

Estimation for Boundary-Value Descriptor Systems¹Ramine Nikoukhah²Milton B. Adams³Alan S. Willsky²Bernard C. Levy⁴Abstract

In this paper we consider models for noncausal processes consisting of discrete-time descriptor dynamics and boundary conditions on the values of the process at the two ends of the interval on which the process is defined. We discuss the general solution and well-posedness of systems of this type and then apply the method of complementary processes to obtain a specification of the optimal smoother in terms of a boundary-value descriptor Hamiltonian system. We then study the implementation of the optimal smoother. Motivated by the Hamiltonian diagonalization results for non-descriptor systems, we show how the descriptor Hamiltonian dynamics can be transformed to two lower-order systems by the use of transformation matrices involving the solution of two generalized Riccati equations. We present several examples illustrating our results and the nature of the smoothing solution and also present equations for covariance analysis of boundary-value descriptor processes including the smoothing error.

¹The research described in this paper was supported in part by the Air Force Office of Scientific Research under Grant AFOSR-88-0032 and in part by the National Science Foundation under Grant ECS-8700903. The work of R. Nikoukhah and A.S. Willsky was also supported in part by Institut National de Recherche en Informatique et en Automatique (INRIA), France.

²Dept. of Electrical Engineering and Computer Science and The Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. 02139. R. Nikoukhah is also with INRIA, Domaine de Voluceau, Rocquencourt, B.P. 105, 78153, Le Chesnay Cedex, France.

³The Charles Stark Draper Laboratory, 555 Technology Square, Cambridge, MA. 02139.

⁴Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616.

I. Introduction

The class of descriptor systems was introduced by Luenberger [1] to describe the dynamics of certain linear systems for which standard state space representations are not particularly natural or appropriate. Since their introduction numerous studies have been performed to investigate the properties of these systems and the solution of control problems for them (see, for example, [2] - [9], [20], [21] and the references cited therein). The fundamental property that all of these studies have had to deal with, in some form or another, is the fact that the system function matrix for such a system is not proper, leading to impulsive behavior in continuous-time and giving rise to noncausal responses in discrete-time. The noncausality of these models makes them a natural choice for modeling spatially (rather than temporally-)varying phenomena, and in this context it is natural to consider descriptor models with general boundary conditions rather than with initial conditions or the special constrained forms for boundary conditions found in the literature. Indeed, if one considers generalizations of descriptor models to more than one independent variable, one finds that these models, together with appropriate boundary conditions, arise in many contexts such as in describing random fields, electromagnetic problems, gravitational anomalies, etc.

The investigation of standard (i.e. not descriptor) boundary-value models in one independent continuous variable was initiated by Krener [12] - [14] who has investigated many of their fundamental properties. Adams, et al. [10] developed a general approach to estimation for boundary-value models and

applied it in [11] to develop efficient estimation algorithms for processes described by the model introduced by Krener. In this paper we extend our estimation methodology to two-point boundary-value descriptor systems (TPBVDS's), i.e. discrete-time descriptor models in one independent variable and with general boundary conditions. To our knowledge this represents the first study of descriptor models devoted to estimation, and as we will see, our analysis uncovers both some important similarities and differences with estimation problems for standard state space models and several important problems whose solutions remain for the future. These questions have in fact inspired the development of a system theory for TPBVDS's [25,29-31], several elements of which will be used in the present development. Furthermore, in another paper [15] we use the results developed here in our investigation of efficient estimation algorithms for random fields describable in terms of a particular class of boundary-value descriptor systems in two-independent variables.

In the next section we introduce the class of TPBVDS's and perform some preliminary analysis. In particular, we discuss the well-posedness of such a system and a general method of solution for TPBVDS's. In Section III we apply the results of [10,26] to the fixed-interval smoothing problem for an n th-order TPBVDS. As we show, aside from a boundary effect which can be dealt with separately, the resulting smoother is itself naturally described as TPBVDS, in this case of dimension $2n$. In Section IV we address the question of implementation of the smoother. Motivated by the "Hamiltonian diagonalization" results in [11,22] for non-descriptor systems, we investigate two procedures for forward-backward diagonalization of the smoother equations.

These procedures, which are illustrated in Section V, point out connections with other work on descriptor systems and also lead to a number of additional questions related to generalizations of causal system-theoretic concepts to TPBVDS's. These are briefly discussed in Section VII following our analysis of the smoothing error in Section VI.

II. Two-Point Boundary-Value Descriptor Systems

The TPBVDS considered in this paper satisfies the difference equation

$$Ex(k+1) = Ax(k) + Bu(k) \quad (2.1)$$

with the two-point boundary condition

$$V_0x(0) + V_Kx(K) = v \quad (2.2)$$

Here $u(k)$ is an $m \times 1$ input sequence defined on the discrete-time interval $[0, k-1]$, $x(k)$ is the n -dimensional boundary value process, v is the n -vector of boundary values, and E , A , B , V_0 , and V_K are matrices of appropriate dimensions. Furthermore we assume that $\{E, A\}$ form a regular pencil (i.e. $|zE-A| = 0$).

As in [2], we can rewrite (2.1), (2.2) as a single set of equations

$$\mathcal{S}x = \mathcal{B}u \quad (2.3)$$

where

$$x' = (x'(0), \dots, x'(K)) \quad (2.4a)$$

$$u' = (u'(0), \dots, u'(K-1), v') \quad (2.4b)$$

$$\mathcal{S} = \begin{bmatrix} -A & E & 0 & \dots & 0 \\ 0 & -A & E & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & 0 & -A & E \\ V_0 & 0 & \dots & 0 & V_K \end{bmatrix} \quad (2.5a)$$

$$\mathcal{B} = \text{diag} (B, \dots, B, I) \quad (2.5b)$$

We see from this immediately that the well-posedness of (2.1), (2.2) is equivalent to the invertibility of \mathcal{S} . Much more can be said about

well-posedness and the solution of (2.1), (2.2), and we refer the reader to [25,29] for details. We limit ourselves here to describing one method for solving (2.1), (2.2) that provides us with an alternate well-posedness condition and with a method for the implementation of the smoother developed in Sections III and IV.

To begin, from Kronecker's canonical form for a regular pencil [17] we can find nonsingular matrices T and F so that⁵

$$FET^{-1} = \begin{bmatrix} I & 0 \\ 0 & A_b \end{bmatrix} \quad (2.6)$$

$$FAT^{-1} = \begin{bmatrix} A_f & 0 \\ 0 & I \end{bmatrix} \quad (2.7)$$

and so that all of the eigenvalues of A_f and A_b have magnitudes no larger than 1. Furthermore if $|zE-A|$ has no zeros on the unit circle, then all of the eigenvalues of A_f and A_b are strictly inside the unit circle. In this case we will say that $\{E,A\}$ is forward-backward stable.

⁵The decomposition in [17] splits the pencil $zE-A$ into forward dynamics corresponding to a pencil of the form $zI-\tilde{A}_f$ and backward dynamics corresponding to $z^{-1}I-\tilde{A}_b$ where \tilde{A}_b is nilpotent. The only difference in (2.6), (2.7) is that the unstable forward modes of \tilde{A}_f have been shifted into the backward dynamics A_b .

