

Minimal Realization and Dynamic Properties of Optimal Smoothers

Augusto Ferrante and Giorgio Picci, *Fellow, IEEE*

Abstract—Smoothing algorithms of various kinds have been around for several decades. However, some basic issues regarding the dynamical structure and the minimal dimension of the steady-state algorithm are still poorly understood. It seems fair to say that the subject has not yet reached a definitive form. In this paper, we derive a realization of minimal dimension of the optimal smoother for a signal admitting a state-space description of dimension n . It is shown that the dimension of the smoothing algorithm can vary from n to $2n$, depending on the zero structure of the signal model. The dynamics (pole structure) of the steady-state smoother is also characterized explicitly and is related to the zero structure of the model.

We use several recent ideas from stochastic realization theory. In particular, a minimal Markovian representation of the smoother is derived, which requires solving a nonsymmetric Wiener–Hopf factorization problem. In this way, the smoother is naturally expressed as the cascade of a whitening filter and a linear filter of least possible dimension, whose state space is a minimal Markovian subspace containing the smoothed estimate $\hat{x}$. This, among other aspects, affords a very simple calculation of the error covariance matrix of the smoother. A reduced-order two-filter implementation of the Mayne–Fraser type is obtained by solving a Riccati equation of reduced dimension, which is in general smaller than the dimension of the Riccati equations considered in the literature.

Index Terms—Filtering, Riccati equation, smoothing, stochastic realization.

I. INTRODUCTION AND PROBLEM FORMULATION

SMOOTHING of linear stochastic systems is a classical subject with a long history. Due to the availability of ultrafast computers, its relative importance with respect to the classical Kalman-type algorithms has been growing in the recent years, since real-time processing can now often be done more efficiently by off-line algorithms, which process the data in batches of finite length.

Among the early references on smoothing we may quote the paper by Rauch *et al.* [26], the celebrated two-filter formula of Mayne [20] and Fraser [10], and the papers by Kailath and Frost [13], Sidhu and Desai [28], and Weinert and Desai [29]. We quote also the important paper by Badawi *et al.* [5] where stochastic realization theory was shown to be the natural frame-

work for the problem. The computational aspects are surveyed in the paper by Park and Kailath [23].

Other recent work on smoothing has been motivated by the two-point boundary value formulation of Adams *et al.* [1] and Levy *et al.* [14]; however, we shall not need to consider this type of framework here. Treatments of smoothing from various points of view are also found in textbooks as [4], [11], and [15].

In our opinion, notwithstanding the vast literature existing on this subject, the theory of smoothing has not yet crystallized into a standard universally accepted format as, for example, causal Kalman filtering. The basic structure of the filter, its implementation and the analysis of its steady-state behavior, do not appear to have reached a definitive form. For example, a basic issue like describing the poles of the steady-state smoother does not seem to have been answered. Also, in virtually all traditional treatments of smoothing, it is given for granted that the smoother should be a dynamical system whose dimension is equal to twice the dimension n of the signal model. Only recently has it been discovered that, instead, the dimension of the optimal smoother can vary from n to $2n$, depending on the zero-structure of the signal model transfer function. This fact was first pointed out by geometric arguments in [18]. This reference, however, does not deal specifically with smoothing, and the characterization of the dimension of the smoother is not explicit and is buried in a wealth of other results related to stochastic modeling.

In this paper, we shall attempt to provide a clear and hopefully definitive picture, at least for the steady-state behavior of the smoother. We shall first derive by elementary computations a minimal realization of the smoother and show that its dimension is between n and $2n$ and can be related to the zero structure of the transfer function of the given model. This fact was shown by more abstract arguments in [18]. In order to understand its dynamic structure, we address the problem of computing a minimal stochastic realization (i.e., a Markovian representation) of the smoother. This problem is reduced to a (nonsymmetric) Wiener–Hopf factorization of a rational matrix function, the cross spectral density matrix of the state and output processes. The factorization need not be minimal in the classical sense, but some other minimality constraints must be satisfied. Exploiting well-known spectral factorization theory, we relate this problem to the solution of an algebraic Riccati equation (ARE). The solution that leads to the requested minimal factorization is characterized as the unique positive definite solution of a generally smaller dimensional Riccati equation. In this way, a numerically stable, two-filter-type implementation of the optimal filter requires the solution of an ARE of dimension that can vary from zero to n , depending on the zero structure of the model.

Manuscript received May 29, 1998; revised December 19, 1999. Recommended by Associate Editor, J. Spall. This work was supported in part by the Ministry of Higher Education of Italy (MURST) under Project *Identification and Control of Industrial Systems*.

A. Ferrante is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Milano 20133, Italy (e-mail: ferrante@elet.polimi.it).

G. Picci is with the Dipartimento di Elettronica e Informatica, Università di Padova, Padova 35131, Italy (e-mail: picci@dei.unipd.it).

Publisher Item Identifier S 0018-9286(00)09443-5.

Our basic assumption will be that the signal model is time invariant. The steady-state smoothing problem can be formulated as follows: we are given a linear stochastic model

$$\dot{x} = Ax + Bw \quad y = Cx + Dw \quad (1.1)$$

driven by a p -dimensional normalized white noise w . Observe that this model is more general than the classical two-noises model sometimes considered in the literature

$$\dot{x} = Ax + B_1 w_1 \quad y = Cx + D_1 w_2 \quad (1.2)$$

where w_1 and w_2 are uncorrelated white noises, since the latter may be viewed as a particular case of the former by setting $B = [B_1 | 0]$ and $D = [0 | D_1]$ in (1.1).

The $n \times n$ matrix A is assumed stable, i.e., all the eigenvalues of A lie in the open left-half complex plane ($\sigma(A) \subset \mathbb{C}_-$). This assumption is made for convenience only. What is really needed here is that no eigenvalue of A lies on the imaginary axis. Under this latter assumption, the seemingly more general framework of arbitrary eigenvalues can be reduced to the one we are considering here; see [25]. This implies that the system is in statistical steady state and the m -dimensional observed process y , output of (1.1), is a stationary process.

Given observations of $y(t)$ on the interval $-\infty < t < +\infty$, we denote by $H(y_{-\infty}^{+\infty})$ the Hilbert space spanned by such observations; see e.g., [27] and [18] for a precise mathematical definition. We want to compute the minimum-variance linear steady-state estimate (wide-sense conditional expectation)

$$\hat{x}(t) = E[x(t) | H(y_{-\infty}^{+\infty})] \quad (1.3)$$

of the (n components of the) state $x(t)$. It is well known [27] that this estimate is the limit in mean square of the finite-interval estimate

$$\hat{x}_{[t_0, t_1]}(t) = E[x(t) | y(s); t_0 \leq s \leq t_1], \quad t_0 \leq t \leq t_1$$

as $t - t_0$ and $t_1 - t$ tend to ∞ and is often used as a constant-parameter approximation of the latter.¹

We assume that the model is *minimal* both in the sense that (A, B) is controllable and (C, A) is observable and in the sense that the transfer function

$$W(s) = C(sI - A)^{-1}B + D \quad (1.4)$$

is a spectral factor of the spectral density of y

$$\Phi_y(s) = W(s)W^T(-s) \quad (1.5)$$

¹Some may argue that this stationary approximation may not be of much value if the observation interval is small, since the optimal finite-interval smoother (which is time-varying) may not get close enough to the steady-state filter. However, in case of a small observation interval (i.e., one consisting of very few data points), there is really no need of *recursive* filters, since the computation of the estimate can be done by one-shot algorithms of static estimation theory. Efficient algorithms of this kind have been available in the literature for a while, an early reference being, e.g., [22]. So the critique really refers to a situation that is of little interest to dynamic smoothing since it is naturally dealt with by different algorithms.

of *minimal degree*.² Moreover, we shall adopt the standard assumption that $\Phi_y(s)$ is *coercive*, i.e., there exist $c > 0$ such that

$$\Phi_y(i\omega) \geq cI, \quad \forall \omega \in \mathbb{R}. \quad (1.6)$$

This assumption implies that the matrix D has full (row) rank so that $R := DD^T = \Phi(\infty)$ is nonsingular. Without loss of generality, we can choose a basis in the input space of (1.1) such that $D = [R^{1/2} | 0]$, where $R^{1/2}$ is, say, the (unique) symmetric square root of R . We partition B conformly as $B = [B_1 | B_2]$.

It is required that the solution $\hat{x}(t)$ be computable recursively as the output of a (generally *noncausal*) dynamical system of transfer function $S(s)$ (the *smoother*). Acausal linear filters and the interpretation of acausal transfer functions are discussed briefly in Appendix A.

$$y(t) \longrightarrow \boxed{S(s)} \longrightarrow \hat{x}(t)$$

We shall require that the smoother is implemented by a numerically stable algorithm of least complexity. We shall come back and discuss the meaning of these specifications in more detail later.

It is well known that the orthogonality principle of linear estimation theory provides the condition

$$\Phi_{xy}(s) = \Phi_{\hat{x}y}(s) \quad (1.7)$$

where $\Phi_{xy}(s)$ is the cross spectral density of the processes x and y and $\Phi_{\hat{x}y}(s)$ is the cross spectral density of the processes $\hat{x}$ and y . From this condition, the well-known relation for the transfer function $S(s)$ of the smoother readily follows:

$$S(s) = \Phi_{xy}(s)\Phi_y^{-1}(s). \quad (1.8)$$

Observe that $\Phi_{xy}(s)$ and $\Phi_y(s)$ may be expressed in terms of the data as

$$\begin{aligned} \Phi_{xy}(s) &= (sI - A)^{-1}BW^T(-s) \\ \Phi_y(s) &= W(s)W^T(-s) \end{aligned} \quad (1.9)$$

so that (1.8) can be written a little more explicitly as

$$S(s) = (sI - A)^{-1}BW^*(s)\Phi_y(s)^{-1}$$

where we have adopted the notation $W^*(s) := W(-s)^T$. These expressions involve several pole-zero cancellations and the dynamics (i.e., the location of the poles) of $S(s)$ is not easy to figure out.

In this paper, similarly to what is done in the classical steady-state analysis of the Kalman filter, we would like to describe the dynamics of the steady-state smoother in terms, say, of the original spectral data of the problem. One basic question is that of describing the poles of $S(s)$. This will be answered in Section V, essentially in the following terms.

Theorem 1.1: The poles (including multiplicity) of the steady-state smoother $S(s)$ are the subset of the zeros of the spectral density $\Phi_y(s)$ obtained by deleting the elements which are also zeros of $W^*(s)$.

²“Degree” is always understood to mean MacMillan degree.

This paper is organized as follows. In Section II, we compute the minimal dimension (McMillan degree) of the optimal smoother. In Section III, in the spirit of the Wiener–Kolmogorov theory of filtering, the smoother is realized as a cascade of a withering filter and a shaping filter. In Section IV, the relation between the zeros of $W(s)$ and the dynamics of the smoother is investigated; moreover, the family of minimal realizations of the smoother is parametrized and the structure of the error covariance is analyzed. Section V discusses the smoother implementation. In Section VI, the discrete-time counterpart of the results of the previous sections (which deal with the continuous-time case) are outlined. In Section VII, some simulation results are described. In Section VIII, we finally draw some conclusions. Appendixes A, B, and C deal with some technical issues.

II. A MINIMAL REALIZATION OF THE OPTIMAL SMOOTHER

In this section, we shall express the smoother as a dynamical system in state space form and compute the relative system matrix and the dimension of a minimal realization of $S(s)$.

State-space realizations of $\Phi_{xy}(s)$ and $\Phi_y(s)$ are easily obtained from the model (1.1) as follows:

$$\begin{aligned}\Phi_{xy}(s) &= (sI - A)^{-1}BW^T(-s) \\ &= [I \ 0] \left(sI - \begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix} \right)^{-1} \begin{bmatrix} B_1 R^{1/2} \\ -C^T \end{bmatrix} \quad (2.1)\end{aligned}$$

$$\begin{aligned}\Phi_y(s) &= W(s)W^T(-s) \\ &= R + [C \ R^{1/2}B_1^T] \\ &\quad \cdot \left(sI - \begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix} \right)^{-1} \begin{bmatrix} B_1 R^{1/2} \\ -C^T \end{bmatrix}. \quad (2.2)\end{aligned}$$

From the latter, employing a well-known formula for the inversion of a rational matrix function, we get

$$\begin{aligned}\Phi_y^{-1}(s) &= R^{-1} - [R^{-1}C \ R^{-1/2}B_1^T](sI - \Lambda)^{-1} \\ &\quad \cdot \begin{bmatrix} B_1 R^{-1/2} \\ -C^T R^{-1} \end{bmatrix}. \quad (2.3)\end{aligned}$$

Here Λ is the Hamiltonian matrix

$$\Lambda := \begin{bmatrix} \Gamma & B_2 B_2^T \\ C^T R^{-1} C & -\Gamma^T \end{bmatrix} \quad (2.4)$$

with Γ being defined as

$$\Gamma := A - B_1 R^{-1/2} C. \quad (2.5)$$

From this a realization of the transfer function of the smoother can be computed.

Theorem 2.1: The transfer function (1.8) of the steady-state smoother has the realization

$$S(s) = [I \ 0] (sI - \Lambda)^{-1} \begin{bmatrix} B_1 R^{-1/2} \\ -C^T R^{-1} \end{bmatrix} \quad (2.6)$$

where Λ is the Hamiltonian matrix (2.4). The McMillan degree n_0 of the smoother is

$$n_0 = 2n - \nu \quad (2.7)$$

where ν is the dimension of the unobservable space of the pair (B_2^T, Γ^T) .

To prove this theorem we shall use the following technical lemma, which is a straightforward consequence of the fact that the unobservable subspace of a pair (H, F) is the largest F -invariant subspace contained in $\ker H$ (the nullspace of H). The proof will be skipped.

Lemma 2.1: Let N be a matrix whose columns form a basis for the unobservable subspace of the pair (H, F) . Then there exists a matrix J such that

$$\begin{bmatrix} F \\ H \end{bmatrix} N = \begin{bmatrix} NJ \\ 0 \end{bmatrix}. \quad (2.8)$$

Conversely, if N is a matrix such that (2.8) holds, then the columns of N belong to the unobservable subspace of (H, F) .

Proof of Theorem 2.1: Equation (2.6) follows by multiplying together (2.1) and (2.3) and employing Lemma B.1 to compute the product of $\Phi_{xy}(s)$ with the strictly proper part of $\Phi_y^{-1}(s)$.

Next, note that the McMillan degree of $\Phi_y(s)$ is $2n$ since $W(s)$ is a minimal spectral factor. Hence the McMillan degree of $\Phi_y(s)^{-1}$ is also $2n$ and, in view of (2.3), the pair

$$\left(\Lambda, \begin{bmatrix} B_1 R^{-1/2} \\ -C^T R^{-1} \end{bmatrix} \right)$$

is controllable. Therefore, from (2.6), it is apparent that n_0 , the McMillan degree of the smoother $S(s)$, is equal to the dimension of the observability space of the pair $([I \ 0], \Lambda)$.

