Linear Estimation in Krein Spaces—Part II: Applications

Babak Hassibi, Ali H. Sayed, *Member, IEEE*, and Thomas Kailath, *Fellow, IEEE*

***Abstract*—We show that several interesting problems in H^∞-filtering, quadratic game theory, and risk sensitive control and estimation follow as special cases of the Krein-space linear estimation theory developed in [1]. We show that all these problems can be cast into the problem of calculating the stationary point of certain second-order forms, and that by considering the appropriate state space models and error Gramians, we can use the Krein-space estimation theory to calculate the stationary points and study their properties. The approach discussed here allows for interesting generalizations, such as finite memory adaptive filtering with varying sliding patterns.**

I. Introduction

CLASSICAL results in linear least-squares estimation and Kalman filtering are based on an L_2 or H_2 criterion and require *a priori* knowledge of the statistical properties of the noise signals. In some applications, however, one is faced with model uncertainties and a lack of statistical information on the exogenous signals which has led to an increasing interest in minimax estimation (see, e.g., [2]–[9] and the references therein) with the belief that the resulting so-called H^∞ algorithms will be more robust and less sensitive to parameter variations.

Furthermore, while the statistical Kalman-filtering algorithm can be viewed as a recursive procedure that minimizes a certain quadratic cost function, there has also been increasing interest in an alternative so-called exponential (LEQG) cost function [13]–[16] which is risk-sensitive in the sense that it depends on a real parameter that determines whether more or less weight should be given to higher or smaller errors. The corresponding filters have been termed risk-sensitive and include the Kalman filter as a special case. We show in this paper that the H^∞ and risk-sensitive filters can both be obtained by using appropriate Krein space–Kalman filters, based on the theory developed in Part I [1].

H^∞ and risk-sensitive estimation and control problems, quadratic games, and finite memory adaptive filtering problems lead almost by inspection to indefinite deterministic quadratic forms. Following [1], we solve these problems by constructing the corresponding Krein-space "stochastic" problems for which the Kalman-filter solutions can be written down immediately; moreover, the conditions for a minimum can also be expressed in terms of quantities easily related to the basic Riccati equations of the Kalman filter. This approach also explains the many similarities between, say, the H^∞ solutions and the classical LQ solutions and in addition marks out their key differences.

The paper is organized as follows. In Section II we introduce the H^∞ estimation problem, state the conventional solution, and discuss its similarities with and differences from the conventional Kalman filter. In Section III we reduce the H^∞ estimation problem to guaranteeing the positivity of a certain indefinite quadratic form. We then relate this quadratic form to a certain Krein state-space model which allows us to use the results of the companion paper [1] to derive conditions for its positivity and to show that projection in the Krein space allows us to solve the H^∞ estimation problem. In this context we derive the H^∞ *a posteriori*, *a priori*, and smoothing filters, and show that H^∞ estimation is essentially Kalman filtering in Krein space; we also obtain a natural parameterization of all H^∞ estimators. One advantage of our approach is that it suggests how well-known conventional Kalman-filtering algorithms, such as square root arrays and Chandrasekhar recursions, can be extended to the H^∞ setting.

In Section IV we describe the problem of risk-sensitive estimation [13]–[15] and show that a risk-sensitive estimator is one that computes the stationary point of a certain second-order form, provided that this second-order form has a minimum over a certain set of variables. By considering a corresponding Krein state-space model, we use the results of [1] to derive conditions for the existence of the minimum and to show that the Krein-space projection also solves the risk-sensitive estimation problem. We then derive risk-sensitive *a posteriori*, *a priori*, and smoothing filters parallel to what was done in Section II. We also use this parallel to stress the connection between H^∞ and risk-sensitive estimation that was first discovered in [21], using different arguments. Before concluding with Section VI, we describe the finite memory adaptive filtering problem in Section V and use the Krein-space approach to solve this problem and to connect it to state-space approaches to adaptive filtering.

As was done in the companion paper [1], we shall use bold letters for elements in a Krein space, and normal letters for corresponding complex numbers. Also, we shall use $\check{z}$ to denote the estimate of z (according to some criterion), and $\hat{z}$ to denote the Krein-space projection, thereby stressing the fact

Manuscript received March 4, 1994; revised June 16, 1995. Recommended by Associate Editor at Large, B. Pasik-Duncan. This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command under Contract AFOSR91-0060, in part by the Army Research Office under Contract DAAL03-89-K-0109, and in part by a grant from NSF under award MIP-9409319.

B. Hassibi and T. Kailath are with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA.

A. H. Sayed is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA.

Publisher Item Identifier S 0018-9286(96)00408-4.

that they need not coincide. Many of the results discussed here were obtained earlier by several other authors using different methods and arguments. Our approach, we believe, provides a powerful unification with immediate insights to various extensions.

II. H^∞ Estimation

Several H^∞-filtering algorithms have been recently derived by a variety of methods in both the continuous and discrete-time cases (see, e.g., [2]–[9] and the references therein).

Many authors have noticed some formal similarities between the H^∞ filters and the conventional Kalman filter, however, we shall further clarify this connection by showing that H^∞ filters are nothing more than certain Krein space–Kalman filters. In other words, the H^∞ filters can be viewed as recursively performing a (Gram–Schmidt) orthogonalization (or projection) procedure on a convenient set of observation data that obey a state-space model whose state evolves in an indefinite metric space. This is of significance since it yields a geometric derivation of the H^∞ filters, and because it unifies H^2- and H^∞-estimation in a simple framework. Moreover, once this connection has been made explicit, many known alternative and more efficient algorithms, such as square-root arrays and Chandrasekhar equations [24], can be applied to the H^∞-setting as well. Also our results deal directly with the time-varying scenario. Finally, we note that although we restrict ourselves here to the discrete-time case, the continuous time analogs follow the same principles.

A. Formulation of the H^∞-Filtering Problem

Consider a time-variant state-space model of the form

$$\begin{cases} x_{i+1} = F_i x_i + G_i u_i, & x_0 \\ y_i = H_i x_i + v_i, & i \geq 0 \end{cases} \tag{1}$$

where $F_i \in \mathcal{C}^{n\times n}$, $G_i \in \mathcal{C}^{n\times m}$ and $H_i \in \mathcal{C}^{p\times n}$ are known matrices, x_0, $\{u_i\}$, and $\{v_i\}$ are unknown quantities, and y_i is the measured output. We can regard v_i as a measurement noise and u_i as a process noise or driving disturbance. We make no assumption on the nature of the disturbances (e.g., normally distributed, uncorrelated, etc). In general, we would like to estimate some arbitrary linear combination of the states, say

$$z_i = L_i x_i$$

where $L_i \in \mathcal{C}^{q\times n}$ is given, using the observations $\{y_j\}$. Let $\check{z}_{i|i} = \mathcal{F}_f(y_0, y_1, \cdots, y_i)$ denote the estimate of z_i given observations $\{y_j\}$ from time 0 to, and including, time i, and $\check{z}_i = \mathcal{F}_p(y_0, y_1, \cdots, y_{i-1})$ denote the estimate of z_i given observations $\{y_j\}$ from time 0 to time $i-1$. We then have the following two estimation errors: the filtered error

$$e_{f,i} = \check{z}_{i|i} - L_i x_i \tag{2}$$

and the predicted error

$$e_{p,i} = \check{z}_i - L_i x_i. \tag{3}$$

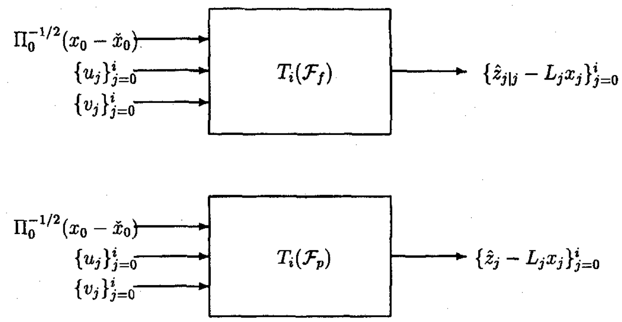

Fig. 1. Transfer matrix from disturbances to filtered and predicted estimation errors.

As depicted in Fig. 1, let $T_i(\mathcal{F}_f)$ and $T_i(\mathcal{F}_p)$ denote the transfer operators that map the unknown disturbances $\{\Pi_0^{-1/2}(x_0 - \check{x}_0), \{u_j\}_{j=0}^i, \{v_j\}_{j=0}^i\}$ (where $\check{x}_0$ denotes an initial guess for x_0, and Π_0 is a given positive definite matrix) to the filtered and predicted errors $\{e_{f,j}\}_{j=0}^i$ and $\{e_{p,j}\}_{j=0}^i$, respectively. The problem is to choose the functionals $\mathcal{F}_f(\cdot)$ and $\mathcal{F}_p(\cdot)$ so as to respectively minimize the H^∞ norm of the transfer operators $T_i(\mathcal{F}_f)$ and $T_i(\mathcal{F}_p)$.

Definition 1: The H^∞ norm of a transfer operator T is defined as

$$\|T\|_\infty = \sup_{u\in h_2, u\neq 0} \frac{\|Tu\|_2}{\|u\|_2}$$

where $\|u\|_2$ is the h_2-norm of the causal sequence $\{u_k\}$, i.e., $\|u\|_2^2 = \sum_{k=0}^\infty u_k^* u_k$.

The H^∞ norm thus has the interpretation of being the maximum energy gain from the input u to the output y. Our problem may now be formally stated as follows.

Problem 1 (Optimal H^∞ Problem): Find H^∞-optimal estimation strategies $\check{z}_{i|i} = \mathcal{F}_f(y_0, y_1, \cdots, y_i)$ and $\check{z}_i = \mathcal{F}_p(y_0, y_1, \cdots, y_{i-1})$ that respectively minimize $\|T_i(\mathcal{F}_f)\|_\infty$ and $\|T_i(\mathcal{F}_p)\|_\infty$ and obtain the resulting

$$\begin{aligned}\gamma_{f,o}^2 &= \inf_{\mathcal{F}_f} \|T_i(\mathcal{F}_f)\|_\infty^2 \\ &= \inf_{\mathcal{F}_f} \sup_{x_0, u\in h_2, v\in h_2} \frac{\sum_{j=0}^i e_{f,j}^* e_{f,j}}{(x_0 - \check{x}_0)^* \Pi_0^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^i u_j^* u_j + \sum_{j=0}^i v_j^* v_j}\end{aligned} \tag{4}$$

and

$$\begin{aligned}\gamma_{p,o}^2 &= \inf_{\mathcal{F}_p} \|T_i(\mathcal{F}_p)\|_\infty^2 \\ &= \inf_{\mathcal{F}_p} \sup_{x_0, u\in h_2, v\in h_2} \frac{\sum_{j=0}^i e_{p,j}^* e_{p,j}}{(x_0 - \check{x}_0)^* \Pi_0^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^{i-1} u_j^* u_j + \sum_{j=0}^{i-1} v_j^* v_j}\end{aligned} \tag{5}$$

where Π_0 is a positive definite matrix that reflects *a priori* knowledge as to how close x_0 is to the initial guess $\check{x}_0$.

Note that the infinum in (5) is taken over all strictly causal estimators $\mathcal{F}_p$, whereas in (4) the estimators $\mathcal{F}_f$ are only causal since they have additional access to y_i. This is relevant since the solution to the H^∞ problem, as we shall see, depends on the structure of the information available to the estimator.

The above problem formulation shows that H^∞ optimal estimators guarantee the smallest estimation error energy over all possible disturbances of fixed energy. They are, therefore, over-conservative, which results in a better robust behavior to disturbance variation.

A closed form solution to the optimal H^∞ estimation problem is available only in some special cases (see, e.g., [25]), and so it is common in the literature to settle for a suboptimal solution.

Problem (Sub-Optimal H^∞ Problem): Given scalars $\gamma_f > 0$ and $\gamma_p > 0$, find H^∞ suboptimal estimation strategies $\check{z}_{i|i} = \mathcal{F}_f(y_0, y_1, \cdots, y_i)$ (known as an *a posteriori* filter) and $\check{z}_i = \mathcal{F}_p(y_0, y_1, \cdots, y_{i-1})$ (known as an *a priori* filter) that respectively achieve $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$ and $\|T_i(\mathcal{F}_p)\|_\infty < \gamma_p$. In other words, find strategies that respectively achieve

$$\sup_{x_0, u \in h_2, v \in h_2} \frac{\sum_{j=0}^{i} e_{f,j}^* e_{f,j}}{(x_0 - \check{x}_0)^* \Pi_0{}^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^{i} u_j^* u_j + \sum_{j=0}^{i} v_j^* v_j} < \gamma_f^2 \tag{6}$$

and

$$\sup_{x_0, u \in h_2, v \in h_2} \frac{\sum_{j=0}^{i} e_{p,j}^* e_{p,j}}{(x_0 - \check{x}_0)^* \Pi_0{}^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^{i-1} u_j^* u_j + \sum_{j=0}^{i-1} v_j^* v_j} < \gamma_p^2. \tag{7}$$

This clearly requires checking whether $\gamma_f \geq \gamma_{f,o}$ and $\gamma_p \geq \gamma_{p,o}$.