Define

$$\begin{bmatrix} x_f(k) \\ x_b(k) \end{bmatrix} = Tx(k) \quad (2.8)$$

Then, we obtain

$$x_f(k+1) = A_f x_f(k) + B_f u(k) \quad (2.9a)$$

$$x_b(k) = A_b x_b(k+1) - B_b u(k) \quad (2.9b)$$

where

$$\begin{bmatrix} B_f \\ B_b \end{bmatrix} = FB \quad (2.10)$$

and (2.9a), (2.9b) are asymptotically stable recursions if $\{E, A\}$ is forward-backward stable. Finally, given the transformation (2.8), the boundary condition (2.2) takes the form

$$\begin{bmatrix} V_{f,0} & V_{b,0} \end{bmatrix} \begin{bmatrix} x_f(0) \\ x_b(0) \end{bmatrix} + \begin{bmatrix} V_{f,K} & V_{b,K} \end{bmatrix} \begin{bmatrix} x_f(K) \\ x_b(K) \end{bmatrix} = v \quad (2.11)$$

$$\begin{bmatrix} V_{f,0} & V_{b,0} \end{bmatrix} = V_0 \Gamma^{-1}, \quad \begin{bmatrix} V_{f,k} & V_{b,k} \end{bmatrix} = V_K \Gamma^{-1} \quad (2.12)$$

Employing the forward/backward representation (2.9) of the dynamics, a general solution to (2.1), (2.2) is derived as follows. Let $x_f^0(k)$ denote the solution to (2.9a) with zero initial condition, and let $x_b^0(k)$ denote the solution of (2.9b) with zero final condition. Then

$$x_f(k) = A_f^k x_f(0) + x_f^0(k) \quad (2.13a)$$

$$x_b(k) = A_b^{K-k} x_b(K) + x_b^0(k) \quad (2.13b)$$

Substituting (2.13) into (2.11) and solving for $x_f(0)$ and $x_b(K)$ yields

$$\begin{bmatrix} x_f(0) \\ x_b(K) \end{bmatrix} = H^{-1} \{v - v_{f,K} x_f^0(K) - v_{b,0} x_b^0(0)\} \quad (2.14)$$

where

$$H = \begin{bmatrix} v_{f,0} + v_{f,K} A_f^K & v_{b,0} A_b^K + v_{b,K} \end{bmatrix} = v_0^T T^{-1} (FET^{-1})^K + v_K^T T^{-1} (FAT^{-1})^K \quad (2.15)$$

Finally, substituting (2.14) into (2.13) we obtain

$$\begin{bmatrix} x_f(k) \\ x_b(k) \end{bmatrix} = \begin{bmatrix} A_f^k & 0 \\ 0 & A_b^{K-k} \end{bmatrix} H^{-1} \{v - v_{f,K} x_f^0(K) - v_{b,0} x_b^0(0)\} + \begin{bmatrix} x_f^0(k) \\ x_b^0(k) \end{bmatrix} \quad (2.16)$$

The solution in the original basis can then be obtained by inverting (2.8).

Assuming that $\{E, A\}$ is forward-backward stable, the solution procedure is just described consists of stable, forward/backward recursive computations for x_f^0 , x_b^0 , followed by the correction for the actual boundary conditions given by the first term on the right-hand side of (2.16). Note also that this

procedure also provides us with another necessary and sufficient condition for the well-posedness of (2.1), (2.2), namely the invertibility of H in (2.15). This condition is the analog of that described by Krener [12] - [14] for standard boundary-value problems. Note that, as one would expect, not all choices of boundary conditions lead to well-posed problems, and the conditions that V_0 and V_K must satisfy depend heavily on the structure of E and A . For example, as is well known, the initial value problem ($V_0 = I$, $V_K = 0$) is not well-posed if E is singular. This can easily be seen from (2.15) or from (2.5a), since the last block of columns then is not of full rank.

III. The Optimal Smoother

Consider now a stochastic process $x(k)$ satisfying (2.1), (2.2) (which we assume is well-posed) where $u(k)$ and v are independent, zero mean and Gaussian, v has covariance Π_v , and $u(k)$ is a white sequence with covariance Q . In this section we examine the estimation of $x(k)$ given the interior observations

$$y(k) = Cx(k) + r(k), \quad k \in [1, K-1] \quad (3.1)$$

and the boundary measurements

$$y_b = W_0 x(0) + W_K x(K) + r_b \quad (3.2)$$

Here $r(k)$, r_b , $u(\ell)$, and v are mutually independent, r_b is zero mean Gaussian with covariance Π_b , and $r(k)$ is zero mean, Gaussian, and white with covariance R .

In order to derive the optimal smoother, we introduce notation analogous to (2.4), (2.5)

$$y = \mathcal{C}x + r \quad (3.3)$$

where

$$y' = [y'(1), y'(2), \dots, y'(K-1), y_b'] \quad (3.4a)$$

$$r' = [r'(1), r'(2), \dots, r'(K-1), r_b'] \quad (3.4b)$$

$$\mathcal{C} = \begin{bmatrix} 0 & C & 0 \dots \dots \dots 0 & 0 \\ 0 & 0 & C & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & C & 0 \\ W_0 & 0 & 0 \dots \dots \dots 0 & W_K \end{bmatrix} \quad (3.5)$$

Also, the covariances of u in (2.4b) and r in (3.4b) are given by

$$Q = \text{diag}(Q, \dots, Q, \Pi_v) \quad (3.6a)$$

$$\mathcal{R} = \text{diag}(R, \dots, R, \Pi_b) \quad (3.6b)$$

Our problem, then is to estimate x given y , and the approach we adopt is the method of complementary processes introduced in [26] and elaborated upon in [10,11]. Specifically, suppose that we can construct a random vector z that is complementary to y in the sense that (i) it is independent of y and (ii) the transformation from (u,r) to (y,z) is linear and invertible. Then we can write x explicitly as a linear function of y and z , and, thanks to (i) can obtain \hat{x} simply by setting z to zero. In the present context, since x is specified implicitly by (2.3), we also obtain an implicit representation for z . Specifically, as we verify below, z is given by the following

$$\mathcal{S}'\lambda = \mathcal{E}'\mathcal{R}^{-1}r \quad (3.7)$$

$$z = -\mathcal{B}'\lambda + Q^{-1}u \quad (3.8)$$

where

$$\lambda' = [\lambda'(1), \dots, \lambda'(K), \lambda'(0)] \quad (3.8)$$

(the reason for our particular choice of labeling of components in (3.8) will be made clear shortly). Note that (3.7) also has an interpretation as a TPBVDS, but we defer discussion of this until our related discussion of the smoother itself.

As a first step in verifying (3.7), (3.8) note that (3.7) is well-posed since \mathcal{S}' is invertible. Next note that the independence of y and z can be obtained by direct computation:

$$\begin{aligned} E\{yz'\} &= E\{[\mathcal{E}^{-1}\mathcal{B}u + r][-\mathcal{B}'(\mathcal{S}')^{-1}\mathcal{E}'\mathcal{R}^{-1}r + Q^{-1}u]'\} \\ &= \mathcal{E}^{-1}\mathcal{B} - \mathcal{E}^{-1}\mathcal{B} = 0 \end{aligned} \quad (3.9)$$

We next show that we can compute x and λ from y and z . Specifically, using (3.8) to eliminate u and (3.3) to eliminate r , we find that (2.1), (3.7) are equivalent to

$$\begin{bmatrix} \mathcal{P} & -\mathcal{BQ}\mathcal{B}' \\ \mathcal{C}'\mathcal{R}^{-1}\mathcal{C} & \mathcal{P}' \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathcal{BQ}z \\ \mathcal{C}'\mathcal{R}^{-1}y \end{bmatrix} \quad (3.10)$$

The matrix on the left-hand side of (3.10) can be shown to be invertible as follows. Since \mathcal{P} is invertible, we need only show that the Schur complement

$$D = \mathcal{P}' + \mathcal{C}'\mathcal{R}^{-1}\mathcal{C}\mathcal{P}^{-1}\mathcal{BQ}\mathcal{B}' \quad (3.11)$$

is invertible. Note that

$$D(\mathcal{P}')^{-1} = I + ML \quad (3.12)$$

where $M = \mathcal{C}'\mathcal{R}^{-1}\mathcal{C} \geq 0$ and $L = \mathcal{P}^{-1}\mathcal{BQ}\mathcal{B}'(\mathcal{P}')^{-1} \geq 0$. The invertibility of D then follows from the fact that ML cannot have negative eigenvalues.⁶ Finally, once we have recovered x and λ from y and z , u and r can be obtained from (3.8) and (3.3), respectively.

Next, by setting z to zero in (3.10) we obtain the implicit equations defining the optimal smoothed estimate \hat{x} :

$$\begin{bmatrix} \mathcal{P} & -\mathcal{BQ}\mathcal{B}' \\ \mathcal{C}'\mathcal{R}^{-1}\mathcal{C} & \mathcal{P}' \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{C}'\mathcal{R}^{-1}y \end{bmatrix} \quad (3.13)$$

⁶Suppose $MLv = \lambda v$. Then $v'L'MLv = \lambda v'L'v$, so that $\lambda = (v'L'MLv)/(v'L'v) \geq 0$.

