty$ performance is lost and divergence can occur (see the simulations in Section VII. Note also that the larger the value of γ the easier it is to satisfy (74). However, larger values of γ reduce the robustness of the filter.

C. Set-Valued Estimation

The set-valued approach to state-estimation [18] is based on determining ellipsoids around the state estimates that are consistent with the measured data. The centers of the ellipsoids are taken as estimates for the states. There have been numerous works in the literature on extending this approach to deal with uncertain models, e.g., of the form (71)–(73)—see [20], [21]–[46]. In particular, the following recursions are from [46] (we are intentionally using notation that is similar to ours although the variables $\{P_{i|i},\,K_{\nu,i},\,R_{\nu,i}\}$ below have of course different meanings from the ones we introduced in Table III for our recursions):

$$\begin{aligned}
P_{i+1|i+1}^{-1}\hat{x}_{i+1|i+1}&=H_{i+1}^TR_{i+1}^{-1}y_{i+1}\\
&\quad+\left(F_i^{-T}-K_{\nu,i}R_{\nu,i}^{\dagger}\begin{bmatrix}G_i^T\\M_i^T\end{bmatrix}F_i^{-T}\right)P_{i|i}^{-1}\hat{x}_{i|i}\\
P_{i+1|i+1}^{-1}&=F_i^{-T}P_{i|i}^{-1}F_i^{-1}-K_{\nu,i}R_{\nu,i}^{\dagger}K_{\nu,i}^T\\
&\quad+H_{i+1}^TR_{i+1}^{-1}H_{i+1}-E_{f,i+1}^TE_{f,i+1}\\
K_{\nu,i}&=F_i^{-T}P_{i|i}^{-1}F_i^{-1}[\,G_i\quad M_i\,]\\
R_{\nu,i}&=\begin{bmatrix}Q_i^{-1}&\\&I\end{bmatrix}+\begin{bmatrix}G_i^T\\M_i^T\end{bmatrix}F_i^{-T}P_{i|i}^{-1}F_i^{-1}[\cdot].
\end{aligned}$$

Some straightforward algebra will show that the recursion for the state estimate given above is equivalent to

$$\begin{aligned}
&P_{i+1|i+1}^{-1}\hat{x}_{i+1|i+1}\\
&\quad=H_{i+1}^TR_{i+1}^{-1}y_{i+1}\\
&\qquad+\Big[\Big(P_{i+1|i+1}^{-1}-H_{i+1}^TR_{i+1}^{-1}H_{i+1}+E_{f,i+1}^TE_{f,i+1}\Big)\\
&\qquad\quad\cdot F_iP_{i|i}\Big]P_{i|i}^{-1}\hat{x}_{i|i}.
\end{aligned}$$

The above equations can now be compared with those in Table III and the differences will become evident.

One issue here is that the above recursions will constitute a robust set-valued estimation solution only under certain conditions on the data. For example, it is necessary that, for all i,

$$R_{\nu,i}\geq 0\quad\text{and}\quad\mathcal{N}(R_{\nu,i})\subset\mathcal{N}(K_{\nu,i}) \tag{75}$$

where $\mathcal{N}(\cdot)$ denotes the nullspace of its argument. These two conditions may sometimes be difficult to satisfy in practice. Note, for example, that a negative term $-E_{f,i+1}^TE_{f,i+1}$ is subtracted from the right-hand side of the recursion for $P_{i+1|i+1}^{-1}$ above. This term can be large and it can lead to indefinite initial values for $R_{\nu,i}$ that ultimately cause filter breakdown, as can be seen from the simulations of Fig. 1.

D. Guaranteed-Cost Designs

The idea of guaranteed-cost designs is to develop filters that guarantee an upper bound on the variance of the estimation error. Such designs have been studied mostly for quadratically-stable time-invariant models in steady-state operation (e.g., [14], [23], [24]), though, as mentioned in the introduction, extensions exist to time-variant and finite-horizon scenarios. The guaranteed-cost filters in [14], [24] have different forms, which is expected since each tries to enforce an upper bound in a particular manner.

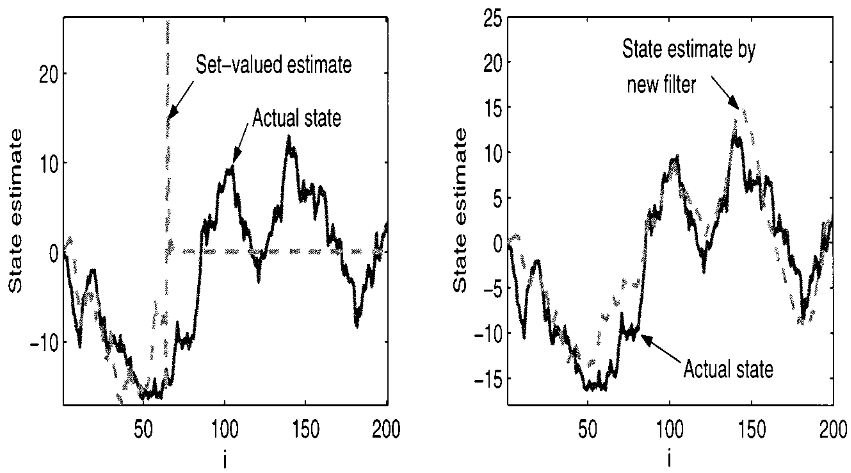

Fig. 1. The top plots show three trajectories: actual state (solid line), state estimate by robust set-valued algorithm (dashed line in the leftmost plot), and state estimate by the new filter (dashed line in rightmost plot).