Note that the solutions to Problem 1 can be obtained to desired accuracy by iterating on the γ_f and γ_p of Problem 2. From here on we shall be only dealing with Problem 2.

Note that the problems defined above are finite-horizon problems. So-called infinite-horizon problems can be considered if we define $T(\mathcal{F}_f)$ and $T(\mathcal{F}_p)$ as the transfer operators that map $\{x_0 - \check{x}_0, \{u_j\}_{j=0}^\infty, \{v_j\}_{j=0}^\infty\}$ to $\{e_{f,j}\}_{j=0}^\infty$ and $\{e_{p,j}\}_{j=0}^\infty$, respectively. Then by guaranteeing $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$ and $\|T_i(\mathcal{F}_p)\|_\infty < \gamma_p$ for all i, we can solve the infinite-horizon problems $\|T(\mathcal{F}_f)\|_\infty \leq \gamma_f$ and $\|T(\mathcal{F}_p)\|_\infty \leq \gamma_p$, respectively. Direct solutions, however, are also possible.

B. Solution of the Suboptimal H^∞ Filtering Problem

We now present the existing solutions (see, e.g., [4], [7]) to the suboptimal H^∞ filtering problem and note that they are intriguingly similar in several ways to the conventional Kalman filter. It was this similarity in structure that led us to extend Kalman filtering to Krein spaces (see [1]); in effect, H^∞ filters are just Kalman filters in Krein space.

Theorem 1 (An H^∞ A Posteriori Filter) [7]*:* For a given $\gamma > 0$, if the $[F_j \quad G_j]$ have full rank, then an estimator that achieves $\|T_i(\mathcal{F}_f)\|_\infty < \gamma$ exists if, and only if

$$P_j^{-1} + H_j^* H_j - \gamma^{-2} L_j^* L_j > 0, \quad j = 0, \cdots, i \tag{8}$$

where $P_0 = \Pi_0$ and P_j satisfies the Riccati recursion

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j [H_j^* \quad L_j^*] R_{e,j}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j F_j^* \tag{9}$$

with

$$R_{e,j} = \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j [H_j^* \quad L_j^*]. \tag{10}$$

If this is the case, then one possible level-γ H^∞ filter is given by

$$\check{z}_{j|j} = L_j \hat{x}_{j|j}$$

where $\hat{x}_{j|j}$ is recursively computed as

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + K_{s,j+1}(y_{j+1} - H_{j+1} F_j \hat{x}_{j|j})$$
$$\hat{x}_{-1|-1} = \text{initial guess} \tag{11}$$

and

$$K_{s,j+1} = P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1}. \tag{12}$$

Theorem 2 (An H^∞ A Priori Filter) [7]*:* For a given $\gamma > 0$, if the $[F_j \quad G_j]$ have full rank, then an estimator that achieves $\|T_i(\mathcal{F}_p)\|_\infty < \gamma$ exists if, and only if

$$\tilde{P}_j^{-1} = P_j^{-1} - \gamma^{-2} L_j^* L_j > 0, \quad j = 0, \cdots, i \tag{13}$$

where P_j is the same as in Theorem 1. If this is the case, then one possible level-γ H^∞ filter is given by

$$\check{z}_j = L_j \hat{x}_j \tag{14}$$

$$\hat{x}_{j+1} = F_j \hat{x}_j + K_{a,j}(y_j - H_j \hat{x}_j)$$
$$\hat{x}_0 = \text{initial guess} \tag{15}$$

where

$$K_{a,j} = F_j \tilde{P}_j H_j^* (I + H_j \tilde{P}_j H_j^*)^{-1}. \tag{16}$$

Comparisons with the Kalman Filter: The Kalman-filter algorithm for estimating the states in (1), assuming that the $\{u_i\}$ and $\{v_i\}$ are uncorrelated unit variance white noise processes, is

$$\hat{x}_{j+1} = F_j \hat{x}_j + F_j P_j H_j^* (I + H_j P_j H_j^*)^{-1} (y_j - H_j \hat{x}_j)$$
$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1} \times (y_{j+1} - H_{j+1} \hat{x}_{j+1})$$

where

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j (I + H_j P_j H_j^*)^{-1} P_j F_j^*, \quad P_0 = \Pi_0. \tag{17}$$

As several authors have noted, the H^∞ solutions are very similar to the conventional Kalman filter. The major differences are the following:

- The structure of the H^∞ estimators depends, via the Riccati recursion (9), on the linear combination of the states that we intend to estimate (i.e., the L_i). This is as opposed to the Kalman filter where the estimate of any linear combination of the state is given by that linear combination of the state estimate. Intuitively, this means that the H^∞ filters are specifically tuned toward the linear combination $L_i x_i$.

- We have additional conditions, (8) or (13), that must be satisfied for the filter to exist; in the Kalman filter problem the L_i would not appear, and the P_i would be positive definite so that (8) and (13) would be immediate.
- We have indefinite (covariance) matrices, e.g., $\begin{bmatrix} I & \mathbf{o} \\ \mathbf{o} & -\gamma^2 I \end{bmatrix}$ versus just I in the Kalman filter.
- As $\gamma \to \infty$, the Riccati recursion (9) reduces to the Kalman filter recursion (17). (This suggests that the H^∞ norm of the conventional Kalman filter may be quite large, and that it may have poor robustness properties. Note also that condition (13) is more stringent than condition (8), indicating that the existence of an *a priori* filter of level γ implies the existence of an *a posteriori* filter of level γ, but not necessarily vice versa.)

Despite these differences, we shall show by applying the results of the companion paper [1] that the filters of Theorems 1 and 2 can in fact be obtained as certain Kalman filters, not in an H^2 (Hilbert) space, but in a certain indefinite vector space called a Krein space. The indefinite covariances and the appearance of L_i in the Riccati equation will be explained easily in this framework. The additional condition (8) will be seen to arise from the fact that in Krein space, unlike as in the usual Hilbert space context, quadratic forms need not always have minima or maxima unless certain additional conditions are met. Moreover, our approach will provide a simpler and more general alternative to the tests (8) and (13).

III. Derivation of the H^∞ Filters

As shown in the companion paper [1], the first step is to associate an indefinite quadratic form with each of the (level γ) *a posteriori* and *a priori* filtering problems. This will lead us to construct an appropriate (so-called partially equivalent) Krein space state-space model, the Kalman filter which will allow us to compute the stationary points for the H^∞ quadratic forms; conditions that these are actually minima will be deduced from the general results of Part I and shown to be just (8) and (13). Simpler equivalent conditions will also be noted.

Therefore we begin by examining the structure of the H^∞ problem in more detail. The goal will be to relate the problem to an indefinite quadratic form. We shall first consider the *a posteriori* filtering problem.

A. The Suboptimal H^∞ Problem and Quadratic Forms

Referring to Problem 2, we first note that $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$, implies that for all nonzero $\{x_0, \{u_j\}_{j=0}^i, \{v_j\}_{j=0}^i\}$

$$\frac{\sum_{j=0}^i |\check{z}_{j|j} - L_j x_j|^2}{(x_0-\check{x}_0)^*\Pi_0^{-1}(x_0-\check{x}_0)+\sum_{j=0}^i |u_j|^2+\sum_{j=0}^i |y_j - H_j x_j|^2} < \gamma_f^2. \tag{18}$$

Moreover, (18) implies that for all $k \le i$, we must have

$$\frac{\sum_{j=0}^k |\check{z}_{j|j} - L_j x_j|^2}{(x_0-\check{x}_0)^*\Pi_0^{-1}(x_0-\check{x}_0)+\sum_{j=0}^k |u_j|^2+\sum_{j=0}^k |y_k - H_i x_i|^2} < \gamma_f^2. \tag{19}$$

We remark that if the $\{y_j\}_{j=0}^i$ are all zero, then it is easy to see that the $\{\check{z}_{j|j}\}$ must all be zero as well. Therefore we need only consider the case where $\{y_j\}_{j=0}^i$ is a nonzero sequence. We shall then prove the following result, relating the condition $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$ to the positivity of a certain indefinite quadratic form. From now on, without loss of generality, we assume $\check{x}_0 = 0$; a nonzero $\check{x}_0 = 0$ will only change the initial condition of the filter.

Lemma 1 (Indefinite Quadratic Form): Given a scalar $\gamma_f > 0$, then $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$ if, and only if, there exists $\check{z}_{k|k} = \mathcal{F}_f(y_0, \cdots, y_k)$ (for all $0 \le k \le i$) such that for all complex vectors x_0, for all causal sequences $\{u_j\}_{j=0}^i$, and for all nonzero causal sequences $\{y_j\}_{j=0}^i$, the scalar quadratic form

$$\begin{aligned} J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k) &= x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^k u_j^* u_j \\ &+ \sum_{j=0}^k (y_j - H_j x_j)^*(y_j - H_j x_j) \\ &- \gamma_f^{-2}\sum_{j=0}^k (\check{z}_{j|j} - L_j x_j)^*(\check{z}_{j|j} - L_j x_j) \end{aligned} \tag{20}$$

satisfies

$$J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k) > 0 \quad \text{for all} \quad 0 \le k \le i. \tag{21}$$

Proof: Assume there exists a solution $\check{z}_{k|k}$ (for all $k \le i$) that achieves $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$. Then if we multiply both sides of (19) by the positive denominator on the left-hand side (LHS), we obtain (21).

Conversely, if there exists a solution $\check{z}_{k|k}$ (for all $k \le i$) that achieves (21), we can divide both sides of (21) by the positive quantity

$$x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^k u_j^* u_j + \sum_{j=0}^k (y_j - H_j x_j)^*(y_j - H_j x_j)$$

to obtain (19), and thereby $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$. □

Remark: Lemma 1 is a straightforward restatement of (19) which is required of all suboptimal H^∞ *a posteriori* filters with level γ_f. The statement of Lemma 2 given below, however, is a key result, since it shows how to check the conditions of Lemma 1 by computing the stationary point of the indefinite quadratic form $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ and checking its condition for a minimum. This is in the spirit of the approach taken in [1].

Note that since the $\check{z}_{k|k}$ are functions of the $\{y_j\}_{j=0}^k$, $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ is really a function of only $\{x_0, u_0, \cdots, u_k, y_0, \cdots, y_k\}$. Moreover, since the $\{y_j\}$ are fixed observations, the only free variables in $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ are the disturbances $\{x_0, u_0, \cdots, u_k\}$. We then have the following result.

Lemma 2 (Positivity Condition): The scalar quadratic forms $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ satisfy the conditions (21), iff, for all $0 \le k \le i$

i) $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ has a minimum with respect to $\{x_0, u_0, u_1, \cdots u_k\}$.

ii) The $\{\check{z}_{k|k}\}_{k=0}^{i}$ can be chosen such that the value of $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ at this minimum is positive, viz.

$$\min_{\{x_0, u_0, \cdots, u_k\}} J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k) > 0.$$

Proof: Assume $J_{f,k}(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k}) > 0$, then condition i) is clearly satisfied because if $J_{f,k}$ $(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k})$ does not have a minimum over $\{x_0, u_0, \cdots, u_k\}$, then it is always possible to choose $\{x_0, u_0, \cdots, u_k\}$ to make $J_{f,k}$ $(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k})$ arbitrarily small and negative. Moreover, the existence of a minimum, along with $J_{f,k}(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k}) > 0$, guarantees condition ii) since the value at the minimum must be positive.

Conversely, if i) and ii) hold, then it follows that $J_{f,k}$ $(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k}) > 0$. □

B. A Krein Space State-Space Model

To apply the methodology of Part I, we first identify the indefinite quadratic form $J_{f,k}$ as a special case of the general form studied in Theorem 6 of [1] by rewriting it as

$$\begin{aligned} J_{f,k}&(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k) \\ &= x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{k} u_j^* u_j \\ &+ \sum_{j=0}^{k} \left(\begin{bmatrix} y_i \\ \check{z}_{j|j} \end{bmatrix} - \begin{bmatrix} H_j \\ L_j \end{bmatrix} x_j \right)^* \\ &\times \begin{bmatrix} I & 0 \\ 0 & -\gamma_f^{-2} I \end{bmatrix} \left(\begin{bmatrix} y_j \\ \check{z}_{j|j} \end{bmatrix} - \begin{bmatrix} H_j \\ L_j \end{bmatrix} x_j \right). \end{aligned} \tag{22}$$

Then by Lemmas 6 and 7 in [1], we can introduce the following Krein-space system

$$\begin{cases} \boldsymbol{x}_{j+1} = F_j \boldsymbol{x}_j + G_j \boldsymbol{u}_j \\ \begin{bmatrix} \boldsymbol{y}_j \\ \check{\boldsymbol{z}}_{j|j} \end{bmatrix} = \begin{bmatrix} H_j \\ L_j \end{bmatrix} \boldsymbol{x}_j + \boldsymbol{v}_j \end{cases} \tag{23}$$

with

$$\left\langle \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{u}_j \\ \boldsymbol{v}_j \end{bmatrix}, \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{u}_k \\ \boldsymbol{v}_k \end{bmatrix} \right\rangle = \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & I\delta_{jk} & 0 \\ 0 & 0 & \begin{bmatrix} I & 0 \\ 0 & -\gamma_f^2 I \end{bmatrix} \delta_{jk} \end{bmatrix}. \tag{24}$$

Note that $Q_j = I$, $S_j = 0$, $\Pi_0 > 0$, and that we must consider a Krein space since

$$R_j = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix} \tag{25}$$

is indefinite.