This again defines a well-posed TPBVDS, but to obtain the most illuminating form of this system requires a permutation of the equations and variables in (3.13). Specifically, it is straightforward to verify that (3.13) is equivalent to

$$\mathcal{F}\xi = \eta \quad (3.14)$$

where

$$\xi' = [(\hat{x}'(0), \hat{\lambda}'(0)), (x'(1), \lambda'(1)), \dots, (\hat{x}'(K), \hat{\lambda}'(K))] \quad (3.15a)$$

$$\eta' = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ W_0' \Pi_b^{-1} y_b & C'R^{-1}y(1) & \dots & C'R^{-1}y(K-1) & W_K' \Pi_b^{-1} y_b \end{bmatrix} \quad (3.15b)$$

$$\mathcal{F} = \begin{bmatrix} \gamma_{11} & \varepsilon & 0 & \dots & 0 & \dots & 0 & \gamma_{12} \\ 0 & -\mathcal{A} & \varepsilon & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & -\mathcal{A} & \varepsilon & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \dots & -\mathcal{A} & \varepsilon \\ \gamma_{21} & 0 & 0 & 0 & \dots & \dots & 0 & \gamma_{22} \end{bmatrix} \quad (3.16)$$

with

$$\varepsilon = \begin{bmatrix} E & -BQB' \\ 0 & -A' \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} A & 0 \\ -C'R^{-1}C & -E' \end{bmatrix} \quad (3.17)$$

$$\gamma_{11} = \begin{bmatrix} -A & 0 \\ W_0' \Pi_b^{-1} W_0 & V_0' \end{bmatrix}, \quad \gamma_{12} = \begin{bmatrix} 0 & 0 \\ W_0' \Pi_b^{-1} W_K & 0 \end{bmatrix} \quad (3.18a)$$

$$\gamma_{21} = \begin{bmatrix} V_0 & -\Pi_V \\ W_K' \Pi_b^{-1} W_0 & V_K' \end{bmatrix}, \quad \gamma_{22} = \begin{bmatrix} V_K & 0 \\ W_K' \Pi_b^{-1} W_K & E' \end{bmatrix} \quad (3.18b)$$

Comparing the form of \mathcal{F} in (3.16) to that of \mathcal{G} in (2.3), we see that (3.14) is almost a standard TPBVDS except for the top row of equations - i.e. the fact that \mathcal{V}_{11} in (3.16) appears rather than $-\mathcal{A}$ and that \mathcal{V}_{12} is present at all. This is a consequence of the discrete nature of the time index and the intrinsic asymmetry of the model (2.1), (2.2).⁷ We can, however, reduce these equations to a standard TPBVDS by means of a basic technique in the analysis of boundary-value systems [14,25,29,30]. Specifically, we can think of (3.14) as a TPBVDS with boundary values consisting of $(\hat{x}'(0), \hat{\lambda}'(0))'$ and $(\hat{x}'(K), \hat{\lambda}'(K))'$. Because of the well-posedness of (3.14) it is possible to eliminate some of the variables from (3.14) by solving for them in terms of the remaining variables. More specifically, it is possible to move the boundary values inward by eliminating boundary values at one end of the interval, the other, or both. One can iterate this process, and in fact this type of recursion forms the basis for a notion of state for boundary value systems [14,25,29,30]. For our purposes here, however, we need only consider a single step of this type.

⁷Note that $u(k)$ is defined on $[0, K-1]$, while $x(k)$ is defined on $[0, K]$.

Referring to [10], it is not possible in the discrete index case to define the domain on which x and u are defined and the boundary of that domain so that either the boundary is contained in or disjoint from the domain.

Specifically, the invertibility of \mathcal{F} implies that

$$\begin{bmatrix} \gamma_{11} \\ \gamma_{21} \end{bmatrix}$$

has full column rank and thus that we can eliminate $(\hat{x}'(0), \hat{\lambda}(0))'$ as follows.

We construct matrices M_1 and M_2 such that $[M_1, M_2]$ has full row rank and

$$[M_1, M_2] \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \end{bmatrix} = 0 \quad (3.19)$$

If we then premultiply (3.14) by the following full-rank matrix

$$\begin{bmatrix} 0 & I & 0 & \dots & 0 & 0 \\ 0 & 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I & 0 \\ M_1 & 0 & 0 & \dots & 0 & M_2 \end{bmatrix}$$

we obtain a TPBVDS of a form exactly as in (2.1), (2.2). Specifically, this computation yields

$$\mathcal{E} \begin{bmatrix} \hat{x}(k+1) \\ \hat{\lambda}(k+1) \end{bmatrix} = \mathcal{A} \begin{bmatrix} \hat{x}(k) \\ \hat{\lambda}(k) \end{bmatrix} + \begin{bmatrix} 0 \\ C'R^{-1}y(k) \end{bmatrix}, \quad k=1, \dots, K-1 \quad (3.20)$$

with boundary conditions

$$\begin{aligned}
M_1 \begin{bmatrix} \hat{x}(1) \\ \hat{\lambda}(1) \end{bmatrix} + [M_1 \gamma_{12} + M_2 \gamma_{22}] \begin{bmatrix} \hat{x}(K) \\ \hat{\lambda}(K) \end{bmatrix} \\
= M_1 \begin{bmatrix} 0 \\ W_0 \Pi_b^{-1} y_b \end{bmatrix} + M_2 \begin{bmatrix} 0 \\ W_K \Pi_b^{-1} y_b \end{bmatrix}
\end{aligned} \tag{3.21}$$

By construction we know that this system is well-posed. Also, once we have computed $\hat{x}(k)$, $\hat{\lambda}(k)$, $k=1, \dots, K$, we can determine the previously eliminated boundary values $\hat{x}(0)$, $\hat{\lambda}(0)$:

$$\begin{aligned}
\begin{bmatrix} \hat{x}(0) \\ \hat{\lambda}(0) \end{bmatrix} = D \gamma_{11} \begin{bmatrix} 0 \\ W_0 \Pi_b^{-1} y_b \end{bmatrix} + \gamma_{21} \begin{bmatrix} 0 \\ W_K \Pi_b^{-1} y_b \end{bmatrix} \\
-\gamma_{11} \begin{bmatrix} \hat{x}(1) \\ \hat{\lambda}(1) \end{bmatrix} - [-\gamma_{11} \gamma_{12} + \gamma_{21} \gamma_{22}] \begin{bmatrix} \hat{x}(K) \\ \hat{\lambda}(K) \end{bmatrix}
\end{aligned} \tag{3.22a}$$

where

$$D = [\gamma_{11} \gamma_{11} + \gamma_{21} \gamma_{21}]^{-1} \tag{3.22b}$$

As a final comment, we note that on examination of (3.20), (3.21) and the form of \mathcal{E} and \mathcal{A} in (3.17), we see that what we have derived is a generalization of the Hamiltonian form of the optimal smoother for causal systems (see, e.g. [11,22]). This immediately suggests the possibility of

generalizing methods for solving smoothing equations such as diagonalization of the Hamiltonian dynamics [11,22] to produce forward and backward recursions. Such an approach is described in the next section.

IV. Implementation of the Smoother

In this section we discuss several approaches to solving the smoothing TPBVDS (3.20), (3.21). One obvious method of solution is the direct application of the method described in Section 2 for solving general TPBVDS's. The question that then arises is the construction of the similarity transformations that block-diagonalize \mathcal{E} and \mathcal{A} as in (2.6), (2.7). One obvious answer to this is to use the general procedure in [17] for the computation of the Kronecher form of $(\mathcal{E}, \mathcal{A})$. A second is to consider generalizations of Hamiltonian diagonalization procedures, which are developed in the following two subsections. In the first of these we closely parallel the approach used for non-descriptor systems and are led to descriptor Riccati equations and decoupled descriptor dynamics. As we will see, this approach does not always work, and this leads us to an alternative approach in Section 4.2 involving a different type of generalized Riccati equation and producing decoupled non-descriptor dynamics. As we discuss in this and in subsequent sections, this approach points the way to developing the relationship between system - theoretic concepts such as reachability and observability and properties and eigenstructure of the smoother.