Define then N to be a matrix whose columns form a basis for the unobservable subspace of the pair $([I \ 0], \Lambda)$. Partition N as $N = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}$, where the blocks N_1 and N_2 have n rows. In view of Lemma 2.1, there exists a matrix J such that

$$\begin{bmatrix} \Lambda & \begin{bmatrix} B_1 R^{-1/2} \\ -C^T R^{-1} \end{bmatrix} \\ [I \ 0] \end{bmatrix} N = \begin{bmatrix} \Gamma & B_2 B_2^T \\ C^T R^{-1} C & -\Gamma^T \\ I & 0 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = \begin{bmatrix} NJ \\ 0 \end{bmatrix} \quad (2.9)$$

which immediately yields $N_1 = 0$, so that (2.9) implies

$$\begin{bmatrix} \Gamma^T \\ B_2^T \end{bmatrix} N_2 = \begin{bmatrix} N_2(-J) \\ 0 \end{bmatrix}. \quad (2.10)$$

In view of Lemma 2.1, this implies that the columns of N_2 belong to the unobservable subspace of the pair (B_2^T, Γ^T) .

Conversely, if N_2 is a matrix whose columns form a basis for the unobservable subspace of the pair (B_2^T, Γ^T) , then the columns of $\begin{bmatrix} 0 \\ N_2 \end{bmatrix}$ belong to the unobservable subspace of the pair $([I \ 0], \Lambda)$.

In conclusion, the unobservable subspaces of the pairs (B_2^T, Γ^T) and $([I \ 0], \Lambda)$ are isomorphic (being expressible as the column-span of N_2 and the column-span of $\begin{bmatrix} 0 \\ N_2 \end{bmatrix}$, respectively) and hence have the same dimension ν . We have already shown that the McMillan degree n_0 of $S(s)$ is equal to the dimension of the observability space of the pair $([I \ 0], \Lambda)$. Therefore

$$n_0 = 2n - \nu \quad (2.11)$$

and the theorem is proved. ■

We are led to conclude that the $2n$ -dimensional realization (2.6) is not necessarily minimal. A minimal realization of $S(s)$ may be obtained from (2.6) by deleting the ν -dimensional unob-

servable subspace (this may be done using standard techniques; see, e.g., [12]).

1) *Relation with Zeros*: As pointed out in [17] (see also [18]), the unobservable subspace of the pair (B_2^T, Γ^T) turns out to be the vector space of *zero directions* for the transfer function (1.4) of the model (1.1), commonly denoted by the symbol $\mathcal{V}^*$. More precisely, the unobservable subspace of the pair $(B_2^T, \Gamma^T)^3$ coincides with the so-called *maximal output nulling subspace* $\mathcal{V}^*(A^T, C^T, B^T, D^T)$ for the dual system (A^T, C^T, B^T, D^T) and plays a basic role in the study of the zero dynamics of spectral factors by means of geometric control theory [17], [6], [31]. Its dimension ν is the number of invariant zeros of $W(s)$, counted with multiplicity. As shown in [17], in force of the coercivity condition (1.6), the invariant zeros of $W(s)$ coincide with the eigenvalues of Γ^T restricted to the invariant subspace $\mathcal{V}^*$. It is evident from (2.6), and in particular from the dimension formula (2.7), that the structure of the steady-state smoother is intimately related to the zero structure of $W(s)$. In particular, only when $W(s)$ has no zeros is the dimension of the smoother $2n$, a fact often claimed to be true in general in the literature.

Observe that the poles of the smoother $S(s)$ are a subset of the eigenvalues of the matrix Λ , and then, by coercivity of $\Phi(s)$, $S(s)$ is analytic on an open strip containing the imaginary axis. Hence it admits “two-filter” type decompositions as a sum of a causal and an anticausal filter. The two-filter structure will be examined in the next sections.

III. MARKOVIAN REPRESENTATION OF THE SMOOTHER

In this section, we shall address the problem of expressing the smoother $S(s)$ in the form

$$S(s) = V_0(s)W_0^{-1}(s) \quad (3.1a)$$

where

$$W_0(s) \text{ is a minimal square spectral factor of } \Phi_y(s) \quad (3.1b)$$

$$V_0(s) \text{ has minimal McMillan degree.} \quad (3.1c)$$

Such a factorization is in the spirit of Wiener–Kolmogorov theory of filtering and prediction and is motivated by the following considerations: the filter $W_0^{-1}(s)$ driven by the observation y is clearly a *whitening filter*. For, the inverse of any $W_0(s)$ satisfying (3.1b) transforms y into a white noise process w_0 (of the same dimension m). Hence the dynamical system $V_0(s)$, whose output is the estimate $\hat{x}(t)$, is driven by a white noise. This implies that the state $x_0(t)$ of any realization of $V_0(s)$

$$V_0(s) = C_0(sI - A_0)^{-1}B_0 \quad (3.2)$$

driven by the white noise process w_0 is a *Markov process*.⁴ Note that $\hat{x}(t)$ is not Markov in general and does not satisfy any differential equation driven by white noise (recursive filter) while $x_0(t)$ instead does by construction.

³Or, equivalently, the orthogonal complement of the reachable subspace for (Γ, B_2) .

⁴Note that this statement has no implications on the stability of A_0 . We refer the reader again to [25] for a discussion of acausal models of Markov processes.

Finally, (3.1c) implies, assuming the realization $\{C_0, A_0, B_0\}$ to be minimal, that the components of the smoothed estimate $\hat{x}(t) = C_0x_0(t)$ are expressed at any time t , as a linear combination of the minimal possible number of state variables. In other words, the Markov process $x_0(t)$ is the state of a dynamical system of minimal dimension among those with the property of serving as a dynamic memory for the smoother. Equivalently, this can be expressed by saying that the Markovian space of random variables spanned by the components of $x_0(t)$

$$\mathbf{X}_t := \text{span}(x_0^i(t)), \quad i = 1, 2, \dots, n \quad (3.3)$$

is a minimal Markovian subspace containing the (components of) the estimate $\hat{x}(t)$.

In conclusion, expressing the smoother as the cascade of a whitening filter W_0^{-1} and a filter $V_0(s)$ of minimal McMillan degree yields a minimal recursive filter with output the smoothed estimate $\hat{x}(t)$. This representation was introduced in a previous publication [18], where, however, the explicit calculation of a state-space realization of the smoother was not addressed.

In order to solve the minimal factorization problem, we shall analyze all solution pairs $(V(s), W(s))$ of (3.1a) and (3.1b) and compare the McMillan degrees of $V(s)$. Since, in view of (1.8), a factorization $S(s) = V(s)W^{-1}(s)$ is equivalent to the factorization

$$\Phi_{xy}(s) = V(s)W^*(s) \quad (3.4)$$

of the cross spectral density, the search for a V of minimal degree is made in the set $\{V(s) = \Phi_{xy}(s)W^{-*}(s) | W(s) \in \mathcal{W}\}$, where $\mathcal{W}$ is the set of all minimal square spectral factors. Note that in principle, we should search the whole set of *not necessarily stable* spectral factors $W(s)$ of $\Phi_y(s)$. However, it is proven in Appendix C⁵ that one can, without loss of generality, restrict the search to the class of *analytic* spectral factors. In fact, it is shown in Appendix C that the transfer function $V(s) = \Phi_{xy}(s)W^{-*}(s)$ corresponding to an arbitrary minimal spectral factor (not necessarily stable) $W(s)$ has the same McMillan degree of the transfer function $V_a(s) = \Phi_{xy}(s)W_a^{-*}(s)$ corresponding to the unique minimal square *analytic* spectral factor $W_a(s)$, which has the same zero structure of $W(s)$. Hence, by restricting attention to analytic spectral factors only, we do not lose in generality.

Observe that, given a pair $(W_0(s), V_0(s))$ solving (3.1), any other pair of the form $(W_0(s)T, V_0(s)T)$, where T is an arbitrary orthogonal matrix, is also a solution of (3.1). We shall regard two such solutions as equivalent and choose as a representative the pair $(W_0(s), V_0(s))$, which has $W_0(\infty) = R^{1/2}$.

A. Riccati Equations and Spectral Factorization

We shall need to review below some classical results of spectral factorization theory, mostly due to Anderson [3], [2]. These results describe a parametrization of the minimal analytic square spectral factors of $\Phi_y(s)$ in terms of the solutions of a certain algebraic Riccati equation.

⁵We advise the reader that Appendix C uses definitions and notations introduced up to the end of Proposition 3.1 below.

Let P be the state covariance of model (1.1), i.e., the unique solution of the Lyapunov equation

$$AP + PA^T = -BB^T. \quad (3.5)$$

Then, a change of basis in the state space of the realization (2.2) of $\Phi_y(s)$ induced by the matrix $T := \begin{bmatrix} I & -P \\ 0 & I \end{bmatrix}$ splits the spectral density $\Phi_y(s)$ in the form

$$\Phi_y(s) = \Phi_{y+}(s) + \Phi_{y+}^*(s) \quad (3.6)$$

where

$$\Phi_{y+}(s) = C(sI - A)^{-1}\bar{C}^T + \frac{1}{2}R \quad (3.7)$$

the matrix $\bar{C}^T$ being given by

$$\bar{C}^T := PC^T + B_1R^{1/2}.$$

It follows from (3.7) that the matrices $(C, A, \bar{C}^T)$ yield a minimal realization of the spectrum $\Phi_y(s)$ and hence are the same (i.e., invariants) for the class of all minimal models (1.1) representing the process y (in the given basis). Then a central result of stochastic system theory (see, e.g., [18]) states that the state covariance matrix X of any other stationary minimal realization of y of the form (1.1) satisfies the algebraic Riccati inequality

$$AX + XA^T + (\bar{C} - CX)^T R^{-1}(\bar{C} - CX) \leq 0 \quad (3.8)$$

and, conversely to each symmetric solution X of (3.8), there corresponds an essentially unique minimal system (A, B, C, D) generating y as the output of a “shaping filter” of the type (1.1). In particular, the state covariance P of our signal model satisfies

$$AP + PA^T + (\bar{C} - CP)^T R^{-1}(\bar{C} - CP) = -B_2B_2^T \quad (3.9)$$

which is just an equivalent way of writing the Lyapunov equation (3.5). The particular solutions P_0 of (3.8) with the equality sign, i.e., the solutions of

$$AP_0 + P_0A^T + (\bar{C} - CP_0)^T R^{-1}(\bar{C} - CP_0) = 0 \quad (3.10)$$

correspond to (minimal) stochastic realizations with the smallest number of input noise components, i.e., to *minimal square spectral factors* W_0 of Φ_y . By subtracting (3.10) from (3.9) and rearranging terms, one obtains another Riccati equation for the difference $\Sigma = P - P_0$

$$\Gamma\Sigma + \Sigma\Gamma^T - \Sigma C^T R^{-1}C\Sigma + B_2B_2^T = 0. \quad (3.11)$$

Clearly, since P is fixed, the solutions Σ of this equation are in one-to-one correspondence with those of (3.10) by the relation

$$\Sigma = P - P_0. \quad (3.12)$$

Some of these facts are collected in the following lemma, which will be used repeatedly in the sequel.

Lemma 3.1 (Anderson): Let Σ be a symmetric solution of the algebraic Riccati equation (3.11). Define

$$B_\Sigma := B_1 + \Sigma C^T R^{-1/2}. \quad (3.13)$$

Then

$$W_\Sigma(s) := C(sI - A)^{-1}B_\Sigma + R^{1/2} \quad (3.14)$$

is a minimal analytic square spectral factor of $\Phi_y(s)$. Conversely, to any minimal analytic square spectral factor $W(s)$ of $\Phi_y(s)$ there corresponds a symmetric solution Σ of (3.11) such that $W(s) = W_\Sigma(s)$ has the form specified by (3.14).

Hence, given a symmetric solution Σ of the algebraic Riccati equation (3.11), we have a corresponding transfer function $V(s)$ of the smoother, defined by

$$\begin{aligned} V_\Sigma(s) &:= \Phi_{xy}(s)W_\Sigma^{-*}(s) \\ &= (sI - A)^{-1}BW_\Sigma^{-*}(s)W_\Sigma^{-*}(s). \end{aligned} \quad (3.15)$$

B. A Family of Smoothing Filters

We shall introduce the following technical assumption.

Assumption 3.1: The pole and the zero sets of the spectral density $\Phi_y(s)$ are disjoint:

$$\begin{aligned} \sigma\left(\begin{bmatrix} A & BB^T \\ 0 & -A^T \end{bmatrix}\right) \\ \cap \sigma\left(\begin{bmatrix} A - B_1R^{-1/2}C & B_2B_2^T \\ C^T R^{-1}C & -A^T + C^T R^{-1/2}B_1^T \end{bmatrix}\right) = \emptyset. \end{aligned} \quad (3.16)$$

Many of the results described below continue to hold even if this assumption does not hold, however their proofs would be overburdened by technicalities risking to hide the meaning of the results. It is worth observing that the condition (3.16) is always satisfied when $\Phi(s)$ is scalar and holds generically in the multivariable case.

The following proposition provides a realization of $V_\Sigma(s)$ and a formula yielding its McMillan degree. Note that the eigenvalues of the matrix Γ_Σ introduced in (3.17) are the zeros of the square spectral factor $W_\Sigma(s)$. For this reason Γ_Σ is sometimes called the *numerator matrix* of $W_\Sigma(s)$.

Proposition 3.1: Let Σ be a solution of (3.11) and $V_\Sigma(s)$ be the corresponding transfer function defined by (3.15). Let

$$\Gamma_\Sigma := A - B_\Sigma R^{-1/2}C = \Gamma - \Sigma C^T R^{-1}C. \quad (3.17)$$

Then

$$V_\Sigma(s) = [-\Sigma \quad I] \left(sI - \begin{bmatrix} -\Gamma_\Sigma^T & 0 \\ 0 & A \end{bmatrix} \right)^{-1} \begin{bmatrix} C^T R^{-1/2} \\ B_\Sigma \end{bmatrix} \quad (3.18)$$

and, under Assumption 3.1⁶

$$\deg(V_\Sigma(s)) = 2n - n_\Sigma \quad (3.19)$$

where n_Σ is the dimension of the unobservable space $\mathcal{V}$ of the pair $(\Sigma, \Gamma_\Sigma^T)$.

Proof: We have

$$\begin{aligned} W_\Sigma^{-*}(s) &:= [W_\Sigma(s)^{-1}]^* \\ &= R^{-1/2}B_\Sigma^T(sI + \Gamma_\Sigma^T)^{-1}C^T R^{-1/2} + R^{-1/2} \end{aligned} \quad (3.20)$$

⁶Here and in the following, $\deg(\cdot)$ denotes the McMillan degree.

where Γ_Σ is defined in (3.17). Then, taking into account that $C^T R^{-1/2} B_\Sigma^T = A^T - \Gamma_\Sigma^T$ and using Lemma B.1, we easily get

$$W^*(s)W_\Sigma^{-*}(s) = \begin{bmatrix} I \\ 0 \end{bmatrix} + \begin{bmatrix} R^{-1/2} C \Sigma \\ -B_\Sigma^T \end{bmatrix} (sI + \Gamma_\Sigma^T)^{-1} C^T R^{-1/2}. \quad (3.21)$$

From this, using again Lemma B.1, we compute $V_\Sigma(s) = (sI - A)^{-1} B W^*(s) W_\Sigma^{-*}(s)$ and obtain (3.18).