C. Proof of Theorem 1

To focus the discussion, we briefly review the procedure of the proof.

- Referring to Lemma 2, we first need to check the whether $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ has a minimum with respect to $\{x_0, u_0, u_1, \cdots u_k\}$. This is done via the Krein space–Kalman filter corresponding to (23) and (24) and yields the condition (8) along with several equivalent conditions.
- Next we need to choose the $\{\check{z}_{k|k}\}_{k=0}^{i}$ such that the value of $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ is positive at its minimum. Now according to Theorem 6 in Part I, the value at the minimum is $J_{f,k}(\min) = \sum_{j=0}^{k} e_j R_{e,j}^{-1} e_j$, where e_j is the innovation corresponding to (23) and (24). We can then compute the $\{e_j\}$ using the Krein space–Kalman filter, and thereby choose the appropriate $\{\check{z}_{k|k}\}_{k=0}^{i}$ which yields the desired *a posteriori* filter.

A remark on the strong regularity of the model (23), (24): In what follows, we would like to use the Krein space–Kalman filter corresponding to the state-space model (23), (24). This of course requires the strong regularity of its output Gramian matrix which we denote by $R_{\substack{y \\ z}}$ (since the output of (23) consists of both a $\boldsymbol{y}$ and a $\boldsymbol{z}$ component).

If $R_{\substack{y \\ z}}$ is strongly regular, then the Krein space–Kalman filter may be applied to check for the positivity of $J_{f,k}$ for each $0 \le k \le i$. But what if $R_{\substack{y \\ z}}$ is not strongly regular? Then it turns out that $J_{f,k}$ cannot be positive for all $0 \le k \le i$. To see why, suppose that $J_{f,k} > 0$ for some arbitrary k. Then $J_{f,k}$ must have a minimum, and according to Lemma 9 in [1], the leading $k \times k$ block submatrices of $R_{\substack{y \\ z}}$ and $R - S^*QS = R$ must have the same inertia. Now due to (25), all leading submatrices of R are nonsingular, and since k was arbitrary, the same will be true of $R_{\substack{y \\ z}}$. Therefore $R_{\substack{y \\ z}}$ will be strongly regular.

To summarize, we may use the Krein space–Kalman filter to check the positivity of $J_{f,k}$. If one of the $R_{e,k}$ becomes singular (so that $R_{\substack{y \\ z}}$ is no longer strongly regular), $J_{f,k}$ will lose its positivity by default.

Proof of Existence Condition (8): The Riccati recursion corresponding to (23) is the exact same Riccati recursion that was given by (9) in Theorem 1. We can now apply any of the conditions for a minimum developed in [1] to check whether a minimum exists for $J_{f,k}(x_0, u_0, \cdots, u_k, y_0 \cdot \cdots, y_k)$ for all $0 \le k \le i$. If we assume that the $[F_k \quad G_k]$ have full rank, then according to Lemma 13 in [1], $J_{f,k}(x_0, u_0, \cdots, u_k, y_0 \cdot \cdots, y_k)$ will have a minimum for all $0 \le k \le i$, iff

$$P_{j|j}^{-1} = P_j^{-1} + [H_j^* \quad L_j^*] \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix} > 0$$

which yields the condition (8).

Since we still need to satisfy the second condition of Lemma 2, this, of course, only shows that (8) is a necessary condition for the existence of an H^∞ *a posteriori* filter of level γ_f. We shall later show, however, that if the minimum condition is satisfied, then the second condition of Lemma 2 can also be satisfied. Therefore (8) is indeed necessary and sufficient for the existence of the filter.

Other Existence Conditions: Using the results of [1], we can obtain alternative conditions for the existence of H^∞ *a posteriori* filters of level γ_f. If we use Lemma 12 in [1], we have the following condition.

Lemma 3 (Alternative Test for Existence): The condition (8) can be replaced by the condition that

$$R_j = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix}$$

and

$$R_{e,j} = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j [H_j^* \quad L_j^*]$$

have the same inertia for all $0 \le j \le i$. We no longer require that $[F_j \quad G_j]$ have full rank, and the size of the matrices involved is generally smaller than in (8).

Using a block triangular factorization of $R_{e,j}$ and the fact that when we have a minimum, P_j is positive definite, we can show the following result.

Corollary 1 (Alternative Test for Existence): The condition of Lemma 3 is equivalent to

$$I + H_j P_j H_j^* > 0$$

and

$$-\gamma_f^2 I + L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^* < 0 \tag{26}$$

for all $0 \le j \le i$.

The test of Lemma 3 has various advantages over (8) that are mentioned in the discussions following Lemma 13 in [1]. In particular, Lemma 3 allows us to go to a square-root form of the H^∞-filtering algorithm, where there is no need to explicitly check for the existence condition; these conditions are built into the square-root recursions themselves so that a solution exists iff the algorithm can be performed [24].

Many alternative existence conditions can also be obtained. Here is one that follows Lemma 14 in [1].

Lemma 4 (Alternative Test for Existence): If the $\{F_j\}$ are nonsingular, an H^∞ *a posteriori* filter of level γ_f exists, iff

$$P_{i+1} > 0$$

and

$$I - G_j^* P_{j+1}^{-1} G_j > 0 \quad j = 0, 1, \cdots, i.$$

Construction of the H^∞ A Posteriori Filters: To complete the proof of Theorem 1 we still need to show that if a minimum over $\{x_0, u_0, \cdots, u_k\}$ exists for all $0 \le k \le i$, then we can find the estimates $\{\check{z}_{k|k}\}_{k=0}^{i}$ such that the value of $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ at its minimum is positive.

According to Theorem 6 in [1], the minimum value of $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ is

$$\sum_{j=0}^{k} [e_{y,j}^* \quad e_{z,j}^*] R_{e,j}^{-1} \begin{bmatrix} e_{y,j} \\ e_{z,j} \end{bmatrix} = \sum_{j=0}^{k} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{z}_{j|j} - \hat{z}_{j|j-1} \end{bmatrix}^* \times \begin{bmatrix} I + H_j P_j H_j^* & H_j P_j L_j^* \\ L_j P_j H_j^* & -\gamma_f^2 I + L_j P_j L_j^* \end{bmatrix}^{-1} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{z}_{j|j} - \hat{z}_{j|j-1} \end{bmatrix}$$

where $\hat{y}_{j|j-1}$ and $\hat{z}_{j|j-1}$ are obtained from the Krein-space projections of $\boldsymbol{y}_j$ and $\check{\boldsymbol{z}}_{j|j}$ onto $\mathcal{L}\left\{\{\boldsymbol{y}_l\}_{l=0}^{j-1}, \{\check{\boldsymbol{z}}_{l|l}\}_{l=0}^{j-1}\right\}$, respectively. Thus $\hat{z}_{j|j-1}$ is a linear function of $\{y_l\}_{l=0}^{j-1}$.

Using the block triangular factorization of the $R_{e,j}$ we may rewrite the above as

$$\sum_{j=0}^{k} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{z}_{j|j} - \hat{z}_{j|j} \end{bmatrix}^* \times \begin{bmatrix} I + H_j P_j H_j^* & 0 \\ 0 & -\gamma_f^2 I + L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^* \end{bmatrix}^{-1} \times \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{z}_{j|j} - \hat{z}_{j|j} \end{bmatrix}$$

where

$$\begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{z}_{j|j} - \hat{z}_{j|j} \end{bmatrix} \triangleq \begin{bmatrix} I & 0 \\ -L_j P_j H_j^* (I + H_j P_j H_j^*)^{-1} & I \end{bmatrix} \times \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{z}_{j|j} - \hat{z}_{j|j-1} \end{bmatrix}. \tag{27}$$

Note that $\hat{z}_{j|j}$ is obtained from the Krein-space projection of $\check{\boldsymbol{z}}_{j|j}$ onto $\mathcal{L}\left\{\{\boldsymbol{y}_l\}_{l=0}^{j}, \{\check{\boldsymbol{z}}_{l|l}\}_{l=0}^{j-1}\right\}$ and is, therefore, a linear function of $\{y_l\}_{l=0}^{j}$. Recall from Corollary 1 that

$$\begin{aligned} &I + H_j P_j H_j^* > 0 \quad \text{and} \\ &-\gamma_f^2 I + L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^* < 0. \end{aligned} \tag{28}$$

Therefore all we must do is choose some $\check{z}_{j|j}$ such that

$$\begin{aligned} &\sum_{j=0}^{k} (y_j - \hat{y}_{j|j-1})^* (I + H_j P_j H_j^*)^{-1} (y_j - \hat{y}_{j|j-1}) \\ &+ \sum_{j=0}^{k} (\check{z}_{j|j} - \hat{z}_{j|j})^* (-\gamma_f^2 I + L_j (P_j^{-1} + H_j H_j^*)^{-1} L_j^*)^{-1} \\ &\times (\check{z}_{j|j} - \hat{z}_{j|j}) > 0. \end{aligned} \tag{29}$$

There are many such choices, but in view of (28), the simplest is

$$\check{z}_{j|j} = \hat{z}_{j|j} = L_j \hat{x}_{j|j} \quad (j \le k \le i)$$

where $\hat{x}_{j|j}$ is given by the Krein-space projection of the state $\boldsymbol{x}_j$ onto $\left\{\{\boldsymbol{y}_l\}_{l=0}^{j}, \{\check{\boldsymbol{z}}_{l|l}\}_{l=0}^{j-1}\right\}$. We may now utilize the filtered form of the Krein space–Kalman filter corresponding to the state-space model (23) to recursively compute $\hat{x}_{j|j}$ (see [1, Corollary 4])

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + P_{j+1} [H_{j+1}^* \quad L_{j+1}^*] \times R_{e,j+1}^{-1} \begin{bmatrix} y_{j+1} - \hat{y}_{j+1|j} \\ \check{z}_{j+1|j+1} - \hat{z}_{j+1|j} \end{bmatrix}. \tag{30}$$

Using $\hat{y}_{j+1|j} = H_{j+1} F_j \hat{x}_{j|j}$ and the above-mentioned triangular factorization of $R_{e,j+1}$, we have the equation shown at the bottom of the next page. Choosing $\check{z}_{j+1|j+1} = \hat{z}_{j+1|j+1}$ yields the desired recursion of Theorem 1

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1} \times (y_{j+1} - H_{j+1} F_j \hat{x}_{j|j}).$$

□

D. Parameterization of All H^∞ A Posteriori Filters

The filter of Theorem 1 is one among many possible filters with level γ. All filters that guarantee $J_{f,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k) > 0$ are represented by (28) and (29). We may use these expressions to obtain a more explicit characterization of all possible estimators. Similar results appear in [4], [7], and [11].

Theorem 3 (All H^∞ A Posteriori Estimators): All H^∞ *a posteriori* estimators that achieve a level γ_f (assuming they exist) are given by

$$\begin{aligned}\check{z}_{j|j} &= L_j\hat{x}_{j|j} + [\gamma_f^2 I - L_j(P_j^{-1} + H_j^* H_j)^{-1} L_j^*]^{\frac{1}{2}} \\ &\quad\times \mathcal{S}_j((I + H_j P_j H_j^*)^{\frac{1}{2}}(y_j - H_j\hat{x}_{j|j}), \cdots, \\ &\quad (I + H_0 P_0 H_0^*)^{\frac{1}{2}}(y_0 - H_0\hat{x}_{0|0}))\end{aligned} \tag{31}$$

where $\hat{x}_{j|j}$ satisfies the recursion

$$\begin{aligned}\hat{x}_{j+1|j+1} &= F_j\hat{x}_{j|j} + K_{s,j+1}(y_{j+1} - H_{j+1}F_j\hat{x}_{j|j}) \\ &\quad - K_{c,j}(\check{z}_{j|j} - L_j\hat{x}_{j|j})\end{aligned} \tag{32}$$

with $K_{s,j+1}$ the same as in Theorem 1

$$\begin{aligned}K_{c,j} &= (I + P_{j+1}H_{j+1}H_{j+1}^*)^{-1} \\ &\quad\times F_j(P_j^{-1} + H_j H_j^* - \gamma_f^{-2} L_j L_j^*)^{-1} L_j^*\end{aligned} \tag{33}$$

and

$$\mathcal{S}(a_j, \cdots, a_0) = \begin{bmatrix} \mathcal{S}_0(a_0) \\ \mathcal{S}_1(a_1, a_0) \\ \vdots \\ \mathcal{S}_j(a_j, \cdots, a_0)\end{bmatrix}$$

is any (possibly nonlinear) contractive causal mapping, i.e.,

$$\sum_{j=0}^{k} |\mathcal{S}_j(a_j, \cdots, a_0)|^2 < \sum_{j=0}^{k} |a_j|^2 \quad \text{for all } k = 0, 1, \cdots, i.$$

Remark: Note that when the contraction of Theorem 3 is chosen as $\mathcal{S} = 0$, then we have $\check{z}_{j|j} = L_j\hat{x}_{j|j}$, and (32) reduces to the recursion of Theorem 1.