4.1 Hamiltonian Diagonalization: Method 1

The general concept of Hamiltonian diagonalization is as follows. We seek two sequences of matrices, $M(k)$ and $N(k)$ so that

$$M(k) \otimes N^{-1}(k+1) = \begin{bmatrix} E_f(k) & 0 \\ 0 & A_b(k) \end{bmatrix} \quad (4.1)$$

and

$$M(k) \otimes N^{-1}(k) = \begin{bmatrix} A_f(k) & 0 \\ 0 & E_b(k) \end{bmatrix} \quad (4.2)$$

In this case the $2n$ -dimensional descriptor dynamics of (3.20) can be decoupled into two n -dimensional descriptor systems (coupled, of course, through the boundary conditions).

The choice of the sequences $M(k)$ and $N(k)$ is far from unique, and the general algebraic equations that the $n \times n$ blocks of $M(k)$ and $N(k)$ must satisfy are presented in [18] and [19]. In this subsection we present one choice that is the direct counterpart of the method used in [11] for non-descriptor continuous-time boundary value processes and that involves descriptor Riccati equations that have appeared elsewhere in the literature. Specifically, suppose that $P(k)$ and $\theta(k)$ are invertible matrix sequences satisfying, respectively, the following forward and backward descriptor Riccati recursions:

$$EP(k+1)E' = A[P^{-1}(k) + C'R^{-1}C]^{-1}A' + BQB' \quad (4.3)$$

$$E'\theta(k)E = A'[\theta^{-1}(k+1) + BQB']^{-1}A + C'R^{-1}C \quad (4.4)$$

In the case of causal systems (with $E = I$), (4.3) is the recursion satisfied by the one-step forward prediction error variance, while (4.4) is the

recursion satisfied by the inverse of the backward filtered error variance.⁸

Also, define

$$Z(k) = E'\theta(k)E + P^{-1}(k) \quad (4.5)$$

In the causal case and with appropriate choices of initial condition for $P(k)$ and final condition for $\theta(k)$, $Z(k)$ is the inverse of the smoothing error variance.

Define

$$M(k) = \left[\begin{array}{c|c} I & A[P^{-1}(k)+C'R^{-1}C]^{-1} \\ \hline Z^{-1}(k)A'[\theta^{-1}(k+1)+BQB']^{-1} & -Z^{-1}(k) \end{array} \right] \quad (4.6)$$

$$N^{-1}(k) = \begin{bmatrix} Z^{-1}(k) & P(k)E' \\ -\theta(k)EZ^{-1}(k) & I \end{bmatrix} \quad (4.7)$$

Some algebraic manipulations verify that $M(k)$ and $N(k)$ are invertible if $P(k)$, $\theta(k)$, and $Z(k)$ are, and if we perform the computations involved in (4.1),

(4.2) and define

$$\begin{bmatrix} \xi(k) \\ \eta(k) \end{bmatrix} \triangleq N(k) \begin{bmatrix} \hat{x}(k) \\ \hat{\lambda}(k) \end{bmatrix} \quad (4.8)$$

⁸The actual quantities $P(k)$ and $\theta^{-1}(k)$ have these interpretations only if the initial and final conditions $P(0)$ and $\theta(K)$ are appropriately chosen. In this case $[P^{-1}(k) + C'R^{-1}C]^{-1}$ is the forward filtered error covariance, while $\theta^{-1}(K) + BQB'$ is the one-step backward prediction error covariance.

the smoother dynamics (3.20) decouple into

$$EP(k+1)\xi(k+1) = A[P^{-1}(k)+C'R^{-1}C]^{-1} [\xi(k)+C'R^{-1}y(k)] \quad (4.9)$$

$$P(k)E'\eta(k) = [P^{-1}(k)+C'R^{-1}C]^{-1}A'\eta(k+1)+Z^{-1}(k)C'R^{-1}y(k) \quad (4.10)$$

The boundary conditions in the transformed coordinates can be determined from (3.21), (3.10), and (3.12).

As a simpler alternative one might consider constant transformations M and N as in (4.6) and (4.7), but using solutions to the steady-state descriptor Riccati equations

$$EPE' = A[P^{-1}+C'R^{-1}C]^{-1}A' + BQB' \quad (4.11)$$

$$E' \theta E = A'[\theta^{-1} + BQB']^{-1}A + C'R^{-1}C \quad (4.12)$$

Note that in this case the transformed smoother dynamics

$$EP\xi(k+1) = A[P^{-1}+C'R^{-1}C]^{-1} [\xi(k)+C'R^{-1}y(k)] \quad (4.13a)$$

$$PE'\eta(k) = [P^{-1}+C'R^{-1}C]^{-1}A'\eta(k+1)+Z^{-1}C'R^{-1}y(k) \quad (4.13b)$$

involves two pencils $\{E_1, A_1\} \stackrel{\Delta}{=} \{EP, A[P^{-1}+C'R^{-1}C]^{-1}\}$ and $\{E_2, A_2\} \stackrel{\Delta}{=} \{PE', [P^{-1}+C'R^{-1}C]^{-1}A'\}$ that are transposes of one another. In this case if we follow the solution procedure outlined in Section 2.1, if the matrices F_1 and T_1 transform $\{E_1, A_1\}$ into the form shown in (2.6), (2.7), then $F_2 = T_1'$ and $T_2 = F_1'$ do the same for $\{E_2, A_2\}$.

The descriptor Riccati equations we have introduced have appeared in the literature. In the case in which E is nonsingular, which was studied by Laub in [24], it is clear that these are no difficulties in solving (4.3), (4.4) or equivalent versions not involving inversions of P and θ) nor in obtaining controllability and observability conditions under which (4.11), (4.12) have unique positive definite solutions. Furthermore in this case it is also possible to parallel the approach in [11] (for the non-descriptor case) in

choosing boundary conditions $P(0)$ and $\theta(K)$ for (4.3), (4.4) so that the boundary conditions associated with (4.10) are minimally coupled. Similarly, in the case in which A is invertible, we can do something analogous, leading to a pair of dual Riccati equations, essentially by reversing time ($k \rightarrow K-k$) thereby interchanging the roles of A and E . While the approach outlined in this section (or its dual) works when either A or E is invertible,⁹ the difficulty arises when both E and A are singular. As pointed out by Bender, singularity can cause equations such as (4.3) to fail to have solutions for particular initial conditions. Also, as we illustrate through an example in the next section, when E and A are both singular (4.11), (4.12) have solutions only in an uninteresting case. What is therefore required is a different approach. Previous studies of control problems for continuous or discrete descriptor systems [9], [21], [23] have circumvented this difficulty by deriving and dealing with lower-order standard Riccati equations (of dimension equal to the rank of E). In our case, however, we are interested in diagonalizing the Hamiltonian dynamics. As we develop in the next section, this is possible if we introduce equations that are not quite standard Riccati equations but are far closer to them than (4.11), (4.12).

⁹Note that \mathcal{E} and \mathcal{A} are both singular if either E or A is, so that the procedure in this section does work on a class of nontrivial Hamiltonian descriptor dynamics. See [27], [28] for investigations of discrete-time algebraic Riccati equations by examination of the pencil defined by \mathcal{E} and \mathcal{A} when $E = I$.

4.2 Hamiltonian Diagonalization: Method 2

In this subsection we focus completely on time-invariant versions of the transformations (4.1), (4.2). The key to the transformations are the generalized Riccati equations

$$\theta = A'(E\theta^{-1}E' + BQB')^{-1}A + C'R^{-1}C \quad (4.14)$$

$$\psi = A(E'\psi^{-1}E + C'R^{-1}C)^{-1}A' + BQB' \quad (4.15)$$

Note that these equations are "almost" standard Riccati equations, except for the presence of E and E' multiplying θ^{-1} and ψ^{-1} in the terms in parentheses. While there appears to be some asymmetry in the roles played by E and A, this is an illusion, as can be seen by introducing an additional pair of matrices. Specifically, if we define

$$S = E\theta^{-1}E' + BQB' \quad (4.16a)$$

we see that

$$\theta = A'S^{-1}A + C'R^{-1}C \quad (4.16b)$$

Similarly, by introducing

$$T = E'\psi^{-1}E + C'R^{-1}C \quad (4.17a)$$

we obtain

$$\psi = AT^{-1}A' + BQB' \quad (4.17b)$$

Consequently, we can view (4.16) and (4.17) individually as pairs of equations to be solving for (S, θ) and (T, ψ), respectively. We assume throughout this section that positive definite solutions for these four quantities exist. As in the previous section, if either E or A is invertible, we can reduce these equations to standard Riccati equations and therefore can obtain the usual type of reachability and observability conditions for existence of such solutions. Also, as we illustrate in the next section these equations admit

positive definite solutions even in cases in which both E and A are singular. General conditions for existence and uniqueness of positive definite solutions and relationships with the notions of reachability and observability for TPBVDS's will be described in a future paper [33] (See Section VII and also [32] for a sketch of the basic results).

