Modeling and Estimation of Discrete-Time Gaussian Reciprocal Processes

BERNARD C. LEVY, SENIOR MEMBER, IEEE, RUGGERO FREZZA, AND ARTHUR J. KRENER, FELLOW, IEEE

Abstract—**In this paper, discrete-time Gaussian reciprocal processes are characterized in terms of a second-order two-point boundary-value nearest-neighbor model driven by a locally correlated noise whose correlation is specified by the model dynamics. This second-order model is the analog for reciprocal processes of the standard first-order state-space models for Markov processes. It is used to obtain a solution to the smoothing problem for reciprocal processes. The resulting smoother obeys second-order equations whose structure is similar to that of the Kalman filter for Gauss–Markov processes. Finally, it is shown that the smoothing error is itself a reciprocal process.**

I. Introduction

Noncausal random processes, or random fields, occur in many areas of science and engineering. These processes are usually indexed by space, instead of time. In this context, the concepts and models which were developed to study causal processes, such as the Markov property, Markov diffusions, or first-order Gauss–Markov state-space models, are no longer applicable. It is of interest to acquire equivalent tools for noncausal systems. In several dimensions, the concept of Markov random field [1]–[3] does not require any causality assumption. Specifically, a random field $x(\vec{r})$ with $\vec{r} \in \boldsymbol{R}^q$ is said to have the *Markov property* if, given an arbitrary closed domain D and its boundary Γ, the values of $x(\vec{r})$ inside D are conditionally independent of the values outside D, given the values of $x(\cdot)$ on the boundary, i.e., given $\{x(\vec{s});\ \vec{s} \in \Gamma\}$.

Somewhat surprisingly, Markov random fields do not reduce in one dimension to Markov processes, but to *reciprocal processes*. These processes were introduced in 1932 by Bernstein [4] and have been studied in detail in [5]–[11]. In the discrete-time case, a process $x(k)$ defined over $\boldsymbol{Z}$ is reciprocal if, given an arbitrary interval $[K, L]$, the values of $x(\cdot)$ in the interior and exterior of this interval are conditionally independent given $x(K)$ and $x(L)$. A consequence of this definition is that if $x(k)$ is a Markov process, it is necessarily reciprocal [6], but the converse is not true. Examples of reciprocal processes which do not have the Markov property can be found in [5]–[12].

The first significant step towards the development of stochastic models for reciprocal processes was undertaken recently in [12], where a theory of reciprocal diffusions was proposed. In this context, it was shown that unlike Markov diffusions, which are described by first-order stochastic differential equations driven by the uncorrelated increments of a Wiener process, reciprocal diffusions are modeled by *second-order* equations driven by locally correlated noise increments. However, a number of issues related to the structure of the driving noise and the selection of boundary conditions were not completely resolved in [12]. One goal of the present paper is to shed some light on these unresolved issues by considering the simple case of discrete-time Gaussian reciprocal processes. The case of continuous-time Gaussian reciprocal processes presents a few added difficulties, which are examined in [13].

To motivate the study of reciprocal processes from an engineering point of view, it is worth noting that the solutions of 1-D stochastic boundary-value systems of the type considered in [14]–[16] are all reciprocal processes. This means, for example, that the steady-state distribution of the temperature along a heated ring or a beam subjected to random loads along its length can be modeled in terms of reciprocal processes. The ship surveillance problem considered in [17] provides another good example of the applicability of reciprocal processes. In this context, given a Gauss–Markov state-space model of the ship's trajectory, it was desired to assign a probability distribution not only to the initial state, but also to the final state, corresponding to some predictive information about the ship's destination. This had the effect of modeling the trajectory as a reciprocal process, a feature which was exploited only indirectly in [17]. This suggests that the results of this paper will be of use for a wide variety of noncausal 1-D estimation problems.

In this paper, it is shown in Section II that discrete-time Gaussian reciprocal processes can be described by a second-order nearest-neighbor model driven by a first-order moving-average process whose correlation structure is determined by the model dynamics. This model is the analog for Gaussian reciprocal processes of first-order state-space models for Gauss–Markov processes. It also has the same form as a model proposed in [3] for discrete Markov random fields. This second-order model is used in Section III to obtain a complete characterization of Gaussian reciprocal processes. The special case of Markov processes is examined in Section IV, where the second-order model satisfied by these processes is constructed directly from their first-order state-space model. A recursive solution procedure for second-order reciprocal models is developed in Section V. This procedure relies on a white-noise representation for the driving noise and on the fact that a reciprocal process, when conditioned with respect to its final value, is a Markov process and can therefore be generated causally.

To illustrate the applications of second-order nearest-neighbor models of Gaussian reciprocal processes, we consider the fixed-interval smoothing problem in Section VI. It is shown that the smoothed estimates satisfy second-order recursions whose structure is similar to that of the Kalman filter for Gauss–Markov processes. These recursions are noncausal, but by factoring the smoother dynamics into first-order components which are, respectively, forward and backward stable, it is shown that the smoother admits a double sweep implementation similar to the Rauch–Tung–Striebel smoothing formula [18] for Markov processes. Finally, a second-order model is constructed for the smoothing error, which is used to show that the error is itself a reciprocal process.

II. Nearest-Neighbor Models of Reciprocal Processes

A. Model Construction

Consider a zero-mean, Gaussian, reciprocal process $x(k)$ defined over the interval $I = [0, N]$ and taking values in $\boldsymbol{R}^n$. Since

Manuscript received March 31, 1989; revised December 2, 1989. Paper recommended by Associate Editor, A. Benveniste. This work was supported by the National Science Foundation under Grant ECS-8700903 and by the Office of Naval Research under Grant USN-N00014-89-J-3153.

B. C. Levy is with the Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616.

R. Frezza was with the Department of Mathematics and the Institute of Theoretical Dynamics, University of California, Davis, CA 95616. He is now with the Department of Electrical Engineering, University of Padua, Padua, Italy.

A. J. Krener is with the Department of Mathematics and the Institute of Theoretical Dynamics, University of California, Davis, CA 95616.

IEEE Log Number 9037207.

0018-9286/90/0900-1013$01.00

$x(\cdot)$ is reciprocal, $x(k)$ must be conditionally independent of the values of $x(\cdot)$ in the exterior of the interval $[k-1, k+1]$, given $x(k-1)$ and $x(k+1)$. Thus,

$$\begin{aligned} E[x(k)|x(s), s \in I - \{k\}] &= E[x(k)|x(k-1), x(k+1)] \\ &= F_-(k)x(k-1) + F_+(k)x(k+1) \quad (2.1) \end{aligned}$$

for $1 \leq k \leq N-1$. This identity implies that the residual process

$$d(k) = x(k) - F_-(k)x(k-1) - F_+(k)x(k+1) \quad (2.2)$$

has the *orthogonality property*

$$d(k) \perp x(s) \qquad \text{for } s \neq k. \quad (2.3)$$

In the above construction, it is interesting to note that the matrices $F_\pm(k)$ and the residual process $d(k)$ are independent of the interval I, as long as this interval contains the point k and its two nearest neighbors $k-1$ and $k+1$.

The relation (2.2) can be viewed as specifying a nearest-neighbor model for $x(k)$, where the driving noise is the residual process $d(k)$. In this context, the orthogonality property (2.3) implies that the noise covariance $D(k, l) = E[d(k)d^T(l)]$ is such that

1) $D(k, l) = 0 \qquad \text{for } |k-l| > 1 \quad (2.4)$

2) $$\begin{aligned} D(k, k+1) &= -D(k, k)F_-^T(k+1) \\ &= -F_+(k)D(k+1, k+1). \quad (2.5) \end{aligned}$$

The property 1) indicates that $d(k)$ is a first-order *moving average process* and 2) imposes a constraint on the projections matrices $F_\pm(\cdot)$ and noise variance $D(k, k)$, since it shows that they cannot be specified independently of each other. The proof of (2.5) relies on the observation that

$$\begin{aligned} E[d(k)d^T(k+1)] &= -E[d(k)x^T(k)]F_-^T(k+1) \\ &= -F_+(k)E[x(k+1)d^T(k+1)] \end{aligned}$$

where

$$D(k, k) = E[d(k)d^T(k)] = E[d(k)x^T(k)].$$

The second-order nearest-neighbor model (2.2) differs from state-space models of Markov processes by the fact that the driving noise $d(k)$ is not white, but locally correlated. The relation (2.5) shows also that the model dynamics, i.e., $F_\pm(\cdot)$ and the noise variance $D(k, k)$ specify entirely the local correlation $D(k, k+1)$ of the driving noise. These features may appear surprising at first, but it is worth noting that a model with precisely the same structure was proposed in [3] for Markov random fields.