Proof of Theorem 3: Expression (29) may be rewritten as

$$\begin{aligned}&\sum_{j=0}^{k}(y_j - H_j\hat{x}_j)^*(I + H_j P_j H_j^*)^{-1}(y_j - H_j\hat{x}_j) \\ &\quad + \sum_{j=0}^{k}(\check{z}_{j|j} - L\hat{x}_{j|j})^*\left(-\gamma_f^2 I + L_j(P_j^{-1} + H_j H_j^*)^{-1} L_j^*\right)^{-1} \\ &\quad \times (\check{z}_{j|j} - L_j\hat{x}_{j|j}) > 0\end{aligned} \tag{34}$$

where $\hat{x}_j$ and $\hat{x}_{j|j}$ denote the Krein-space projections of $\boldsymbol{x}_j$ onto $\left\{\{\boldsymbol{y}_l\}_{l=0}^{j-1}, \{\check{z}_{l|l}\}_{l=0}^{j-1}\right\}$ and $\left\{\{\boldsymbol{y}_l\}_{l=0}^{j}, \{\check{z}_{l|l}\}_{l=0}^{j-1}\right\}$, respectively. Therefore $\hat{x}_j$ and $\hat{x}_{j|j}$ are related through one additional projection onto y_j, and we may write

$$\hat{x}_{j|j} = \hat{x}_j + P_j H_j^*(I + H_j P_j H_j^*)^{-1}(y_j - H_j\hat{x}_j). \tag{35}$$

Therefore

$$\begin{aligned}y_j - H_j\hat{x}_{j|j} &= \left(I - H_j P_j H_j^*(I + H_j P_j H_j^*)^{-1}\right)(y_j - H_j\hat{x}_j) \\ &= (I + H_j P_j H_j^*)^{-1}(y_j - H_j\hat{x}_j)\end{aligned}$$

so that

$$(y_j - H_j\hat{x}_j) = (I + H_j P_j H_j^*)(y_j - H_j\hat{x}_{j|j}).$$

Now (34) can be written as

$$\begin{aligned}&\sum_{j=0}^{k}(y_j - H_j\hat{x}_{j|j})^*(I + H_j P_j H_j^*)(y_j - H_j\hat{x}_{j|j}) \\ &\quad + \sum_{j=0}^{k}(\check{z}_{j|j} - L_j\hat{x}_{j|j})^*\left(-\gamma_f^2 I + L_j(P_j^{-1} + H_j H_j^*)^{-1} L_j^*\right)^{-1} \\ &\quad \times (\check{z}_{j|j} - L_j\hat{x}_{j|j}) > 0\end{aligned}$$

or equivalently

$$\begin{aligned}&\|(\gamma_f^2 I - L_j(P_j^{-1} + H_j H_j^*)^{-1} L_j^*)^{-\frac{1}{2}}(\check{z}_{j|j} - \hat{z}_{j|j})\|_2^2 \\ &\quad < \|(I + H_j P_j H_j^*)^{\frac{1}{2}}(y_j - H_j\hat{x}_{j|j})\|_2^2.\end{aligned}$$

Since $\check{z}_{j|j}$ is a causal function of the observations y_j, then $(\check{z}_{j|j} - \hat{z}_{j|j})$ will also be a causal function of $(y_j - H_j\hat{x}_{j|j})$. Therefore using the above expression, we can write

$$\begin{bmatrix}(\gamma_f^2 I - L_0(P_0^{-1} + H_0 H_0^*)^{-1} L_0^*)^{-\frac{1}{2}}(\check{z}_{0|0} - \hat{z}_{0|0}) \\ \vdots \\ (\gamma_f^2 I - L_i(P_i^{-1} + H_i H_i^*)^{-1} L_i^*)^{-\frac{1}{2}}(\check{z}_{i|i} - \hat{z}_{i|i})\end{bmatrix} = \mathcal{S}\begin{bmatrix}(I + H_0 P_0 H_0^*)^{\frac{1}{2}}(y_0 - H_0\hat{x}_{0|0}) \\ \vdots \\ (I + H_i P_i H_i^*)^{\frac{1}{2}}(y_i - H_i\hat{x}_{i|i})\end{bmatrix}$$

for some causal contractive mapping $\mathcal{S}$. Equation (31) now readily follows.

Finally, we must show (32). To this end, recall from the proof of Theorem 1 [see (30)] that the recursion for $\hat{x}_j$ is given by

$$\begin{aligned}\hat{x}_{j+1|j+1} &= F_j\hat{x}_{j|j} + P_{j+1}[H_{j+1}^* \quad L_{j+1}^*] \\ &\quad \times R_{e,j+1}^{-1}\begin{bmatrix} y_{j+1} - \hat{y}_{j+1|j} \\ \check{z}_{j+1|j+1} - \hat{z}_{j+1|j}\end{bmatrix}.\end{aligned}$$

Using (35) to replace $\hat{x}_j$ by $\hat{x}_{j|j}$ yields, after some algebra, the desired recursion (32). □

Note that although the filter obtained in Theorem 1 is linear, the full parameterization of all H^∞ filters with level γ_f is given by a nonlinear causal contractive mapping $\mathcal{S}$. The filter

$$\begin{aligned}\hat{x}_{j+1|j+1} &= F_j\hat{x}_{j|j} + P_{j+1}[H_{j+1}^* \quad L_{j+1}^*]\begin{bmatrix} I & -(I + H_{j+1}P_{j+1}H_{j+1}^*)^{-1}H_{j+1}P_{j+1}L_{j+1}^* \\ 0 & I\end{bmatrix} \\ &\quad\times\begin{bmatrix} I + H_{j+1}P_{j+1}H_{j+1}^* & 0 \\ 0 & -\gamma_f^2 I + L_{j+1}(P_{j+1}^{-1} + H_{j+1}^* H_{j+1})^{-1} L_{j+1}^*\end{bmatrix}^{-1}\begin{bmatrix} y_{j+1} - H_{j+1}F_j\hat{x}_{j|j} \\ \check{z}_{j+1|j+1} - \hat{z}_{j+1|j+1}\end{bmatrix}\end{aligned}$$

of Theorem 1 is known as the central filter, and as we have seen, corresponds to $\mathcal{S} = 0$. This central filter has a number of other interesting properties. It corresponds, as we shall see in the next section, to the risk-sensitive optimal filter and can be shown to be the maximum entropy filter [10]. Moreover, in the game theoretic formulation of the H^∞ problem, the central filter corresponds to the solution of the game [12]. In our context, the central filter is recognized as the Krein space–Kalman filter corresponding to the state-space model (23).

E. Derivation of the A Priori H^∞ Filter

We shall now turn to the H^∞ *a priori* filter of Problem 2, and our main goal will be to prove the results of Theorem 2. Our approach will follow the one used for the *a posteriori* case, namely we will relate an indefinite quadratic form to the *a priori* problem, construct its corresponding Krein space state-space model, and use the Krein space–Kalman filter to obtain the solution. Since our derivations parallel the ones given earlier, we shall omit several details.

The Suboptimal H^∞ A Priori Problem and Quadratic Forms: Referring to Problem 2, we first note that $\|T_i\ (\mathcal{F}_p)\|_\infty < \gamma_p$ implies that for all nonzero $\{x_0, \{u_j\}_{j=0}^{i-1}, \{v_j\}_{j=0}^{i-1}\}$

$$\frac{\sum_{j=0}^{i} |\check{z}_j - L_j x_j|^2}{x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i-1} |u_j|^2 + \sum_{j=0}^{i-1} |y_j - H_j x_j|^2} < \gamma_p^2 \quad (36)$$

where, without loss of generality, we have assumed $\check{x}_0 = 0$. Moreover, (36) implies that for all $k \le i$, we must have

$$\frac{\sum_{j=0}^{k} |\check{z}_j - L_j x_j|^2}{x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{k-1} |u_j|^2 + \sum_{j=0}^{k-1} |y_j - H_j x_j|^2} < \gamma_p^2. \quad (37)$$

As before, we may easily show the following result.

Lemma 5 (Indefinite Quadratic Form): Given a scalar $\gamma_p > 0$, then $\|T_i\ (\mathcal{F}_p)\|_\infty < \gamma_p$ if, and only if, there exists $\check{z}_k = \mathcal{F}_p\ (y_0, \cdots, y_{k-1})$ (for all $0 \le k \le i$) such that for all complex vectors x_0, for all causal sequences $\{u_j\}_{j=0}^{i-1}$, and for all nonzero causal sequences $\{y_j\}_{j=0}^{i-1}$ the scalar quadratic form

$$\begin{aligned} &J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1}) \\ &= x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{k-1} u_j^* u_j + \sum_{j=0}^{k-1} (y_j - H_j x_j)^*(y_j - H_j x_j) \\ &\quad - \gamma_p^{-2} \sum_{j=0}^{k} (\check{z}_j - L_j x_j)^*(\check{z}_j - L_j x_j) \end{aligned} \quad (38)$$

satisfies

$$J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1}) > 0 \quad \text{for all } 0 \le k \le i. \quad (39)$$

We can also readily obtain the analog of Lemma 2.

Lemma 6 (Positivity Condition): The scalar quadratic forms $J_{p,k}\ (x_0,\ u_0,\ \cdots,\ u_{k-1}\ , y_0, \cdots, y_{k-1})$ satisfy the conditions (39), iff, for all $0 \le k \le i$

i) $J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1})$ has a minimum with respect to $\{x_0, u_0, u_1, \cdots u_{k-1}\}$.
ii) The $\{\check{z}_k\}_{k=0}^{i}$ can be chosen such that the value of $J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1})$ at this minimum is positive, viz.

$$\min_{\{x_0, u_0, \cdots, u_{k-1}\}} J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1}) > 0.$$

F. A Krein Space State-Space Model

Because the summations in $J_{p,k}$ go up to both k and $k-1$ [see (38)], it is slightly more difficult to come up with a Krein state-space model whose corresponding quadratic form is $J_{p,k}$. With some effort, however, we see that the appropriate Krein state-space model is

$$\begin{cases} \boldsymbol{\xi}_{2j+1} = \boldsymbol{\xi}_{2j}, \quad \boldsymbol{\xi}_0 = \boldsymbol{x}_0 \\ \check{\boldsymbol{z}}_j = L_j \boldsymbol{\xi}_{2j} + \boldsymbol{v}_{2j} \\ \boldsymbol{\xi}_{2j+2} = F_j \boldsymbol{\xi}_{2j+1} + G_j \bar{\boldsymbol{u}}_{2j+1} = \boldsymbol{x}_{2i+2} \\ \boldsymbol{y}_j = H_j \boldsymbol{\xi}_{2j+1} + \boldsymbol{v}_{2j+1} \end{cases} \quad j \le 0 \quad (40)$$

where $\Pi_0 > 0$, $Q_{2j} = 0$, $Q_{2j+1} = I$, $R_{2j} = -\gamma_p^2 I$, $R_{2j+1} = I$, and $S_j = 0$. To see why, let us construct the deterministic quadratic form corresponding to (40). Thus

$$\begin{aligned} J_{\xi,2k} &= \xi_0^* \Pi_0^{-1} \xi + \sum_{j=0}^{2k} \bar{u}_j^* Q_j^{-1} \bar{u}_j + \sum_{j=0}^{2k} v_j^* R_j^{-1} v_j \\ &= \xi_0^* \Pi_0^{-1} \xi + \sum_{j=0}^{k-1} \bar{u}_{2j+1}^* \bar{u}_{2j+1} \\ &\quad + \sum_{j=0}^{k-1} v_{2j+1}^* R_{2j+1}^{-1} v_{2j+1} + \sum_{j=0}^{k} v_{2j}^* R_{2j}^{-1} v_{2j} \\ &= \xi_0^* \Pi_0^{-1} \xi + \sum_{j=0}^{k-1} \bar{u}_{2j+1}^* \bar{u}_{2j+1} \\ &\quad + \sum_{j=0}^{k-1} |y_j - H_j \xi_{2j+1}|^2 - \gamma_p^{-2} \sum_{j=0}^{k} |\check{z}_j - L_j \xi_{2j}|^2. \end{aligned}$$

From (40) we see that $\xi_{2j} = \xi_{2j+1} = x_j$. Using this fact, and defining $\bar{u}_{2j+1} = u_j$, we readily see that $J_{\xi,2k} = J_{p,k}$.