Consider next the matrices

$$M = \begin{bmatrix} I & AT^{-1} \\ A'S^{-1} & -I \end{bmatrix} \quad (4.18a)$$

$$N = \begin{bmatrix} E & -\Psi \\ \Theta & E' \end{bmatrix} \quad (4.18b)$$

The invertibility of N is immediate from the invertibility of Θ and the invertibility of the Schur complement

$$-\Psi - E\Theta^{-1}E'$$

Similarly the invertibility of M follows from the invertibility of $-I$ and of the Schur complement

$$I + A'S^{-1}AT^{-1}$$

(which is invertible since $T > 0$, $A'S^{-1}A \geq 0$ so that the eigenvalues of $A'S^{-1}AT^{-1}$ are nonnegative).

It is a straightforward exercise, using (4.16), (4.17) to show that

$$M \&N^{-1} = \begin{bmatrix} I & 0 \\ 0 & A'S^{-1}E\Theta^{-1} \end{bmatrix} \quad (4.19a)$$

$$M \&N^{-1} = \begin{bmatrix} AT^{-1}E'\Psi^{-1} & 0 \\ 0 & I \end{bmatrix} \quad (4.19b)$$

Therefore, if we premultiply (3.20) by M and make the change of coordinates

$$\begin{bmatrix} \hat{\delta}(k) \\ \hat{\gamma}(k) \end{bmatrix} = N \begin{bmatrix} \hat{x}(k) \\ \hat{\lambda}(k) \end{bmatrix} \quad (4.20)$$

the smoother dynamics are transformed into standard non-descriptor recursions:

$$\hat{\delta}(k+1) = AT^{-1}E'\Psi^{-1}\hat{\delta}(k) + AT^{-1}C'R^{-1}y(k) \quad (4.21a)$$

$$\hat{\gamma}(k) = A'S^{-1}E\Theta^{-1}\hat{\gamma}(k+1) + C'R^{-1}y(k) \quad (4.21b)$$

with boundary conditions

$$\begin{aligned} & M_1 \varepsilon N^{-1} \begin{bmatrix} \hat{\delta}(1) \\ \hat{\gamma}(1) \end{bmatrix} + [M_1 \gamma_{12} + M_2 \gamma_{22}] N^{-1} \begin{bmatrix} \hat{\delta}(K) \\ \hat{\gamma}(K) \end{bmatrix} \\ &= M_1 \begin{bmatrix} 0 \\ W_0' \Pi_b^{-1} y_b \end{bmatrix} + M_2 \begin{bmatrix} 0 \\ W_K' \Pi_b^{-1} y_b \end{bmatrix} \end{aligned} \quad (4.22)$$

Note that (4.21) consists of a forward recursion (a) and a reverse recursion (b), with coupled boundary conditions (4.22). The approach outlined in Section II (see (2.13) - (2.16)) can then be used directly to obtain the solution. Once this is accomplished, we can recover $\hat{x}(k)$ and $\hat{\lambda}(k)$, $k=1, \dots, K$ by inverting (4.20), i.e. from the relationship

$$\hat{x}(k) = [\Theta + E'\Psi^{-1}E]^{-1}[\hat{\gamma}(k) + E'\Psi^{-1}\hat{\delta}(k)] \quad (4.23a)$$

$$\hat{\lambda}(k) = \Psi^{-1}E[\Theta + E'\Psi^{-1}E]^{-1}\hat{\gamma}(k) - [\Psi + E\Theta^{-1}E']^{-1}\hat{\delta}(k) \quad (4.23b)$$

and then can recover $\hat{x}(0)$, $\hat{\lambda}(0)$ from (3.22). Note that since one is generally interested only in \hat{x} , it is only necessary to compute $\hat{\lambda}(1)$ and $\hat{\lambda}(K)$ in (4.23b) in order to be able to determine $\hat{x}(0)$ from (3.22).

V. Examples

In this section we first present an example illustrating our smoothing results for TPBVDS's and then introduce the class of cyclic systems in a second example.

Example 5.1: As we indicated in the previous section, the case in which either E or A is invertible can be thought of as a slight generalization of the causal case (perhaps with time reversal), and consequently both of the Riccati-like methods of the previous section (or the dual of the method of Section 4.1) work without difficulty. In this example, we look at a system for which both E and A are singular and first illustrate the problems with the method of Section 4.1 and the apparent superiority of the approach in Section 4.2.

Consider the descriptor system with

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

In this case it is not difficult to check that difficulties arise in solving the time-varying descriptor Riccati equations (4.3), (4.4) or their time-invariant counterparts (4.11), (4.12). For example, let

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix}, \quad [P^{-1} + C'R^{-1}C]^{-1} = \begin{bmatrix} U_{11} & U_{12} \\ U_{12} & U_{22} \end{bmatrix}$$

and consider (4.14) which in this case reduces to

$$\begin{bmatrix} P_{11} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & U_{22} \end{bmatrix} + BQB'$$

which is obviously inconsistent with a positive definite solution for P . Even if one considers indefinite solutions, we see that none can possibly exist if BQB' is not diagonal. Indeed the only case in which any solutions exist to (4.11), (4.12) is when BQB' 's and $C'R^{-1}C$ are both diagonal. In this case P and θ are also diagonal, with the positive diagonal element corresponding to the error covariance of the causal part of the system (the first state component) and the negative element to the negative of the error covariance of the anticausal (second state) component. Furthermore, the diagonal nature of BQB' and $C'R^{-1}C$ implies that independent noises drive each component and independent observations are available for each -- i.e. the problem reduces to the trivial and uninteresting case of two completely decoupled systems.

On the other hand, the generalized Riccati equations (4.14), (4.15) admit solutions in nontrivial cases. For example, if

$$B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad Q = 1, \quad C = R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

the solutions to (4.14), (4.15) are

$$\theta = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \quad \psi = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

and

$$S = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

This example also illustrates the degeneracy that arises in the dynamic portion of the smoother for TPBVDS's whenever either E or A is singular.

Indeed in this case (4.21) reduces to

$$\hat{\delta}(k+1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} y(k) \quad (5.1)$$

$$\hat{\gamma}(k) = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \hat{\gamma}(k+1) + y(k) \quad (5.2)$$

This is of course an extreme example, since the two components, x_1 and x_2 , of x are essentially identical white noise sequences (with a sign inversion and a one unit relative time shift) except for the possible correlation between $x(0)$ and $x(K)$ introduced by the boundary conditions. However, while in general the system matrices in (4.21) will not be nilpotent as they are here, there will always be some rank deficiency if either A or E is singular.

Finally, let us illustrate the rest of the smoothing solution for this example. Even in this degenerate case the one time-step delay between x_1 and x_2 and the nature of the boundary conditions can lead to a nontrivial form for the smoother. In particular, suppose that

$$V_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad V_K = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Pi_v = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (5.3)$$

$$W_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad W_K = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Pi_b = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (5.4)$$

The dynamics plus boundary conditions in this case are

$$x_1(k+1) = u(k)$$

$$k = 1, \dots, K - 1$$

$$x_2(k) = -u(k)$$

with $x_2(0)$ a unit variance random variable independent of u , and with

$$x_2(K) = x_1(0) + u(0)$$

Referring to (3.18), we have

$$\Psi_{11} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad \Psi_{12} = 0$$

$$\Psi_{21} = \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Psi_{22} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

We can then compute M_1 and M_2 satisfying (3.19):

$$M_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 3 & 1 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 2 & 0 & 0 \end{bmatrix}$$

The boundary conditions (4.22) for (5.1), (5.2) then are

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{2}{3} & \frac{5}{3} & -\frac{2}{3} & 0 \end{bmatrix} \begin{bmatrix} \hat{\delta}(1) \\ \hat{\gamma}(1) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{3} \\ 0 & 0 & 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \hat{\delta}(K) \\ \hat{\gamma}(K) \end{bmatrix} = \alpha \quad (5.5)$$

where

$$\alpha = M_1 \begin{bmatrix} 0 \\ W_0' \Pi_b^{-1} y_b \end{bmatrix} + M_2 \begin{bmatrix} 0 \\ W_K' \Pi_b^{-1} y_b \end{bmatrix}$$