Another consequence of the orthogonality property (2.3) is that the covariance $R(k, s) = E[x(k)x^T(s)]$ of the process $x(k)$ satisfies the *second-order difference equation*

$$\begin{aligned} R(k, s) - F_-(k)R(k-1, s) - F_+(k)R(k+1, s) \\ = D(k, k)\delta(k-s) \quad (2.6) \end{aligned}$$

where $\delta(k)$ denotes the Kronecker delta function. This equation is the discrete analog of the second-order differential equation obtained by Krener [12], [13] for continuous-time Gaussian reciprocal processes.

Specializing (2.6) for $s = k-1, k+1$, we find that the projection matrices $F_\pm(k)$ can be obtained by solving

$$[F_-(k) F_+(k)]P(k) = [R(k, k-1) R(k, k+1)] \quad (2.7a)$$

where

$$P(k) = \begin{bmatrix} R(k-1, k-1) & R(k-1, k+1) \\ R(k+1, k-1) & R(k+1, k+1) \end{bmatrix} \quad (2.7b)$$

is the variance of

$$\xi^T(k) = [x^T(k-1) x^T(k+1)].$$

The existence and unicity of $F_\pm(k)$ is therefore guaranteed if the variance matrix $P(k)$ is positive definite for all k. Given $F_\pm(k)$, by setting $s = k$ in (2.6), we find that the error variance is given by

$$\begin{aligned} D(k, k) = R(k, k) - F_-(k)R(k-1, k) \\ - F_+(k)R(k+1, k). \quad (2.8) \end{aligned}$$

From (2.7)–(2.8), we can immediately conclude that when $x(k)$ is *stationary*, i.e., when $R(k, s) = R(k-s)$, the matrices $F_\pm(k)$ and $D(k, k)$ are *constant*, and (2.2) is a linear time-invariant system.

B. Boundary Conditions

The second-order model (2.2) with noise structure (2.4), (2.5) does not specify completely the process $x(k)$ over the interval $I = [0, N]$. Some boundary conditions must also be imposed at both ends of I. It turns out that there is more than one way to select a satisfactory set of boundary conditions. We consider here Dirichlet and cyclic boundary conditions, since both will be of use in subsequent developments.

Dirichlet Conditions: By construction, the residual process $d(k)$ for $1 \leq k \leq N-1$ is uncorrelated with $x(0)$ and $x(N)$. Thus, we can select

$$\begin{bmatrix} x(0) \\ x(N) \end{bmatrix} = b = \begin{bmatrix} b_i \\ b_f \end{bmatrix} \sim N(0, P_b) \quad (2.9a)$$

with

$$P_b = \begin{bmatrix} R(0, 0) & R(0, N) \\ R(N, 0) & R(N, N) \end{bmatrix} \quad (2.9b)$$

where b is independent of the noise $d(k)$ as boundary condition for the model (2.2).

Cyclic Conditions: Since $x(k)$ is reciprocal, $x(0)$ is conditionally independent of the values of $x(\cdot)$ in the interior of interval $[1, N]$, given $x(1)$ and $x(N)$. Thus,

$$\begin{aligned} E[x(0)|x(s), s \in I - \{0\}] \\ = F_-(0)x(N) + F_+(0)x(1). \quad (2.10a) \end{aligned}$$

Similarly, $x(N)$ is conditionally independent of the values of $x(\cdot)$ on the interval $[0, N-1]$, given $x(0)$ and $x(N-1)$, so that

$$\begin{aligned} E[x(N)|x(s), s \in I - \{N\}] \\ = F_-(N)x(N-1) + F_+(N)x(0). \quad (2.10b) \end{aligned}$$

Then, if we introduce the residuals

$$d(0) = x(0) - F_-(0)x(N) - F_+(0)x(1) \quad (2.11a)$$

$$d(N) = x(N) - F_-(N)x(N-1) - F_+(N)x(0) \quad (2.11b)$$

at both ends of interval I, we have the orthogonality relations

$$d(0) \perp x(s), \qquad s \neq 0 \quad (2.12a)$$

$$d(N) \perp x(s), \qquad s \neq N. \quad (2.12b)$$

The relations (2.11)–(2.12) have the effect of extending cyclically the model (2.2), as well as the noise structure (2.4), (2.5), and the second-order covariance equation (2.6), to the whole interval $I = [0, N]$, provided that in these identities, $k-1$ and $k+1$ are defined modulo $N+1$. There exists, however, one important difference between the relation (2.2) for points k in the interior of I, and relations (2.11a) and (2.11b) for the end points. Namely, unlike the case of interior points, the boundary matrices $F_{\pm}(0)$, $F_{\pm}(N)$ and residuals $d(0)$, $d(N)$ depend on I. If the interval I is extended or reduced at either end, *all* the boundary matrices and residuals are affected. In other words, the cyclic conditions (2.11a) and (2.11b) have the effect of wrapping around the interval I onto a circle, but the wrapping procedure needs to be modified as the interval I is increased or decreased.

C. Well-Posedness

Given the model (2.2) with the noise structure (2.4), (2.5), and the Dirichlet boundary conditions (2.9), or cyclic conditions (2.11), it is now possible to determine whether the resulting system is well posed.

For the *Dirichlet* model, the relations (2.2), (2.9) can be rewritten as a single matrix equation

$$\boldsymbol{F_D x} = \boldsymbol{d_D} \tag{2.13}$$

with

$$\boldsymbol{x}^T = [x^T(0) \quad x^T(1) \quad \cdots \quad x^T(N-1) \quad x^T(N)] \tag{2.14a}$$

$$\boldsymbol{d}_D^T = [b_i^T \quad d^T(1) \cdots d^T(N-1) \quad b_f^T] \tag{2.14b}$$

$$\boldsymbol{F_D} = \begin{bmatrix} I & 0 & & & & \\ -F_-(1) & I & -F_+(1) & & 0 & \\ & \vdots & \vdots & \vdots & \vdots & \\ & 0 & & -F_-(N-1) & I & -F_+(N-1) \\ & & & & 0 & I \end{bmatrix}. \tag{2.14c}$$

The well-posedness of the second-order model (2.2), (2.9) is therefore equivalent to the invertibility of $\boldsymbol{F_D}$. But the orthogonality property (2.3) of the residuals implies

$$\boldsymbol{F_D R} = E[\boldsymbol{d_D x}^T] = \Delta_D \tag{2.15a}$$

where $\boldsymbol{R} = E[\boldsymbol{xx}^T]$, and

$$\Delta_D = \begin{bmatrix} R(0,0) & R(0,1) & \cdots & R(0,N-1) & R(0,N) \\ & D(1,1) & & & \\ & & & 0 & 0 \\ 0 & 0 & \ddots & & \\ & & & D(N-1,N-1) & \\ R(N,0) & R(N,1) & \cdots & R(N,N-1) & R(N,N) \end{bmatrix}. \tag{2.15b}$$

When the covariance matrix $\boldsymbol{R} > 0$, the variance $D(k, k)$ of the noise $d(k)$ must be positive definite for all k. Otherwise, there exists a vector $v \neq 0$ such that $v^T d(k) = 0$. But from (2.2), this implies that a linear combination of $x(k-1)$, $x(k)$, and $x(k+1)$ is identically zero, which contradicts the assumption that $\boldsymbol{R} > 0$. This implies the following.

Lemma 2.1: If the covariance matrix $\boldsymbol{R}$ of $x(k)$ is positive definite, the Dirichlet-model (2.2), (2.9) is well posed.

Proof: If $\boldsymbol{R} > 0$, we have shown that $D(k, k) > 0$ in the interior of I. Furthermore, the matrix P_b formed by the four corner blocks of Δ_D is invertible, since it is a principal minor of $\boldsymbol{R}$, so that Δ_D is invertible. From (2.15a), this implies that $\boldsymbol{F_D}$ is invertible. ■

The well-posedness of the *cyclic* second-order model (2.2), (2.11) can be characterized in the same way. In this case, the identity (2.13) takes the form

$$\boldsymbol{F_C x} = \boldsymbol{d_C} \tag{2.16a}$$

with

$$\boldsymbol{d}_C^T = [d^T(0) \quad d^T(1) \quad \cdots \quad d^T(N-1) \quad d^T(N)] \tag{2.16b}$$

$$\boldsymbol{F_C} = \begin{bmatrix} I & -F_+(0) & & & & -F_-(0) \\ -F_-(1) & I & -F_+(1) & & 0 & \\ & \vdots & \vdots & \vdots & \vdots & \\ & 0 & & -F_-(N-1) & I & -F_+(N-1) \\ -F_+(N) & & & & -F_-(N) & I \end{bmatrix} \tag{2.16c}$$

and the orthogonality property of the residuals implies

$$\boldsymbol{F_C R} = \Delta_C \tag{2.17a}$$

with

$$\Delta_C = \mathrm{diag}\{D(k, k)\}. \tag{2.17b}$$

Then, when $\boldsymbol{R} > 0$, it is easy to check that $D(k, k) > 0$ for $0 \le k \le N$, so that the cyclic model (2.2), (2.11) is well posed.