To compute $\deg(V_\Sigma(s))$, we observe that from the minimal realization (3.14), it follows that the pair (A, B_Σ) is reachable, and from the minimal realization (3.20), it follows that the pair $(-\Gamma_\Sigma^T, C^T R^{-1/2})$ is also reachable. Moreover, (3.20) implies that $\sigma(-\Gamma_\Sigma^T)$ is a subset of the zeros of $\Phi_y(s)$; hence from Assumption 3.1 it follows that

$$\sigma(A) \cap \sigma(-\Gamma_\Sigma^T) = \emptyset. \quad (3.22)$$

Then the following is clear.

- 1) The realization (3.18) is reachable.
- 2) The unobservable subspace $\mathcal{N}$ of the realization (3.18) is isomorphic to the unobservable subspace $\mathcal{V}$ of the pair $(\Sigma, \Gamma_\Sigma^T)$. Indeed, $v \in \mathcal{N}$ if and only if $v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$ with $v_1 \in \mathcal{V}$.

These two observations clearly yield (3.19). ■

The following proposition shows that cascading with the withening filter $W_\Sigma(s)^{-1}$ does not change the MacMillan degree of $V_\Sigma(s)$.

Proposition 3.2: Denote by $[\bar{x}_\Sigma(t)^T \ x_\Sigma(t)^T]^T$ the state (Markov) process of the realization (3.18) driven by the input white noise process $w_\Sigma(t)$. Then, for each Σ , the smoothing filter is obtained by applying to (3.18) the state feedback

$$w_\Sigma(t) = -R^{-1/2} C x_\Sigma(t) + R^{-1/2} y(t). \quad (3.23)$$

The transfer function of the smoother can be written as

$$\begin{aligned} S(s) &= V_\Sigma(s) W_\Sigma^{-1}(s) \\ &= [-\Sigma \quad I] \left(sI - \begin{bmatrix} -\Gamma_\Sigma^T & -C^T R^{-1} C \\ 0 & \Gamma_\Sigma \end{bmatrix} \right)^{-1} \\ &\quad \cdot \begin{bmatrix} C^T R^{-1} \\ B_\Sigma R^{-1/2} \end{bmatrix}. \end{aligned} \quad (3.24)$$

Proof: The formula can be obtained by transfer function manipulations of (3.18) and $W_\Sigma(s)$. A more instructive derivation is obtained by noting that the shaping filter of transfer function $W_\Sigma(s)$ can be represented by the state space model

$$\begin{aligned} \dot{x}_\Sigma(t) &= A x_\Sigma(t) + B_\Sigma w_\Sigma(t) \\ y(t) &= C x_\Sigma(t) + R^{1/2} w_\Sigma(t) \end{aligned}$$

where the input and hence the state processes are the same of the second component of (3.18). From this obtain a realization of the withening filter of transfer function $W_\Sigma(s)^{-1}$

$$\begin{aligned} \dot{x}_\Sigma(t) &= (A - B_\Sigma R^{-1/2} C) x_\Sigma(t) + B_\Sigma R^{-1/2} y(t) \\ w_\Sigma(t) &= -R^{-1/2} C x_\Sigma(t) + R^{-1/2} y(t) \end{aligned}$$

and then just substitute the last equation in the state space model corresponding to the realization (3.18). ■

Formula (3.24) describes a *family* of state-space realizations of the smoother, parametrized by Σ . It is easy to show that this formula particularizes to the well-known *two-filters formula* of the smoothing literature, of which it provides a generalized version.

Consider the Lyapunov equation

$$X \Gamma_\Sigma + \Gamma_\Sigma^T X = -C^T R^{-1} C. \quad (3.25)$$

Since the spectral density $\Phi_y(s)$ has been assumed to be coercive [(1.6)], we may pick Σ in such a way that

$$\sigma(\Gamma_\Sigma) \cap \sigma(-\Gamma_\Sigma^T) = \emptyset \quad (3.26)$$

i.e., the spectra of Γ_Σ and of $-\Gamma_\Sigma^T$ are disjoint (which is commonly called “unmixed spectrum” condition). In this case, (3.25) has a unique symmetric solution which we denote by Δ^{-1} for convenience. This solution is nonsingular since $(\Gamma_\Sigma^T, C^T R^{-1/2})$ is reachable, as noted in the proof of Proposition 3.1. For example, choosing for Σ the maximal solution of the ARE (3.11), $\Sigma_+ := P - P_-$, where P_- is the minimal solution of the algebraic Riccati equation (3.10); then $\Gamma_{\Sigma_+} := \Gamma_+$ is a stability matrix and Δ turns out to be precisely the difference between the maximal and minimal solutions of (3.10), namely, $\Delta = P_+ - P_- > 0$, which is sometimes called the *gap* of the Riccati equation [30], [9].

Proposition 3.3 (Two-Filters Formula): Let Σ be any symmetric solution of the ARE (3.11) such that the numerator matrix Γ_Σ has unmixed spectrum. Denote by Δ^{-1} the corresponding unique symmetric solution (necessarily invertible) of the Lyapunov equation (3.25). Then the transfer function of the smoother (3.24) admits the additive decomposition

$$\begin{aligned} S(s) &= (I - \Sigma \Delta^{-1})(sI - \Gamma_\Sigma)^{-1} B_\Sigma R^{-1/2} \\ &\quad + \Sigma \Delta^{-1} (sI - \bar{\Gamma}_\Sigma)^{-1} \bar{B}_\Sigma R^{-1/2} \end{aligned} \quad (3.27)$$

where

$$\begin{aligned} \bar{\Gamma}_\Sigma &:= \Gamma_\Sigma - \Delta C^T R^{-1} C = -\Delta \Gamma_\Sigma^T \Delta^{-1} \\ \bar{B}_\Sigma &= B_\Sigma - \Delta C^T R^{-1/2}. \end{aligned} \quad (3.28)$$

If $\Sigma = \Sigma_+ = P - P_-$, then $\Delta = P_+ - P_- > 0$, and the smoothed estimate can be written in the form

$$\begin{aligned} \hat{x}(t) &= [I - (P - P_-)(P_+ - P_-)^{-1}] \hat{x}_+(t) \\ &\quad + (P - P_-)(P_+ - P_-)^{-1} \hat{x}_-(t) \end{aligned} \quad (3.29)$$

where $\hat{x}_+(t)$ is the steady-state causal (forward) Kalman Filter estimate of $x(t)$ and $\hat{x}_-(t)$ is the steady-state anticausal (backward) Kalman Filter estimate of $x(t)$.

Proof: We show that there is a similarity transformation of the form $T := \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$, which block-diagonalizes the realization (3.24), if and only if there is a symmetric solution X of the Lyapunov equation (3.25). In fact, since $T^{-1} := \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$, we have

$$\begin{aligned} T^{-1} \begin{bmatrix} -\Gamma_\Sigma^T & -C^T R^{-1} C \\ 0 & \Gamma_\Sigma \end{bmatrix} T &= \begin{bmatrix} -\Gamma_\Sigma^T & -X \Gamma_\Sigma - \Gamma_\Sigma^T X - C^T R^{-1} C \\ 0 & \Gamma_\Sigma \end{bmatrix} \end{aligned}$$

and the northeast corner of the transformed matrix is zero if and only if X solves (3.25). The similarity transformation induced by any such X gives

$$T^{-1} \begin{bmatrix} -\Sigma & I \\ C^T R^{-1} & B_\Sigma R^{-1/2} \end{bmatrix} T = \begin{bmatrix} -\Sigma & I - \Sigma X \\ C^T R^{-1} - X B_\Sigma R^{-1/2} & B_\Sigma R^{-1/2} \end{bmatrix}$$

and consequently provides the following decomposition of $S(s)$:

$$S(s) = \Sigma(-sI - \Gamma_\Sigma^T)^{-1}(C^T R^{-1} - X B_\Sigma R^{-1/2}) + (I - \Sigma X)(sI - \Gamma_\Sigma)^{-1} B_\Sigma R^{-1/2}.$$

Now, substituting $X = \Delta^{-1}$ and rearranging terms, one obtains

$$S(s) = (I - \Sigma \Delta^{-1})(sI - \Gamma_\Sigma)^{-1} B_\Sigma R^{-1/2} + \Sigma \Delta^{-1}(-sI - \Delta \Gamma_\Sigma^T \Delta^{-1})^{-1} \cdot (\Delta C^T R^{-1} - B_\Sigma R^{-1/2})$$

which is (3.27). That the two expressions of $\bar{\Gamma}_\Sigma$ are the same follows since Δ^{-1} solves the Lyapunov equation (3.25) and hence $-\Delta \Gamma_\Sigma^T \Delta^{-1} = \Gamma_\Sigma - \Delta C^T R^{-1} C$. Finally, in case $\Sigma = \Sigma_+ = P - P_-$, we obtain the well-known formulas for the backward Kalman filter; see [16] and [9, p. 124]. ■

The last part of this proposition can be obtained as a corollary of [5, Theorem 4.1]. Note that the parameters of the anticausal Kalman filter are

$$\bar{\Gamma}_{\Sigma_+} = -\Delta \Gamma_+^T \Delta^{-1} = \Gamma_{\Sigma_-} := \Gamma_- \\ \bar{B}_{\Sigma_+} = B_{\Sigma_-} := B_-.$$

Compare [16] and [9].

Yet another useful form of the smoother that is easily derived from (3.24) is the *cascade decomposition* described in the following proposition, which to our knowledge seems not to be in the literature.

Proposition 3.4 (Cascade Decomposition): To each nonsingular solution Σ of the Riccati equation (3.11), there corresponds a cascade decomposition of the transfer function of the smoother, $S(s) = W_1(s)W_2(s)$, where

$$W_1(s) = \Sigma(-sI - \Gamma_\Sigma^T)^{-1} \Sigma^{-1} \quad (3.30)$$

$$W_2(s) = B_2 B_2^T \Sigma^{-1} (sI - \Gamma_\Sigma)^{-1} B_\Sigma R^{-1/2} - B_1 R^{-1/2}. \quad (3.31)$$

If $\Sigma = \Sigma_+$, then $W_1(s)$ is anticausal and $W_2(s)$ is causal, while for $\Sigma = \Sigma_-$, the opposite is true.

Proof: Since S is strictly proper, we look for factorizations where one of the factors is also strictly proper. Suppose we want to find factorizations of $S(s)$ with $W_1(s) = H_1(sI - F_1)^{-1} G_1$ (strictly proper) and $W_2(s) = H_2(sI - F_2)^{-1} G_2 + D_2$. If such a factorization exists, the cascade realization defined by the triplet

$$[H_1 \quad 0], \quad \begin{bmatrix} F_1 & G_1 H_2 \\ 0 & F_2 \end{bmatrix}, \quad \begin{bmatrix} G_1 D_2 \\ G_2 \end{bmatrix}$$

should be equivalent to a realization of the family (3.24). We try a similarity transformation of block-upper triangular form $T := \begin{bmatrix} I & K \\ 0 & I \end{bmatrix}$, whereby

$$[H_1 \quad 0] = [-\Sigma \quad I] T = [-\Sigma \quad I] \begin{bmatrix} I & K \\ 0 & I \end{bmatrix}$$

implies, assuming Σ is invertible, that $K = \Sigma^{-1}$ and $H_1 = -\Sigma$. Computing the upper right off-diagonal block (of place 12) in

$$\begin{bmatrix} I & -\Sigma^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} -\Gamma_\Sigma^T & -C^T R^{-1} C \\ 0 & \Gamma_\Sigma \end{bmatrix} \begin{bmatrix} I & \Sigma^{-1} \\ 0 & I \end{bmatrix}$$

we find $-\Sigma^{-1} \Gamma_\Sigma - \Gamma_\Sigma^T \Sigma^{-1} - C^T R^{-1} C$. Now Σ is a solution of the algebraic Riccati equation (3.11), which can be rewritten

$$\Gamma_\Sigma \Sigma + \Sigma \Gamma_\Sigma^T + \Sigma C^T R^{-1} C \Sigma + B_2 B_2^T = 0 \quad (3.32)$$

and this implies that

$$G_1 H_2 = \Sigma^{-1} B_2 B_2^T \Sigma^{-1}.$$

Finally, imposing that

$$\begin{bmatrix} I & -\Sigma^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} C^T R^{-1} \\ B_\Sigma R^{-1/2} \end{bmatrix} = \begin{bmatrix} G_1 D_2 \\ G_2 \end{bmatrix}$$

we obtain $G_2 = B_\Sigma R^{-1/2}$ and

$$G_1 D_2 = \Sigma^{-1} (\Sigma C^T R^{-1} - B_\Sigma R^{-1/2}) = \Sigma^{-1} (-B_1) R^{-1/2}$$

where the last equality is based on (3.13). It follows that for each invertible solution Σ of (3.11) the transfer functions (3.30) and (3.31) provide a cascade factorization of $S(s)$. ■

Since the two-filter formula (3.27) and the cascade decomposition above are just a decomposition of the transfer function (3.24), which has the same MacMillan degree of $V_\Sigma(s)$ in (3.18), it is clear that these formulas provide a minimal recursive implementation of the smoother if and only if $(\Sigma, \Gamma_\Sigma^T)$ is observable, which is the same of saying that $n_\Sigma = 0$, or equivalently

$$\deg(V_\Sigma(s)) = 2n.$$

This is really a condition on the *zero structure* of $W(s)$, as stated in the following theorem.

Theorem 3.1: The two-filter realization (3.27) or the cascade decomposition of Proposition 3.4 are of minimal possible dimension $2n$ if and only if $W(s)$ has no zeros.

For the proof of this theorem, we shall need the following result.

Lemma 3.2: Let Σ be any solution of (3.11); then the unobservable subspace $\mathcal{V}$ is equal to $\ker \Sigma$ and $\mathcal{V}$ is invariant for both Γ_Σ^T and Γ^T . In fact, $\mathcal{V} \subset \mathcal{V}^*$ is a zero-direction subspace and

$$\Gamma_\Sigma^T|_{\mathcal{V}} = \Gamma^T|_{\mathcal{V}} \quad (3.33)$$

so that the square factor $W_\Sigma(s)$ shares $n_\Sigma = \dim \mathcal{V}$ zeros (counted with multiplicity) with $W(s)$. Hence

$$n_\Sigma \leq \nu \quad (3.34)$$

where ν is the dimension of $\mathcal{V}^*$, the zero direction space of $W(s)$, and

$$\deg(V_\Sigma(s)) \geq 2n - \nu \quad (3.35)$$

for all Σ solving (3.11).