Then applying (2.16) to (5.1), (5.2), (5.5) (with an adjustment for the fact that the smoother (5.1), (5.2) runs from 1, rather than 0, to K), we find that

$$\hat{\delta}(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{2}{11} & 0 & \frac{6}{11} & \frac{3}{11} \end{bmatrix} \alpha + \begin{bmatrix} 1 & 0 \\ -\frac{2}{11} & 0 \end{bmatrix} \hat{\gamma}(1)$$

$$\hat{\delta}(k) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} y(k-1), \quad 2 \leq k \leq K$$

$$\gamma(k) = y(k) + \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} y(k-1), \quad 1 \leq k \leq K-2$$

$$\hat{\gamma}(K-1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \alpha + y(K-1)$$

$$\hat{\gamma}(K) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{6}{11} & 0 & \frac{15}{11} & \frac{9}{11} \end{bmatrix} \alpha + \begin{bmatrix} 0 & 0 \\ \frac{6}{11} & 0 \end{bmatrix} \hat{\gamma}(1)$$

Finally, using (4.23) we compute

$$\hat{x}(k) = \frac{1}{3} \left\{ \hat{\gamma}(k) + \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} \hat{\delta}(k) \right\} \quad k=1, \dots, K$$

$$\hat{\lambda}(k) = \begin{bmatrix} \frac{2}{3} & 0 \\ -\frac{1}{3} & 0 \end{bmatrix} \hat{\gamma}(k) - \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \hat{\delta}(k)$$

and, from (3.22)

$$\hat{x}(0) = \begin{bmatrix} \frac{5}{8} & \frac{1}{16} & \frac{3}{16} & \frac{1}{4} \\ \frac{1}{16} & \frac{13}{32} & \frac{7}{32} & -\frac{1}{8} \end{bmatrix} \beta - \begin{bmatrix} \frac{1}{16} & -\frac{1}{4} \\ \frac{13}{32} & \frac{1}{8} \end{bmatrix} \hat{\lambda}(1) - \begin{bmatrix} 0 & \frac{1}{8} \\ 0 & \frac{3}{32} \end{bmatrix} \hat{x}(K)$$

where

$$\beta = \mathcal{V}_{11}' \begin{bmatrix} 0 \\ W_0' \Pi_b^{-1} y_b \end{bmatrix} + \mathcal{V}_{21}' \begin{bmatrix} 0 \\ W_K' \Pi_b^{-1} y_b \end{bmatrix}$$

Example 5.2: In this example we introduce the class of cyclic TPBVDS's for which the boundary condition (2.2) takes the special form

$$x(0) = x(K)$$

Equivalently we can think of a cyclic system as being defined on $[0, K-1]$ with the boundary condition

$$Ex(0) - Ax(K-1) = Bu(K-1)$$

(so that \mathcal{F} in (2.5a) is block-circulant).

Consider the smoothing problem for such a system when the boundary measurements are

$$y_b = Cx(0) + r_b$$

with $\Pi_b = R$. It is not difficult to check that in this case \mathcal{F} is also block-circulant (i.e. $\mathcal{V}_{11} = \mathcal{V}_{22} = -A$, $\mathcal{V}_{12} = 0$, $\mathcal{V}_{21} = \mathcal{E}$) so that the smoother is also a cyclic TPBVDS over $[0, K-1]$ (with no need to move the boundary in one step as in (3.19) - (3.22)). If we then follow the procedure described in Section 4.2, we obtain two non-descriptor cyclic systems

$$\hat{\delta}(k+1) = F_{\delta} \hat{\delta}(k) + G_{\delta} y(k) \quad , \quad \hat{\delta}(0) = \hat{\delta}(K) \quad (5.6)$$

$$\hat{\gamma}(k) = F_{\gamma} \hat{\gamma}(k+1) + G_{\gamma} y(k) \quad , \quad \hat{\gamma}(0) = \hat{\gamma}(K) \quad (5.7)$$

where the F's and G's are specified in (4.21) and we have adopted the notation $y(0) = y(K) = y_b$.

Obviously the symmetry of the cyclic case leads to some simplifications. In fact, note that the two systems (5.6), (5.7), including boundary conditions, are completely decoupled. This greatly simplifies their solution, which we can write as cyclic convolutions:

$$\hat{\delta}(k) = [I - F_{\delta}^K]^{-1} \sum_{\ell=0}^{K-1} F_{\delta}^{\ell} G_{\delta} y(k-\ell-1) \quad (5.8)$$

$k=0, 1, \dots, K-1$

$$\hat{\gamma}(k) = [I - F_{\gamma}^K]^{-1} \sum_{\ell=0}^{K-1} F_{\gamma}^{\ell} G_{\gamma} y(k+\ell) \quad (5.9)$$

where we extend $y(k)$ periodically (i.e. $y(k+K) = y(k)$). The estimate $\hat{x}(k)$ can then be computed from (4.23a) again without any need to determine $\hat{x}(0)$ separately since we did not need to move the smoother boundary.

VI. The Smoothing Error for TPBVDS's

Recall from the development in Section III that we obtained the form of the optimal smoother by expressing x and λ in terms of y and z as in (3.10) and then setting z to zero. Thanks to the orthogonality of z and y we can similarly obtain an expression for the smoothing error by setting z to zero in (3.10):

$$\begin{bmatrix} \mathcal{Y} & -\mathcal{BQ}\mathcal{B}' \\ \mathcal{C}'\mathcal{R}^{-1}\mathcal{C} & \mathcal{Y}' \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{\lambda} \end{bmatrix} = \begin{bmatrix} \mathcal{BQ}z \\ 0 \end{bmatrix} \quad (6.1)$$

where $\tilde{x} = x - \hat{x}$, $\tilde{\lambda} = \lambda - \hat{\lambda}$. If we then use these relationships, together with (3.7), (3.8) we obtain

$$\begin{bmatrix} \mathcal{Y} & -\mathcal{BQ}\mathcal{B}' \\ \mathcal{C}'\mathcal{R}^{-1}\mathcal{C} & \mathcal{Y}' \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \mathcal{B} & 0 \\ 0 & -\mathcal{C}'\mathcal{R}^{-1} \end{bmatrix} \begin{bmatrix} u \\ r \end{bmatrix} \quad (6.2)$$

As in Section III, this is equivalent to

$$\mathcal{E} \begin{bmatrix} \tilde{x}(k+1) \\ -\hat{\lambda}(k+1) \end{bmatrix} = \mathcal{A} \begin{bmatrix} \tilde{x}(k) \\ -\hat{\lambda}(k) \end{bmatrix} + \begin{bmatrix} \mathcal{B} & 0 \\ 0 & \mathcal{C}'\mathcal{R}^{-1} \end{bmatrix} \begin{bmatrix} u(k) \\ r(k) \end{bmatrix} \quad (6.3)$$

$k=1, \dots, N-1$

with boundary conditions

$$\begin{aligned}
& M_1 \varepsilon \begin{bmatrix} \tilde{\mathbf{x}}(1) \\ -\hat{\lambda}(1) \end{bmatrix} + [M_1 \gamma_{12} + M_2 \gamma_{22}] \begin{bmatrix} \tilde{\mathbf{x}}(N) \\ -\hat{\lambda}(N) \end{bmatrix} \\
& = M_1 \begin{bmatrix} \text{Bu}(0) \\ W_0 \Pi_b^{-1} r_b \end{bmatrix} + M_2 \begin{bmatrix} v \\ W_K \Pi_b^{-1} r_b \end{bmatrix} \tag{6.4}
\end{aligned}$$

Examining (6.3), (6.4), we see that the evaluation of the covariance of the estimation error $\tilde{\mathbf{x}}(k)$ corresponds to the computation of (the upper left-hand block of) the covariance of the TPBVDS (6.3), (6.4) driven by white noise $(u'(k), r'(k))$ and with independent boundary conditions. In the Appendix we describe one method for performing this computation for the original TPBVDS introduced in Section II. This calculation is somewhat more complicated than the corresponding one for causal systems since $\mathbf{x}(k)$ in (2.1) is not Markov and in fact is not independent of future values of $u(k)$. We refer the reader to [25,31] for more on the properties and calculation of the covariance and correlation function of such processes.