From (2.17), we see also that if $x(k)$ is a nonsingular (i.e., $\boldsymbol{R} > 0$) reciprocal process over $[0, N]$, its covariance $\boldsymbol{R}$ has a *cyclic block tridiagonal inverse*. This property will be used below to obtain a characterization of reciprocal processes.

III. Characterization of Reciprocal Processes

Up to this point, we have shown that if a process is reciprocal, it admits a nearest-neighbor model (2.2) with noise structure (2.4), (2.5), and with either Dirichlet or cyclic boundary conditions. However, we have not yet shown that the model (2.2) and (2.4), (2.5) captures completely the structure of reciprocal processes. In other words, we have not proved that the solution $x(k)$ of such a model is necessarily reciprocal. It turns out that this is the case, but instead of considering directly the nearest-neighbor model (2.2), (2.4), and (2.5), we introduce a renormalized version of this model which is simpler to analyze.

A. Second-Order Descriptor Model

Specifically, we consider the second-order descriptor system

$$M_0(k)x(k) - M_-(k)x(k-1) - M_+(k)x(k+1) = e(k) \tag{3.1}$$

such that

$$M_+(k) = M_-^T(k+1) \tag{3.2}$$

with Dirichlet boundary conditions

$$\begin{bmatrix} x(0) \\ x(N) \end{bmatrix} = b = \begin{bmatrix} b_i \\ b_f \end{bmatrix} \sim N(0, P_b) \tag{3.3}$$

and where the input noise $e(k)$ is a Gaussian process uncorrelated with b, whose covariance $E(k,l) = E[e(k)e^T(l)]$ is such that

1) $E(k,l) = 0 \quad \text{for } |k-l| > 1$ (3.4a)
2) $E(k,k) = M_0(k), \quad E(k,k+1) = -M_+(k).$ (3.4b)

In the above equations, $x(k) \in R^n$, and M_0, $M_\pm$ are square matrices of size n. Furthermore, it is assumed that the descriptor model (3.1)–(3.3) is *well posed*, i.e., that it admits a unique solution.

To see how the descriptor model (3.1)–(3.4) is related to the nearest-neighbor model (2.2) with noise structure (2.4), (2.5), assume that $x(k)$ is a reciprocal process with full rank residuals, i.e., such that $D(k,k)$ is invertible for all k in the interior of I. We have seen earlier that this is ensured by $R > 0$. Then, multiplying (2.2) on the left by $D^{-1}(k,k)$, and identifying

$$M_0(k) = D^{-1}(k,k), \quad M_\pm(k) = D^{-1}(k,k)F_\pm(k) \tag{3.5a}$$

$$e(k) = D^{-1}(k,k)d(k) \tag{3.5b}$$

the nearest-neighbor model (2.2), (2.9) is transformed into the descriptor model (3.1), (3.3). Furthermore, the noise structure (2.4), (2.5) for $d(k)$ implies that the descriptor model satisfies (3.2) and that the covariance of the normalized noise $e(k)$ obeys (3.4). This shows that any nonsingular reciprocal process $x(k)$ admits a well-posed descriptor model of the form (3.1)–(3.4).

The descriptor model (3.1)–(3.4) has several advantages over (2.2). The first one is that the noise structure is displayed more clearly. Specifically, the relations (3.4a) and (3.4b) show that the noise covariance $E(k,l)$ is totally specified by the descriptor dynamics. The matrix $M_0(k)$ is positive definite, and (3.2) indicates that the matrix functions $M_+(\cdot)$ and $M_-(\cdot)$ can be specified from each other. Thus, as in the Markov case, only two functions, say $M_0(\cdot)$ and $M_+(\cdot)$, are needed to describe the model. Another advantage of (3.1)–(3.4) is that the constraint (3.2) between $M_-(\cdot)$ and $M_+(\cdot)$ implies that the second-order difference operator associated to (3.1) is *self-adjoint*, as will be shown below. This last feature is, in fact, the primary motivation for the introduction of the renormalization (3.5).

Now that we have shown that a nonsingular reciprocal process $x(k)$ admits a well-posed descriptor model (3.1)–(3.4), our goal in this section is to prove the converse, i.e., that the solution $x(k)$ of a well-posed descriptor system (3.1)–(3.4) is necessarily reciprocal.

B. Solution

The first step of our analysis consists of introducing a Green's identity and Green's function, which will be used to express $x(k)$ in terms of the input noise $e(k)$ and boundary vector b. Thus, let

$$\Lambda = M_0(k)I - M_-(k)Z^{-1} - M_+(k)Z \tag{3.6a}$$

be the difference operator associated to (3.1), where Z denotes the forward time-shift

$$Zf(k) = f(k+1). \tag{3.6b}$$

Λ maps n-vector functions defined over the interval $I = [0, N]$ into n-vector functions defined over the interior $[1, N-1]$ of this interval. Consider now the space S of n-vector functions defined over the interior of I, with the inner product

$$\langle r, e \rangle = \sum_{s=1}^{N-1} r^T(s)e(s). \tag{3.7}$$

Then, the following *Green's identity* is satisfied:

$$\langle r, \Lambda x \rangle = \langle \Lambda r, x \rangle + [r^T(0) \quad r^T(1)]E_i \begin{bmatrix} x(0) \\ x(1) \end{bmatrix} + [r^T(N) \quad r^T(N-1)]E_f \begin{bmatrix} x(N) \\ x(N-1) \end{bmatrix} \tag{3.8}$$

with

$$E_i = \begin{bmatrix} 0 & M_+(0) \\ -M_-(1) & 0 \end{bmatrix}$$

$$E_f = \begin{bmatrix} 0 & M_-(N) \\ -M_+(N-1) & 0 \end{bmatrix} \tag{3.9}$$

where, as was indicated earlier, a consequence of (3.2) is that the operator Λ is *self-adjoint*.

The *Green's function* associated to Λ is given by

$$\Lambda\Gamma(k,s) = I\delta(k-s) \tag{3.10a}$$

$$\Gamma(0,s) = \Gamma(N,s) = 0 \tag{3.10b}$$

with $1 \le s \le N-1$, where since Λ is self-adjoint

$$\Gamma(k,s) = \Gamma^T(s,k). \tag{3.11}$$

Note that $\Gamma(k,s)$ depends on the interval of definition $I = [0, N]$, so that it should be really denoted as $\Gamma(k, s; I)$.

Solution: Applying the Green's identity (3.8) to the case where $x(s)$ satisfies (3.1)–(3.3), and $r(s) = \Gamma(s,k)a$ where a is an arbitrary vector of R^n, we find

$$a^T \left[\sum_{s=1}^{N-1} \Gamma^T(s,k)e(s) + \Gamma^T(1,k)M_-(1)b_i + \Gamma^T(N-1,k)M_+(N-1)b_f - x(k) \right] = 0. \tag{3.12}$$

Since a is arbitrary, and taking into account (3.11), the solution of (3.1)–(3.3) is therefore given by

$$x(k) = \sum_{s=1}^{N-1} \Gamma(k,s)e(s) + \Gamma(k,1)M_-(1)b_i + \Gamma(k,N-1)M_+(N-1)b_f. \tag{3.13}$$

Orthogonality Property: The expression (3.13) can be used to show that if the noise $e(k)$ in (3.1) has the covariance structure (3.4), then the solution $x(k)$ is such that

$$x(k) \perp e(s) \qquad \text{for } k \neq s. \tag{3.14}$$

This can be checked from

$$\begin{aligned} E[x(k)e^T(l)] &= \sum_{s=1}^{N-1} \Gamma(k, s)E(s, l) \\ &= \Gamma(k, l)M_0(l) - \Gamma(k, l-1)M_-^T(l) \\ &\quad - \Gamma(k, l+1)M_+^T(l) \\ &= I\delta(k-l) \end{aligned} \tag{3.15}$$

where we have used the definition (3.10) of the Green's function, as well as the self-adjointness relation (3.11).