The filter equations from [14, p. 44] require $G = I$ and they can be summarized as follows. First, the designer selects a small positive parameter ϵ (as explained below). Then uses

$$\begin{aligned}\hat{x}_{i+1} =& F\left[I + \epsilon\left(P^{-1} + H^T R^{-1} H - \epsilon E_f^T E_f\right)^{-1} E_f^T E_f\right]\hat{x}_i \\ &+ F\left(P^{-1} + H^T R^{-1} H - \epsilon E_f^T E_f\right)^{-1} \\ &\cdot H^T R^{-1}(y_i - H\hat{x}_i)\end{aligned} \tag{76}$$

where P is taken as the positive–definite stabilizing solution of the following Riccati equation:

$$\begin{aligned}P =& F\left(P^{-1} + H^T R^{-1} H - \epsilon E_f^T E_f\right)^{-1} F^T \\ &+ Q + \epsilon^{-1} M M^T.\end{aligned} \tag{77}$$

The value of ϵ is picked from within an open interval $(0, \epsilon^o)$, where $\epsilon^o > 0$ is chosen such that the following additional Riccati equation

$$\overline{P} = F\left(\overline{P}^{-1} - \epsilon^o E_f^T E_f\right)^{-1} F^T + Q + \epsilon^{-o} M M^T$$

has a positive–definite stabilizing solution satisfying

$$\overline{P}^{-1} - \epsilon^o E_f^T E_f > 0.$$

This condition guarantees $P^{-1} - \epsilon E_f^T E_f > 0$ since it can be shown that $P \leq \overline{P}$. [Such an ϵ^o is always guaranteed to exist for quadratically stable models.]

In order to compare our steady-state filter with the above equations, we shall show first how to rewrite our prediction filter in a form that is close to the above. Thus refer to Section V and note first the easily verifiable identity

$$\hat{F}_i P_{i|i} = F\left(P_i^{-1} + H^T \hat{R}_i^{-1} H + \hat{\lambda} E_f^T E_f\right)^{-1} = F\hat{P}_{i|i}$$

so that our steady-state filter can be written in the alternative form

$$\begin{aligned}\hat{x}_{i+1} =& F\left[I - \hat{\lambda}\left(P^{-1} + H^T \hat{R}^{-1} H + \hat{\lambda} E_f^T E_f\right)^{-1} E_f^T E_f\right]\hat{x}_i \\ &+ F\left(P^{-1} + H^T \hat{R}^{-1} H + \hat{\lambda} E_f^T E_f\right)^{-1} \\ &\cdot H^T \hat{R}^{-1}(y_i - H\hat{x}_i)\end{aligned} \tag{78}$$

where P is the stabilizing solution of the Riccati equation

$$P = F\left(P^{-1} + H^T \hat{R}^{-1} H + \hat{\lambda} E_f^T E_f\right)^{-1} F^T + Q \tag{79}$$

and $\hat{R} = R - \hat{\lambda}^{-1} H M M^T H^T$.

Comparing the recursions (76) and (78) for the state estimates, as well as the Riccati equations (77) and (79), we see that there are essentially four differences between a guaranteed-cost design and the proposed steady-state filter.

1) The negative scalar $-\epsilon$ is replaced by a positive scalar $\hat{\lambda}$. The appearance of a negative scalar $-\epsilon$ in the guaranteed-cost DARE (77) imposes a constraint on the selection of ϵ: its value has to be properly selected so as to guarantee a positive–definite difference $P^{-1} - \epsilon E_f^T E_f > 0$.
2) The matrix R is replaced by its corrected version $\hat{R}$.
3) The locations of the factor MM^T in the Riccati equations are different: in (77) it appears added to Q while in (79) it is incorporated into $\hat{R}^{-1}$. In this sense, a guaranteed-cost design can be interpreted as increasing the value of Q while our filter can be interpreted as decreasing the value of R.
4) The parameters $\{\epsilon, \hat{\lambda}\}$ are selected differently.

The guaranteed-cost filter equations from [24] have a different form that relies on two Riccati equations; they are omitted for brevity.