Proof: Let $u \in \ker \Sigma$. By multiplying (3.11) on the left side by u^T and on the right side by u , it readily follows that

$$\ker B_2^T \supseteq \ker \Sigma \quad (3.36)$$

which, using the Riccati equation (3.11) rewritten in the form

$$\Gamma \Sigma + \Sigma \Gamma_\Sigma^T = \Sigma \Gamma^T + \Gamma_\Sigma \Sigma = -B_2 B_2^T \quad (3.37)$$

implies that the subspace $\ker \Sigma$ is an invariant subspace for both Γ^T and Γ_Σ^T . Since $\ker \Sigma$ is Γ_Σ^T -invariant and $\mathcal{V}$ is the largest Γ_Σ^T -invariant subspace of $\ker \Sigma$, it is obvious that $\ker \Sigma = \mathcal{V}$. Formula (3.33) follows from (3.17). Finally, in view of the inclusion (3.36) and the Γ^T -invariance of $\ker \Sigma$, it is obvious that

$$\ker \Sigma \subseteq \ker \begin{bmatrix} B_2^T \\ B_2^T \Gamma^T \\ \vdots \\ B_2^T (\Gamma^T)^{n-1} \end{bmatrix} = \mathcal{V}^* \quad (3.38)$$

which clearly implies $\mathcal{V} \subset \mathcal{V}^*$ and (3.34). ■

Inclusion (3.36) is also proved in [21]; see also [18, Lemma 10.2], and [17, Prop. 4.10] for a slightly different formulation. Note that the lemma establishes a lower bound for the dimension of the smoother ($\deg(V_\Sigma(s))$). We shall show later that this lower bound is always attainable for suitable Σ 's.

Proof of Theorem 3.1: If $W(s)$ has no zeros, $\nu = 0 \Rightarrow n_\Sigma = 0 \Rightarrow \mathcal{V} (= \ker \Sigma) = \{0\}$, i.e., Σ is nonsingular and $(\Sigma, \Gamma_\Sigma^T)$ is trivially observable so that (3.18) is minimal.

Conversely, assume there are minimal filters of dimension $2n$, i.e., such that $\mathcal{V} (= \ker \Sigma) = \{0\}$. This will imply in particular that the maximal solution $\Sigma_+ = P - P_-$ of the ARE is nonsingular and, in fact, positive definite. Since $\Gamma_+ := \Gamma - \Sigma_+ C^T R^{-1} C$ is asymptotically stable, it follows from standard Riccati theory that (Γ, B_2) must be a controllable pair.

Alternative Proof: From standard Riccati theory, Σ_+ is the unique positive semidefinite solution of (3.11). Similarly, $\Sigma_- := P - P_+ < 0$ is the unique negative semidefinite solution of (3.11). Anticipating from Lemma 4.3, which will be stated in the next section, this implies that $\mathcal{V}^* = 0$, since Σ_+ and Σ_- are both nonsingular. See Lemma 4.3 below.

This concludes the proof of Theorem 3.1. ■

IV. A FAMILY OF MINIMAL DEGREE SMOOTHERS

The solution of the factorization problem (3.1) has been reduced to the following question: For what Σ 's is the McMillan degree, $2n - n_\Sigma$ of $V_\Sigma(s)$, minimal? This question will be answered in this section in a series of Lemmas culminating with Theorem 4.1.

In the light of Lemma 3.2, the analytic spectral factors that solve problem (3.1) are the square factors $W_\Sigma(s)$, which share the maximal number of zeros with $W(s)$. In fact, Lemma 3.2

leads to conjecture that the optimal $W_\Sigma(s)$'s should be those which share *all the* ν zeros of $W(s)$.

Indeed, there is a whole family of minimal square factors of Φ_y that share exactly the ν zeros of the (nonsquare) transfer function $W(s)$ of the original model. In order to describe this family, we need to recall the concept of tightest local frame of a solution P of the Riccati inequality (3.8). Let P_{0-} and P_{0+} be, respectively, the maximal solution P_0 of the Riccati equation (3.10) for which $P - P_0 \geq 0$ and the minimal solution of (3.10) for which $P - P_0 \leq 0$ [18]. The *tightest local frame* of P , denoted $[[P_{0-}, P_{0+}]]$, is the subset of solutions $Q = Q^T$ of (3.8) defined by the matrix inequality

$$[[P_{0-}, P_{0+}]] := \{Q \mid P_{0-} \leq Q \leq P_{0+}\}. \quad (4.1)$$

Naturally, $P \in [[P_{0-}, P_{0+}]]$, the inclusion being trivial ($P = P_{0-} = P_{0+}$) in case P itself solves the Riccati equation (3.10). It is shown in [18, Theorem 11.1] and [17] that

$$\begin{aligned} \ker(P - P_{0-}) &= \ker(P_{0+} - P) \\ &= \ker(P_{0+} - P_{0-}) = \mathcal{V}^* \end{aligned} \quad (4.2)$$

where $\mathcal{V}^*$ is the subspace of zero directions of the spectral factor $W(s)$ corresponding to P .

To our purposes, it will be convenient to reparametrize the tightest frame in terms of the solutions Σ of the “centered” algebraic Riccati equation (3.11). Letting

$$\Sigma_{0+} := P - P_{0-} \quad \Sigma_{0-} := P - P_{0+} \quad (4.3)$$

it follows readily from the definition that Σ_{0+} is the minimal positive semidefinite solution of (3.11), and similarly, Σ_{0-} is the maximal negative semidefinite solution of (3.11).

Let Γ_{0-} and Γ_{0+} be the numerator matrices of the “extreme” square factors $W_{0-}(s)$, $W_{0+}(s)$ corresponding to Σ_{0-} and Σ_{0+} , respectively, i.e., $\Gamma_{0-} := \Gamma - \Sigma_{0-} C^T R^{-1} C$, $\Gamma_{0+} := \Gamma - \Sigma_{0+} C^T R^{-1} C$. We have the following result on the zeros of $W(s)$.

Lemma 4.1: There holds

$$\Gamma_{0-}^T |_{\mathcal{V}^*} = \Gamma^T |_{\mathcal{V}^*} = \Gamma_{0+}^T |_{\mathcal{V}^*} \quad (4.4)$$

so that $W_{0-}(s)$, $W_{0+}(s)$ share the ν zeros of $W(s)$.

Proof: This is essentially the same claim as that of Lemma 3.2, with the additional information that in this case $\ker \Sigma = \ker(P - P_{0-}) = \ker(P_{0+} - P) = \mathcal{V}^*$. ■

It is shown in [18] and [17] that the zero-sets of all spectral factors $W_Q(s)$ corresponding to solutions Q of the Riccati inequality (3.8) belonging to the set $[[P_{0-}, P_{0+}]]$ contain the common zeros of the “extreme” square factors $W_{0-}(s)$, $W_{0+}(s)$ (counting multiplicity). When $[[P_{0-}, P_{0+}]]$ is actually the *tightest* frame for Q , in the sense defined above, the zeros of $W_Q(s)$ are *exactly* the common zeros of $W_{0-}(s)$ and $W_{0+}(s)$. Of interest for our problem are the (square) spectral factors attached to the elements P_0 of the tightest frame of P , $[[P_{0-}, P_{0+}]]$, which solve the Riccati equation (3.10). These will be discussed in more detail below.

⁷That is, $P_{0-} \leq Q \leq P_{0+}$, in which case one says that $[[P_{0-}, P_{0+}]]$ is a frame for Q .

Define

$$[\Sigma_{0-}, \Sigma_{0+}] := \{\Sigma \mid \Sigma \text{ solves (3.11) and } \Sigma_{0-} \leq \Sigma \leq \Sigma_{0+}\} \quad (4.5)$$

so that the set $\{P_0 = P - \Sigma \mid \Sigma \in [\Sigma_{0-}, \Sigma_{0+}]\}$ is the subset of the tightest local frame of P , made of all solutions of the Riccati equation (3.10) that belong to $[[P_{0-}, P_{0+}]]$.

A Change of Basis: It will henceforth be convenient to assume, without loss of generality, that a change of basis is introduced in the signal model (1.1), which transforms the pair (Γ, B_2) into the so-called *standard form of controllability*, which is a block structure of the type

$$\Gamma = \begin{bmatrix} F & L \\ 0 & Z \end{bmatrix}, \quad B_2 = \begin{bmatrix} G \\ 0 \end{bmatrix} \quad (4.6)$$

where $F \in \mathbb{R}^{(n-\nu) \times (n-\nu)}$ and the pair (F, G) is reachable. In a basis of this kind, the subspace $\mathcal{V}^*$ is made up of vectors of the form $x = [0 \ v']^T, v \in \mathbb{R}^\nu$.

The $\nu \times \nu$ matrix Z is the restriction of Γ to $\mathcal{V}^*$ and carries the zeros of the signal model (1.1). The change of basis induces a partition of the matrices C and B_1 of the signal model (1.1) in blocks of the same dimension

$$C = [C_1 \ C_2], \quad B_1 = \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix} \quad (4.7)$$

where $C_1 \in \mathbb{R}^{m \times (n-\nu)}, C_2 \in \mathbb{R}^{m \times \nu}, B_{11} \in \mathbb{R}^{(n-\nu) \times m}$, and $B_{12} \in \mathbb{R}^{\nu \times m}$.

Consider then the reduced-order algebraic Riccati equation (RARE) obtained by restricting (3.11) to $(\mathcal{V}^*)^\perp$

$$FY + YF^T - YC_1^T R^{-1} C_1 Y = -GG^T. \quad (4.8)$$

It is immediate to check that for any solution Y of (4.8) the matrix Σ given by

$$\Sigma = \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} \quad (4.9)$$

is a solution of (3.11).

Observe that the pair (C, A) is, by assumption, observable. Hence (C, Γ) , and consequently (C_1, F) are observable. Therefore the reduced-order Riccati equation (4.8) satisfies the system-theoretic conditions ensuring the existence and uniqueness of a maximal (positive definite) solution Y_+ and of a minimal (negative definite) solution Y_- . Each other solution Y of (4.8) is such that

$$Y_- \leq Y \leq Y_+. \quad (4.10)$$

Since the pair (G^T, F^T) is observable by construction, any solution Y of (4.8) is nonsingular; moreover Y_+ is the only positive semidefinite solution and Y_- is the only negative semidefinite solution of (4.8). Therefore, we obtain the following representations for the extremal solutions Σ_{0+}, Σ_{0-} .

Lemma 4.2: In any basis bringing (Γ, B_2) in standard form of controllability, Σ_{0+} and Σ_{0-} have the form

$$\Sigma_{0+} := \begin{bmatrix} Y_+ & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_{0-} := \begin{bmatrix} Y_- & 0 \\ 0 & 0 \end{bmatrix} \quad (4.11)$$

where Y_+ and Y_- are the maximal and the minimal solution of the RARE (4.8).

Returning to the parametrization of minimal analytic square spectral factors in terms of solutions $P_0 = P - \Sigma$ of the Riccati equation (3.10), we have the following characterization.

Lemma 4.3: Let $\Sigma \in [\Sigma_{0-}, \Sigma_{0+}]$ and let Γ_Σ be the numerator matrix of the corresponding minimal square spectral factor $W_\Sigma(s)$ (Lemma 3.1). Then, $\ker \Sigma = \mathcal{V}^*$ and

$$\Gamma_\Sigma^T|_{\mathcal{V}^*} = \Gamma^T|_{\mathcal{V}^*} \quad (4.12)$$

so that all these spectral factors, including the extremes $W_{0-}(s)$ and $W_{0+}(s)$, share all the ν zeros of $W(s)$.

Proof: In view of the relation between $[\Sigma_{0-}, \Sigma_{0+}]$ and $[[P_{0-}, P_{0+}]]$ we have

$$\begin{aligned} \ker((P - \Sigma_{0-}) - (P - \Sigma_{0+})) \\ = \ker(P - (P - \Sigma_{0+})) = \ker(P - (P - \Sigma_{0-})) \\ = \ker \Sigma_+ = \ker \Sigma_- = \mathcal{V}^*. \end{aligned} \quad (4.13)$$

Moreover, by a famous representation theorem of Willems [30], every solution Y of (4.8) has a representation

$$Y = \Pi Y_- + (I - \Pi) Y_+ \quad (4.14)$$

where Π is a projection matrix, so that we also have

$$\Sigma = \tilde{\Pi} \Sigma_- + (I - \tilde{\Pi}) \Sigma_+, \quad \tilde{\Pi} = \begin{bmatrix} \Pi & 0 \\ 0 & I \end{bmatrix} \quad (4.15)$$

from which it clearly follows that $\Sigma a = 0$ for all $a \in \mathcal{V}^*$, i.e., $\ker \Sigma \supseteq \mathcal{V}^*$. On the other hand, by (3.36) in the Proof of Lemma 3.2, we always have $\ker \Sigma \subseteq \mathcal{V}^*$. This proves the claim that $\ker \Sigma = \mathcal{V}^*$. The rest is as usual. ■

The following theorem gives a solution of (3.1) and parametrizes all factorizations $S(s) = V_\Sigma(s) W_\Sigma^{-1}(s)$, where $W_\Sigma(s)$ is a minimal square analytic spectral factor of $\Phi_y(s)$ and $V_\Sigma(s)$ has minimal McMillan degree, in terms of solutions Σ of (3.11) belonging to the set $[\Sigma_{0-}, \Sigma_{0+}]$.

Theorem 4.1: Assume that a basis transformation on the model (1.1) has been chosen so that (Γ, B_2) is in standard controllability form (4.6). Then, the filter $V_\Sigma(s)$ is of minimal McMillan degree

$$\deg(V_\Sigma(s)) = 2n - \nu \quad (4.16)$$

if and only if $\Sigma \in [\Sigma_{0-}, \Sigma_{0+}]$, i.e., has the form (4.9), the matrix Y being a solution of the reduced Riccati equation (4.8). Hence all pairs $(W_\Sigma(s), V_\Sigma(s))$ solving (3.1) are parametrized by the solutions $\Sigma \in [\Sigma_{0-}, \Sigma_{0+}]$ of (3.11).