Covariance Description: The orthogonality property (3.14) has important consequences. In particular, if we substitute the expression (3.13) for the solution $x(k)$ inside the covariance $R(k, s) = E[x(k)x^T(s)]$, and take into account the orthogonality property (3.14), we find that $R(k, s)$ can be expressed in terms of the Green's function $\Gamma(k, s)$ as

$$\begin{aligned} R(k, s) &= \Gamma(k, s) \\ &\quad + [\Gamma(k, 1)M_-(1)\,\Gamma(k, N-1)M_+(N-1)]P_b \\ &\quad \cdot \begin{bmatrix} M_-^T(1)\Gamma^T(s, 1) \\ M_+^T(N-1)\Gamma^T(s, N-1) \end{bmatrix}. \end{aligned} \tag{3.16}$$

This provides the following stochastic interpretation of the Green's function $\Gamma(k, s)$. Let $x_0(k)$ be the pinned process over $[0, N]$ which is obtained by solving (3.1)–(3.4) with zero boundary conditions, i.e., with $b = 0$, or equivalently $P_b = 0$. Then, according to (3.16),

$$E[x_0(k)x_0^T(s)] = R_0(k, s) = \Gamma(k, s) \tag{3.17}$$

so that the Green's function $\Gamma(k, s)$ is the covariance of the pinned process $x_0(k)$. This implies that $\Gamma(k, s)$ is a nonnegative operator. By applying the operator Λ on both sides of (3.16), we also find that the covariance $R(k, s)$ satisfies the second-order difference equation

$$\begin{aligned} M_0(k)R(k, s) - M_-(k)R(k-1, s)& \\ -M_+(k)R(k+1, s) &= I\delta(k-s) \end{aligned} \tag{3.18}$$

which, except for the fact that we are considering here a descriptor model, is identical to (2.6). Note, however, that while (2.6) had been derived under the assumption that $x(k)$ was reciprocal, up to this point, we have only assumed that $x(k)$ is a solution of (3.1)–(3.4).

From the above discussion, we see, therefore, that the Green's function $\Gamma(k, s)$ and covariance $R(k, s)$ of model (3.1)–(3.4) satisfy the same difference equation and differ only by the choice of boundary conditions.

C. Main Result

We are now ready to prove the following.

Theorem 3.1: Let $x(k)$ be a zero-mean Gaussian process whose covariance is nonsingular, i.e., $\boldsymbol{R} > 0$. Then, $x(k)$ is reciprocal if and only if it admits a well-posed second-order descriptor model of the form (3.1)–(3.4).

Proof: Necessity was proved in Section III-A. To prove sufficiency, we must show that the solution $x(k)$ of (3.1)–(3.4) given by (3.13) is reciprocal. Thus, let $J = [K, L]$ be a subinterval of I, and let l and k be two points such that

$$0 \leq l \leq K \leq k \leq L \leq N,$$

i.e., l and k are, respectively, in the exterior and interior of J. To prove that $x(\cdot)$ is reciprocal, we only need to show that

$$\tilde{x}(k) = x(k) - E[x(k)|x(K), x(L)] \perp x(l). \tag{3.19}$$

Solving (3.1), (3.2) over the subinterval J, with boundary conditions $x(K)$ and $x(L)$, yields

$$\begin{aligned} x(k) &= \sum_{s=K+1}^{L-1} \Gamma(k, s; J)e(s) \\ &\quad + \Gamma(k, K+1; J)M_-(K+1)x(K) \\ &\quad + \Gamma(k, L-1; J)M_+(L-1)x(L) \end{aligned}$$

where $\Gamma(k, s; J)$ denotes the Green's function for the system defined over the subinterval J. But the orthogonality property (3.14) implies that $e(s) \perp x(K), x(L)$ for $K < s < L$, so that

$$\begin{aligned} E[x(k)|x(K), x(L)] &= \Gamma(k, K+1; J)M_-(K+1)x(K) \\ &\quad + \Gamma(k, L-1; J)M_+(L-1)x(L) \end{aligned} \tag{3.20a}$$

$$\tilde{x}(k) = \sum_{s=K+1}^{L-1} \Gamma(k, l; J)e(s). \tag{3.20b}$$

Using the orthogonality property (3.14) and expression (3.20b) for $\tilde{x}(k)$, we can conclude that $\tilde{x}(k) \perp x(l)$, so that the process $x(k)$ is reciprocal. □

Remark: It is worth noting at this point that since the nearest-neighbor models of reciprocal processes that we are discussing here can be viewed as a specialization to the 1-D case of Markov random field models considered in [3], the characterization of reciprocal processes that we have obtained in Theorem 3.1 is in some sense implied by the earlier results of [3]. However, the analysis of [3] was restricted to the scalar stationary case, and several issues, such as the solution of the models, were not examined. Furthermore, the application of nearest-neighbor models to the study of estimation problems, which is considered in Section VI, is entirely new.

D. Cyclic Model

The characterization of reciprocal processes that we have just derived is expressed in terms of a second-order descriptor model with Dirichlet conditions. It is possible to obtain an equivalent characterization for a cyclic model, where the Dirichlet conditions (3.3) are replaced by cyclic conditions. This means that the descriptor equations (3.1), (3.2) and the noise structure (3.4) are now assumed to hold for all $k \in I$, provided that we define $k-1$ and $k+1$ modulo $N+1$.

The steps which can be used to convert the characterization of Theorem 3.1 into an equivalent result for cyclic descriptor models are as follows. First, observe that if $x(k)$ is a nonsingular ($\boldsymbol{R} > 0$) reciprocal process, its cyclic nearest neighbor model is well posed, and the noise variance $D(k, k) > 0$ for all $k \in I$. Therefore, by applying the transformation (3.5), we obtain an equivalent well-posed second-order cyclic descriptor model.

Next, consider a well-posed cyclic descriptor model. Its solution $x(k)$ can be computed as follows. Let Λ_C be the operator obtained by replacing Z by the cyclic shift

$$Z_C f(k) = f((k+1)\,\mathrm{mod}(N+1)) \tag{3.21}$$

in (3.6a). Note that the range of Λ_C is now the set of n-vector functions over the whole interval I. Then, if the inner product (3.7) is defined over I instead of its interior, the Green's identity (3.8) becomes

$$\langle r, \Lambda_C x\rangle = \langle \Lambda_C r, x\rangle. \tag{3.22}$$

If we define the cyclic Green's function as the solution of

$$\Lambda_C \Gamma_C(k, s) = I\delta(k - s) \qquad \text{for } k, s \in I \tag{3.23}$$

the solution of the cyclic descriptor model is given by

$$x(k) = \sum_{s=0}^{N} \Gamma_C(k, s)e(s). \tag{3.24}$$

Taking into account the noise structure (3.4), we see immediately that the orthogonality property (3.14) is satisfied for all $k, s \in I$ with $k \neq s$. This property can then be used to show that

$$R(k, s) = \Gamma_C(k, s), \tag{3.25}$$

i.e., the covariance $R(k, s)$ of the solution $x(k)$ is identical to the cyclic Green's function $\Gamma_C(k, s)$. This implies that the covariance $R(k, s)$ satisfies (3.23). Another consequence of the orthogonality property (3.14) is that

$$x(0), x(N) \perp e(s)$$

for s in the interior of I. This shows that the original cyclic descriptor model can be converted back to a Dirichlet model, whose solution is then guaranteed to be reciprocal. Thus, we have proved the following.

Corollary 3.1: Let $x(k)$ be a zero-mean nonsingular Gaussian process. Then, $x(k)$ is reciprocal if and only if it admits a well-posed second-order cyclic descriptor model (3.1), (3.2) with noise structure (3.4).

Since the covariance $R(k, s)$ of a cyclic descriptor model satisfies (3.23), we obtain also the following result.

Theorem 3.2: $R > 0$ is the covariance matrix of a reciprocal process if and only if its inverse covariance R^{-1} has a cyclic block tridiagonal structure, i.e.,

$$R^{-1} = \begin{bmatrix} M_0(0) & -M_+(0) & & & & -M_-(0) \\ -M_-(1) & M_0(1) & -M_+(1) & & 0 & \\ & \vdots & \vdots & \vdots & \vdots & \\ & 0 & & -M_-(N-1) & M_0(N-1) & -M_+(N-1) \\ -M_+(N) & & & & -M_-(N) & M_0(N) \end{bmatrix}. \tag{3.26}$$

This characterization of the covariance of reciprocal processes is extremely convenient. It is the key to the derivation of the smoothing results presented in Section VI, and it is, in fact, the primary motivation for the introduction of cyclic models.

E. Conjugate Process

An interesting feature of the cyclic descriptor model (3.1), (3.2), (3.4) is that its input $e(k)$ is the *conjugate process* of $x(k)$. The notion of conjugate process was originally introduced by Rozanov [19, ch. 2] and Masani [20, sect. 2] for the prediction and interpolation of stationary Gaussian processes, and was subsequently applied to stochastic realization theory in [21], [22]. The conjugate process of a nonsingular process $x(k)$ defined over I is the unique process $e(k)$ with $k \in I$ such that:

1) $E[x(k)e^T(l)] = I\delta(k - l) \qquad \text{for } k, l \in I; \qquad (3.27)$

and 2) $E = X$, where E and X denote the Hilbert spaces of random variables spanned, respectively, by $e(k)$ and $x(k)$ for $k \in I$. For the cyclic descriptor model that we consider here, the biorthogonality relation (3.27) is a consequence of (3.14) and of the renormalization (3.5b) that was used to define $e(k)$, and the equality of X and E is implied by expression (3.24) for $x(k)$.