We close this section with two final observations. First, note that the computation described in the Appendix, when applied to (6.3), (6.4) yields the covariance of $\tilde{\mathbf{x}}(k)$ for $k \geq 1$. In order to compute the covariance of $\tilde{\mathbf{x}}(0)$, we need to examine the counterpart to (3.22):

$$\begin{aligned}
& \begin{bmatrix} \tilde{\mathbf{x}}(0) \\ -\hat{\lambda}(0) \end{bmatrix} = D \gamma_{11} \begin{bmatrix} \text{Bu}(0) \\ W_0 \Pi_b^{-1} y_b \end{bmatrix} + \gamma_{21} \begin{bmatrix} v \\ W_K \Pi_b^{-1} y_b \end{bmatrix} \\
& - \gamma_{11} \varepsilon \begin{bmatrix} \tilde{\mathbf{x}}(1) \\ -\hat{\lambda}(1) \end{bmatrix} - [\gamma_{11} \gamma_{12} + \gamma_{21} \gamma_{22}] \begin{bmatrix} \tilde{\mathbf{x}}(K) \\ -\hat{\lambda}(K) \end{bmatrix} \tag{6.5}
\end{aligned}$$

The calculation of the covariance of the left-hand side of (6.5) then involves the computation of the covariances of and the correlations among the various random vectors appearing on the right-hand side of (6.5). An analogous computation is also carried out in the Appendix.

The second point concerns the diagonalization of (6.2). In particular, assuming that positive definite solutions exist to (4.11), (4.12), we can perform analogous steps to those used in Section 4.2 to transform (6.2) into the non-descriptor, forward and backward pair of equations

$$\tilde{\gamma}(k) = A'S^{-1}E\Theta^{-1}\tilde{\gamma}(k+1) - A'S^{-1}Bu(k) + C'R^{-1}r(k) \quad (6.6a)$$

$$\tilde{\delta}(k+1) = AT^{-1}E'\Psi^{-1}\tilde{\delta}(k) + Bu(k) + AT^{-1}\Theta'R^{-1}r(k) \quad (6.6b)$$

(with corresponding, and generally coupled, boundary conditions) with $\tilde{x}(k)$ and $-\hat{\lambda}(k)$ then obtained from (4.23a, b), respectively, with $\hat{\gamma}$ and $\hat{\delta}$ replaced by $\tilde{\gamma}$ and $\tilde{\delta}$.

Equation (6.6) is extremely useful. In the first place, it provides the forward-backward decomposition needed in the covariance analysis procedure described in the Appendix. More importantly, it provides the basis for a system-theoretic investigation of the smoother, several elements of which are briefly described in the next section.

VII. Discussion

In this paper we have investigated the optimal estimation problem for two-point boundary-value descriptor systems (TPBVDS's). Using the method of complementary processes we developed a generalization of the Hamiltonian form of the optimal smoother for causal systems. This generalized Hamiltonian system is itself a TPBVDS. In addition, we have generalized the notion of Hamiltonian diagonalization as a method for reducing the smoother to two systems of lower order. Both of the approaches described involve generalizations of standard Riccati equations. One of these, corresponding to descriptor Riccati equations that have appeared in the literature, is shown to work only in certain cases and is not appropriate when the system dynamics are intrinsically acausal, i.e. when both system matrices E and A are singular.

However, our second approach, involving what we call generalized Riccati equations, appears to offer much promise. Indeed we have illustrated that it does provide a viable approach in the acausal case. Furthermore, in a forthcoming paper [33] we present results that make clear many of the system-theoretic properties of these new Riccati-like equations and the associated generalized Hamiltonian system. In particular in [25,29,30] we introduce and analyze the properties of weak and strong reachability for TPBVDS's, and in [33] we show that strong reachability and observability guarantee the existence of unique positive definite solutions to (4.14), (4.15) and the forward-backward stability of the smoother. Also, in [31] we investigate a more meaningful notion of stability for TPBVDS's as well as the existence of stationary statistics for such systems. In [33] we show that

strong reachability and observability guarantee that the smoother is stable in this sense and we also use the machinery of [31] to relate the solutions of (4.14), (4.15) to the asymptotic error covariance. Also, as one might expect, additional results can be obtained relating smoother stability to notions of detectability and stabilizability, and a generalized Hamiltonian eigenvector approach can be developed to solve (4.14), (4.15) in a manner analogous to the popular method for standard Riccati equations [27,28]

Appendix: Covariance Analysis for TPBVDS's

In this appendix we develop formulas for covariance analysis of TPBVDS's. As a starting point for this computation, we assume that our TPBVDS has been placed in the forward-backward form given in (2.9), (2.11). The general solution for this system is given in (2.16). Given the independence of the boundary value v and the white sequence $u(k)$, we see that the covariance of $x(k)$ can be expressed in terms of the covariance, Π_v , of v and the three quantities

$$P_f^0(k) = E[x_f^0(k)x_f^0(k)'] \quad (B.1a)$$

$$P_b^0(k) = E[x_b^0(k)x_b^0(k)'] \quad (B.1b)$$

$$P_{fb}^0(n,k) = E[x_f^0(n)x_b^0(k)'] \quad (B.1c)$$

The computations of these quantities are straightforward:

$$P_f^0(k+1) = A_f P_f^0(k) A_f' + B_f Q B_f' \quad P_f^0(0) = 0 \quad (B.2a)$$

$$P_b^0(k-1) = A_b P_b^0(k) A_b' + B_b Q B_b' \quad P_b^0(K) = 0 \quad (B.2b)$$

and

$$P_{fb}^0(n,k) = \begin{cases} 0 & , n \leq k \\ \Pi_{fb}^0(n) (A_b')^{n-k} - A_f^{n-k} \Pi_{fb}^0(k) & , n > k \end{cases} \quad (B.2c)$$

where

$$\Pi_{fb}^0(k+1) = A_f \Pi_{fb}^0(k) A_b' + B_f Q B_b' \quad \Pi_{fb}^0(0) = 0 \quad (B.2d)$$

Given these quantities, we can now determine an expression for

$$\Sigma(k) = E \left\{ \begin{bmatrix} x_f(k) \\ x_b(k) \end{bmatrix} [x_f'(k), x_b'(k)] \right\} \quad (B.3)$$

$$\Sigma(k) = G(k)\Pi_v G'(k) + G(k)\Psi(k) + \Psi'(k)G'(k) + G(k)\Delta G'(k) + \begin{bmatrix} P_f^0 & 0 \\ 0 & P_b^0(k) \end{bmatrix} \quad (\text{B.4})$$

where

$$G(k) = \begin{bmatrix} A_f^k & 0 \\ 0 & A_b^{K-k} \end{bmatrix} H^{-1} \quad (\text{B.5})$$

$$\Psi(k) = -V_{f,K} [A_f^{K-k} P_f^0(k) : P_{fb}^0(K,k)] + V_{b,0} [P_{fb}^0(k,0) : A_b^k P_b^0(k)] \quad (\text{B.6})$$

$$\Delta = [V_{f,K} : V_{b,0}] \begin{bmatrix} P_f^0(K) & P_{fb}^0(K,0) \\ P_{fb}^0(K,0) & P_b^0(0) \end{bmatrix} \begin{bmatrix} V_{f,K} \\ V_{b,0} \end{bmatrix} \quad (\text{B.7})$$

As mentioned in Section 6, the computation of the error covariance at the initial point in the interval of interest involves an additional computation. In the remainder of this Appendix we describe the corresponding calculation for (2.9), (2.11). Specifically, suppose we would like to compute the

covariance of

$$\eta = N_1 \xi + N_2 \begin{bmatrix} x_f(0) \\ x_b(0) \end{bmatrix} + N_3 \begin{bmatrix} x_f(K) \\ x_b(K) \end{bmatrix} \quad (\text{B.8})$$

where ξ is a zero-mean random vector correlated with the boundary condition v but independent of u . Let

$$E[\xi\xi'] = P_\xi \quad E[\xi v'] = P_{\xi v} \quad E[\eta\eta'] = P_\eta \quad (\text{B.9})$$