VII. Simulation Results

For comparison purposes, we employ the following numerical values of a two-dimensional (2-D) time-invariant model used in [14]:

$$F = \begin{bmatrix} 0.9802 & 0.0196 \\ 0 & 0.9802 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$H = [1 \quad -1], \quad M = \begin{bmatrix} 0.0198 \\ 0 \end{bmatrix}, \quad R = 1$$

$$E_f = [0 \quad 5], \quad Q = \begin{bmatrix} 1.9608 & 0.0195 \\ 0.0195 & 1.9605 \end{bmatrix}.$$

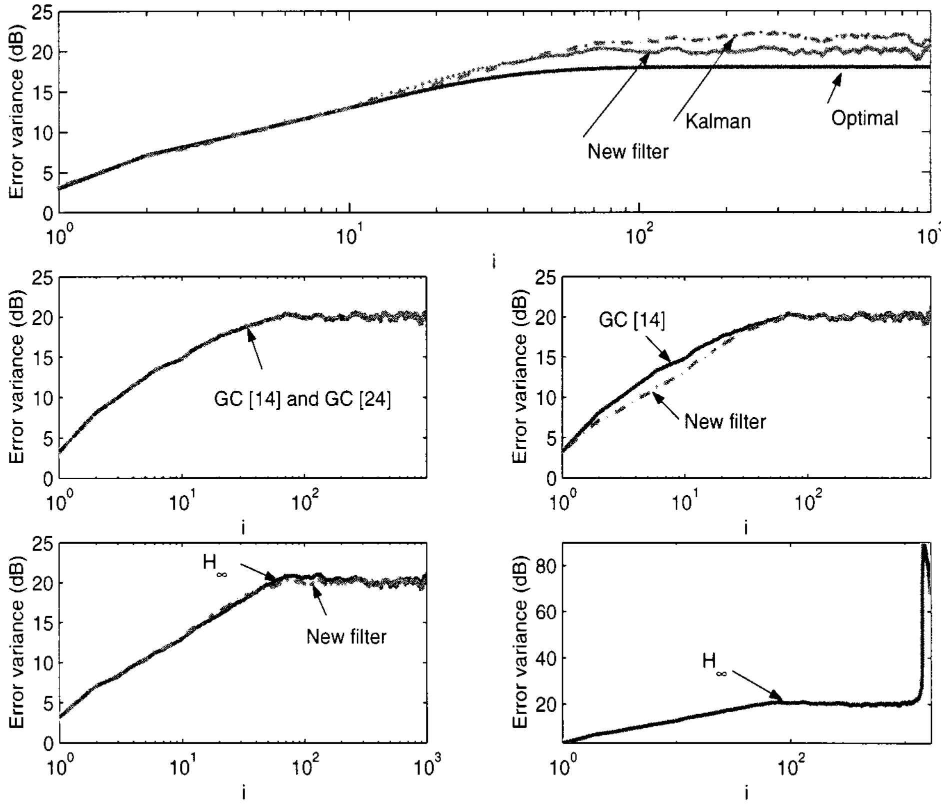

Fig. 2. Error variance curves for different filters with Δ selected uniformly from within the interval [−1, 1].

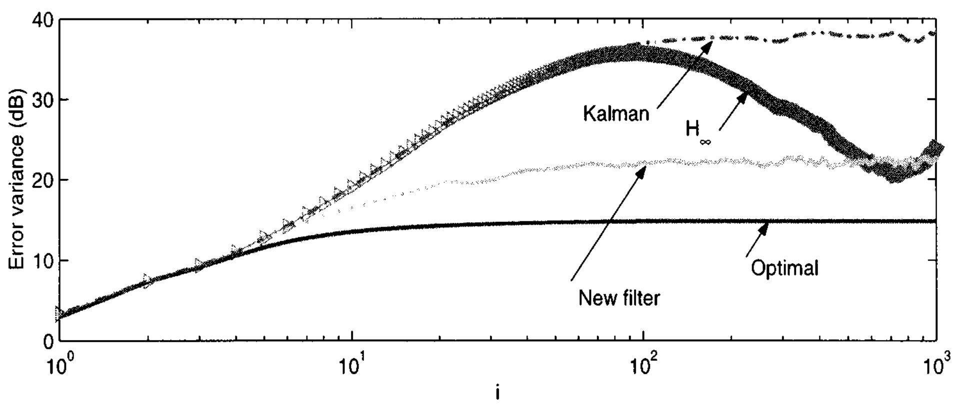

Fig. 3. Error variance curves with Δ selected uniformly from within the interval [−1, 1] and with $M = \mathrm{col}\{0.1980,\ 0\}$.