From (3.27), we see that the covariance of the conjugate process $e(k)$ is R^{-1}, and Theorem 3.2 just states that a zero-mean nonsingular Gaussian process is reciprocal if and only if its conjugate process has a cyclic tridiagonal covariance. The previous characterization of Gaussian reciprocal processes can also be extended to the case where $x(k)$ is a stationary Gaussian process defined over the whole line, instead of a finite interval. In this case, assume that the covariance $R(k)$ of $x(k)$ is summable, so that $x(k)$ admits a spectral density $S(\lambda)$, i.e., $x(k)$ is a regular process. Assume also that

$$\int_{-\pi}^{\pi} \operatorname{tr} S^{-1}(\lambda)\, d\lambda < \infty \tag{3.28}$$

which guarantees that the conjugate process is well defined [19], i.e., that it has a finite variance. The spectral density of the conjugate process is $S^{-1}(\lambda)$ [19]. Then, if we let the length N of the interval I tend to infinity in the cyclic descriptor model (3.1), (3.2), (3.4), and observe that $R(N) \to 0$ as $N \to \infty$ implies that $M_-(0), M_+(N) \to 0$ in (3.26), we find that a zero-mean Gaussian stationary process $x(k)$ is reciprocal if and only if its conjugate process admits a spectral density of the form

$$S^{-1}(\lambda) = M_0 - M_+ e^{-j\lambda} - M_- e^{j\lambda}. \tag{3.29}$$

It turns out that this result is well known for Markov random fields [8], [23]. In the context of reciprocal processes, it is of limited interest, since regular Gaussian reciprocal processes defined over Z are necessarily Markov [8], so that (3.29) is really a characterization of stationary Gauss–Markov processes.

IV. Markov Processes

Since Markov processes are reciprocal, they admit a second-order description, which, as shown below, can be constructed directly from the first-order state-space model that they satisfy. Thus, let $x(k)$ be a zero-mean Gauss–Markov process with state-space model

$$x(k + 1) = A(k)x(k) + w(k) \tag{4.1a}$$

and initial condition

$$x(0) = b_i \sim N(0, \Pi_b) \tag{4.1b}$$

where $w(k)$ is a white Gaussian noise (WGN) uncorrelated with b_i and with intensity $Q(k)$:

$$E[w(k)w^T(l)] = Q(k)\delta(k - l). \tag{4.1c}$$

The solution of (4.1) is given by

$$x(k) = \Phi(k, 0)b_i + \sum_{s=0}^{k-1} \Phi(k, s + 1)w(s) \tag{4.2}$$

where $\Phi(k, s)$ is the state-transition matrix

$$\Phi(k, s) = \begin{cases} A(k - 1) \cdots A(s) & \text{for } k > s \\ I & \text{for } k = s. \end{cases} \tag{4.3}$$

The state variance $\Pi(k) = E[x(k)x^T(k)]$ satisfies

$$\Pi(k+1) = A(k)\Pi(k)A^T(k) + Q(k) \tag{4.4}$$

with initial condition $\Pi(0) = \Pi_b$, and the state covariance $R(k, s)$ is given by

$$R(k, s) = \begin{cases} \Phi(k, s)\Pi(s) & \text{for } k \geq s \\ \Pi(k)\Phi^T(s, k) & \text{for } k \leq s. \end{cases} \tag{4.5}$$

The second-order model of $x(k)$ can be constructed as follows. In (4.1a), the noises $w(k)$ and $w(k-1)$ are uncorrelated with the past states, i.e.,

$$w(k),\ w(k-1) \perp x(s), \qquad 0 \leq s \leq k-1. \tag{4.6}$$

Their correlation with future states is given by

$$E[w(k-1)x^T(s)] = Q(k-1)\Phi^T(s, k) \tag{4.7a}$$

$$E[w(k)x^T(s)] = Q(k)\Phi^T(s, k+1). \tag{4.7b}$$

Consequently, for the full rank noise case where $Q(k) > 0$, the process

$$e(k) = Q^{-1}(k-1)w(k-1) - A^T(k)Q^{-1}(k)w(k) \tag{4.8}$$

has the orthogonality property

$$e(k) \perp x(s) \qquad \text{for } s \neq k \tag{4.9}$$

which, as was seen earlier, characterizes the driving noise of second-order models of reciprocal processes.

Second-Order Model: Substituting (4.1a) inside (4.8) gives

$$\begin{aligned} e(k) = {} & [Q^{-1}(k-1) + A^T(k)Q^{-1}(k)A(k)]x(k) \\ & - Q^{-1}(k-1)A(k-1)x(k-1) \\ & - A^T(k)Q^{-1}(k)x(k+1). \end{aligned} \tag{4.10}$$

Denoting

$$M_0(k) = Q^{-1}(k-1) + A^T(k)Q^{-1}(k)A(k) \tag{4.11a}$$

$$M_+(k) = A^T(k)Q^{-1}(k) \tag{4.11b}$$

$$M_-(k) = Q^{-1}(k-1)A(k-1) \tag{4.11c}$$

and observing from (4.8) that the covariance $E(k, l)$ of $e(k)$ has the structure (3.4a), (3.4b), we find, therefore, that (4.10)–(4.11) is the desired second-order descriptor model of the Markov process $x(k)$.

Consider the difference operator

$$\Omega = I - A(k-1)Z^{-1} \tag{4.12a}$$

and operator $\boldsymbol{Q} = Q(k-1)I$. The dual operator of Ω is

$$\Omega^* = I - A^T(k)Z \tag{4.12b}$$

and

$$\Omega^*\boldsymbol{Q}^{-1}\Omega = \Lambda \tag{4.13}$$

is an anticausal times causal factorization of the operator Λ associated with the descriptor model (4.10)–(4.11). Furthermore, by observing that

$$e(k) = \Omega^*\boldsymbol{Q}^{-1}w(k-1) \tag{4.14}$$

the second-order model (4.8) can be rewritten in operator form as

$$\Omega^*\boldsymbol{Q}^{-1}[\Omega x(k) - w(k-1)] = 0 \tag{4.15}$$

which clearly shows how it is derived from the state-space model (4.1a).

Boundary Conditions: The Dirichlet conditions (2.9a), (2.9b), which were derived for general reciprocal processes, remain valid for Markov processes, provided that we take into account the covariance structure (4.5) inside expression (2.9b) for the boundary variance P_b. To obtain cyclic conditions, we need only to observe that

$$\begin{aligned} e(0) &= \Pi_b^{-1}x(0) - A^T(0)Q^{-1}(0)w(0) \\ &= M_0(0)x(0) + M_+(0)x(1) \end{aligned} \tag{4.16a}$$

$$\begin{aligned} e(N) &= Q^{-1}(N-1)w(N-1) \\ &= M_-(N)x(N-1) + M_0(N)x(N) \end{aligned} \tag{4.16b}$$

with

$$M_0(0) = \Pi_b^{-1} + A^T(0)Q^{-1}(0)A(0), \tag{4.17a}$$

$$M_0(N) = Q^{-1}(N-1) \tag{4.17b}$$

$$M_+(0) = -A^T(0)Q^{-1}(0), \tag{4.18a}$$

$$M_-(N) = -Q^{-1}(N-1)A(N-1) \tag{4.18b}$$

satisfy the orthogonality relation (4.9) for $k = 0, N$, respectively. These boundary conditions are *separable*, in the sense that the states at each end of the interval $[0, N]$ are decoupled, i.e.,

$$M_-(0) = M_+(N) = 0. \tag{4.19}$$

Observing that $M_-(0)$ and $M_+(N)$ are the two matrices appearing in the off-diagonal corners of the cyclic block tridiagonal inverse of $\boldsymbol{R}$ in (3.26), we see, therefore, that the inverse covariance of a Markov process must be *block tridiagonal*. It was shown in [24, sect. 2] that this property characterizes Markov processes. Thus, one difference between Gaussian reciprocal and Markov processes is that their inverse covariances are cyclic block tridiagonal and block tridiagonal, respectively.

Since Gauss–Markov processes satisfy both first- and second-order models, which are, respectively, causal and anticausal, it is of interest to determine when one model should be used instead of the other. As a general rule, the usual first-order model should be used for causal estimation and stochastic control problems, whereas the second-order nearest-neighbor model is better adapted to the study of noncausal estimation problems, such as smoothing (see Section VI), interpolation, or noncausal stochastic control problems.

V. Recursive Solution Procedure

The solution (3.13) of the descriptor model (3.1)–(3.4) has the disadvantage of being nonrecursive. We describe in this section a recursive procedure which relies on decomposing the solution $x(k)$ into a Markov process which can be computed recursively, plus a component depending only on the end boundary condition.