Then, with the help of (2.16) we have

$$\begin{aligned} P_\eta &= N_1 P_\xi N_1' + N_2 \begin{bmatrix} P_f(0) & 0 \\ 0 & P_b(0) \end{bmatrix} N_2' + N_3 \begin{bmatrix} P_f(K) & 0 \\ 0 & P_b(K) \end{bmatrix} N_3' \\ &+ N_1 P_{\xi v} G'(0) N_2' + N_2 G(0) P_{\xi v}' N_1' + N_1 P_{\xi v} G'(K) N_3' \\ &+ N_3 G(K) P_{\xi v}' N_1' + N_2 P'(K,0) N_3' + N_3 P(K,0) N_2' \end{aligned} \quad (\text{B.10})$$

where

$$P(n,k) = E \left\{ \begin{bmatrix} x_f(n) \\ x_b(n) \end{bmatrix} [x_f'(k) \quad x_b'(k)] \right\} \quad , \quad n \geq k \quad (\text{B.11})$$

can be calculated in the same manner as $\Sigma(k)$:

$$P(n,k) = G(n) \Pi_V G'(k) + G(n) \Psi(k) + \Psi'(n) G'(k)$$

$$+ G(n)\Delta G'(k) + \begin{bmatrix} A_f^{n-k} P_f^0(k) & 0 \\ 0 & P_b^0(n)(A_b')^{n-k} \end{bmatrix} \quad (\text{B.12})$$

References

1. D.G. Luenberger, "Dynamic Equations in Descriptor Form," IEEE Trans. Auto. Contr., Vol. AC-22, June 1977, pp. 312-321.
2. D.G. Luenberger, "Time-Invariant Descriptor Systems," Automatica, Vol. 14, 1978, pp. 473-480.
3. G.C. Verghese, B.C. Levy, and T. Kailath, "A Generalized State Space for Singular Systems," IEEE Trans. Auto. Contr., Vol. AC-26, Aug. 1981, pp. 811-831.
4. F.L. Lewis, "Descriptor Systems: Expanded Descriptor Equation and Markov Parameters," IEEE Trans. Auto. Contr., Vol. AC-28, No. 5, 1983, pp. 623-627.
5. D. Cobb, "Descriptor Variable Systems and Optimal State Regulation," IEEE Trans. Auto. Contr., Vol. 28, No. 5, 1983, pp. 601-611.
6. F.L. Lewis, "Adjoint Matrix, Cayley-Hamilton Theorem, and Fadeev's Method for the Matrix Pencil (sE-A)," Proc. 22nd IEEE Conf. Dec. and Contr., San Antonio, TX, 1983.
7. F.L. Lewis, "Descriptor Systems: Fundamental Matrix, Reachability and Observability Matrices, Subspaces," Proc. 23rd IEEE Conf. Dec. and Contr., Las Vegas, NV, 1984.
8. F.L. Lewis, "Descriptor Systems: Decomposition into Forward and Backward Subsystems," IEEE Trans. Auto. Contr., Vol. AC-29, No. 2, 1984, pp. 167-170.
9. D.J. Bender, "Singular Riccati Equations for Singular Systems," Proc. Symp. on Math. Theory of Networks and Systems, Stockholm, June 1985.
10. M.B. Adams, A.S. Willsky, and B.C. Levy, "Linear Estimation of Boundary-Value Stochastic Processes, Part I: The Role and Construction of Complementary Models," IEEE Trans. Auto. Contr., Vol. AC-29, No. 9, Sept. 1984, pp. 803-810.
11. 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. Auto. Contr., Vol. AC-29, No. 9, Sept. 1984, pp. 811-821.
12. A.J. Krener, "Boundary Value Linear Systems," Asterisque, Vol. 75-76, 1980, pp. 149-165.
13. A.J. Krener, "Acausal Linear Systems," Proc. 18th IEEE Conf. Decision and Contr., Ft. Lauderdale, Fl., 1981.

14. A.J. Krener, "Acausal Realization Theory, Part I: Linear Deterministic Systems," preprint.
15. M.B. Adams, B.C. Levy, and A.S. Willsky, "Optimal Smoothing for Nearest-Neighbor 2D Models with Boundary Conditions," in preparation.
16. J. Wall, A.S. Willsky, and N.R. Sandell, Jr., "On the Fixed-Interval Smoothing Problem," Stochastics, Vol. 5, 1981, pp. 1-41.
17. P. Van Dooren, "The Computation of Kronecker's Canonical Form of a Singular Pencil," Linear Algebra and Its Applications, Vol. 23, 1979, pp. 103-140.
18. M.B. Adams, "Linear Estimation of Boundary Value Stochastic Processes," Sc.D. Thesis, Dept. of Aero. and Astro., M.I.T., Cambridge, MA, Jan. 1983.
19. M.B. Adams, B.C. Levy, and A.S. Willsky, "Linear Smoothing for Descriptor Systems," Proc. 23rd Conf. on Dec. and Contr., Las Vegas, NV, Dec. 1984.
20. A.J.J. Van Der Weiden and O.H. Bosgra, "The Determination of Structural Properties of a Linear System by Operations of System Similarity; 2. Non-Proper Systems in Generalized State-Space Form," Int. J. Control, Vol. 32, No. 3, 1980, pp. 489-537.
21. D. Cobb, "Feedback and Pole Placement in Descriptor Variable Systems," Int. J. Control, Vol. 33, No. 6, 1981, pp. 1135-1146.
22. T. Kailath and L. Ljung, "Two-Filter Smoothing Formulas by Diagonalization to the Hamiltonian Equations," Int. J. Control.
23. R.E. Larson, D.G. Luenberger, and D.N. Stengel, "Descriptor Variable Theory and Spatial Dynamic Programming," Systems Control, Inc. Report No. TM 5168-9, Feb. 1978.
24. A.J. Laub, "Schur Techniques in Invariant Imbedding Methods for Solving Two-Point Boundary Value Problems," Proc. 21st IEEE Conf. on Decision and Control, Orlando, Fla., Dec. 1982.

25. R. Nikoukhah, "System Theory for Two-Point Boundary Value Descriptor Systems," S.M. thesis, Dept. of Elec. Eng. and Comp. Sci., MIT, and Lab. for Infor. and Dec. Sys. Rept. LIDS-TH-1559, June 1986.
26. H.L. Weinert and U.B. Desai, "On Complementary Models and Fixed-Interval Smoothing," IEEE Trans. Aut. Control, Vol. AC-26, Aug. 1981.
27. T. Pappas, A.J. Laub, and N.R. Sandell, Jr., "On the Numerical Solution of the Discrete-Time Riccati Equation," IEEE Trans. Aut. Control, Vol. AC-25, 1980, pp. 631-641.
28. C.E. DeSouza, M.R. Gevers, and G.C. Goodwin, "Riccati Equations in Optimal Filtering of Nonstabilizable Systems Having Singular State Transition Matrices," IEEE Trans. Aut. Control, Vol. AC-31, Sept. 1986, pp. 831-838.
29. R. Nikoukhah, A.S. Willsky, and B.C. Levy, "Boundary-Value Descriptor Systems: Well-Posedness, Reachability and Observability," Int. J. Control, Vol. 46, No. 5, 1987, pp. 1715-1737.
30. R. Nikoukhah, A.S. Willsky, and B.C. Levy, "Reachability, Observability, and Minimality for Shift-Invariant Two-Point Boundary-Value Descriptor Systems," in preparation.
31. R. Nikoukhah, B.C. Levy, and A.S. Willsky, "Stability, Stochastic Stationarity, and Generalized Lyapunov Equations for Two-Point Boundary-Value Descriptor Systems," Rept. LIDS-P-1797 (Revised Version), M.I.T. Lab. for Information and Decision Systems, Cambridge, Massachusetts 02139; to be submitted for publication.
32. R. Nikoukhah, A.S. Willsky, and B.C. Levy, "Generalized Riccati Equations for Two-Point Boundary-Value Descriptor Systems," Proc of the 1987 IEEE Conf. on Decision and Control, Los Angeles, CA, Dec. 1987.
33. R. Nikoukhah, A.S. Willsky, and B.C. Levy, "Generalized Riccati Equations: Existence and Uniqueness Results and Smoothing for Boundary-Value Models," in preparation.