Note also that the Riccati recursion for the model (40) is (41) shown at the bottom of the page.

Existence Conditions: Using Lemma 12 from [1], the condition for a minimum is that $R_{e,j}$ and R_j should have the same inertia for all $j = 0, 1, \cdots, 2i$ (since each two time steps in (40) correspond to one time step in $J_{p,k}$). Thus the condition for a minimum is

$$-\gamma_p^2 I + L_j \Sigma_{2j} L_j^* < 0 \quad \text{and} \quad I + H_j \Sigma_{2j+1} H_j^* > 0. \quad (42)$$

$$\begin{cases} \Sigma_{2j+1} &= \Sigma_{2j} - \Sigma_{2j} L_j^* (-\gamma_p^2 I + L_j \Sigma_{2j} L_j^*)^{-1} L_j \Sigma_{2j} \\ \Sigma_{2j+2} &= F_j \Sigma_{2j+1} F_j^* + G_j G_j^* - F_j \Sigma_{2j+1} H_j^* (I + H_j \Sigma_{2j+1} H_j^*)^{-1} H_j \Sigma_{2j+1} F_j^* \end{cases} \quad \Sigma_0 = \Pi_0. \quad (41)$$

The second of the above conditions is obvious since when we have a minimum Σ_j is positive definite. If the $[F_j \quad G_j]$ have full rank, then using Lemma 13 from [1], the condition for a minimum is

$$\Sigma_{2j}^{-1} - \gamma_p^{-2} L_j^* L_j > 0 \quad \text{and} \quad \Sigma_{2j+1}^{-1} + H_j^* H_j > 0 \tag{43}$$

where once more the second of the above conditions is redundant.

To connect with the results of Theorem 2, we may note that by defining $P_j = \Sigma_{2j}$ and combining the coupled pair of Riccati recursions in (41), we can write the following Riccati recursion for P_j

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j [L_j^* \quad H_j^*] \times R_{e,j}^{-1} \begin{bmatrix} L_j \\ H_j \end{bmatrix} P_j F_j^*, \qquad P_0 = \Pi_0 \tag{44}$$

with

$$R_{e,j} = \begin{bmatrix} -\gamma_p^2 I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} L_j \\ H_j \end{bmatrix} P_j [L_j^* \quad H_j^*]. \tag{45}$$

But this is the same Riccati as (9) in Theorem 2. Thus the condition for a minimum (43) becomes

$$P_j^{-1} - \gamma_p^{-2} L_j^* L_j > 0 \tag{46}$$

which is the condition (13) of Theorem 2.

To express the condition (42) in a form that is more similar to that of Lemma 3, we introduce the Krein space state-space model

$$\begin{cases} \boldsymbol{x}_{j+1} = F_j \boldsymbol{x}_j + G_j \boldsymbol{u}_j \\ \begin{bmatrix} \check{\boldsymbol{z}}_{j|j} \\ \boldsymbol{y}_j \end{bmatrix} = \begin{bmatrix} L_j \\ H_j \end{bmatrix} \boldsymbol{x}_j + \boldsymbol{v}_j \end{cases} \tag{47}$$

where $\Pi_0 > 0$, $Q_j = I$, $S_j = 0$ and

$$R_j = \begin{bmatrix} -\gamma_p^2 I & 0 \\ 0 & I \end{bmatrix}.$$

Note that the only difference between the state-space models (23) and (47) is that the order of the output equations has been reversed.

We can now use the state-space model (47) to express the condition (42) in the form of the following Lemma.

Lemma 7 (Alternative Test for Existence): The condition (13) can be replaced by the condition that all leading submatrices of

$$R_j = \begin{bmatrix} -\gamma_p^2 I & \mathbf{o} \\ \mathbf{o} & I \end{bmatrix} \text{ and}$$
$$R_{e,j} = \begin{bmatrix} -\gamma_p^2 I & \mathbf{o} \\ \mathbf{o} & I \end{bmatrix} + \begin{bmatrix} L_j \\ H_j \end{bmatrix} P_j [L_j^* \quad H_j^*]$$

have the same inertia for all $0 \le j \le i$. In other words

$$-\gamma_p^2 I + L_j P_j L_j^* < 0 \quad \text{and} \quad I + H_j \tilde{P}_j H_j^* > 0$$

where $\tilde{P}_j^{-1} = P_j^{-1} - \gamma_p^{-2} L_j^* L_j$. We no longer require that $[F_j \quad G_j]$ have full rank, and the size of the matrices involved is generally smaller than in (8).

Note that compared to Lemma 3, the condition in Lemma 7 is more stringent since it requires that all leading submatrices of R_j and $R_{e,j}$ have the same inertia. This distinction is especially important in square-root implementations of the H^∞ filters [24].

Construction of the H^∞A Priori Filter: To complete the proof of Theorem 2 we still need to show that if a minimum over $\{x_0, u_0, \cdots, u_{k-1}\}$ exists for all $0 \le k \le i$, then we can find the estimates $\{\check{z}_k\}_{k=0}^{i}$ such that the value of $J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1})$ at its minimum is positive.

According to Theorem 6 in [1], the minimum value of $J_{p,k}(x_0, u_0, \cdots, u_{k-1}, y_0, \cdots, y_{k-1})$ is

$$\sum_{j=0}^{k-1} [e_{z,j}^* \quad e_{y,j}^*] R_{e,j}^{-1} \begin{bmatrix} e_{z,j} \\ e_{y,j} \end{bmatrix} + e_{z,k}^* (-\gamma_p^2 I + L_k P_k L_k^*)^{-1} e_{z,k}$$
$$= \sum_{j=0}^{k-1} \begin{bmatrix} \check{z}_j - \hat{z}_{j|j-1} \\ y_j - \hat{y}_{j|j-1} \end{bmatrix}^* \begin{bmatrix} -\gamma_p^2 I + L_j P_j L_j^* & L_j P_j H_j^* \\ H_j P_j L_j^* & I + H_j P_j H_j^* \end{bmatrix}^{-1} \times \begin{bmatrix} \check{z}_j - \hat{z}_{j|j-1} \\ y_j - \hat{y}_{j|j-1} \end{bmatrix} + (\check{z}_k - \hat{z}_{k|k-1})^* \times (-\gamma_p^2 I + L_k P_k L_k^*)^{-1} (\check{z}_k - \hat{z}_{k|k-1}) > 0$$

where $\hat{z}_{j|j-1}$ and $\hat{y}_{j|j-1}$ are obtained from the Krein-space projections of $\check{\boldsymbol{z}}_j$ and $\boldsymbol{y}_j$ onto $\mathcal{L}\{\{\check{z}_l\}_{l=0}^{j-1}, \{\boldsymbol{y}_l\}_{l=0}^{j-1}\}$, respectively. Thus $\hat{z}_{j|j-1}$ is a linear function of $\{y_l\}_{l=0}^{j-1}$. Using the block triangular factorization of the $R_{e,j}$ we may rewrite the above as

$$\sum_{j=0}^{k} (\check{z}_j - \hat{z}_{j|j-1})^* (-\gamma_p^2 I + L_j P_j L_j^*)^{-1} (\check{z}_j - \hat{z}_{j|j-1}) + \sum_{j=0}^{k-1} (y_j - \bar{y}_{j|j-1})^* (I + H_j \tilde{P}_j H_j^*)^{-1} (y_j - \bar{y}_{j|j-1}) > 0 \tag{48}$$

where

$$\begin{bmatrix} \check{z}_j - \hat{z}_{j|j-1} \\ y_j - \bar{y}_{j|j-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -H_j P_j L_j^* (-\gamma_p^2 I + L_j P_j L_j^*)^{-1} & I \end{bmatrix} \times \begin{bmatrix} \check{z}_j - \hat{z}_{j|j-1} \\ y_j - \hat{y}_{j|j-1} \end{bmatrix}. \tag{49}$$

Note that $\bar{y}_{j|j-1}$ is given by the Krein-space projection of $\boldsymbol{y}_j$ onto $\{\{\check{z}_l\}_{l=0}^{j}, \{\boldsymbol{y}_l\}_{l=0}^{j-1}\}$. Recall from Lemma 7 that

$$-\gamma_p^2 I + L_j P_j L_j^* < 0, I + H_j \tilde{P}_j H_j^* > 0.$$

Any choice of $\check{z}_{j|j-1}$ that renders (48) positive will do, and the simplest choice is $\check{z}_{j|j-1} = \hat{z}_{j|j-1} = L_j \hat{x}_{j|j-1}$, where $\hat{x}_{j|j-1}$ is given by the Krein-space projection of $\boldsymbol{x}_j$ onto $\{\{\check{z}_l\}_{l=0}^{j-1}, \{\boldsymbol{y}_j\}_{l=0}^{j-1}\}$. We may now utilize the Krein space–Kalman filter corresponding to the state-space model (47) to recursively compute $\hat{x}_{j|j-1}$, viz.

$$\hat{x}_{j+1|j} = F_j \hat{x}_{j|j-1} + F_j P_j [L_j^* \quad H_j^*] \times R_{e,j}^{-1} \begin{bmatrix} \check{z}_j - L_j \hat{x}_{j|j-1} \\ y_j - H_j \hat{x}_{j|j-1} \end{bmatrix}. \tag{50}$$

Setting $\check{z}_j - L_j \hat{x}_{j|j-1} = 0$ and simplifying, we get the desired recursion for $\hat{x}_{j+1|j}$. □

G. All H^∞ A Priori Filters

The positivity condition (48) gives a full parameterization of all H^∞ *a priori* estimators. We thus have the following result.

Theorem 4 (All H^∞ A Priori Estimators): All H^∞ *a priori* estimators that achieve a level γ_p (assuming they exist) are given by

$$\begin{aligned}\hat{z}_j = L_j\hat{x}_j + (\gamma_p^2 I - L_j P_j L_j^*)^{\frac{1}{2}} \\ \times \mathcal{S}_j\big((I + H_{j-1}\tilde{P}_{j-1}H_{j-1}^*)^{-\frac{1}{2}}(y_{j-1} - H_{j-1}\bar{x}_{j-1}), \cdots, \\ \times (I + H_0\tilde{P}_0 H_0^*)^{-\frac{1}{2}}(y_0 - H_0\bar{x}_0)\big)\end{aligned} \tag{51}$$

where

$$\bar{x}_k = \hat{x}_k + P_k L_k^*(-\gamma_p^2 I + L_k L_k^*)^{-1}(\check{z}_k - L_k\hat{x}_k) \tag{52}$$

$\hat{x}_j$ satisfies the recursion

$$\hat{x}_{j+1|j} = F_j\hat{x}_{j|j-1} + F_j P_j [L_j^* \quad H_j^*] \times R_{e,j}^{-1}\begin{bmatrix}\check{z}_j - L_j\hat{x}_{j|j-1} \\ y_j - H_j\hat{x}_{j|j-1}\end{bmatrix} \tag{53}$$

with P_j, $\tilde{P}_j$, and $R_{e,j}$ given by Theorem 2 and $\mathcal{S}$ is any (possibly nonlinear) contractive causal mapping.

Proof: Referring to (49), we see that the $\bar{y}_{j|j-1} = H_j\bar{x}_j$ differ from $\hat{y}_{j|j-1} = H_j\hat{x}_j$ via the additional projection onto $\check{z}_{j|j-1}$. Thus we can write

$$\bar{x}_{j|j-1} = \hat{x}_{j|j-1} + P_j L_j^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1}(\check{z}_{j|j-1} - \hat{z}_{j|j-1})$$

which proves (52). Moreover, from the proof of Theorem 2 [see (50)], the recursion for $\hat{x}_j$ is given by (53). Condition (48) can now be rewritten as

$$\begin{aligned}\sum_{j=0}^{k}(\check{z}_{j|j-1} - \hat{z}_{j|j-1})^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1}(\check{z}_{j|j-1} - \hat{z}_{j|j-1}) \\ + \sum_{j=0}^{k-1}(y_j - \bar{y}_{j|j-1})^*(I + H_j\tilde{P}_j H_j^*)^{-1}(y_j - \bar{y}_{j|j-1}) > 0\end{aligned} \tag{54}$$

and an argument similar to the one given in the proof of Theorem 3 will yield the desired result. □

H. The H^∞ Smoother

If instead of $e_{f,k}$ and $e_{p,k}$, which correspond to the *a posteriori* and *a priori* filters, respectively, we consider the smoothed error

$$e_{s,k} = \check{z}_{k|i} - L_k x_k, \qquad k \le i$$

where $\check{z}_{k|i} = \mathcal{F}_s(y_0, y_1, \cdots, y_i)$ is the estimate of z_k given all observations $\{y_j\}$ from time 0 until time i; we are led to the so-called H^∞ smoothers. Such estimators guarantee that the maximum energy gain from the disturbances $\{\Pi_0^{-1/2}(x_0 - \check{x}_0), \{u_j\}_{j=0}^i, \{v_j\}_{j=0}^i\}$ to the smoothing errors $\{e_{s,j}\}_{j=0}^i$ is bounded by γ_s, i.e.,

$$\sup_{x_0, u\in h_2, v\in h_2} \times \frac{\sum_{j=0}^i e_{s,j}^* e_{s,j}}{(x_0 - \check{x}_0)^*\Pi_0^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^i u_j^* u_j + \sum_{j=0}^i v_j^* v_j} < \gamma_s^2. \tag{55}$$

Using an argument similar to the ones given before, we are led to the following quadratic form

$$\begin{aligned}J_{s,i}(x_0, u_0, \cdots, u_i, y_0, \cdots, y_i) \\ = x_0\Pi_0^{-1}x_0^* + \sum_{k=0}^i u_k^* u_k + \sum_{k=0}^i (y_k - H_k x_k)^*(y_k - H_k x_k) \\ - \gamma^{-2}\sum_{k=0}^i(\check{z}_{k|i} - L_k x_k)^*(\check{z}_{k|i} - L_k x_k).\end{aligned} \tag{56}$$

Note that the only difference between $J_{s,i}$ and $J_{f,i}$ is that $\check{z}_{k|k}$ has been replaced by $\check{z}_{k|i}$ (i.e., filtered estimates have been replaced by smoothed estimates). Once more it can be shown that an H^∞ smoother of level γ_s will exist if, and only if, there exists some $\check{z}_{k|i}$ such that $J_{s,i} \ge 0$. The rather interesting result shown below, and which has already been pointed out in the literature (see e.g., [4], [7], and [16]), is that one H^∞ smoother is given by the conventional H^2 smoother (which does not even depend on γ_s).