Noise Representation: The first step is to obtain a first-order moving average representation of the form (4.8) for the driving noise process $e(k)$. In this representation, $w(k)$ will be a WGN process with intensity $Q(k)$ which is uncorrelated with the boundary vector b appearing in (3.3). Then, the relations (3.4) for the covariance of $e(k)$ imply that the matrices $A(k)$ and $Q(k)$, appearing in the representation (4.8), must satisfy (4.11a) and (4.11b). Eliminating

$$A(k) = Q(k)M_+^T(k) \tag{5.1}$$

from these two identities, we see that $Q(k)$ must satisfy the *backward Riccati equation*

$$Q^{-1}(k-1) = M_0(k) - M_+(k)Q(k)M_+^T(k) \quad \text{for } 1 \le k \le N-1. \tag{5.2}$$

Since we want the solution $Q(k)$ to be a covariance matrix, it must be nonnegative. To see how this can be achieved, note that if $n(k) = Q^{-1}(k)w(k)$, the noise representation (4.8) can be rewritten as a single matrix equation

$$\boldsymbol{e} = \Omega^T \boldsymbol{n} \tag{5.3}$$

with

$$\boldsymbol{e}^T = [e^T(1)\cdots e^T(k)\cdots e^T(N-1)] \tag{5.4a}$$

$$\boldsymbol{n}^T = [n^T(0)\cdots n^T(k)\cdots n^T(N-1)] \tag{5.4b}$$

$$\Omega^T = \begin{bmatrix} I & -A^T(1) & & & \\ & I & -A^T(2) & 0 & \\ & 0 & \vdots & \vdots & \\ & & & I & -A^T(N-1) \end{bmatrix} \tag{5.4c}$$

where Ω is just the matrix representation of operator (4.12a). Note that in (5.3), $\boldsymbol{n}$ has one more block entry than $\boldsymbol{e}$, and accordingly Ω^T has one more block column than rows. From (3.4), the covariance $\boldsymbol{E} = E[\boldsymbol{ee}^T]$ has the structure

$$\boldsymbol{E} = \begin{bmatrix} M_0(1) & -M_+(1) & & & & \\ -M_-(2) & M_0(2) & -M_+(2) & & 0 & \\ & \vdots & \vdots & \vdots & \vdots & \\ & & & -M_-(N-2) & M_0(N-2) & -M_+(N-2) \\ & & & & -M_-(N-1) & M_0(N-1) \end{bmatrix} \tag{5.5}$$

and from (5.3), we find

$$\boldsymbol{E} = \Omega^T \boldsymbol{Q}^{-1} \Omega \tag{5.6a}$$

with

$$\boldsymbol{Q} = \mathrm{diag}\{Q(k)\} \tag{5.6b}$$

which is the matrix representation of the operator factorization (4.13). The covariance matrix $\boldsymbol{E}$ is positive definite, since it is a principal minor of the inverse covariance $\boldsymbol{R}^{-1}$ given by (3.26). Then, if we select

$$Q(N-1) = 0 \tag{5.7}$$

which corresponds to deleting the last block column of Ω^T, the factorization (5.6a) is just the standard UDL factorization of $\boldsymbol{E}$, so that $Q^{-1}(k) > 0$ for all k. In this context, the recursions (5.1), (5.2) can be viewed as the standard *backward Cholesky recursions* for the UDL factorization of the block tridiagonal matrix $\boldsymbol{E}$. Thus, the proper initialization of the Riccati equation (5.2) is (5.7).

Recursive Solution: The solution of (5.1), (5.2) yields both a representation of the form (4.8) for the noise $e(k)$ and a factorization (4.13) for the operator Λ associated to the descriptor model (3.1), (3.2). Note, however, that this factorization involves only the interior points, and does not include the boundary conditions. In the interior of I, the descriptor system (3.1) can therefore be rewritten as (4.15). But (4.15) can be decomposed into two coupled first-order equations

$$z(k-1) = A^T(k)z(k), \qquad 1 \le k \le N-1 \tag{5.8a}$$

$$x(k+1) = A(k)x(k) + w(k) + Q(k)z(k), \qquad 0 \le k \le N-1 \tag{5.8b}$$

with boundary conditions (3.3). These two equations propagate causally in the backward and forward directions, respectively, but the boundary conditions (3.3) make the overall system noncausal.

Now let $z_M(k)$ and $x_M(k)$ be the solution of (5.8) with

$$z_M(N-1) = 0, \qquad x_M(0) = b_i. \tag{5.9}$$

Since (5.8a) is undriven, we have $z_M(k) = 0$, and consequently, $x_M(k)$ is a *Markov process* which can be computed recursively by propagating (5.8b) in the forward direction. This process satisfies (3.1) and the correct initial condition, but its end value $x_M(N)$ is different from the boundary vector b_f of (3.3). If $\Gamma(k, s)$ is the Green's function defined in (3.10), the solution $x(k)$ of (3.1)–(3.3) is therefore given by

$$x(k) = x_M(k) + \Gamma(k, N-1)M_+(N-1)[b_f - x_M(N)]. \tag{5.10}$$

Thus, $x(k)$ is obtained by first computing recursively the Markov solution $x_M(k)$ and then adding to it a term correcting the mismatch of the end boundary condition.

The above solution procedure relies on the observation that given an initial condition $x(0)$ and a second-order model (3.1), there is exactly one end condition at $k = N$ for which the solution of (3.1), (3.2) has the Markov property [7]. The selection of this specific end condition is accomplished here by setting $z_M(N-1) = 0$.

VI. Smoothing Problem

Consider a reciprocal process $x(k)$ defined over $I = [0, N]$, with descriptor model (3.1), (3.2), noise structure (3.4), and cyclic boundary conditions. We are given the observations

$$y(k) = H(k)x(k) + v(k), \qquad 0 \le k \le N \tag{6.1}$$

where $v(k)$ is a WGN uncorrelated with $e(k)$, with intensity $V(k)$

$$E[v(k)v^T(l)] = V(k)\delta(k-l). \tag{6.2}$$

We seek to compute the *smoothed estimate*

$$\hat{x}(k) = E[x(k)|Y] \tag{6.3}$$

where Y is the Hilbert space of random variables spanned by $\{y(s),\ 0 \le s \le N\}$.

Equation (6.1) can be rewritten in matrix form as

$$\boldsymbol{y} = \boldsymbol{Hx} + \boldsymbol{v} \tag{6.4}$$

where

$$y^T = [y^T(0) \cdots y^T(k) \cdots y^T(N)] \quad (6.5a)$$

$$x^T = [x^T(0) \cdots x^T(k) \cdots x^T(N)] \quad (6.5b)$$

$$v^T = [v^T(0) \cdots v^T(k) \cdots v^T(N)] \quad (6.5c)$$

$$H = \operatorname{diag}\{H(k)\}. \quad (6.5d)$$

Denoting

$$\hat{x}^T = [\hat{x}^T(0) \cdots \hat{x}^T(k) \cdots \hat{x}^T(N)] \quad (6.6a)$$

$$R = E[xx^T], \quad V = E[vv^T] \quad (6.6b)$$

the solution of the static estimation problem (6.4) is given by

$$[R^{-1} + H^T V^{-1} H]\hat{x} = H^T V^{-1} y \quad (6.7)$$

or equivalently

$$R^{-1}\hat{x} = H^T V^{-1}(y - H\hat{x}). \quad (6.8)$$

Second-Order Smoother: But it was shown earlier that R^{-1} has the cyclic tridiagonal structure (3.26). From (6.8), we see, therefore, that the smoothed estimate $\hat{x}(k)$ obeys the cyclic second-order descriptor recursions

$$M_0(k)\hat{x}(k) - M_-(k)\hat{x}(k-1) - M_+(k)\hat{x}(k+1) = H^T(k)V^{-1}(k)[y(k) - H(k)\hat{x}(k)] \quad (6.9)$$

for $0 \le k \le N$, where $k-1$ and $k+1$ are defined modulo $N+1$. The smoother (6.9) has a structure similar to that of the Kalman filter for Gauss–Markov processes, since the left-hand side of (6.9) is obtained by applying the operator Λ to the smoothed estimate $\hat{x}(k)$, and the right-hand side consists of a gain $H^T(k)V^{-1}(k)$ multiplying the smoothing residuals

$$v_s(k) = y(k) - H(k)\hat{x}(k) \quad (6.10)$$

which play here the role of the innovations for the Kalman filter.

The boundary conditions for the smoother (6.9) are cyclic. However, when the process $x(k)$ is observed exactly at both ends of the interval I, i.e., when $y(0) = x(0)$ and $y(N) = x(N)$, we can use instead the Dirichlet conditions

$$\hat{x}(0) = y(0), \qquad \hat{x}(N) = y(N) \quad (6.11)$$

which may be more convenient.

With either cyclic or Dirichlet boundary conditions, the smoother (6.9) is *noncausal*. Nevertheless, by noting that the matrix $R^{-1} + H^T V^{-1} H$ appearing on the left-hand side of (6.7) is positive definite cyclic tridiagonal, efficient solution techniques can be used to solve (6.7). These include, for example, cyclic block reduction [25, sect. 5.5], or iterative techniques such as the SOR or preconditioned conjugate gradient methods.