We also use $\Pi_0 = I$ and $\hat{x}_0 = 0$. The above parameters correspond to a quadratically stable time-invariant model of the form

$$x_{i+1} = (F + M\Delta_i E_f)x_i + u_i \quad y_i = Hx_i + v_i.$$

Observe that Δ_i is a scalar in this case and that the product $M\Delta_i E_f$ is strictly upper triangular so that the eigenvalues of $F + M\Delta_i E_f$ always coincide with those of F. In particular observe that the actual F matrix has the form

$$F_{ac,\,i} = \begin{bmatrix} 0.9802 & 0.0196 + 0.099\Delta_i \\ 0 & 0.9802 \end{bmatrix}$$

with the uncertainty Δ_i affecting only its (1, 2) entry.

A. Set-Valued Estimation

The two top plots in Fig. 1 show a simulation for which $\Delta_i = -0.8508 = \Delta$ for all i. Two curves are shown in each plot and they refer to: the actual trajectory of the top entry of the state vector (solid line), the trajectory that is produced by the robust set-valued estimation algorithm (dashed line in leftmost plot), and the trajectory that is produced by the new robust filter (dashed line in rightmost plot). We see that the set-valued estimates diverge around iteration 64 (the overshoot reaches a peak of approximately 28) and then stay at a constant level thereafter. This behavior was observed repeatedly for random choices of Δ_i. This occurs because of the violation of the existence conditions (75); the smallest eigenvalue of $R_{\nu,\,i}$ becomes negative fast and stays negative for an extended period of time. Due to this

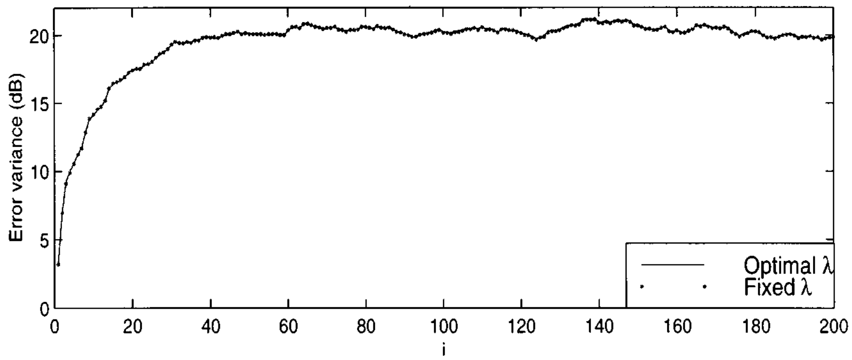

Fig. 4. Comparison of the variance curves for optimal and for approximate $\hat{\lambda}_i's$ for the new filter. In this case, the curves are indistinguishable.

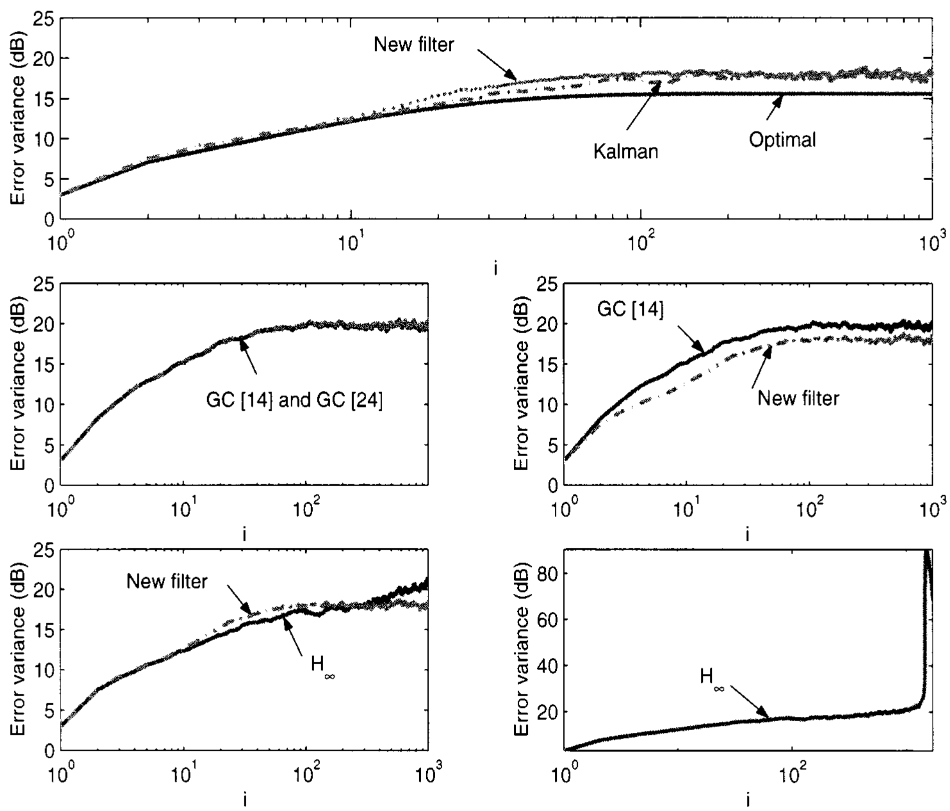

Fig. 5. Error variance curves for different filters with Δ selected uniformly from within the interval $[-1, 0]$.

effect, we are excluding the robust set-valued algorithm from our additional simulations below.