Theorem 5 (H^∞ Smoother): For a given $\gamma_s > 0$, an H^∞ smoother that achieves level γ_s exists, iff the block diagonal matrix

$$R_e = R_{e,0} \oplus R_{e,1} \oplus \cdots \oplus R_{e,i}$$

where

$$R_{e,j} = \begin{bmatrix} I & 0 \\ 0 & -\gamma_s^2\end{bmatrix} + \begin{bmatrix}H_j \\ L_j\end{bmatrix} P_j [H_j^* \quad L_j^*]$$

and P_j is the same as in Theorem 1, has $(i+1)p$ positive eigenvalues and $(i+1)q$ negative eigenvalues. In other words, iff

$$\mathrm{In}[R_e] = [(i+1)p \quad 0 \quad (i+1)q].$$

If this is the case, one possible H_∞ smoother is given by the H^2 smoother.

Proof: The condition for $J_{s,i}(x_0, u_0, \cdots, u_i, y_0, \cdots, y_i)$ to have a minimum is slightly different than the earlier cases since we do not require that $J_{s,k}(x_0, u_0, \cdots, u_k, y_0, \cdots, y_k)$ have a minimum over the disturbances for all past values $k < i$. Thus using Lemma 9 from the companion paper [1], the condition for a minimum over $\{x_0, u_0, \cdots, u_i\}$ is that the matrices

$$R_e \text{ and } R = R_0 \oplus R_1 \oplus \cdots \oplus R_i$$

have the same inertia, where $R_j = I_p \oplus (-\gamma_s^2 I_q)$. But this is precisely the inertia condition given in the statement of the Theorem.

The value of $J_{s,i}$ at its minimum is (see [1])

$$[y^* \quad \check{z}_{|i}^*]\begin{bmatrix}R_y & R_{y\check{z}} \\ R_{\check{z}y} & R_{\check{z}}\end{bmatrix}^{-1}\begin{bmatrix} y \\ \check{z}_{|i}\end{bmatrix}$$

where we have defined

$$\begin{bmatrix} R_y & R_{y\check{z}} \\ R_{\check{z}y} & R_{\check{z}} \end{bmatrix} = \left\langle \begin{bmatrix} \boldsymbol{y} \\ \check{\boldsymbol{z}}_{|i} \end{bmatrix}, \begin{bmatrix} \boldsymbol{y} \\ \check{\boldsymbol{z}}_{|i} \end{bmatrix} \right\rangle \tag{57}$$

with

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_0 \\ \vdots \\ \boldsymbol{y}_i \end{bmatrix} \qquad \check{\boldsymbol{z}}_{|i} = \begin{bmatrix} \check{\boldsymbol{z}}_{0|i} \\ \vdots \\ \check{\boldsymbol{z}}_{i|i} \end{bmatrix}$$

$$y = \begin{bmatrix} y_0 \\ \vdots \\ y_i \end{bmatrix} \qquad \check{z}_{|i} = \begin{bmatrix} \check{z}_{0|i} \\ \vdots \\ \check{z}_{i|i} \end{bmatrix}$$

and where the $\{\boldsymbol{y}_j\}$ and $\{\check{\boldsymbol{z}}_{j|i}\}$ satisfy the Krein state-space model (23). In this case all the entries in $\check{z}_{|i}$ are unknown and there is no causal dependence between the $\{\check{z}_{j|k}\}$ and the $\{y_j\}$. Using a block triangular factorization, or a completion of squares argument, the value at the minimum can be rewritten as

$$y^* R_y^{-1} y + (\check{z}_{|i} - R_{\check{z}y} R_y^{-1} y)^* (R_{\check{z}} - R_{\check{z}y} R_y^{-1} R_{y\check{z}})^{-1} \times (\check{z}_{|i} - R_{\check{z}y} R_y^{-1} y).$$

But $R_y > 0$ (since it is the covariance of a Hilbert space state-space model), and hence one possible choice of $\check{z}_{|i}$ to guarantee $J_{s,i} > 0$ is to choose $\check{z}_{|i} = R_{\check{z}y} R_y^{-1} y = \hat{z}_{|i}$ which is clearly the H^2 smoothed estimate of $\check{z}$. □

The following result is now straightforward.

Theorem 6 (All H^∞ Smoothers): All H^∞ smoothers that achieve a level γ_s (assuming they exist) are given by

$$\check{z}_{|i} = \hat{z}_{|i} + (R_{\check{z}} - R_{\check{z}y} R_y^{-1} R_{y\check{z}})^{1/2} \mathcal{S}(R_y^{-1/2} y) \tag{58}$$

where $\mathcal{S}$ is any (not necessarily causal) contractive mapping, $\hat{z}_{|i}$ is the usual H^2 smoothed estimate, and R_y and $R_{\check{z}} - R_{\check{z}y} R_y^{-1} y R_{y\check{z}}$ are defined in (57).

It is clear from the discussions so far in the paper that the Krein-space estimation formalism provide simple derivations of H^∞ estimators. These estimators turn out to be certain Krein space–Kalman filters, and show that Krein-space estimation yields a unified approach to H^2 and H^∞ problems. To derive such filters and to solve other related problems as discussed ahead, all one essentially needs is to identify an indefinite quadratic form and to construct a convenient auxiliary state-space model with the appropriate Gramians. Two further applications of this approach are discussed next.

IV. Risk-Sensitive Estimation Filters

The so-called risk-sensitive (or exponential cost) criterion was introduced in [13] and further studied in [14]–[16]. Glover and Doyle [21] noticed their close connection to the H^∞ filters discussed earlier. We shall make this connection in a different way by bringing in an appropriate quadratic form.

A. The Exponential Cost Function

We again start with a state-space model of the form

$$\begin{cases} \boldsymbol{x}_{j+1} = F_j \boldsymbol{x}_j + G_j \boldsymbol{u}_j, & j \geq 0 \\ \boldsymbol{y}_j = H_j \boldsymbol{x}_j + \boldsymbol{v}_j. \end{cases} \tag{59}$$

We now assume, however, that $\boldsymbol{x}_0$, $\{\boldsymbol{u}_j\}$, and $\{\boldsymbol{v}_j\}$ are independent zero mean Gaussian random variables with covariances Π_0, Q_i, and R_i, respectively. We further assume that the $\{\boldsymbol{u}_j\}$ and $\{\boldsymbol{v}_j\}$ are white-noise processes. Conventional H^2 estimators, such as the Kalman filter, estimate the quantity $\boldsymbol{z}_i = L_i \boldsymbol{x}_i$ from the observations $\{\boldsymbol{y}_j\}$ by performing the following minimization (see e.g., [1], [22], and [23])

$$\min_{\{\check{\boldsymbol{z}}_{j|l}\}} E[\boldsymbol{C}_i] \tag{60}$$

where $\boldsymbol{C}_i = \sum_{j=0}^{i} (\check{\boldsymbol{z}}_{j|l} - L_j \boldsymbol{x}_j)^* (\check{\boldsymbol{z}}_{j|l} - L_j \boldsymbol{x}_j)$, $\check{\boldsymbol{z}}_{j|l}$ denotes the estimate of $\boldsymbol{z}_j$ given the observations up to and including time l, and $E[\cdot]$ denotes expectation. As we have seen earlier, $l = j$, $l = j - 1$, and $l = i$ correspond to the *a posteriori*, *a priori*, and smoothed estimation problems, respectively. Moreover, the expectation is taken over the Gaussian random variables $\boldsymbol{x}_0$ and $\{\boldsymbol{u}_i\}$ whose joint conditional distribution is given by

$$p(x_0, u_0, \cdots, u_i \mid y_0, \cdots, y_i) \propto \exp\left[-\frac{1}{2} J_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\right] \tag{61}$$

where the symbol $\propto$ stands for "proportional to," and $J_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)$ is equal to (using the fact that $\boldsymbol{x}_0$, $\{\boldsymbol{u}_j\}$, and $\{\boldsymbol{v}_j\}$ are independent, and that $v_j = y_j - H_j x_j$)

$$x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i} u_j^* Q_j^{-1} u_j + \sum_{j=0}^{i} (y_j - H_j x_j)^* R_j^{-1} (y_j - H_j x_j). \tag{62}$$

In the terminology of [15], the filter that minimizes (60) is known as a risk-neutral filter.

An alternative criterion that is risk-sensitive has been extensively studied in [13]–[16] and corresponds to the minimization problem

$$\min_{\{\check{\boldsymbol{z}}_{j|l}\}} \mu_i(\theta) = \min_{\{\check{\boldsymbol{z}}_{j|l}\}} \left(-\frac{2}{\theta} \log \left[E \exp(-\frac{\theta}{2} \boldsymbol{C}_i)\right]\right). \tag{63}$$

The criterion in (63) is known as an exponential cost criterion, and any filter that minimizes $\mu_i(\theta)$ is referred to as a risk-sensitive filter. The scalar parameter θ is correspondingly called the risk-sensitivity parameter. Some intuition concerning the nature of this modified criterion is obtained by expanding $\mu_i(\theta)$ in terms of θ and writing

$$\mu_i(\theta) = E(\boldsymbol{C}_i) - \frac{\theta}{4} \mathrm{Var}(\boldsymbol{C}_i) + O(\theta^2).$$

The above equation shows that for $\theta = 0$, we have the risk-neutral case (i.e., conventional H^2 estimation). When $\theta > 0$, we seek to maximize $E \exp(-\frac{\theta}{2} \boldsymbol{C}_i)$ which is convex and decreasing in $\boldsymbol{C}_i$. Such a criterion is termed risk-seeking (or optimistic) since larger weights are on small values of $\boldsymbol{C}_i$, and hence we are more concerned with the frequent occurrence of moderate values of $\boldsymbol{C}_i$ than with the occasional occurrence of large values. When $\theta < 0$, we seek to minimize $E \exp(-\frac{\theta}{2} \boldsymbol{C}_i)$ which is convex and increasing in $\boldsymbol{C}_i$. Such a criterion is termed risk-averse (or pessimistic) since large weights are on large values of $\boldsymbol{C}_i$, and hence we are more concerned with the occasional occurrence of large values than with the frequent occurrence of moderate ones. In what follows, we shall see that in the risk-averse case $\theta < 0$, the limit at which minimizing (63) makes sense is the optimal H^∞ criterion.

B. Minimizing the Risk-Sensitive Criterion

Using the conditional distribution density function (61), we can easily verify that

$$E\left(\exp\left(-\frac{\theta}{2}\boldsymbol{C}_i\right)\right) \propto \int \exp\left(-\frac{\theta}{2}C_i\right) \times \exp\left[-\frac{1}{2}J_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\right] \, dx_0 \, du_0 \cdots du_i$$

which shows that the risk-sensitive criterion (63) can be alternatively written as

i) $\theta > 0$: $\max_{\{\check{z}_{j|l}\}} \int \exp\left[- \frac{\theta}{2}C_i - \frac{1}{2}J_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\right] \, dx_0 \, du_0 \cdots du_i$.

ii) $\theta < 0$: $\min_{\{\check{z}_{j|l}\}} \int \exp\left[- \frac{\theta}{2}C_i - \frac{1}{2}J_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\right] \, dx_0 \, du_0 \cdots du_i$.