Proof: We only need to show that if a solution Σ of (3.11) has the structure

$$\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_2 \end{bmatrix} \quad (4.17)$$

where at least one of the two matrices Σ_{12} and Σ_2 is different from zero, then

$$n_\Sigma < \nu \quad (4.18)$$

strictly. To this aim, observe that if Σ is given by (4.17), then there exists a vector v such that

$$\Sigma \begin{bmatrix} 0 \\ v \end{bmatrix} = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_2 \end{bmatrix} \begin{bmatrix} 0 \\ v \end{bmatrix} \neq 0. \quad (4.19)$$

But we also have

$$B_2^T \begin{bmatrix} 0 \\ v \end{bmatrix} = [G^T \quad 0] \begin{bmatrix} 0 \\ v \end{bmatrix} = 0 \quad (4.20)$$

and

$$\begin{aligned} B_2^T (\Gamma^T)^l \begin{bmatrix} 0 \\ v \end{bmatrix} &= B_2^T (\Gamma^T)^{l-1} \begin{bmatrix} F^T & 0 \\ \Gamma_{12}^T & Z^T \end{bmatrix} \begin{bmatrix} 0 \\ v \end{bmatrix} \\ &= B_2^T (\Gamma^T)^{l-1} \begin{bmatrix} 0 \\ v_1 \end{bmatrix} \\ &= [G^T \quad 0] \begin{bmatrix} 0 \\ v_l \end{bmatrix} = 0, \quad \forall l > 0. \end{aligned} \quad (4.21)$$

This argument, taking into account (3.38), proves the *strict* inclusion

$$\ker \Sigma \subset \ker \begin{bmatrix} B_2^T \\ B_2^T \Gamma^T \\ \vdots \\ B_2^T (\Gamma^T)^{n-1} \end{bmatrix} = \mathcal{V}^* \quad (4.22)$$

which is equivalent to (4.18). ■

Remark: As pointed out before, the dimension (McMillan degree) of a minimal smoother is equal to the McMillan degree of $V_\Sigma(s)$, $\Sigma \in [\Sigma_{0-}, \Sigma_{0+}]$, which is in turn equal to the dimension of a minimal Markovian subspace containing the estimate $\hat{x}(\cdot)$. Theorem 4.1 states that this dimension is $n_0 = 2n - \nu$. ■

A. Error Covariance of the Optimal Smoother

Once Γ is put in standard controllability form (4.6), the numerator matrix Γ_Σ of W_Σ , defined in (3.17), has, for any $\Sigma \in [\Sigma_{0-}, \Sigma_{0+}]$, the partitioned form

$$\Gamma_\Sigma = \Gamma - \Sigma C^T R^{-1} C = \begin{bmatrix} F_Y & L_Y \\ 0 & Z \end{bmatrix} \quad (4.23)$$

where $F_Y := F - Y C_1^T R^{-1} C_1$, $L_Y := L - Y C_1^T R^{-1} C_2$. Because of $\mathcal{V}^*$ -invariance, the lower block $[\Gamma_\Sigma]_{22}$ is independent of Σ and equal to Z . The eigenvalues of Z are the zeros of $W(s)$, which are shared by all W_Σ 's in the chosen family. They are (fixed and) independent of Σ .

In the chosen basis, the unobservable part of (3.18) can be deleted by inspection obtaining a minimal realization of $V_\Sigma(s)$

$$\begin{aligned} V_\Sigma(s) &= \begin{bmatrix} -Y & I & 0 \\ 0 & 0 & I \end{bmatrix} \left(sI - \begin{bmatrix} -F_Y^T & 0 \\ 0 & A \end{bmatrix} \right)^{-1} \\ &\quad \cdot \begin{bmatrix} C_1^T R^{-1/2} \\ B_\Sigma \end{bmatrix}. \end{aligned} \quad (4.24)$$

In this formula, it is evident that the “additional dynamics” ($-F_Y^T$) of the stochastic realization of the smoother depends on the particular choice of the whitening filter W_Σ^{-1} , i.e., on the choice of Σ . Now, choose $\Sigma_{0-} = P - P_{0+} = \text{diag}\{Y_-, 0\}$, where Y_- is the minimal (negative definite) solution of the RARE, and denote by $F_- := F - Y_- C_1^T R^{-1} C_1$ the upper left

block of the corresponding numerator matrix $\Gamma_{\Sigma_{0-}}$. Since Y_- satisfies the reduced Riccati equation, which can be written (in the form (3.32)) as

$$F_- Y_- + Y_- F_- = -G G^T - Y_- C_1^T R^{-1} C_1 Y_- \quad (4.25)$$

a standard Lyapunov-type argument proves that all the eigenvalues of F_- lie in the open right-half plane or, equivalently, $-F_-$ is a stability matrix. For this choice of Σ , the filter $V_\Sigma(s)$ is analytic on $\{\text{Re}[s] > 0\}$ having a minimal realization

$$\begin{aligned} V_{\Sigma_{0-}}(s) &= \begin{bmatrix} -Y_- & I & 0 \\ 0 & 0 & I \end{bmatrix} \left(sI - \begin{bmatrix} -F_-^T & 0 \\ 0 & A \end{bmatrix} \right)^{-1} \\ &\quad \cdot \begin{bmatrix} C_1^T R^{-1/2} \\ B_{\Sigma_{0-}} \end{bmatrix} \end{aligned} \quad (4.26)$$

where the eigenvalues of the state matrix lie all in $\{\text{Re}[s] < 0\}$.

Since $V_{\Sigma_{0-}}(s)$ is an analytic filter driven by white noise, its state process $x_0(t)$ is a stationary Markov process, and its covariance P_0 satisfies the Lyapunov equation

$$\begin{aligned} \begin{bmatrix} -F_-^T & 0 \\ 0 & A \end{bmatrix} P_0 + P_0 \begin{bmatrix} -F_-^T & 0 \\ 0 & A \end{bmatrix}^T \\ + \begin{bmatrix} C_1^T R^{-1/2} \\ B_{\Sigma_{0-}} \end{bmatrix} \begin{bmatrix} C_1^T R^{-1/2} \\ B_{\Sigma_{0-}} \end{bmatrix}^T = 0. \end{aligned} \quad (4.27)$$

Partitioning P_0 in four blocks conformly with the state matrix of $V_{\Sigma_{0-}}(s)$

$$P_0 = \begin{bmatrix} P_{01} & P_{012}^T \\ P_{012} & P_{02} \end{bmatrix} \quad (4.28)$$

equation (4.27) can be decoupled into the three independent equations

$$-F_-^T P_{01} - P_{01} F_- + C_1^T R^{-1} C_1 = 0 \quad (4.29a)$$

$$A P_{012} - P_{012} F_- + B_{\Sigma_{0-}} R^{-1/2} C_1 = 0 \quad (4.29b)$$

$$A P_{02} + P_{02} A^T + B_{\Sigma_{0-}} B_{\Sigma_{0-}}^T = 0. \quad (4.29c)$$

It is standard, and not difficult to check directly, that the solution of (4.29a) is given by

$$\begin{aligned} P_{01} &= (Y_+ - Y_-)^{-1} \\ &= \left([I \quad 0] (P_{0+} - P_{0-}) \begin{bmatrix} I \\ 0 \end{bmatrix} \right)^{-1}. \end{aligned} \quad (4.30)$$

Taking into account (3.17), it is easy to verify that the (unique) solution of (4.29b) is $P_{012} = \begin{bmatrix} -I \\ 0 \end{bmatrix}$, while the solution of (4.29c) is clearly $P_{02} = P_{0+} = P - \Sigma_{0-}$.

On the other hand, the optimal smoothed estimate $\hat{x}(t)$ is expressed by a linear function of the Markov process $x_0(t)$

$$\hat{x}(t) = \begin{bmatrix} -Y_- & I & 0 \\ 0 & 0 & I \end{bmatrix} x_0(t) \quad (4.31)$$

so that its covariance matrix is computed as

$$\begin{aligned} \hat{P} &= \begin{bmatrix} -Y_- & I & 0 \\ 0 & 0 & I \end{bmatrix} P_0 \begin{bmatrix} -Y_- & I & 0 \\ 0 & 0 & I \end{bmatrix}^T \\ &= P_{0+} + \begin{bmatrix} Y_-(Y_+ - Y_-)^{-1} Y_- + 2Y_- & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \quad (4.32)$$

By standard properties of the orthogonal projection, the error process $\tilde{x}(t) = x(t) - \hat{x}(t)$ is orthogonal to $\hat{x}(t)$, so that its covariance is given by $\tilde{P} = P - \hat{P}$. Recalling that $P = P_{0+} + \Sigma_{0-}$, by straightforward computations, we obtain the following formula.

Proposition 4.1: In a basis bringing (Γ, B_2) in standard controllability form (4.6), the error covariance matrix $\tilde{P}$ of the smoothed estimate $\hat{x}(t)$, is given by

$$\tilde{P} = \begin{bmatrix} Y_+ - Y_+(Y_+ - Y_-)^{-1}Y_+ & 0 \\ 0 & 0 \end{bmatrix}. \quad (4.33)$$

where Y_+ and Y_- are the extreme solutions of the reduced-order ARE (4.8).

This formula is remarkably similar to the one derived in [5]. The difference is that in [5], the extremal solutions P_+ and P_- of the full Riccati equation were needed while here we only require the extreme solutions Y_+ and Y_- of a reduced Riccati equation of dimension $n - \nu$.

Incidentally, (4.33) shows that the optimal estimate $\hat{x}(t)$ is exact, i.e., not affected by errors, along the zero-directions space $\mathcal{V}^*$, the smoothing error occurring only in the directions of the orthogonal complement $(\mathcal{V}^*)^\perp$. This fact agrees with the geometric property of the “output-induced subspace” of the state space of the smoother, discussed in [17].

In the “extreme,” yet scarcely interesting, case when the given model is internal [$W(s)$ square spectral factor and P is solution of (3.10)] $P_{0+} = P_{0-} = P$, $\tilde{P} = 0$, and the estimate is, as expected, not affected by errors. Moreover there is no need of solving Riccati equations.

V. SMOOTHER IMPLEMENTATION

The smoother has, in general, an acausal structure, and a numerically stable implementation of the algorithm requires a causal–anticausal decomposition of its transfer function (this is, in fact, the motivation of “two-filter” formulas). In this section, we shall address the problem of computing causal–anticausal decompositions of the smoother in the general case when $W(s)$ may have an arbitrary number of zeros.

In a basis in which (Γ, B_2) has the control canonical form (4.6), Σ has the block diagonal form (4.9) and Γ_Σ has the block structure (4.23). In this basis, the unobservable part of (3.24) can be deleted by inspection, yielding a minimal realization

$$S(s) = \begin{bmatrix} -Y & I & 0 \\ 0 & 0 & I \end{bmatrix} \cdot \left(sI - \begin{bmatrix} -F_Y^T & -C_1^T R^{-1}C_1 & -C_1^T R^{-1}C_2 \\ 0 & F_Y & L_Y \\ 0 & 0 & Z \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} C_1^T R^{-1} \\ B_{1Y} R^{-1/2} \\ B_{12} R^{-1/2} \end{bmatrix} \quad (5.1)$$

where the two row-blocks $B_{1Y} = B_{11} + Y C_1^T R^{-1/2}$ of dimension $(n - \nu) \times m$, and B_{12} of dimension $\nu \times m$, are the partitioning of $B_\Sigma = B_1 + \Sigma C^T R^{-1}$, induced by the partitioning $B_1 = [B_{11}^T \ B_{12}^T]^T$ in the standard controllability form.

From this realization, we can obtain a family of minimal “two-filter” or “cascade” formulas of the type seen in Propositions 3.3 and 3.4. In particular, a minimal causal–anticausal two-filters implementation is described in the theorem below.

Theorem 5.1 (Reduced Two-Filter Formula): Assume that the signal model (1.1) is transformed by a change of basis in the state space, bringing (Γ, B_2) into a standard controllability form of the type (1.1).

Let $Y_+ = P - P_{0-}$ be the maximal symmetric solution of the RARE (4.8), and let F_+ be the corresponding numerator matrix with spectrum in the left-half plane. Consider the reduced Lyapunov equation

$$X F_+ + F_+^T X = -C_1^T R^{-1} C_1 \quad (5.2)$$

and denote by Δ_0^{-1} its unique symmetric solution, necessarily invertible, where $\Delta_0 = Y_+ - Y_- > 0$.

Then the smoothing filter (5.1) has a minimal realization described by the following state-space equations:

$$\dot{\zeta} = Z\zeta + B_{12}R^{-1/2}y \quad (5.3)$$

$$\dot{\xi}_+ = F_+\xi_+ + L_+\zeta + B_{1+}R^{-1/2}y \quad (5.4)$$

$$\dot{\xi}_- = F_-\xi_- + L_-\zeta + B_{1-}R^{-1/2}y \quad (5.5)$$

$$\hat{x}(t) = \begin{bmatrix} \Pi_+\xi_+(t) + \Pi_-\xi_-(t) \\ \zeta(t) \end{bmatrix} \}_{n-\nu}^{\nu} \quad (5.6)$$

where

$$\Pi_+ = -Y_-(Y_+ - Y_-)^{-1} \quad \Pi_- = Y_+(Y_+ - Y_-)^{-1} \quad (5.7)$$

and

$$F_+ = F - Y_+ C_1^T R^{-1} C_1$$

$$F_- := F_+ + \Delta_0 C_1^T R^{-1} C_1 = -\Delta_0 F_+^T \Delta_0^{-1} \quad (5.8)$$

$$B_{1+} = B_{11} + Y_+ C_1^T R^{-1/2}$$

$$B_{1-} = B_{1+} - \Delta_0 C_1^T R^{-1/2} \quad (5.9)$$

$$L_+ = L - Y_+ C_1^T R^{-1} C_2$$

$$L_- = L_+ + \Delta_0 C_1^T R^{-1} C_2. \quad (5.10)$$

The state variables of the filter have the following interpretation. Partition the state of the signal model (1.1) in two subvectors as $x^T(t) = [\xi(t)^T \ \zeta(t)^T]$, where ζ is the ν -dimensional output-induced component. Then $\hat{\xi}_+(t)$ is the steady-state (forward) Kalman filter estimate of $\xi(t)$ and $\hat{\xi}_-(t)$ is the steady-state backward Kalman filter estimate of $\xi(t)$.

The error covariance matrix of the estimate is given by (4.33).

Proof: Clearly ζ is also the subvector formed by the last ν components of the state vector in the realization (5.1). It is immediate that ζ satisfies (5.3) and that it stays unchanged under projection onto the space spanned by y . Therefore, ζ is also the subvector of the last ν components of the output, as in (5.6). Next consider the $2(n - \nu)$ -dimensional subsystem obtained by extracting the first two blocks of (5.1) with transfer function

$$S_\nu(s) = [-Y \ I] \left(sI - \begin{bmatrix} -F_Y^T & -C_1^T R^{-1}C_1 \\ 0 & F_Y \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} C_1^T R^{-1} & -C_1^T R^{-1}C_2 \\ B_{1Y} R^{-1/2} & L_Y \end{bmatrix}.$$

The output of this system is the smoothed estimate $\hat{\xi}$ of the state subvector ξ . Note that the input to this subsystem is the “augmented” input variable $\begin{bmatrix} y \\ \zeta \end{bmatrix}$, as it follows from the block-triangular structure of the realization (5.1). By a change of basis T of the same upper triangular form as used in the Proof of Proposition 3.3, with X a solution of the reduced Lyapunov equation

$$XF_Y + F_Y^T X = -C_1^T R^{-1} C_1 \quad (5.11)$$

(here we assume F_Y has unmixed spectrum), the realization of $S_\nu(s)$ above is transformed into one of the form

$$\begin{aligned} \begin{bmatrix} \dot{\bar{\xi}}_Y(t) \\ \dot{\xi}_Y(t) \end{bmatrix} &= \begin{bmatrix} -F_Y^T & 0 \\ 0 & F_Y \end{bmatrix} \begin{bmatrix} \bar{\xi}_Y(t) \\ \xi_Y(t) \end{bmatrix} \\ &+ \begin{bmatrix} C_1^T R^{-1} - X B_{1Y} R^{-1/2} & -C_1^T R^{-1} C_2 - X L_Y \\ B_{1Y} R^{-1/2} & L_Y \end{bmatrix} \cdot \begin{bmatrix} y \\ \zeta \end{bmatrix} \\ \hat{\xi}(t) &= [-Y \quad I - YX] \begin{bmatrix} \bar{\xi}_Y(t) \\ \xi_Y(t) \end{bmatrix}. \end{aligned}$$

Choosing $Y = Y_+$, this is easily rewritten in the state-space form of the theorem. ■

The transfer function of the smoother can be written compactly as

$$\begin{aligned} S(s) &= \begin{bmatrix} S_\nu(s) \\ 0 \quad I \end{bmatrix} \begin{bmatrix} I \\ J(s) \end{bmatrix} \\ J(s) &= (sI - Z)^{-1} B_{12} R^{-1/2} \end{aligned} \quad (5.12)$$

where $J(s)$ is the transfer function $y \rightarrow \zeta$ describing the output-induced subvector of the state $x(t)$.