Double-Sweep Solution: Alternatively, we can use the following double-sweep implementation, which relies on the same operator factorization approach as the solution technique of Section V. The first step is to note that in the interior of I, the smoother (6.9) can be rewritten in operator form as

$$\Lambda_s \hat{x}(k) = H^T(k)V^{-1}(k)y(k) \quad (6.12)$$

where

$$\Lambda_s = M_{0s}(k)I - M_-(k)Z^{-1} - M_+(k)Z \quad (6.13a)$$

with

$$M_{0s}(k) = M_0(k) + H^T(k)V^{-1}(k)H(k). \quad (6.13b)$$

Then, using the factorization procedure of Section V, the operator Λ_s can be represented as

$$\Lambda_s = \Omega_s^* Q_s^{-1} \Omega_s \quad (6.14a)$$

with

$$\Omega_s = I - A_s(k-1)Z^{-1}, \qquad Q_s = Q_s(k-1)I. \quad (6.14b)$$

This implies that (6.12) can be decomposed into the two first-order equations

$$z_s(k-1) = A_s^T(k)z_s(k) + H^T(k)V^{-1}(k)y(k), \qquad 1 \le k \le N-1 \quad (6.15a)$$

$$\hat{x}(k+1) = A_s(k)\hat{x}(k) + Q_s(k)z_s(k), \qquad 0 \le k \le N-1 \quad (6.15b)$$

which propagate, respectively, in the backward and forward directions. Let now $z_{0s}(k)$ and $\hat{x}_0(k)$ be the solution of (6.15) with boundary conditions

$$z_{0s}(N-1) = 0, \qquad \hat{x}_0(0) = 0. \quad (6.16)$$

This solution can be computed recursively by performing a double sweep of interval $[0, N-1]$, where (6.15a) and (6.15b) are propagated in succession in the backward and forward directions, respectively.

The smoothed estimate $\hat{x}_0(k)$ which is obtained by this approach does not satisfy the cyclic boundary conditions obtained by setting $k = 0, N$ in (6.9). However, it satisfies (6.9) in the interior of I. Denote by

$$\epsilon(0) = H^T(0)V^{-1}(0)y(0) - [M_{0s}(0)\hat{x}_0(0) - M_-(0)\hat{x}_0(N) - M_+(0)\hat{x}_0(1)] \quad (6.17a)$$

$$\epsilon(N) = H^T(N)V^{-1}(N)y(N) - [M_{0s}(N)\hat{x}_0(N) - M_-(N)\hat{x}_0(N-1) - M_+(N)\hat{x}_0(0)] \quad (6.17b)$$

the mismatch between the true boundary conditions at the ends of I, and the values associated to $\hat{x}_0(k)$. Then, if $\Gamma_{Cs}(k, s)$ is the cyclic Green's function associated to the smoother (6.9), we have

$$\hat{x}(k) = \hat{x}_0(k) + \Gamma_{Cs}(k, 0)\epsilon(0) + \Gamma_{Cs}(k, N)\epsilon(N). \quad (6.18)$$

In other words, the true smoothed estimate $\hat{x}(k)$ is obtained by adding to $\hat{x}_0(k)$ some correction terms representing the mismatch of the cyclic boundary conditions at both ends of I.

The double-sweep recursions (6.15)–(6.16) are also applicable to the case when the Dirichlet conditions (6.11) are imposed on the smoother. However, the correction procedure (6.17)–(6.18) for handling the boundary conditions needs to be modified. Specifically, the mismatch between the Dirichlet conditions (6.11) and the values specified for $\hat{x}_0(k)$ is now given by

$$\epsilon(0) = y(0) - \hat{x}_0(0), \qquad \epsilon(N) = y(N) - \hat{x}_0(N) \quad (6.19)$$

and the correction formula (6.18) remains valid, provided that the cyclic Green's function Γ_{Cs} is replaced by the Dirichlet Green's function Γ_s for (6.9).

To summarize, the double-sweep smoothing technique requires, first, the solution of a Riccati equation to solve the factorization problem (6.14) for $Q_s(\cdot)$ and $A_s(\cdot)$. Then, the backward and forward filters (6.15a) and (6.15b) are propagated in succession. Finally, the true boundary conditions are taken into account by performing the correction (6.18), which requires the precomputation of $\Gamma_{Cs}(k, 0)$ and $\Gamma_{Cs}(k, N)$, or of the corresponding values of Γ_s. As mentioned earlier, this smoothing procedure can be viewed as an extension of the Rauch–Tung–Striebel smoothing method for Markov processes, which relies on a similar double sweep idea. A smoothing procedure of this type was also derived recently for first-order two-point boundary-value descriptor systems driven by white noise [26].

Smoothing Error: Subtracting (6.9) from the second-order descriptor model (3.1) for $x(k)$, we find that the smoothing error

$$\tilde{x}(k) = x(k) - \hat{x}(k) \tag{6.20}$$

has the cyclic second-order model

$$M_{0s}(k)\tilde{x}(k) - M_-(k)\tilde{x}(k-1) - M_+(k)\tilde{x}(k+1) = e_s(k) \tag{6.21a}$$

with driving noise

$$e_s(k) = e(k) - H^T(k)V^{-1}(k)v(k). \tag{6.21b}$$

Since $e(k)$ and $v(k)$ are uncorrelated, it is easy to check that $e_s(k)$ has the covariance structure (3.4) for model (6.21a). This implies the following.

Theorem 6.1: The smoothing error process $\tilde{x}(k)$ associated to a reciprocal process $x(k)$ with full rank noise is itself reciprocal.

This result extends to reciprocal processes a result which was only shown a few years ago [27], [28], [24] for Markov processes, namely, that their smoothing error is a Markov process.

Markov Processes: If $x(k)$ is a Markov process described by the first-order state-space model (4.1), it is known [15, p. 190] that the smoothed estimate $\hat{x}(k)$ satisfies the *Hamiltonian system*

$$\begin{bmatrix} I & -Q(k) \\ 0 & A^T(k) \end{bmatrix} \begin{bmatrix} \hat{x}(k+1) \\ \hat{\lambda}(k+1) \end{bmatrix} = \begin{bmatrix} A(k) & 0 \\ H^T(k)V^{-1}(k)H(k) & I \end{bmatrix} \cdot \begin{bmatrix} \hat{x}(k) \\ \hat{\lambda}(k) \end{bmatrix} + \begin{bmatrix} 0 \\ -H^T(k)V^{-1}(k) \end{bmatrix} y(k) \tag{6.22}$$

for $0 \le k \le N-1$, with boundary conditions

$$\Pi_b^{-1}\hat{x}(0) - \hat{\lambda}(0) = 0 \tag{6.23a}$$

$$\hat{\lambda}(N) = H^T(N)V^{-1}(N)[y(N) - H^T(N)V^{-1}(N)\hat{x}(N)]. \tag{6.23b}$$

Assuming that the noise variance $Q(k)$ is invertible, and eliminating $\hat{\lambda}(k)$ from (6.22)–(6.23) yields the second-order smoother (6.9), where $M_0(k)$ and $M_\pm(k)$ are the matrices obtained in (4.11) and (4.17) for the cyclic descriptor model of a Markov process. Thus, the smoothing results that we have derived for reciprocal processes are consistent with earlier results for Markov processes.

Interpolation Problem: The smoothing results that we have obtained previously can also be used to solve the interpolation problem [19, sect. 2.11], [29] for $x(k)$. Thus, assume that we do not have any observation $y(k)$ over a subinterval $J = [K, L]$ of I. The interpolation problem consists of finding the estimate $\hat{x}(k)$ over this subinterval based on the remaining observations. It is easy to see that the solution of this problem is obtained by setting $H(k) = 0$ in the smoothing equation (6.9) over the subinterval J. Note, that in this case, the double sweep solution technique still needs to be applied to the whole interval I. This approach should be contrasted with the one employed in [29] for Markov processes, which relied on a two-filter approach. Specifically, it was shown that the interpolated estimate $\hat{x}(k)$ at a point of J is a linear combination of the forward filtered estimate $\hat{x}_f(K-1)$ and the backward filtered estimate $\hat{x}_b(L+1)$. The extension of this implementation to reciprocal processes is currently under investigation.

VII. Conclusions

In this paper, a characterization of reciprocal processes has been obtained in terms of second-order nearest-neighbor or descriptor models driven by locally correlated noise. This last feature is consistent with earlier results of Krener [12] for continuous-time reciprocal processes, and of Woods [3] for discrete Markov random fields. By observing that the inverse covariance of a reciprocal process has a cyclic block tridiagonal structure, we have also been able to derive a second-order smoother whose structure is similar to that of the Kalman filter for Markov processes, and which can be solved recursively by a double sweep procedure.