This suggests that we define the second-order scalar form

$$\begin{aligned}
&\bar{J}_i(x_0, u_0, \cdots, u_{i,0}, \cdots, y_i) \\
&= J_i(x_0, u_0, \cdots, u_{i,0}, \cdots, y_i) + \theta C_i \\
&= x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i} u_j^* Q_j^{-1} u_j + \sum_{j=0}^{i} (y_j - H_j x_j)^* \\
&\quad \times R_j^{-1}(y_j - H_j x_j) + \theta \sum_{j=0}^{i} (\check{z}_{j|j-1} - L_j x_j)^* \\
&\quad \times (\check{z}_{j|j-1} - L_j x_j).
\end{aligned}$$

Before proceeding with the extremizations in i) and ii), we need to ensure that the integrals in i) and ii) are finite. The condition is given by the following lemma, which is easy to prove.

Lemma 8 (Finiteness Condition): The integral

$$\int \exp\left[-\frac{1}{2}\bar{J}_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\right] \, dx_0 \, du_0 \cdots du_i$$

is finite iff $\bar{J}_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)$ has a minimum over $\{x_0, u_0, \cdots, u_i\}$. In that case it is proportional to

$$\exp\left\{-\frac{1}{2}\min_{x_0, u} \bar{J}_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\right\}.$$

The above lemma thus reduces the risk-sensitive problem to one of finding the minimum of a second-order scalar form. More precisely, the criterion becomes

i) $\theta > 0$: $\min_{\{\check{z}_{j|l}\}}\{\min_{x_0,u} \bar{J}_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\}$.

ii) $\theta < 0$: $\max_{\{\check{z}_{j|l}\}}\{\min_{x_0,u} \bar{J}_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)\}$.

Note that the second of the above problems is a quadratic game problem [12]. Though we shall not consider quadratic games here, it is also possible to solve them using the approach given in this and the companion paper.

To solve the above problem, we can introduce the following auxiliary Krein state-space model that corresponds to the (possibly indefinite) quadratic form $\bar{J}_i(x_0, u_0, \cdots, u_{i,0}, \cdots, y_i)$

$$\begin{cases} \boldsymbol{x}_{j+1} = F_j \boldsymbol{x}_j + G_j \boldsymbol{u}_j \quad j \geq 0 \\ \begin{bmatrix} \boldsymbol{y}_i \\ \check{\boldsymbol{z}}_{i|l} \end{bmatrix} = \begin{bmatrix} H_i \\ L_i \end{bmatrix} \boldsymbol{x}_i + \boldsymbol{v}_i \end{cases} \tag{64}$$

with

$$\left\langle \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{u}_j \\ \boldsymbol{v}_j \end{bmatrix}, \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{u}_k \\ \boldsymbol{v}_k \end{bmatrix} \right\rangle = \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & Q_i \delta_{ij} & 0 \\ 0 & 0 & \begin{bmatrix} R_j & 0 \\ 0 & \theta^{-1} I \end{bmatrix} \delta_{jk} \end{bmatrix}. \tag{65}$$

We can now readily use the state-space model (64) and the results of the companion paper [1] to check for the condition of a minimum over $\{x_0, u_0, \cdots, u_i\}$ and to compute the value at the minimum. Then the resulting quadratic form can be further extremized via a Krein-space projection. The details will not be repeated here, since they are the same as those given in the derivation of the H^∞ filter. We shall just note that the quadratic form $\bar{J}_i(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)$ is exactly the same as the quadratic forms $J_{f,i}$ $(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)$, $J_{p,i}$ $(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)$, and $J_{s,i}$ $(x_0, u_0, \cdots, u_i; y_0, \cdots, y_i)$, when we choose $\theta = -\gamma_f^{-2}$, $\theta = -\gamma_p^{-2}$, and $\theta = -\gamma_s^{-2}$, and when the estimate is chosen as a filtered, predicted, and smoothed estimate, respectively. Therefore the derivations of the risk-sensitive filters follow exactly the same derivation of the H^∞ filters discussed earlier. We thus have the following results.

Theorem 7 (A Posteriori Risk-Sensitive Filter): For a given $\theta > 0$, the risk-sensitive estimation problem always has a solution. For a given $\theta < 0$, a solution exists iff

$$\begin{bmatrix} R_j & 0 \\ 0 & \theta^{-1} I \end{bmatrix} \text{ and } R_{e,j} = \begin{bmatrix} R_j & 0 \\ 0 & \theta^{-1} I \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j [H_j^* \quad L_j^*]$$

have the same inertia for all $j = 0, 1, \cdots i$, where $P_0 = \Pi_0$ and

$$\begin{aligned} P_{j+1} &= F_j P_j F_j^* + G_j Q_j G_j^* \\ &\quad - F_j P_j [H_j^* \quad L_j^*] R_{e,j}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_i F_i^*. \end{aligned}$$

In both cases the optimal risk-sensitive filter with parameter θ is given by

$$\begin{aligned} \check{z}_{i|i} &= L_i \hat{x}_{i|i} \\ \hat{x}_{i+1|i+1} &= F_i \hat{x}_{i|i} + K_{s,i+1}(y_{i+1} - H_{i+1} F_i \hat{x}_{i|i}), \; \hat{x}_{-1|-1} = 0 \end{aligned}$$

and

$$K_{s,i+1} = P_{i+1} H_{i+1}^* (I + H_{i+1} P_{i+1} H_{i+1}^*)^{-1}.$$

Proof: The proof is exactly that of Theorem 1. We only note here that for $\theta > 0$ a solution always exists since $\begin{bmatrix} \theta^{-1} I & 0 \\ 0 & R_i \end{bmatrix} > 0$ and the state-space model reduces to the usual Hilbert-space setting.

Theorem 8 (A Priori Risk-Sensitive Filter): For a given $\theta > 0$, the *a priori* risk-sensitive estimation problem always has a solution. For a given $\theta < 0$, a solution exists iff all leading submatrices of

$$\begin{bmatrix}\theta^{-1}I & 0\\ 0 & R_j\end{bmatrix} \text{ and } R_{e,j} = \begin{bmatrix}\theta^{-1}I & 0\\ 0 & R_j\end{bmatrix} + \begin{bmatrix}L_j\\ H_j\end{bmatrix} P_j [L_j^* \quad H_j^*]$$

have the same inertia for all $j = 0, 1, \cdots i$, where P_j is the same as in the *a posteriori* case. In both cases the *a priori* risk-sensitive filter with parameter θ is given by

$$\check{z}_{i|i-1} = L_i \hat{x}_{i|i-1}$$
$$\hat{x}_{i+1|i} = F_i \hat{x}_{i|i-1} + K_{a,i}(y_i - H_i \hat{x}_{i|i-1}), \; \hat{x}_{0|-1} = 0$$

where

$$K_{a,i} = F_i \tilde{P}_i H_i^* (I + H_i \tilde{P}_i H_i^*)^{-1}.$$

Theorem 9 (Risk-Sensitive Smoother): For a given $\theta > 0$, the risk-sensitive smoother always has a solution. For a given $\theta < 0$, a solution exists iff the block diagonal matrix

$$R_e = R_{e,0} \oplus R_{e,1} \oplus \cdots \oplus R_{e,i}$$

where

$$R_{e,j} = \begin{bmatrix}R_j & 0\\ 0 & \theta^{-1}I\end{bmatrix} + \begin{bmatrix}H_j\\ L_j\end{bmatrix} P_j [H_j^* \quad L_j^*]$$

and P_j is the same as in the *a posteriori* case, has $(i+1)p$ positive eigenvalues and $(i+1)q$ negative eigenvalues. In other words, iff

$$\text{In}[R_e] = [(i+1)p \quad 0 \quad (i+1)q].$$

In both cases the risk-sensitive smoother is the H^2 smoother.

We can now state the striking resemblances between the H_∞ and the risk-sensitive filters. The H^∞ filters obtained earlier are essentially risk-sensitive filters with parameter $\theta = -\gamma^{-2}$. Note, however, that at each level γ, the H^∞ filters are not unique, whereas for each θ, the risk-sensitive filters are unique. Also, the risk-sensitive filters generalize to the $\theta > 0$ case. It is also noteworthy that the optimal H_∞ filter corresponds to the risk-sensitive filter with $\bar{\theta} = -\gamma_o^{-2}$, and that $\bar{\theta}$ is that value for which the minimizing property of $\bar{J}_i$ breaks down and $\mu_i(\theta)$ becomes infinite. This relationship between the optimal H_∞ filter and the corresponding risk-sensitive filter was first noted in [21].

V. Finite Memory Adaptive Filtering

We now consider an application of the Krein space–Kalman filter to the problem of finite memory (or sliding window) adaptive filtering. It has been recently shown [20] that a unified derivation of adaptive filtering algorithms and their corresponding fast versions can be obtained by properly recasting the adaptive problem into a standard state-space estimation problem. We now verify that if we further allow the elements of the state-space model to belong to a Krein space, then the so-called sliding window problem can also be handled within the same framework. In fact, this framework also allows us to easily consider more general sliding patterns with windows of varying lengths, as explained ahead. Moreover, we shall obtain a physical interpretation of innovations with negative Gramian, and see that it corresponds to the loss of information.

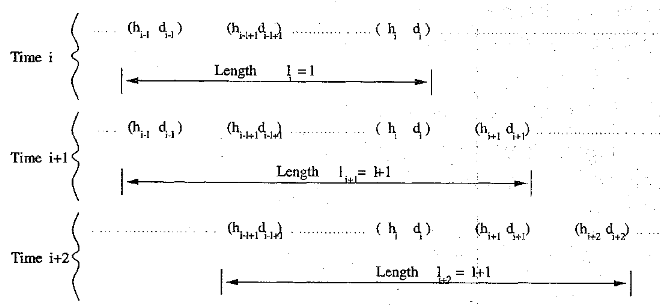

Fig. 2. Sliding window with varying window length.

A. The Standard Problem

The finite memory adaptive filtering problem can be formulated as follows: given the input–output pairs $\{h_j, d_j\}$ where $h_j \in \mathcal{C}^{1\times n}$ is a known input vector and $d_j \in \mathcal{C}$ is a known output scalar, recursively determine estimates of an unknown weight vector $w \in \mathcal{C}^n$, such that the scalar second-order form

$$J_i(w, d_{i-l_i+1}, \cdots, d_i, h_{i-l_i+1}, \cdots, h_i) = w^* \Pi_0^{-1} w + \sum_{j=i-l_i+i}^{i} (d_j - h_j w)^* (d_j - h_j w) \quad (66)$$

where $\Pi_0 > 0$, is minimized for each i.

Since J_i is a function of the pairs $\{h_j, d_j\}_{j=i-l_i+i}^{i}$, at each time instant i, we are interested in determining the estimate of w using only the data given over an interval of length l_i. The quantity $l_i \geq 0$ is therefore referred to as the (memory) length of the sliding window.

Note that we allow for a time-variant window length. To clarify this point, consider the example of Fig. 2 where at time i we have a window length of $l_i = l$. At the next time instant we add the data point $\{h_{i+1}, d_{i+1}\}$, so that the window length changes to $l_{i+1} = l + 1$. At time $i+2$ we add the data point $\{h_{i+2}, d_{i+2}\}$ and drop the data point $\{h_{i-l}, d_{i-l}\}$ so that the window length remains $l_{i+2} = l+1$. In a similar fashion, more general sliding window patterns can be considered as well.

To recast expression (66) into the usual quadratic form considered in this paper, the lower index of the summation term needs to start at the fixed time 0. For this purpose, we rewrite J_i as follows

$$\begin{aligned} &J_i(w, d_{i-l_i+1}, \cdots, d_i, h_{i-l_i+1}, \cdots, h_i)\\ &= w^* \Pi_0^{-1} w + \sum_{j=i-l_i+1}^{i} (d_j - h_j w)^* (d_j - h_j w)\\ &\quad + \sum_{j=0}^{i-l_i} (d_j - h_j w)^* (d_j - h_j w)\\ &\quad - \sum_{j=0}^{i-l_i} (d_j - h_j w)^* (d_j - h_j w) \end{aligned} \quad (67)$$

where we have added and subtracted identical terms. We now invoke a change of variables and substitute the time index i by another time index k that allows us to replace J_i by a $\bar{J}_k$. The new index k has the property whenever a new data point is added (i.e., i is incremented), then k is incremented. Whenever a data point is discarded from the window, however, k is incremented as well. Thus if at time i the length of the window is l_i, then the index k will run from 0 to $2i - l_i + 1$ (since there will have been i data points added and $i-l_i+1$ data points removed). To be more specific, the change of variables is as follows.

a) At each time i, since the data point $\{h_i, d_i\}$ is added, we increment the index k and define

$$\bar{d}_k = d_i, \quad \bar{h}_k = h_i \quad \text{and} \quad \bar{R}_k = 1. \tag{68}$$

b) If at time i the data point $\{h_{i-l_i}, d_{i-l_i}\}$ is removed, we increment the index k once more and define

$$\bar{d}_k = d_{i-l_i}, \quad \bar{h}_k = h_{i-l_i} \quad \text{and} \quad \bar{R}_k = -1. \tag{69}$$

With this convention we may write the quadratic form $J_i(w, d_{i-l_i+1}, \cdots, d_i, h_{i-l_i+1}, \cdots, h_i)$ as

$$\begin{aligned} J_i &= \bar{J}_k(w, d_0, \cdots, d_k, h_0, \cdots, h_k) \\ &= w^*\Pi_0^{-1}w + \sum_{j=0}^{k}(\bar{d}_j - \bar{h}_j w)^*(\bar{d}_j - \bar{h}_j w) \end{aligned} \tag{70}$$

which is of the form that we have been considering in this and the companion paper [1]. Note that the quadratic form $\bar{J}_k(w, d_0, \cdots, d_k, h_0, \cdots, h_k)$ is indefinite, since whenever a data point is dropped we have $\bar{R}_k = -1$. We can therefore use Krein-space methods to solve the problem.