A reduced cascade decomposition of the same structure of (3.30) and (3.31) of Proposition 3.4 also holds for $S_\nu(s)$. The proof of this result is identical and will be omitted.

Theorem 5.2: To each solution Y of the reduced Riccati equation (4.8), there corresponds a cascade decomposition of the “reduced” transfer function of the smoother $S_\nu(s)$ of the form $S_\nu(s) = W_{\nu+}(s)W_{\nu-}(s)$, where

$$W_{\nu+}(s) = Y(-sI - F_Y^T)^{-1} Y^{-1} \quad (5.13)$$

$$\begin{aligned} W_{\nu-}(s) &= GG^T Y^{-1} (sI - F_Y)^{-1} [B_{1Y} R^{-1/2} \quad L_Y] \\ &\quad - [B_{11} R^{-1/2} \quad L]. \end{aligned} \quad (5.14)$$

If $Y = Y_+ = P - P_{0-} > 0$, then $F_{Y_+} = F_+$ is a stability matrix so that $W_{\nu+}(s)$ is anticausal while $W_{\nu-}(s)$ is causal. If instead we choose $Y = Y_- = P - P_{0+} < 0$, then $F_{Y_-} = F_-$ is antistable so that $W_{\nu+}(s)$ is causal while $W_{\nu-}(s)$ is anticausal.

We see that in any case the dynamics of the minimal smoother (5.1) splits into three decoupled subsystems.

- 1) A causal part governed by the eigenvalues of F_+ .
- 2) An anticausal part governed by the eigenvalues of F_- (which is similar to $-F_+^T$).
- 3) An invariant subsystem, which provides the output-induced subvector of the state. This part is governed by the eigenvalues of Z , i.e., the zeros of $W(s)$. It follows from coercivity of the spectrum that these eigenvalues

may be located anywhere in the complex plane except on the imaginary axis.

What computations are needed to implement the minimal smoother?

One should first perform a basis transformation on the model (1.1) in order to bring (Γ, B_2) in standard controllability form (4.6). This may be obtained from the data of the problem at a modest computational cost employing one of several algorithms existing in the literature, and no solution of Riccati equations are required.

Further, in order to obtain the reduced causal–anticausal decomposition of the smoother, the stable/unstable eigenspaces of the matrix

$$\begin{bmatrix} -F^T & C_1^T R^{-1} C_1 & C_1^T R^{-1} C_2 \\ GG^T & F & L \\ 0 & 0 & Z \end{bmatrix} \quad (5.15)$$

need to be computed.

This may, in turn, be decoupled into two separate subproblems.

- 1) Compute the stable/unstable eigenspace of the $\nu \times \nu$ matrix Z carrying the zero structure of $W(s)$. This decomposition has to be dealt with on a case-by-case basis. If $W(s)$ is minimum phase or maximum phase, no decomposition is needed.
- 2) Compute the stable/unstable eigenspace of the matrix

$$H := \begin{bmatrix} -F^T & C_1^T R^{-1} C_1 \\ GG^T & F \end{bmatrix} \quad (5.16)$$

which has dimension $2(n - \nu) \times 2(n - \nu)$.

This latter matrix is Hamiltonian, and computing its stable eigenspace is equivalent to the solution of the reduced algebraic Riccati equation (4.8) of dimension $n - \nu$.

VI. DISCRETE-TIME RESULTS

We shall list in this section the discrete-time versions of the main results obtained in the previous sections for the continuous-time problem. The derivations are in principle the same, although the calculations are often more involved than in the continuous-time case and will not be reported here.

We shall consider the following discrete-time linear stochastic model with constant coefficients:

$$x(t+1) = Ax(t) + Bw(t) \quad (6.1)$$

$$y(t) = Cx(t) + Dw(t) \quad (6.2)$$

driven by a p -dimensional normalized white Gaussian noise w .

We seek the linear minimum-variance estimate

$$\hat{x}(t) = E[x(t) | H(y_{-\infty}^{+\infty})] \quad (6.3)$$

of the state $x(t)$ given the whole time history of the observations $\{y(t); t \in \mathbb{Z}\}$. The transfer function of this smoothing filter, formally given by the well-known expression

$$\Phi_{xy}(z)\Phi_y(z)^{-1}$$

is in general noncausal. We want to decompose it in a combination (either parallel or cascade) of causal–anticausal filters of *minimal dimension*.

We shall make the following assumptions.

- 1) (A, B) is reachable and (C, A) is observable.
- 2) $\sigma(A) \in \mathbb{C}_< := \{z \in \mathbb{C}: |z| < 1\}$; as in the continuous-time case, this assumption is not strictly necessary, and we could only assume (at the price of some complications) that none of the eigenvalues of A has modulus one.
- 3) $W(z)$ is a minimal spectral factor, i.e., $W(z)$ is a solution of $\Phi_y(z) := W(z)W^*(z)$ of minimal degree [the notation $W^*(z)$ now stands for $W^T(z^{-1})$].
- 4) $\Phi_y(z)$ is coercive, i.e., $\Phi_y(e^{i\omega}) > 0, \forall \omega \in \mathbb{R}$.
- 5) $\Phi_y(\infty)$ is finite and nonsingular. This assumption, which in the continuous-time case is implied by coercivity, is known as *regularity* [24], [7]. It implies that the minimal spectral factors of $\Phi_y(z)$ have zeros neither at the origin nor at infinity [24]. In particular, if $W(z)$ is a minimal square spectral factor, $W(\infty)$ is nonsingular and the numerator matrix Γ of an arbitrary minimal realization of $W(z)$ has no zero eigenvalues, i.e., is also nonsingular.

Since $W(\infty)$ is nonsingular, the matrix D in (6.1) may be assumed to be in the form $D = [R^{1/2}|0]$, with R being square and nonsingular. We partition B conformly as $B = [B_1 \ B_2]$.

Note that the regularity assumption implies that the numerator matrix $\Gamma := A - B_1 R^{-1/2} C$ is nonsingular.

A calculation in the same spirit of that in the Proof of Theorem 2.1 leads to the following realization of the steady-state smoother:

$$S(z) = [I \ 0] \left(\begin{bmatrix} zI & 0 \\ 0 & z^{-1}I \end{bmatrix} - \begin{bmatrix} \Gamma & B_2 B_2^T \\ -C^T R^{-1} C & \Gamma^T \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} B_1 R^{-1/2} \\ C^T R^{-1} \end{bmatrix}. \quad (6.4)$$

By nonsingularity of Γ , this realization can also be rewritten in the familiar “forward difference” form, with z in place of z^{-1} . The price to pay for this operation is somewhat more complicated formulas.

As in the continuous-time case, it is not difficult to check that the realization (6.4) is reachable but not necessarily observable, and its unobservable subspace is isomorphic to $\mathcal{V}^*$, the unobservable subspace of the pair (B_2^T, Γ^T) , so that the result (2.11) remains true in the discrete-time case as well.

A. Discrete-Time Stochastic Realization and ARE

The state covariance P of the model (6.1) is the unique solution of the discrete-time Lyapunov equation

$$P = APA^T + BB^T. \quad (6.5)$$

By the same block triangular change of basis, the spectral density $\Phi_y(z)$ may be decomposed in the form

$$\Phi_y(z) = \Phi_{y+}(z) + \Phi_{y+}^*(z) \quad (6.6)$$

where

$$\Phi_{y+}(z) = C(zI - A)^{-1} \bar{C}^T + \frac{1}{2} \Lambda_0 \quad (6.7)$$

with

$$\bar{C}^T := APC^T + B_1 R^{1/2}, \quad \Lambda_0 := R + CPC^T. \quad (6.8)$$

These two quantities are *invariants* of the output process in a chosen basis. From the discrete-time version of the positive real lemma (see, e.g., [7] and [19]), it follows along the same lines of the continuous-time case that the set of minimal square stable (i.e., analytic in $\{z \in \mathbb{C}: |z| > 1\}$) spectral factors of $\Phi_y(z)$ can be parametrized in terms of the symmetric solutions Σ of the discrete-time Riccati equation

$$\Sigma = \Gamma \Sigma \Gamma^T - \Gamma \Sigma C^T (R + C \Sigma C^T)^{-1} C \Sigma \Gamma^T + B_2 B_2^T. \quad (6.9)$$

There is a one-to-one correspondence that makes any such Σ correspond to the minimal square spectral factor

$$W_\Sigma(z) := C(zI - A)^{-1} B_\Sigma + D_\Sigma \quad (6.10)$$

where

$$\begin{cases} D_\Sigma := (R + C \Sigma C^T)^{1/2} \\ B_\Sigma := (B_1 R^{1/2} + A \Sigma C^T) D_\Sigma^{-1} \\ \quad = B_1 R^{-1/2} D_\Sigma + \Gamma \Sigma C^T D_\Sigma^{-1}. \end{cases} \quad (6.11)$$

Define $R_\Sigma := (R + C \Sigma C^T) = D_\Sigma D_\Sigma^T$. Notice that the regularity assumption guarantees that R_Σ is nonsingular. The numerator matrix of the spectral factor (6.10) is

$$\begin{aligned} \Gamma_\Sigma &:= A - B_\Sigma D_\Sigma^{-1} C \\ &= \Gamma - \Gamma \Sigma C^T (R + C \Sigma C^T)^{-1} C = \Gamma K \end{aligned} \quad (6.12)$$

where $K := I - \Sigma C^T (R + C \Sigma C^T)^{-1} C$. In view of the regularity assumption, Γ_Σ , and hence Γ and K , are nonsingular.

Arguing as in the continuous-time case, it is possible to show that a Markovian space containing the optimal estimates is the state space of the filter

$$V_\Sigma(z) := \Phi_{xy}(z) W_\Sigma^{-*}(z) \quad (6.13)$$

driven by the output of the whitening filter $W_\Sigma^{-1}(z)$. Thus, we have to face again the problem of characterizing the solutions Σ for which $V_\Sigma(z)$ has minimal McMillan degree. It is possible to show that $V_\Sigma(z)$ has the realization

$$V_\Sigma(z) = [-\Sigma \ I] \left(zI - \begin{bmatrix} \Gamma_\Sigma^{-T} & 0 \\ 0 & A \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} \Gamma_\Sigma^{-T} C^T D_\Sigma^{-1} \\ B_\Sigma \end{bmatrix} \quad (6.14)$$

and that the smoothing filter is obtained by state feedback from this realization so that

$$S(z) = [-\Sigma \ I] \left(zI - \begin{bmatrix} \Gamma_\Sigma^{-T} & -\Gamma_\Sigma^{-T} C^T R_\Sigma^{-1} C \\ 0 & \Gamma_\Sigma^T \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} \Gamma_\Sigma^{-T} C^T R_\Sigma^{-1} \\ B_\Sigma D_\Sigma^{-1} \end{bmatrix}. \quad (6.15)$$

Let us now assume that the pair (Γ, B_2) is in canonical form of controllability (4.6), let C be partitioned conformly as in

(4.6), and consider the reduced-order algebraic Riccati equation obtained by restricting (6.9) to $(V^*)^\perp$

$$Y = FYF^T - FYC_1^T(R + C_1YC_1^T)^{-1}C_1YF^T + GG^T. \quad (6.16)$$

It is immediate to check that for any solution Y of (6.16), the matrix Σ given by

$$\Sigma = \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} \quad (6.17)$$

is a solution of (6.9). We define the sets $[\Sigma_{0-}, \Sigma_{0+}]$ and $[Y_-, Y_+]$ exactly as in the continuous-time case. Moreover, we set

$$\begin{aligned} R_Y &:= R + C_1YC_1^T, & D_Y &:= R_Y^{1/2} \\ F_Y &:= F - FYC_1^TR_Y^{-1}C_1 \\ L_Y &:= L - FYC_1^TR_Y^{-1}C_2 \end{aligned}$$

and

$$B_Y = \begin{bmatrix} B_{1Y} \\ B_{2Y} \end{bmatrix} = \begin{bmatrix} B_{11}R^{-1/2}D_Y + FYC_1^TD_Y^{-1} \\ B_{12}R^{-1/2}D_Y \end{bmatrix}$$

where B_{11} and B_{12} are defined in (4.7).

The following discrete-time version of Theorem 4.1 holds.

Theorem 6.1: Assume that the pole and zero sets of the spectral density $\Phi_y(z)$ are disjoint, and (Γ, B_2) is in standard controllability form. Then the filter $V_\Sigma(s)$ is of minimal McMillan degree

$$\deg(V_\Sigma(s)) = 2n - \nu \quad (6.18)$$

if and only if Σ has the form (6.17), the matrix Y being a solution of the reduced Riccati equation (6.16) belonging to the tightest frame $[Y_-, Y_+]$. Hence the minimal state-space realizations of the optimal smoother are parametrized by the solutions $Y \in [Y_-, Y_+]$ of (6.16).

To the minimum solution Y_- of (6.16) [or equivalently, to the solution $\Sigma_{0-} = \text{diag}\{Y_-, 0\}$ of (6.9)] there corresponds a stable filter $V_{Y_-}(z)$, which has the following minimal realization:

$$V_{Y_-}(z) = \begin{bmatrix} -Y_- & I & 0 \\ 0 & 0 & I \end{bmatrix} \left(zI - \begin{bmatrix} F_-^T & 0 \\ 0 & A \end{bmatrix} \right)^{-1} \begin{bmatrix} F_-^TC_1^TD_{Y_-}^{-1} \\ B_{Y_-} \end{bmatrix} \quad (6.19)$$

where $F_- := F_{Y_-}$ has all eigenvalues outside of the unit circle so that F_-^T is a (discrete-time) stability matrix.