Since reciprocal processes are 1-D Markov random fields, the results of this paper can be extended almost directly to multidimensional Markov random fields. In fact, in the 2-D case, by considering 2-D nearest-neighbor models of Markov random fields with locally correlated noise of the type considered here and in [3], we have been able to show recently [30] that the corresponding smoother satisfies a second-order nearest-neighbor equation of the form examined in Section VI. This result should be contrasted with the one obtained in [15], [31] for 2-D nearest-neighbor models driven by *white noise*, where it was shown that the smoother satisfies a Hamiltonian system consisting of two coupled second-order equations, thus yielding effectively a fourth-order smoother. This means that in spite of their apparent complexity, which is associated with the local correlation structure of the driving noise, 2-D nearest-neighbor models of Markov random fields lead, in fact, to simple estimators.

Acknowledgment

The authors would like to express their gratitude to Dr. M. Clark of Imperial College for his participation in lively discussions on the properties of reciprocal processes during his stay at UCD in 1988–1989. They also thank an anonymous reviewer who pointed out the interpretation of the residual process $e(k)$ in (3.1) as the conjugate process of $x(k)$.

References

[1] P. Lévy, "A special problem of Brownian motion, and a general theory of Gaussian random functions," in *Proc. 3rd Berkeley Symp. on Mathematical Statistics and Probability*, Univ. Calif. Press, Berkeley, CA, 1956, vol. 2, pp. 133–175.

[2] E. Wong, "Homogeneous Gauss-Markov random fields," *Ann. Math. Statist.*, vol. 40, no. 5, pp. 1625–1634, 1969.

[3] J. W. Woods, "Two-dimensional discrete-Markovian fields," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 2, pp. 232–240, Mar. 1972.

[4] S. Bernstein, "Sur les liaisons entre les grandeurs aleatoires," in *Proc. Int. Cong. of Math.*, Zurich, Switzerland, 1932, pp. 288–309.

[5] D. Slepian, "First passage time for a particular Gaussian process," *Ann. Math. Statist.*, vol. 32, pp. 610–612, 1961.

[6] B. Jamison, "Reciprocal processes: The stationary Gaussian case," *Ann. Math. Statist.*, vol. 41, no. 5, pp. 1624–1630, 1970.

[7] ——, "Reciprocal processes," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 30, pp. 65–86, 1974.

[8] S. C. Chay, "On quasi-Markov random fields," *J. Multivariate Anal.*, vol. 2, pp. 14–76, 1972.

[9] J. Abrahams and J. B. Thomas, "Some comments on conditionally Markov and reciprocal Gaussian processes," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 4, pp. 523–525, July 1981.

[10] J. P. Carmichael, J. C. Masse, and R. Theodorescu, "Processus Gaussiens stationaires réciproques sur un intervalle," *Comptes Rendus Acad. Sci., Paris, Serie I*, vol. 295, pp. 291–293, Sept. 1982.

[11] ——, "Multivariate reciprocal stationary Gaussian processes," *J. Multivariate Anal.*, vol. 23, pp. 47–66, 1987.

[12] A. J. Krener, "Reciprocal diffusions and stochastic differential equations of second order," *Stochastics*, vol. 24, pp. 393–422, 1988.

[13] A. J. Krener, R. Frezza, and B. C. Levy, "Gaussian reciprocal processes and selfadjoint differential equations of second order," preprint, Nov. 1989.

[14] A. J. Krener, "Reciprocal processes and the stochastic realization problem for acausal systems," in *Modeling, Identification and Control*, C. I. Byrnes and A. Lindquist, Eds. Amsterdam, The Netherlands: North-Holland, 1986, pp. 197–211.

[15] M. B. Adams, "Linear estimation of boundary value stochastic processes," Ph.D. dissertation, Dep. Aeronaut. Astronaut., and Lab. Inform. Decision Syst., M.I.T., Cambridge, MA, Rep. LIDS-TH-1295, Jan. 1983.

[16] M. B. Adams, A. S. Willsky, and B. C. Levy, "Linear estimation of boundary-value stochastic processes, Parts 1 and 2," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 803-821, Sept. 1984.

[17] D. A. Castanon, B. C. Levy, and A. S. Willsky, "Algorithms for the incorporation of predictive information in surveillance theory," *Int. J. Syst. Sci.*, vol. 16, pp. 367-382, Mar. 1985.

[18] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear systems," *J. AIAA*, vol. 3, no. 8, pp. 1445-1450, 1965.

[19] Yu. A. Rozanov, *Stationary Random Processes*. San Francisco, CA: Holden-Day, 1967.

[20] P. Masani, "The prediction theory of multivariate stochastic processes, III: Unbounded spectral densities," *Acta Mathematica*, vol. 104, pp. 141-162, 1960.

[21] G. Ruckebusch, "Représentations Markoviennes de processus Gaussiens stationaires," Thèse de 3ème cycle, Université de Paris VI, 1975.

[22] M. Pavon, "The conjugate process in stochastic realization theory," *Math. Programming Study*, vol. 18, pp. 12-26, 1982.

[23] L. D. Pitt, "A Markov property for Gaussian processes with a multi-dimensional parameter," *Arch. Rat. Mech. Anal.*, vol. 43, pp. 367-391, 1971.

[24] R. Ackner and T. Kailath, "Discrete-time complementary models and smoothing," *Int. J. Contr.*, vol. 49, no. 5, pp. 1665-1682, May 1989.

[25] G. H. Golub and C. H. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1983.

[26] R. Nikoukhah, "A deterministic and stochastic theory for two-point boundary-value descriptor systems," Ph.D. dissertation, Dep. Elec. Eng. Comput. Sci., and Lab. Inform. Decision Syst., M.I.T., Cambridge, MA, Rep. LIDS-TH-1820, Sept. 1988.

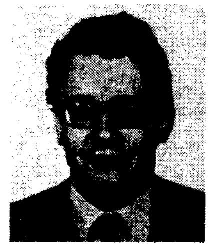
[27] F. A. Badawi, A. Lindquist, and M. Pavon, "A stochastic realization approach to the smoothing problem," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 878-887, Dec. 1979.

[28] M. G. Bello, A. S. Willsky, B. C. Levy, and D. A. Castanon, "Smoothing error dynamics and their use in the solution of smoothing and mapping problems," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 4, pp. 483-495, July 1986.

[29] M. Pavon, "Optimal interpolation for linear stochastic systems," *SIAM J. Contr. Optimiz.*, vol. 22, no. 4, pp. 618-629, July 1984.

[30] B. C. Levy, "Noncausal estimation for discrete Gauss-Markov random fields," to be published in *Proc. Int. Symp. Math. Theory Networks Syst.* (*MTNS-89*), M. A. Kaashoek, A. C. M. Ran, and J. H. Van Schuppen, Eds. Cambridge, MA: Birkhauser Boston, Dec. 1990.

[31] B. C. Levy, M. B. Adams, and A. S. Willsky, "Solution and linear estimation of 2-D nearest-neighbor models," *Proc. IEEE*, vol. 78, no. 4, pp. 627-641, Apr. 1990.

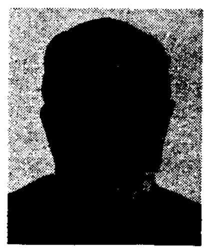


Bernard C. Levy (S'74–S'78–M'78–SM'90) was born in Princeton, NJ, on July 31, 1951. He received the diploma of Ingénieur Civil des Mines from the Ecole Nationale Supérieure des Mines, Paris, France, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

While at Stanford University, he held an INRIA Fellowship, and worked also as Teaching Assistant, Research Assistant, and Instructor. From June 1979 to June 1987, he was Assistant, and then Associate Professor in the Department of Electrical Engineering and Computer Science at M.I.T. Since July 1987, he has been with the University of California, Davis, where he is currently Professor of Electrical Engineering. His research interests are in the areas of multidimensional and statistical signal processing, inverse problems, estimation, detection, and scientific computing.

Ruggero Frezza was born in Pordenone, Italy, on November 25, 1961. He received the Laurea in electrical engineering from the University of Padua, Padua, Italy, in 1985, and the M.S. and Ph.D. degrees in applied mathematics in 1987 and 1990, respectively, from the University of California, Davis.

From September 1985 to April 1990, he was employed at the University of California, Davis, as a Teaching Assistant, Associate, and Post Graduate Researcher. Since May 1990, he has been a Researcher in the Department of Electrical Engineering at the University of Padua, Italy. His research interests are in the areas of systems and control theory, stochastic processes, and estimation.



Arthur J. Krener (M'77–SM'82–F'90) was born in Brooklyn, NY, in 1942. He received the B.S. degree from Holy Cross College in 1964 and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1967 and 1971, respectively, all in mathematics.

At present, he is Professor and Chair of the Department of Mathematics at the University of California, Davis. His current research interests include linear and nonlinear systems and stochastic processes. He has held visiting positions at Harvard University, the University of Rome, Imperial College, NASA Ames Research Center, the University of California at Berkeley, and the University of Paris.