B. Ensemble-Average Error Variance Curves

In Fig. 2 we generated more elaborate performance curves as follows. Each point in each curve is the average over 500 experiments. Each experiment j fixes Δ at a value that is selected randomly between -1 and 1 and generates 1000-long random measurements $\{y_i\}$. The data is then filtered by a particular algorithm leading to an estimated trajectory $\{\hat{x}_i^{(j)}\}$ for the experiment j. At the end of the 500 experiments, we have 500 such trajectories (of length 1000 points each) for each algorithm and we can use them to approximate the actual error variance curve by computing the ensemble-average

$$E\|x_i - \hat{x}_i\|^2 \approx \frac{1}{500}\sum_{j=1}^{500}\left\|x_i - \hat{x}_i^{(j)}\right\|^2.$$

The top part of the figure highlights the degradation in performance by the Kalman filter due to modeling errors (approximately 4 dB for this example). More specifically, the smooth lower curve (termed optimal) refers to the error variance that is obtained when the actual model is used (about 18 dB in steady-state). The highest curve is the error variance by a Kalman filter using the nominal model (about 22 dB in

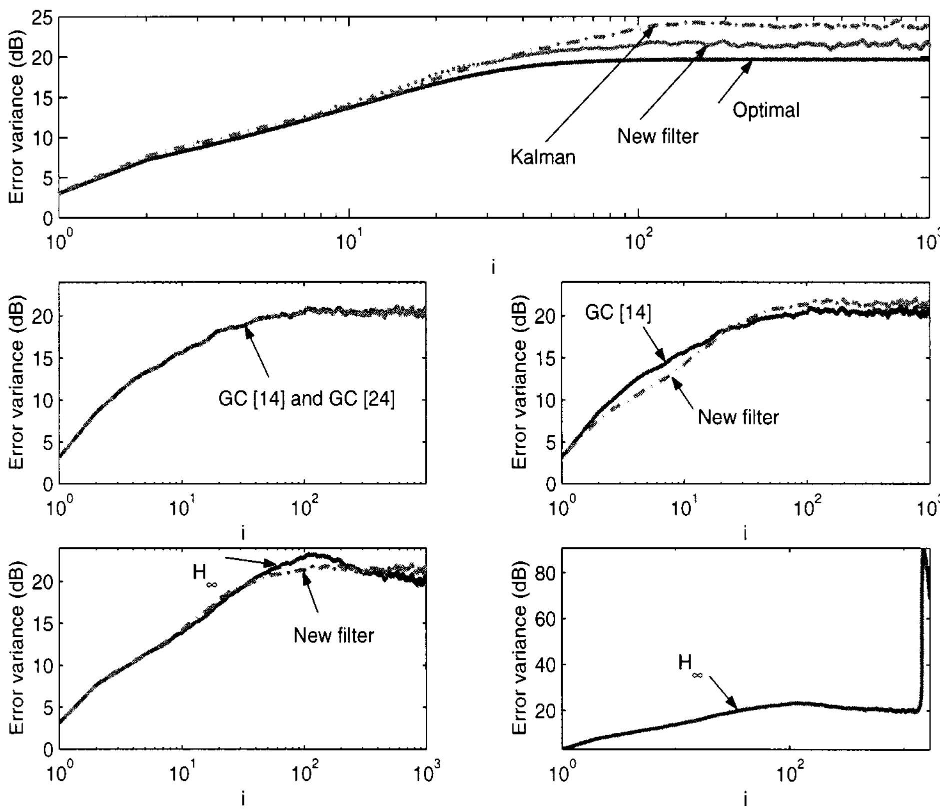

Fig. 6. Error variance curves for different filters with Δ selected uniformly from within the interval [0, 1].

steady-state). The curve below it indicates the performance of the new filter (about 20 dB in steady-state). It is of course not hard to find other examples where the performance of the Kalman filter is significantly worse. For instance, if we change the value of M to $M = \mathrm{col}\{0.1980,\, 0\}$ we obtain the performance shown in Fig. 3, where the Kalman filter exhibits a degradation of approx. 23 dB in performance.

The plots in the middle row of Fig. 2 compare the performance of the new robust filter to that of optimized guaranteed-cost designs. These designs choose ϵ by minimizing an upper bound on the error variance. The value of ϵ used for (76) is 5.98×10^{-6}. We see that in this example, by averaging the performance over all experiments (which corresponds to averaging over random selections of Δ in the interval $[-1, 1]$), the filters tend to exhibit a similar steady-state performance (recall that the proposed filter achieves this performance via regularization).