The covariance P_0 of the state process of $V_{\Sigma_{0-}}(z)$ can then be obtained by solving a Lyapunov equation, and, similar to what

happens in the continuous-time case, this allows the computation of the smoothing error covariance $\tilde{P}$, which turns out to be given by the same expression of the continuous-time case so that (4.33) remains valid in the discrete time too. We shall skip the details, referring the reader to [7, pp. 95–96] for a guideline on discrete-time computations.

As in the continuous-time case, by choosing Σ to be of the form (6.17) and by deleting the unobservable part of (6.15), we get the following minimal realization of the smoother as shown in (6.20) at the bottom of the page, which, particularized for $Y = Y_+$, leads to the following discrete-time version of Theorem 5.1.

Theorem 6.2 (Reduced Two-Filter Formula, Discrete Case): Assume that the signal model (6.1) is transformed by a change of basis in the state space, bringing (Γ, B_2) into a standard controllability form of the type (4.6).

Let $Y_+ = P - P_{0-}$ be the maximal symmetric solution of the RARE (6.16), and let F_+ be the corresponding numerator matrix with spectrum in the unit circle. Consider the reduced Lyapunov equation

$$X = F_+^TXF_+ + C_1^TR_{Y_+}^{-1}C_1 \quad (6.21)$$

and denote by Δ_0^{-1} its unique symmetric solution, necessarily invertible, where $\Delta_0 = Y_+ - Y_- > 0$.

Then the smoothing filter (6.20) has a minimal realization described by the following state-space equations:

$$\zeta(t+1) = Z\zeta(t) + B_{2Y_+}D_{Y_+}^{-1}y(t) \quad (6.22)$$

$$\xi_+(t+1) = F_+\xi_+(t) + L_+\zeta(t) + B_{1+}D_{Y_+}^{-1}y(t) \quad (6.23)$$

$$\begin{aligned} \xi_-(t-1) &= \bar{F}_-\xi_-(t) - \bar{L}_-\zeta(t-1) \\ &\quad - \bar{B}_{1-}D_{Y_+}^{-1}y(t-1) \end{aligned} \quad (6.24)$$

$$\hat{x}(t) = \begin{bmatrix} \Pi_+\xi_+(t) + \Pi_-\xi_-(t) \\ \zeta(t) \end{bmatrix} \begin{matrix} \} n - \nu \\ \} \nu \end{matrix} \quad (6.25)$$

where

$$\begin{aligned} \Pi_+ &= -Y_-(Y_+ - Y_-)^{-1} \\ \Pi_- &= Y_+(Y_+ - Y_-)^{-1} \end{aligned} \quad (6.26)$$

and

$$\begin{aligned} F_+ &:= F_{Y_+} = F - FY_+C_1^TR_{Y_+}^{-1}C_1 \\ \bar{F}_- &:= \Delta_0F_+^T\Delta_0^{-1} \end{aligned} \quad (6.27)$$

$$\begin{aligned} B_{1+} &:= B_{1Y_+} = B_{11}R^{-1/2}D_{Y_+} + FY_+C_1^TD_{Y_+}^{-1} \\ \bar{B}_{1-} &:= F_-B_{1+} - \Delta_0C_1^TD_{Y_+}^{-1} \end{aligned} \quad (6.28)$$

$$\begin{aligned} L_+ &:= L_{Y_+} = L - FY_+C_1^TR_{Y_+}^{-1}C_2 \\ \bar{L}_- &:= F_-L_+ + \Delta_0C_1^TR_{Y_+}^{-1}C_2. \end{aligned} \quad (6.29)$$

$$S(z) = \begin{bmatrix} -Y & I & 0 \\ 0 & 0 & I \end{bmatrix} \left(zI - \begin{bmatrix} F_Y^T & -F_Y^TC_1^TR_Y^{-1}C_1 & -F_Y^TC_1^TR_Y^{-1}C_2 \\ 0 & F_Y & L_Y \\ 0 & 0 & Z \end{bmatrix} \right)^{-1} \begin{bmatrix} F_Y^TC_1^TR_Y^{-1} \\ B_{1Y}D_{Y_+}^{-1} \\ B_{12}D_{Y_+}^{-1} \end{bmatrix} \quad (6.20)$$

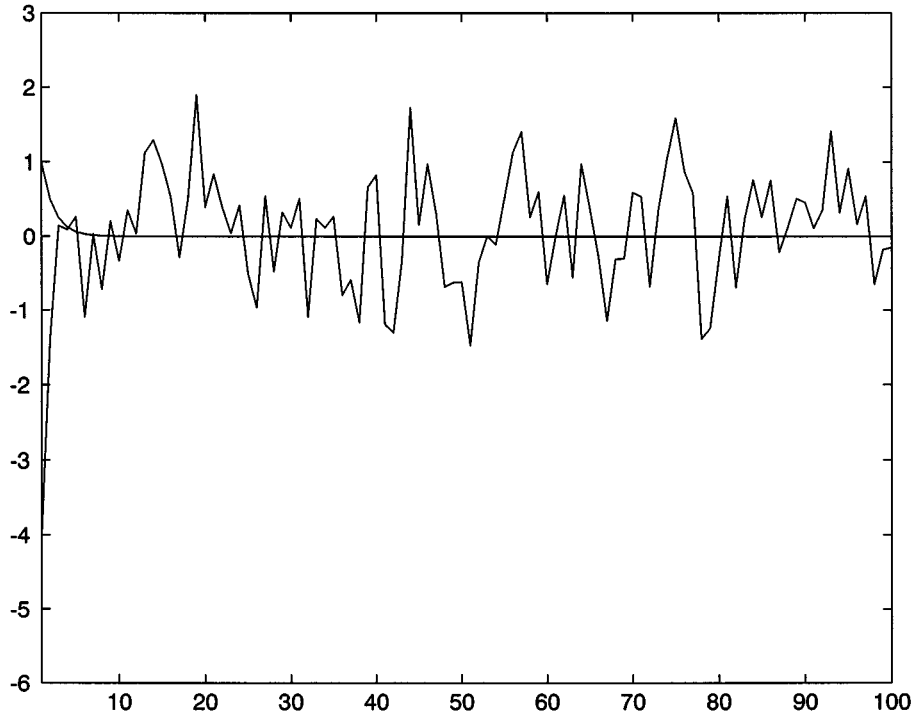


Fig. 1. State estimation errors.

The state variables of the filter have the following interpretation. Partition the state of the signal model (6.1) in two subvectors as $x^T(t) = [\xi^T(t) \quad \zeta^T(t)]$ where ζ is the ν -dimensional output-induced component. Then $\hat{\xi}_+(t)$ is the steady-state (forward) Kalman filter estimate of $\xi(t)$ and $\hat{\xi}_-(t)$ is the steady-state backward Kalman filter estimate of $\xi(t)$.

The error covariance matrix of the estimate is given by (4.33).

A reduced cascade decomposition of the same structure of Theorem 5.2 holds also in the discrete-time case.

Theorem 6.3: The transfer function of the smoother can be written as

$$S(z) = \begin{bmatrix} S_\nu(z) \\ 0 \quad I \end{bmatrix} \begin{bmatrix} I \\ J(z) \end{bmatrix} \quad (6.30)$$

$$J(z) = (zI - Z)^{-1} B_{2Y_+} D_{Y_+}^{-1}$$

where $J(z)$ is the transfer function $y \rightarrow \zeta$ describing the output-induced subvector of the state $x(t)$. Moreover, $S_\nu(z)$ has a cascade decomposition of the form $S_\nu(z) = W_{\nu-}(z)W_{\nu+}(z)$, where, employing the same notations of Theorem 6.2

$$W_{\nu-}(z) = -Y_+(zI - F_+^T)^{-1} F_+^T \quad (6.31)$$

$$W_{\nu+}(z) = (Y_+^{-1} - C_1^T R_{Y_+}^{-1} C_1 - F_+^T Y_+ F_+) \cdot (zI - F_+)^{-1} [B_1 + D_{Y_+}^{-1} L_+] + [C_1^T R_{Y_+}^{-1} - F_+^T Y_+^{-1} B_+ D_{Y_+}^{-1} | C_1^T R_{Y_+}^{-1} C_2 - F_+^T Y_+^{-1} L_+] \quad (6.32)$$

where $F_+ = F_{Y_+}$ is a (discrete-time) stability matrix so that $W_{\nu-}(s)$ is anticausal while $W_{\nu+}(s)$ is causal.

VII. SIMULATIONS

In this section, we present a very simple example of application of the reduced algorithm to simulated data.

We have driven a two-dimensional discrete-time system of the form (6.1) with white Gaussian noise and computed the smoothed estimates of the state employing the reduced two-filter formula of Theorem 6.2. The model has the following parameters:

$$A = -\frac{1}{2}I_2, \quad B = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$C = I_2, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that the system has dimension two but, thanks to the reduction process, we only need to solve a one-dimensional ARE corresponding to $F = 1/2$, $C_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $G = 1$, whose solutions are $Y_+ = (1 + \sqrt{65}/8)$ and $Y_- = (1 - \sqrt{65}/8)$.

Fig. 1 shows the two components of the estimation error $\tilde{x}$. The estimation error of the second state variable (the output-induced component), which should be zero on an infinite time interval, for a finite smoothing interval converges to zero very fast. The sample error covariance of the first state variable computed on 100 samplepoints is 0.5, which should be compared with the theoretical value $Y_+ - Y_+(Y_+ - Y_-)^{-1}Y_+ = 0.4961 \dots$

The smoothed estimate represented in Fig. 1 is obtained considering the steady-state process. Hence in the extremes of the interval, it is indeed a suboptimal estimate. Let then $\hat{x}_{\text{opt}}(t)$ be the optimal estimate (obtained with the time-varying filter). The difference $\hat{x}_{\text{opt}}(t) - \hat{x}(t)$ is appreciably different from zero only in correspondence of small intervals at the two extremes, as represented in Fig. 2.

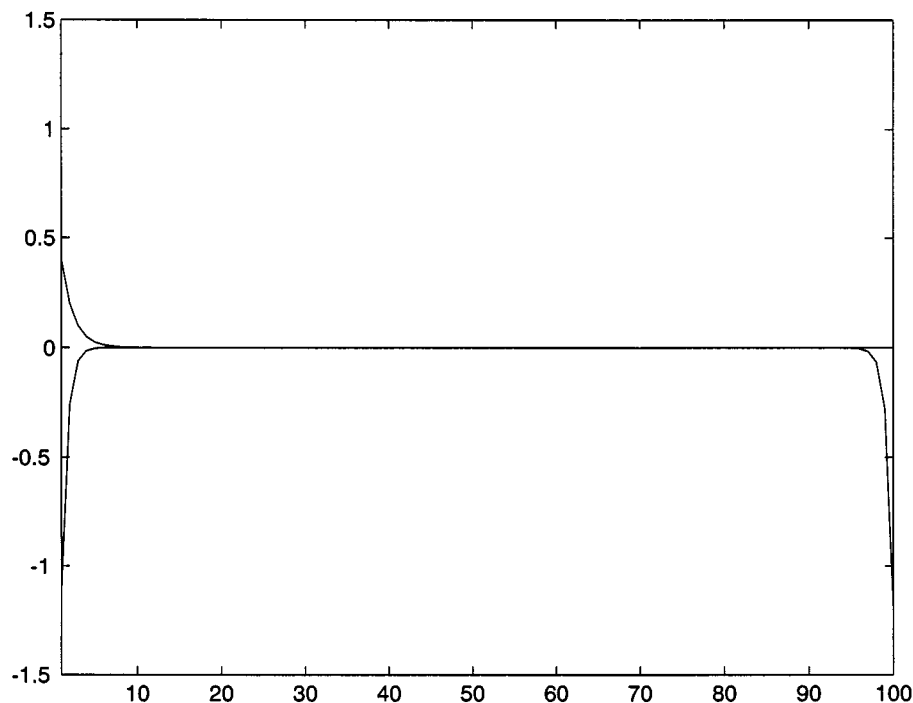


Fig. 2. Difference between the optimal estimate and the steady-state suboptimal estimate of the state.

It may be worth noticing that the estimate $\hat{x}_2$ of the second state variable is optimal in the left extreme of the interval too. This is not surprising since it is given by the forward Kalman filter, and after a brief transient, its error covariance vanishes.

Similar results may be obtained also in the continuous-time case.

VIII. CONCLUSIONS

In this paper, we have provided a thorough analysis of the steady-state smoothing problem for linear signal models. The dynamic structure of the smoother has been elucidated, and a simple computational procedure for constructing the minimal smoother has been proposed. In the construction of a state-space realization, the minimal smoother does not require the solution of Riccati equations. The solution of a Riccati equation of reduced order is needed only for the decomposition of the filter into a causal and an anticausal part.

APPENDIX A NONCAUSAL TRANSFER FUNCTIONS

All transfer functions of this paper represent linear operations on stationary processes defined on the whole time axis. The underlying mathematical theory is called “spectral representation theory of stationary processes” and can be found in the classical literature on stationary processes, for example, in [27, ch. I, Sect. 8]. Here we shall just recall the essential facts.

A (not necessarily causal) linear filter operating on the stationary process $\{y(t)\}$ is a convolution operator with a kernel function $s(t)$ (the *impulse response* of the filter) such that the

sum $\int_{-\infty}^{+\infty} s(t-\tau)y(\tau) d\tau$ converges in mean square. For purely nondeterministic processes (like the ones we consider in this paper) this is equivalent to $s(t)$ being square integrable and hence to the Fourier transform

$$S(j\omega) = \int_{-\infty}^{+\infty} s(t)e^{-j\omega t} dt$$

being also square integrable on the imaginary axis (Parseval theorem). If $S(\cdot)$ is rational, it can be extended (by letting $s = j\omega$) to the whole complex plane. The extension $S(s)$ is referred to as the *transfer function* of the filter. Notice that in general some of the poles of $S(s)$ may lie in the right-half plane, although square integrability implies that no poles can lie on the imaginary axis.

Every rational noncausal filter can be decomposed into the sum of a *causal* and an *anticausal* part by just decomposing $S(s)$ as

$$S(s) = S_+(s) + S_-(s)$$

where $S_+(s)$ is analytic on the right, while $S_-(s)$ is analytic on the left-half plane. The corresponding impulse responses $s_+(t)$ and $s_-(t)$ are zero for $t < 0$, i.e., causal, and zero for $t > 0$, i.e., anticausal, respectively (Paley–Wiener theorem). Therefore the operation of convolution of $s(t)$ with a stationary input process y splits into a sum of a causal and an anticausal convolution operators, whereby the “stable” modes s_+ of the impulse response are integrated forward in time and provide a causal functional of y , while the “unstable” modes s_- are integrated *backwards in time* and involve instead the future history of the input process. This is explained in more detail in [18, pp. 298–299] and [25].

APPENDIX B PRODUCT AND INVERSE OF TRANSFER FUNCTIONS

In this appendix, we present in form of lemmas two formulas that are useful for the computation of the product of two transfer functions and of the inverse of a transfer function. Such formulas are used several times in the paper. The first formula generalizes a well-known trick first used by Popov.