Using the same approach that we have used so far, we now construct the partially equivalent state-space model to the indefinite quadratic form $\bar{J}_k$. Thus

$$\begin{cases} \boldsymbol{x}_{j+1} = \boldsymbol{x}_j, \quad \boldsymbol{x}_0 = \boldsymbol{w} \quad 0 \le j \le k \\ \bar{d}_j = \bar{h}_j \boldsymbol{x}_k + \boldsymbol{v}_k \end{cases} \tag{71}$$

with $\Pi > 0$, $Q_j = 0$, $S_j = 0$, and $\bar{R}_j$ as in (68) and (69).

We can now state the following result.

Theorem 10 (Finite Memory Adaptive Filter): The finite memory adaptive filter is given by the following recursions.

a) For updating the data point $\{h_i, d_i\}$ at time i, we have

$$\begin{aligned} \hat{w}_{|i:i-l_{i-1}} &= \hat{w}_{|i-1:i-l_{i-1}} \\ &\quad + K_{p,k}(d_i - h_i \hat{w}_{|i-1:i-l_{i-1}}) \end{aligned} \tag{72}$$

where $\hat{w}_{|i:j}$ is the estimate when the sliding window encompasses all the data from time j to time i, and

$$\begin{aligned} K_{p,k} &= P_k h_i^* R_{e,k}^{-1} \\ R_{e,k} &= 1 + h_i P_k h_i^* \end{aligned} \tag{73}$$

and where P_k satisfies the recursion

$$P_{k+1} = P_k - K_{p,k} R_{e,k} K_{p,k}^*, \quad P_0 = \Pi_0. \tag{74}$$

b) For downdating the data point $\{h_{i-l_i}, d_{i-l_i}\}$ at time i, we have

$$\hat{w}_{|i:i-l_i+1} = \hat{w}_{|i:i-l_i} + K_{p,k}(d_{i-l_i} - h_{i-l_i}\hat{w}_{|i:i-l_i}) \tag{75}$$

where

$$\begin{aligned} K_{p,k} &= P_k h_{i-l_i}^* R_{e,k}^{-1} \\ R_{e,k} &= -1 + h_{i-l_i} P_k h_{i-l_i}^* \end{aligned} \tag{76}$$

and P_k satisfies the recursion

$$P_{k+1} = P_k - K_{p,k} R_{e,k} K_{p,k}^*. \tag{77}$$

Moreover, the above solutions for $\hat{w}_{|i:j}$ always correspond to a minimum, and in particular

$$R_{e,k} = 1 + h_i P_k h_i^* > 0 \tag{78}$$

when we are updating, and

$$R_{e,k} = -1 + h_{i-l_i} P_k h_{i-l_i}^* < 0 \tag{79}$$

when we are downdating.

Proof: The solutions given by a) and b) in the above theorem are simply the Krein space–Kalman filter recursions for the state-space model (71) which we know computes the stationary point of $\bar{J}_k$ over w. This stationary point is always a minimum, however, since

$$\frac{\partial^2 \bar{J}_k}{\partial w^2} = \frac{\partial^2 J_i}{\partial w^2} = \Pi_0^{-1} + \sum_{j=i-l_i+1}^{i} h_j^* h_j > 0$$

(recall that $\Pi_0^{-1} > 0$). □

Using Lemma 12 in the companion paper [1], having a minimum means that $R_{e,k}$ and R_k have the same inertia for all k. Thus the statements (78) and (79) readily follow.

The fact that whenever we drop data we have $R_{e,k} < 0$ has an interesting interpretation. Consider the equation

$$P_{k+1} = P_k - K_{p,k} R_{e,k} K_{p,k}.$$

If we drop data at step k we would expect P_{k+1} to get larger (more positive-definite) than P_k. This can only happen if $R_{e,k} < 0$. Thus, we may infer that innovations with negative Gramian correspond to a loss of information.

The above discussion puts the problem of finite memory adaptive filtering into the same state-space estimation framework as conventional adaptive filtering techniques (see [20]). Therefore the various algorithmic extensions discussed there may be applied to finite memory problems, albeit that we now need to consider a Krein space. We shall not give the details here, but shall just mention that when the elements of the input vectors $\{h_i\}$ form a time sequence, viz.

$$h_i = [u_i \quad u_{i-1} \quad \cdots \quad u_{i-n+1}]$$

and when the window length is constant, i.e., $l_i \equiv l$, then the state-space model (71) is periodic with period $T = 2$, and we may speed up the estimation algorithm by a so-called Chandrasekhar-type recursion. Similar results, obtained via a different approach, have been reported in [26].

VI. CONCLUSION

Certain studies in least-squares estimation, adaptive filtering, and H^∞ filtering motivated us to develop a theory for linear estimation in certain indefinite metric spaces, called Krein spaces. The main difference from the conventional Hilbert-space framework for Kalman filtering and LQG control are that projections in Krein spaces may not necessarily exist or be unique, and that quadratic forms may have stationary points that are not necessarily extreme points (i.e., minima or maxima). We showed that these simple but fundamental differences explain both the unexpected similarities and differences between the well-known Kalman-filter solution for stochastic state-space systems and the solution for the completely nonstochastic H^∞ filtering problem.

The main points are the following. There are many problems whose solution can be reduced to the recursive minimization of some indefinite quadratic form. A stationary point, when it exists, of the quadratic form can be computed as follows: set up a (partially equivalent) problem of projecting a vector in a Krein space onto a certain subspace. The advantage is that when there is state-space structure, this projection can be recursively computed by using the innovations approach to derive a Krein space–Kalman filter. The equivalence is only partial because the Krein-space projection only defines the stationary point of the quadratic form and further conditions need to be checked to determine if this point is also a minimum. It turns out that this checking can also be done recursively using quantities arising in the Kalman-filtering algorithms.

Apart from quite straightforward derivations of known results in H^2, H^∞, and risk-sensitive estimation and control, the above approach allows us to extend to the H^∞ setting some of the huge body of results and insights developed over the last three decades in the field of Kalman filtering (and LQG control). A first bonus is the derivation (see [24]) of square-root and (fast) Chandrasekhar algorithms for H^∞ estimation and control, a possibility that is much less obvious in current approaches. These square-root algorithms, which are now increasingly standard in H^2 Kalman filtering, have two advantages over the earlier H^∞ algorithms: they eliminate the need for explicitly checking the existence conditions of the filters and have various potential numerical and implementational advantages.

Application of the Krein-space formulation to adaptive filtering arises from the approach in [20] where it was shown how to recast adaptive filtering problems as state-space estimation problems. If we further allow the elements of the state-space model to belong to a Krein space, then we can solve finite memory and H^∞ adaptive filtering problems. In the finite memory case, this allows us to consider general sliding patterns with windows of varying lengths. In the H^∞ adaptive case, this has allowed us to establish that the famed LMS (or stochastic gradient) algorithm is an optimal H^∞ filter [25].

We also remark that, although not pursued here, it is also possible to construct dual (rather than partially equivalent) Krein state-space models (via the concept of a dual basis) which can be used to extend the methods of this paper to the solution of H^2 and H^∞ control problems.

We may finally remark that a major motivation for the Krein-space formulation is that it provides geometric insights into various estimation and control problems. Such geometric insights are also useful elsewhere. For example, they can be used to provide a stochastic interpretation and a geometric proof of the KYP (Kalman–Yacobovich–Popov) Lemma [27]–[29].

ACKNOWLEDGMENT

The authors would like to thank P. P. Khargonekar and D. J. N. Limebeer for helpful discussions during the preparation of this manuscript. Seminars by P. Park on the KYP Lemma were also helpful in leading us to begin the research.

REFERENCES

[1] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in Krein spaces—Part I: Theory," *IEEE Trans. Automat. Contr.*, this issue, pp. 18–33.
[2] H. Kwakernaak, "A polynomial approach to minimax frequency domain optimization of multivariable feedback systems," *Int. J. Contr.*, vol. 44, pp. 117–156, 1986.
[3] J. C. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Trans. Automat. Contr.*, vol. 34, no. 8, pp. 831–847, Aug. 1989.
[4] P. P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an H^∞-setting," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 151–166, 1991.
[5] T. Basar, "Optimum performance levels for minimax filters, predictors and smoothers," *Syst. Contr. Lett.*, vol. 16, pp. 309–317, 1991.
[6] D. J. Limebeer and U. Shaked, "New results in H^∞-filtering," in *Proc. Int. Symp. MTNS*, June 1991, pp. 317–322.
[7] U. Shaked and Y. Theodor, "H_∞-optimal estimation: A tutorial," in *Proc. IEEE Conf. Decision Contr.*, Tucson, AZ, Dec. 1992, pp. 2278–2286.
[8] I. Yaesh and U. Shaked, "H_∞-optimal estimation—The discrete time case," in *Proc. Int. Symp. MTNS*, June 1991, pp. 261–267.
[9] M. J. Grimble, "Polynomial matrix solution of the H^∞ filtering problem and the relationship to Riccati equation state-space results," *IEEE Trans. Signal Processing*, vol. 41, no. 1, pp. 67–81, Jan. 1993.
[10] K. Glover and D. Mustafa, "Derivation of the maximum entropy H^∞ controller and a state space formula for its entropy," *Int. J. Contr.*, vol. 50, pp. 899–916, 1989.
[11] M. Green and D. J. N. Limebeer, *Linear Robust Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[12] T. Basar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Boston, MA: Birkhauser, 1991.
[13] D. H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic games," *IEEE Trans. Automat. Contr.*, vol. AC-18, no. 2, Apr. 1973.
[14] J. Speyer, J. Deyst, and D. H. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria," *IEEE Trans. Automat. Contr.*, vol. 19, pp. 358–366, Aug. 1974.
[15] P. Whittle, *Risk Sensitive Optimal Control*. New York: Wiley, 1990.
[16] J. L. Speyer, C. Fan, and R. N. Banavar, "Optimal stochastic estimation with exponential cost criteria," in *Proc. IEEE Conf. Decision Contr.*, Tucson, AZ, Dec. 1992, pp. 2293–2298.
[17] N. Young, *An Introduction to Hilbert Space*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
[18] B. A. Francis, *A Course in H^∞ Control Theory*. New York: Springer-Verlag, 1987.
[19] P. Henrici, *Applied and Computational Complex Analysis*, vol. 1. New York: Wiley, 1974.
[20] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Mag.*, pp. 18–60, July 1994.
[21] K. Glover and J. C. Doyle, State-space formulae for all stabilizing controllers that satisfy an H^∞-norm bound and relations to risk sensitivity," *Syst. Contr. Lett.*, vol. 11, pp. 167–172, 1988.
[22] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, vol. 64, *Mathematics in Science and Engineering*. New York: Academic, 1970.

[23] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
[24] B. Hassibi, A. H. Sayed, and T. Kailath, "Square root arrays and Chandrasekhar recursions for H^∞ problems," submitted to *IEEE Trans. Automat. Contr*; also in *Proc. 33rd IEEE Conf. Dec. Contr.*, 1994, pp. 2237–2243.
[25] ——, "H^∞ optimality of the LMS algorithm," *IEEE Trans. Signal Processing*, to appear. Also in *Proc. 32rd IEEE Conf. Dec. Contr.*, 1993, pp. 74–80.
[26] A. Houacine, "Regularized fast recursive least squares algorithms for finite memory filtering," *IEEE Trans. Signal Processing*, vol. 40, no. 4, pp. 758–769, Apr. 1992.
[27] R. E. Kalman, "Lyapunov functions for the problem of Lur'e in automatic control," *Proc. Nat. Acad. Sci.*, vol. 49, no. 2, pp. 201–205, 1963.
[28] V. A. Yakubovich, "The solution of certain matrix inequalities in automatic control theory," *Doklady Akademii Nauk, SSSR*, vol. 143, pp. 1304–1307, 1962.
[29] V. M. Popov, "Hyperstability and optimality of automatic systems with several control functions," *Rev. Roumaine Sci. Tech.—Elektrotechn. et Energe.*, vol. 9, no. 4, pp. 629–690, 1964.

Babak Hassibi, for a photograph and biography, see p. 33 of this issue of this TRANSACTIONS.

Ali H. Sayed (S'90–M'92), for a photograph and biography, see p. 33 of this issue of this TRANSACTIONS.

Thomas Kailath (S'57–M'62–F'70), for a photograph and biography, see p. 33 of this issue of this TRANSACTIONS.