The plots in the last row of Fig. 2 compare the performance of the new robust filter with that of an $\mathcal{H}_\infty$ filter. While the leftmost plot suggests a good performance by the $\mathcal{H}_\infty$ filter, this result can be deceiving. First, we had to employ a large value of γ, $\gamma = 70$, in order to guarantee that the existence conditions (74) are not violated over the first 1000 iterations. Second, if we extend the filter operation beyond $N = 1000$, we find that the conditions (74) are violated starting at iteration 1441. This degrades the performance of the $\mathcal{H}_\infty$ filter considerably. To see this, the resulting error variance is shown in the bottom rightmost plot of Fig. 2, with a peak error variance of about 90 dB. All other filters maintain their performance over the extended operation period. We should say that we employed $L = [1 \;\; 0]$ in the $\mathcal{H}_\infty$ recursions, since with $L = I$ an even larger value of γ is needed. Moreover, for this $\mathcal{H}_\infty$ filter, it turns out that the value $\gamma = 71$ guarantees filter performance over arbitrarily long perdios of time (since it meets the sufficient conditions for feasibility and convergence from [17]). Still, it is clear that the choice of γ represents an important design tradeoff: while large values of γ can guarantee filter operation, this is usually achieved at the expense of robustness.

In the above simulations, we employed the approximation (60) by choosing $\alpha = 0.5$. Fig. 4 shows that this approximation provides a good alternative for this example over the implementation that is based on computing the optimal $\hat{\lambda}_i$ at each iteration (the curves were obtained in this case by averaging over 200 experiments). (We may remark that we have omitted from our comparisons the robust minimum-variance filter of [28] due to divergence problems.)

There are some interesting distinctions in performance between all filters. Fig. 5 shows the variance curves that correspond to the case in which Δ is selected uniformly from within the interval $[-1, 0]$, while Fig. 6 corresponds to the case in Δ is selected uniformly from within the interval [0, 1]. In the former case we see that the performance of the Kalman filter is comparable to, or even better than, the other filters, while it is noticeably worse in the latter case. The performance of the proposed filter in the latter case can also be improved by increasing the value of α.

Finally, Fig. 7 demonstrates the case in which Δ_i is allowed to vary randomly during each experiment.

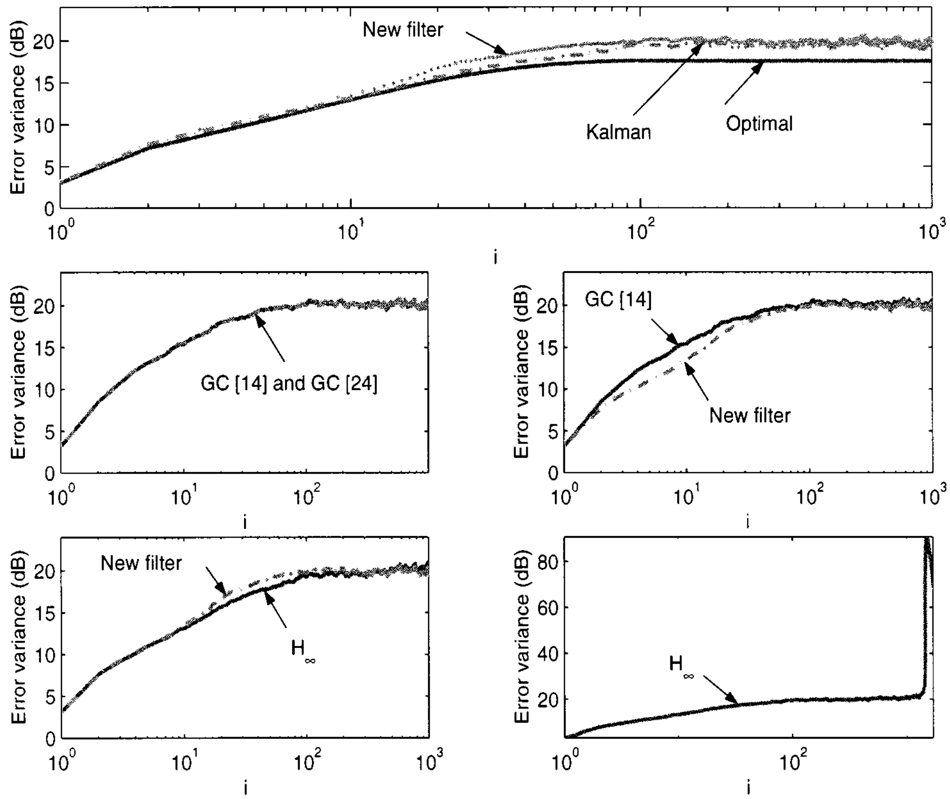

Fig. 7. Error variance curves for all filters with Δ_i allowed to vary randomly at each iteration.