Lemma B.1: Let $W_i(s) = H_i(sI - F_i)^{-1}G_i$, $i = 1, 2$ and assume that $G_1H_2 = F_2 - F_1$. Then

$$W_1(s)W_2(s) = H_1(sI - F_2)^{-1}G_2 - H_1(sI - F_1)^{-1}G_2. \quad (\text{B.1})$$

Proof:

$$\begin{aligned} W_1(s)W_2(s) &= H_1(sI - F_1)^{-1}G_1H_2(sI - F_2)^{-1}G_2 \\ &= H_1(sI - F_1)^{-1}(F_2 - F_1)(sI - F_2)^{-1}G_2 \quad (\text{B.2}) \end{aligned}$$

$$\begin{aligned} &= H_1(sI - F_1)^{-1}((sI - F_1) \\ &\quad - (sI - F_2))(sI - F_2)^{-1}G_2 \quad (\text{B.3}) \end{aligned}$$

$$= H_1(sI - F_2)^{-1}G_2 - H_1(sI - F_1)^{-1}G_2. \quad (\text{B.4})$$

■

Lemma B.2: Let $W(s) = H(sI - F)^{-1}G + J$ be a minimal realization of a square matrix function with J nonsingular. Then

$$W^*(s) = -G^T(sI + F^T)^{-1}H^T + J^T \quad (\text{B.5})$$

$$\begin{aligned} W^{-1}(s) &= -J^{-1}H(sI - (F - GJ^{-1}H))^{-1} \\ &\quad \cdot GJ^{-1} + J^{-1} \quad (\text{B.6}) \end{aligned}$$

$$\begin{aligned} W^{-*}(s) &= J^{-T}G^T(sI - (-F^T + H^TJ^{-T}G^T))^{-1} \\ &\quad \cdot H^TJ^{-T} + J^{-T} \quad (\text{B.7}) \end{aligned}$$

are minimal realizations of $W^*(s)$, $W^{-1}(s)$, and $W^{-*}(s)$, respectively.

The proof is straightforward.

APPENDIX C RULING OUT NONANALYTIC SPECTRAL FACTORS

In this appendix, we show that to solve Problem (3.1) we can restrict attention to the set of analytic spectral factors.

The set $\mathcal{W}$ of minimal square (not necessarily stable) spectral factors $W(s)$ of $\Phi_y(s)$ can be parametrized in terms of two n -dimensional algebraic Riccati equations (see, e.g., [8] or [25]). In particular, we shall need the following result which we recall from [8]:

Lemma C.1: Let Σ be a symmetric solution of the ARE (3.11) and X be a solution of the homogeneous ARE

$$A^T X + XA + XB_\Sigma B_\Sigma^T X = 0 \quad (\text{C.1})$$

with

$$B_\Sigma := B_1 + \Sigma C^T R^{-1/2}. \quad (\text{C.2})$$

Moreover, let

$$A_0 := A + B_\Sigma B_\Sigma^T X \quad (\text{C.3})$$

and

$$K_X(s) := I + B_\Sigma^T X(sI - A_0)^{-1}B_\Sigma. \quad (\text{C.4})$$

Then $K_X(s)$ is unitary on the imaginary axis, i.e., $K_X(s) = K_X^{-*}(s)$, and

$$W_{\Sigma X}(s) := W_\Sigma(s)K_X(s) \quad (\text{C.5})$$

is a minimal square spectral factor of $\Phi_y(s)$. Conversely, to any minimal square spectral factor $W_1(s)$ of $\Phi_y(s)$, there correspond a solution Σ of (3.11) and a solution X of (C.1) such that $W_1(s)$ has the form specified by (C.5).

To each pair Σ, X of solutions of (3.11) and (C.1), we can therefore associate a unique minimal spectral factor $W_{\Sigma X}(s)$ and define the corresponding transfer function

$$V_{\Sigma X}(s) := \Phi_{xy}(s)W_{\Sigma X}^{-*}(s) = V_\Sigma(s)K_X(s) \quad (\text{C.6})$$

where $V_\Sigma(s)$ is defined by $V_\Sigma(s) := \Phi_{xy}(s)W_\Sigma^{-*}(s)$ and the last equality of (C.6) readily follows from the relation $K_X^{-*}(s) = K_X(s)$ [8].

Proposition C.1: Let Σ be a solution of (3.11), X be a solution of (C.1), and $V_{\Sigma X}(s)$ be defined by (C.6).

Then, under Assumption 3.1

$$\deg(V_{\Sigma X}(s)) = \deg(V_\Sigma(s)). \quad (\text{C.7})$$

To prove this proposition we need the following preliminary result.

Lemma C.2: For any pair (Σ, X) solving (3.11) and (C.1), respectively, the pair $((I + \Sigma X), A_0)$ is observable.

Proof: Taking into account (3.11) and the definition (C.2) of B_Σ , we get

$$B_\Sigma B_\Sigma^T = BB^T + \Sigma A^T + \Lambda \Sigma \quad (\text{C.8})$$

so that (C.1) may be rewritten as

$$(I + X\Sigma)A^T X + XA(I + \Sigma X) + XBB^T X = 0. \quad (\text{C.9})$$

Now, let v be a vector in the unobservable subspace of the pair $((I + \Sigma X), A_0)$, i.e.,

$$A_0 v = \lambda v \quad (\text{C.10a})$$

$$(I + \Sigma X)v = 0. \quad (\text{C.10b})$$

Multiplying (C.9) on the left by v^T and on the right by v and taking into account (C.10b), we get

$$B^T X v = 0. \quad (\text{C.11})$$

Now write (C.1) in the form

$$A^T X = -XA_0 \quad (\text{C.12})$$

which, multiplied on the right by v , yields

$$A^T X v = -\lambda X v \quad (\text{C.13})$$

where we have employed (C.10a). Since the pair (B^T, A^T) is observable, (C.11) and (C.13) imply $Xv = 0$ and, in view of (C.10b), $v = 0$. ■

Proof of Proposition C.1: Let us compute now $V_{\Sigma X}(s)$ and its McMillan degree $\deg(V_{\Sigma X}(s))$. To this aim, we observe that the matrix transfer function $K_X(s)$ introduced in Lemma C.1 is given by

$$\begin{aligned} K_X(s) &= I + B_{\Sigma}^T X(sI - A_0)^{-1} B_{\Sigma} \\ &= I + [B_{\Sigma}^T \quad 0] \\ &\quad \cdot \left(sI - \begin{bmatrix} A_0 & 0 \\ -\Gamma_{\Sigma}^T X - X A_0 & -\Gamma_{\Sigma}^T \end{bmatrix} \right)^{-1} \\ &\quad \cdot \begin{bmatrix} B_{\Sigma} \\ C^T R^{-1/2} \end{bmatrix}. \end{aligned} \quad (C.14)$$

Then, in view of the identity

$$\begin{aligned} &\begin{bmatrix} B_{\Sigma} B_{\Sigma}^T X & 0 \\ C^T R^{-1/2} B_{\Sigma}^T X & 0 \end{bmatrix} \\ &= \begin{bmatrix} A_0 & 0 \\ -\Gamma_{\Sigma}^T X - X A_0 & -\Gamma_{\Sigma}^T \end{bmatrix} - \begin{bmatrix} A & 0 \\ 0 & -\Gamma_{\Sigma}^T \end{bmatrix} \end{aligned} \quad (C.15)$$

we can employ Lemma B.1, which yields

$$\begin{aligned} V_{\Sigma X}(s) &= V_{\Sigma}(s) K_X(s) \\ &= [I \quad -\Sigma] \left(sI - \begin{bmatrix} A_0 & 0 \\ -\Gamma_{\Sigma}^T X - X A_0 & -\Gamma_{\Sigma}^T \end{bmatrix} \right)^{-1} \\ &\quad \cdot \begin{bmatrix} B_{\Sigma} \\ C^T R^{-1/2} \end{bmatrix} \\ &= [I + \Sigma X \quad -\Sigma] \left(sI - \begin{bmatrix} A_0 & 0 \\ 0 & -\Gamma_{\Sigma}^T \end{bmatrix} \right)^{-1} \\ &\quad \cdot \begin{bmatrix} B_{\Sigma} \\ X B_{\Sigma} + C^T R^{-1/2} \end{bmatrix}. \end{aligned} \quad (C.16)$$

From (3.14), (C.4), and (C.5), it easily follows that

$$W_{\Sigma X}(s) = (C + R^{1/2} B_{\Sigma}^T X)(sI - A_0)^{-1} B_{\Sigma} + R^{1/2} \quad (C.17)$$

which implies that the pair (A_0, B_{Σ}) is reachable; moreover, it is easy to check that

$$\begin{aligned} W_{\Sigma X}^{-*}(s) &= R^{-1/2} B_{\Sigma}^T (sI + \Gamma_{\Sigma}^T)^{-1} (X B_{\Sigma} + C^T R^{-1/2}) \\ &\quad + R^{-1/2} \end{aligned} \quad (C.18)$$

which implies that the pair $(-\Gamma_{\Sigma}^T, (X B_{\Sigma} + C^T R^{-1/2}))$ is also reachable. Finally, Lemma C.2 states the observability of the pair $((I + \Sigma X), A_0)$ and, in view of Assumption 3.1

$$\sigma(A_0) \cap \sigma(-\Gamma_{\Sigma}^T) = \emptyset. \quad (C.19)$$

Then, arguing as in Proposition 3.1, we conclude that

$$\deg V_{\Sigma X}(s) = 2n - n_{\Sigma} = \deg V_{\Sigma}(s) = 2n - n_{\Sigma} \quad (C.20)$$

where $n_{\Sigma}(s)$ is defined in Proposition 3.1. ■

REFERENCES

- [1] M. B. Adams, A. S. Willsky, and B. C. Levy, "Linear estimation of boundary value stochastic processes—Part II: 1-D smoothing problems," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 811–821, 1984.
- [2] B. D. O. Anderson, "A system theory criteria for positive real matrices," *SIAM J. Contr.*, vol. 5, pp. 171–182, 1967.
- [3] —, "The inverse problem of stationary covariance generation," *J. Statist. Phys.*, vol. 1, pp. 133–147, 1969.
- [4] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
- [5] F. Badawi, A. Lindquist, and M. Pavon, "A stochastic realization approach to the smoothing problem," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 878–888, 1979.
- [6] G. Basile and G. Marro, "Controlled and conditioned invariant subspaces in linear system theory," *J. Optim. Theory Appl.*, vol. 3, pp. 306–316, 1973.
- [7] P. Faure, M. Clerget, and F. Bermain, *Operateurs Rationnels Positifs*. Dunod, Paris, France, 1979.
- [8] A. Ferrante, "A parametrization of minimal stochastic realizations," *IEEE Trans. Automat. Contr.*, vol. 39, no. 10, pp. 2122–2126, 1994.
- [9] L. Finesso and G. Picci, "On the structure of minimal square spectral factors," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 122–127, 1982.
- [10] D. C. Fraser, "A new technique for optimal smoothing of data," Sc.D. dissertation, MIT, Cambridge, MA, 1967.
- [11] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
- [12] T. Kailath, *Linear System Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
- [13] T. Kailath and P. A. Frost, "An innovations approach to least-squares estimation—Part II: Linear smoothing in additive noise," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 655–660, 1968.
- [14] B. C. Levy, R. Frezza, and A. J. Krener, "Modeling and estimation of discrete-time Gaussian reciprocal processes," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 1013–1023, Sept. 1990.
- [15] F. L. Lewis, *Applied Optimal Control & Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [16] A. Lindquist and G. Picci, "On the stochastic realization problem," *SIAM J. Contr. Optim.*, vol. 17, pp. 365–389, 1979.
- [17] A. Lindquist, Gy. Michaletzky, and G. Picci, "Zeros of spectral factors, the geometry of splitting subspaces, and the algebraic Riccati inequality," *SIAM J. Contr. Optim.*, vol. 33, no. 401, p. 365, 1995.
- [18] A. Lindquist and G. Picci, "A geometric approach to modeling and estimation of linear stochastic systems," *J. Math. Syst., Estim. Contr.*, vol. 1, pp. 241–333, 1991.
- [19] —, "Geometric methods for state-space identification," in *Identification, Adaptation, Learning*, S. Bittanti and G. Picci, Eds. Berlin, Germany: Springer-Verlag, 1996.
- [20] D. Q. Mayne, "A solution of the smoothing problem for linear dynamic systems," *Automatica*, vol. 4, pp. 73–92, 1966.
- [21] B. P. Molinari, "The time-invariant linear-quadratic optimal-control problem," *Automatica*, vol. 13, pp. 347–357, 1977.
- [22] C. C. Paige and M. A. Sauders, "Least-squares estimation of discrete linear dynamical systems using orthogonal transformations," *SIAM J. Numer. Anal.*, vol. 14, pp. 180–193, 1977.
- [23] P. Park and T. Kailath, "New square-root smoothing algorithms," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 727–732, 1996.
- [24] M. Pavon, "Stochastic realization and invariant directions of the matrix Riccati equation," *SIAM J. Contr. Optim.*, vol. 18, pp. 155–180.
- [25] G. Picci and S. Pinzoni, "Acausal models and balanced realizations of stationary processes," *Linear Algebra Appl.*, vol. 205–206, pp. 997–1043, 1994.
- [26] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA J.*, vol. 3, pp. 1445–1450, 1965.
- [27] Yu. Rozanov, *Stationary Random Processes*. San Francisco, CA: Holden-Day, 1967.
- [28] G. S. Sidhu and U. B. Desai, "New smoothing algorithms based on reversed-time lumped models," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 538–541, 1976.
- [29] H. L. Weinert and U. B. Desai, "On complementary models and fixed-interval smoothing," *IEEE Trans. Automat. Contr.*, vol. AC-26, 1981.
- [30] J. C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 621–634, 1971.
- [31] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. Berlin, Germany: Springer-Verlag, 1977.



Augusto Ferrante was born on August 5, 1967. He received the "Laurea," *cum laude*, in electrical engineering in 1991 and the Ph.D. degree in control systems engineering in 1995, both from the University of Padua, Italy.

In June 1995, he became a Junior Faculty Member in the "Dipartimento di Ingegneria Elettrica, Gestionale e Meccanica," of the University of Udine, Italy. Since November 1998, he has been Associate Professor in the "Dipartimento di Elettronica e Informazione" of the "Politecnico di Milano," Italy.

His research interests are in the areas of stochastic realization, optimal control and filtering, and Riccati Equation.



Giorgio Picci (S'67–M'70–SM'91–F'94) holds a full professorship with the University of Padova, Italy, Department of Electronics and Informatics, since 1980. He graduated *cum laude* from the University of Padova in 1967 and since then has held several visiting appointments with various American and European universities among which M.I.T., Arizona State University, C.W.I. in Amsterdam, the Royal Institute of Technology, Stockholm Sweden, Kyoto University, Japan, and Washington University, St. Louis, MO.

He has been contributing to Systems and Control theory mostly in the area of modeling, estimation, and identification of stochastic systems. Since 1992, he has been active also in the field of dynamic vision and scene and motion reconstruction from monocular vision.