VIII. Concluding Remarks

In this paper, we proposed a framework for robust state-space estimation that is based on minimizing, at each filter iteration, the worst-possible (regularized) squared residual norm. The resulting recursions were presented in three equivalent forms: a time- and measurement-update form, a prediction form, and an information form. All forms share similar characteristics with the corresponding forms in Kalman filtering with the distinction that in the robust context, the recursions rely on corrected parameters rather than nominal parameters. The filters were also shown, under certain detectability and stabilizability assumptions, to tend to stable steady-state estimators. In addition, for models that are quadratically stable, the filters were further shown to guarantee bounded error variances.

The new recursions were also compared with other robust filters, namely, $\mathcal{H}_\infty$ filters, guaranteed-cost filters, and set-valued estimation filters. In particular, it was shown that the new filters do not require existence conditions and that they apply to time-variant as well as time-invariant models. They also apply to finite-horizon and infinite-horizon scenarios.

There are several issues that deserve further investigation. One issue is extension of the results to other classes of model uncertainties, such as replacing (41) with conditions of the form $\|\delta F_i\| \leq \eta_{f,i}$ and $\|\delta G_i\| \leq \eta_{g,i}$ for some known bounds $\{\eta_{n,i}, \eta_{f,i}\}$. This corresponds to a different choice of the function $\phi(x)$ in (3). Other issues include a closer examination of the stochastic properties of the developed filters, a more explicit characterization of the error variance, and a more detailed study of the optimality properties of the filters over extended intervals, rather than locally. Another issue is the development of array variants, in addition to fast algorithms. The former would tend to exhibit better numerical properties while the latter would be more appropriate for large-scale problems.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Upper Saddle River, NJ: Prentice-Hall, 1979.
[2] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.
[3] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
[4] H. W. Sorenson, Ed., *Kalman Filtering: Theory and Application*. New York: IEEE Press, 1985.
[5] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1993.
[6] P. P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an $\mathcal{H}_\infty$-setting," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 151–166, 1991.
[7] T. Basar and P. Bernhard, *$\mathcal{H}_\infty$-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Boston, MA: Birkhauser, 1991.
[8] U. Shaked and Y. Theodor, "$\mathcal{H}_\infty$-optimal estimation: A tutorial," in *Proc. Conf. Decision Control*, Tucson, AZ, Dec. 1992, pp. 2278–2286.
[9] M. Green and D. J. N. Limebeer, *Linear Robust Control*. Upper Saddle River, NJ: Prentice-Hall, 1995.
[10] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ: Prentice-Hall, 1996.
[11] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Theories*. Philadelphia, PA: SIAM, 1999.
[12] L. Xie, C. E. de Souza, and M. Fu, "$\mathcal{H}_\infty$ estimation for discrete-time linear uncertain systems," *Int. J. Rob. Nonlinear Control*, vol. 1, pp. 111–123, 1991.
[13] M. Fu, C. E. de Souza, and L. Xie, "$\mathcal{H}_\infty$ estimation for uncertain systems," *Int. J. Rob. Nonlinear Control*, vol. 2, pp. 87–105, 1992.

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

[15] A. H. Sayed, B. Hassibi, and T. Kailath, "Inertia conditions for the minimization of quadratic forms in indefinite metric spaces," in *Operator Theory: Advances and Applications*, I. Gohberg, P. Lancaster, and P. N. Shivakumar, Eds. Basel, Switzerland: Springer-Verlag, 1996, vol. 87, pp. 309–347.

[16] P. Bolzern, P. Colaneri, and G. De Nicolao, "Transient and asymptotic analysis of discrete-time $\mathcal{H}_\infty$ filters," *Eur. J. Control*, vol. 3, pp. 317–324, 1997.

[17] P. Bolzern and M. Maroni, "New conditions for the convergence of $\mathcal{H}_\infty$ filters and predictors," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1564–1568, Aug. 1999.

[18] 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.

[19] A. Garulli, A. Vicino, and G. Zappa, "Conditional central algorithms for worst case set-membership identification and filtering," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 14–23, Jan. 2000.

[20] M. Milanese *et al.*, Ed., *Bounding Approaches to System Identification*, NY: Plenum, 1996.

[21] B. N. Jain, "Guaranteed error estimation in uncertain systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 230–232, Feb. 1975.

[22] A. V. Savkin and I. R. Petersen, "Recursive state estimation for uncertain systems with an integral quadratic constraint," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1080–1083, June 1995.

[23] I. R. Petersen and D. C. McFarlane, "Robust state estimation for uncertain systems," in *Proc. Conf. Decision Control*, Brighton, U.K., 1991, pp. 2630–2631.

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

[25] 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.

[26] D. S. Bernstein, W. M. Haddad, and D. Mustafa, "Mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ regulation and estimation: The discrete-time case," *Syst. Control Lett.*, vol. 16, pp. 235–248, 1991.

[27] W. M. Haddad and D. S. Bernstein, "The optimal projection equations for reduced-order, discrete-time state estimation for linear systems with multiplicative white noise," *Syst Control Lett.*, vol. 8, pp. 381–388, 1987.

[28] Y. Theodor and U. Shaked, "Robust discrete-time minimum-variance filtering," *IEEE Trans. Signal Processing*, vol. 44, pp. 181–189, Feb. 1996.

[29] P. Bolzern, P. Colaneri, and G. De Nicolao, "Guaranteed-cost prediction of discrete-time systems: The finite and infinite-horizon case," in *Proc. 2nd IFAC Symp. Robust Control Design*, Budapest, Hungary, 1997, pp. 471–476.

[30] A. H. Sayed and V. H. Nascimento, "Design criteria for uncertain models with structured and unstructured uncertainties," in *Robustness in Identification and Control*, A. Garulli, A. Tesi, and A. Vicino, Eds. London, U.K.: Springer-Verlag, 1999, vol. 245, pp. 159–173.

[31] A. H. Sayed, V. H. Nascimento, and S. Chandrasekaran, "Estimation and control with bounded data uncertainties," *Linear Alg. Appl.*, vol. 284, pp. 259–306, Nov. 1998.

[32] L. E. Ghaoui and H. Hebret, "Robust solutions to least-squares problems with uncertain data," *SIAM J. Matrix Anal. Appl.*, vol. 18, no. 4, pp. 1035–1064, 1997.

[33] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, "Parameter estimation in the presence of bounded data uncertainties," *SIAM J. Matrix Analy. Appl.*, vol. 19, no. 1, pp. 235–252, Jan. 1998.

[34] A. H. Sayed and S. Chandrasekaran, "Parameter estimation in the presence of multiple sources and levels of uncertainties," *IEEE Trans. Signal Processing*, pp. 680–692, Mar. 2000.

[35] Y. Cheng and B. L. de Moor, "Robustness analysis and control system design for hydraulic servo system," *IEEE Trans. Contr. Syst. Technol.*, vol. 2, pp. 183–198, 1994.

[36] V. H. Nascimento and A. H. Sayed, "Optimal state regulation for uncertain state-space models," in *Proc. Amer. Control Conf.*, vol. 1, San Diego, June 1999, pp. 419–424.

[37] A. H. Sayed, V. H. Nascimento, and F. Cipparrone, "A regularized robust design criterion for uncertain data," , submitted for publication.

[38] A. H. Sayed and H. Chen, "A uniqueness result concerning a robust regularized least-squares solution," , submitted for publication.

[39] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*. New York: Taylor & Francis, 1975. revised printing.

[40] R. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1957.

[41] A. E. Bryson, *Dynamic Optimization*. Upper Saddle River, NJ: Prentice-Hall, 1999.

[42] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.

[43] R. C. K. Lee, *Optimal Estimation, Identification, and Control*. Cambridge, MA: MIT Press, 1964.

[44] A. Packard and J. Doyle, "Quadratic stability with real and complex perturbations," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 198–201, 1990.

[45] P. P. Khargonekar, I. R. Petersen, and K. Zhou, "Robust stabilization of uncertain systems and $\mathcal{H}^\infty$ optimal control," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 356–361, Mar. 1990.

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

[47] U. Shaked and C. E. de Souza, "Robust minimum variance filtering," *IEEE Trans. Signal Processing*, vol. 43, pp. 474–2483, 1995.

Ali H. Sayed (S'90–M'92–SM'99–F'01) received the Ph.D. degree in electrical engineering in 1992 from Stanford University, Stanford, CA.

He is currently Associate Professor of Electrical Engineering at University of California, Los Angeles. He has over 150 journal and conference publications, is coauthor of the research monograh *Indefinite Quadratic Estimation and Control* (Philadelphia, PA: SIAM, 1999) and of the graduate-level textbook *Linear Estimation* (Upper Saddle River, NJ: Prentice-Hall, 2000). He is also coeditor of the volume *Fast Reliable Algorithms for Matrices with Structure* (Philadelphia, PA: SIAM, 1999). He is a member of the editorial boards of the *SIAM Journal on Matrix Analysis and Its Applications*, of the *International Journal of Adaptive Control and Signal Processing*, and has served as coeditor of special issues of the journal *Linear Algebra and Its Applications*. He has contributed several articles to engineering and mathematical encyclopedias and handbooks, and has served on the program committees of several international meetings. He has also consulted to several companies in the areas of adaptive filtering, adaptive equalization, and echo cancellation. His research interests span several areas including adaptive and statistical signal processing, filtering and estimation theories, equalization techniques for communications, interplays between signal processing and control methodologies, and fast algorithms for large-scale problems. To learn more about his work, visit the website of the UCLA Adaptive Systems Laboratory at http://www.ee.ucla.edu/asl

Dr. Sayed is a recipient of the 1996 IEEE Donald G. Fink Award. He is Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.