UC Irvine
UC Irvine Previously Published Works

Title
Optimal Estimation With Missing Observations via Balanced Time-Symmetric Stochastic Models

Authors
Georgiou, Tryphon T
Lindquist, Anders

Peer reviewed

Optimal estimation with missing observations via balanced time-symmetric stochastic models

Tryphon T. Georgiou, *Fellow, IEEE* and Anders Lindquist, *Life Fellow, IEEE*

Abstract—We consider data fusion for the purpose of smoothing and interpolation based on observation records with missing data. Stochastic processes are generated by linear stochastic models. The paper begins by drawing a connection between time reversal in stochastic systems and all-pass extensions. A particular normalization (choice of basis) between the two time-directions allows the two to share the same *orthonormalized* state process and simplifies the mathematics of data fusion. In this framework we derive symmetric and balanced Mayne-Fraser-like formulas that apply simultaneously to smoothing and interpolation.

I. INTRODUCTION

Data fusion is the process of integrating different data sets, or statistics, into a more accurate representation for a quantity of interest. A case in point in the context of systems and control is provided by the Mayne-Fraser two-filter formula [1], [2] in which the estimates generated by two different filters are merged into a combined more reliable estimate in fixed-interval smoothing. The purpose of this paper is to develop such a two-filter formula that is universally applicable to smoothing and interpolation based on general records with missing observations.

In [3], [4] the Mayne-Fraser formula was analyzed in the context of stochastic realization theory and was shown that it can be formulated in terms a forward and a backward Kalman filter. In a subsequent series of papers, Pavon [5], [6] addressed in a similar manner the hitherto challenging problem of interpolation [7], [8], [9], [10]. This latter problem consists of reconstructing missing values of a stochastic process over a given interval. In departure from the earlier statistical literature, [5], [6] considered a stationary process with rational spectral density and, therefore, reliazable as the output of a linear stochastic system. Interpolation was then cast as seeking an estimate of the state process based on an incomplete observation record. A basic tool in these works is the concept of time-reversal in stochastic systems which has been central in stochastic realization theory (see, e.g., [11], [12], [13], [14], [5], [6], [15], [16], [17]). For a recent overview of smoothing and interpolation theory in the context of stochastic realization theory see [18, Chapter 15].

In the present paper we are taking this program several steps further. Given intermittent observations of the output of a linear stochastic system over a finite interval, we want to determine the linear least-squares estimate of the state of the system in an arbitrary point in the interior of the interval, which may either be in a subinterval of missing data or in one where observations are available. Hence, this combines smoothing and interpolation over general patterns of available observations. Our main interest is in continuous-time (possibly time-varying) systems. However, the absence of data over subintervals, depending on the information pattern, may necessitate a hybrid approach involving discrete-time filtering steps.

In studying the statistics of a process over an interval, it is natural to decompose the interface between past and future in a time-symmetric manner. This gives rise to systems representations of the process running in either time direction, forward or backward in time. This point was fundamental in early work in stochastic realization; see [18] and references therein. In a different context [19] a certain duality between the two time-directions in modeling a stochastic process was introduced in order to characterize solutions to moment problems. In this new setting the noise-process was general (not necessarily white), and the correspondence between the driving inputs to the two time-opposite models was shown to be captured by suitable dual all-pass dynamics.

Here, we begin by combining these two sets of ideas to develop a general framework where two time-opposite stochastic systems model a given stochastic process. We study the relationship between these systems and the corresponding processes. In particular, we recover as a special case certain results of stochastic realization theory [11], [5], [6], [4] from the 1970's using a novel procedure. This theory provides a normalized and balanced version of the forward-backward duality which is essential for our new formulation of the two-filter Mayne-Fraser-like formula uniformly applicable to intervals with or without observations.

The paper is structured as follows. In Section II we explain how a lifting of state-dynamics into an all-pass system allows direct correspondence between sample-paths of driving

generating processes, in opposite time-directions, via causal and anti-causal mappings, respectively. This is most easily understood and explained in discrete-time and hence we begin with that. In Section III we utilize this mechanism in the context of general output processes and, similarly, introduce a pair of time-opposite models. These two introductory sections, II and III, deal with stationary models for simplicity and are largely based on [20]. The corresponding generalizations to time-varying systems are given in Section IV and in the appendix, in continuous and discrete-time, respectively. In Section V we explain Kalman filtering for problems with missing information in the continuous-time setting. In this, we first consider the case where increments of the output process across intervals of no information are unavailable as a simplified preliminary, after which we focus on the central problem where the output process is the object of observation. Section VI deals with the geometry of information fusion. In Section VII we present a generalized balanced two-filter formula that applies uniformly over intervals where data is or is not available. We summarize the computational steps of this approach in Section VIII. Finally, we highlight the use of the two-filter formula with a numerical example given in Section IX and provide concluding remarks in Section X.

II. STATE DYNAMICS AND ALL-PASS EXTENSION

In this paper we consider discrete-time as well as continuous-time stochastic linear state-dynamics. We begin by explaining basic ideas in a stationary setting. In discrete-time systems take the form of a set of difference equations

$$x(t+1) = Ax(t) + Bw(t) \tag{1}$$

where $t \in \mathbb{Z}$, $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times p}$, A has all eigenvalues in the open unit disc $\mathbb{D} = \{z \mid |z| < 1\}$, and $w(t), x(t)$ are (centered) stationary vector-valued stochastic processes with $w(t)$ normalized white noise; i.e.,

$$\mathrm{E}\{w(t)w(s)'\} = I_p \delta_{ts}, \tag{2}$$

where E denotes mathematical expectation. The system of equations is assumed to be reachable, i.e.,

$$\mathrm{rank}\left[B, AB, \dots A^{n-1}B\right] = n. \tag{3}$$

In continuous-time, state-dynamics take the form of a system of stochastic differential equations

$$dx(t) = Ax(t)dt + Bdw(t) \tag{4}$$

where, here, $x(t)$ is a stationary continuous-time vector-valued stochastic process and $w(t)$ is a vector-valued process with orthogonal increments with the property

$$\mathrm{E}\{dwdw'\} = I_p dt, \tag{5}$$

where I_p is the $p \times p$ identity matrix. Reachability of the pair (A, B) is also assumed throughout and the condition for

this is identical to the one for discrete-time given above (as is well known). In continuous time, stability of the system of equations is equivalent to A having only eigenvalues with negative real part.

In either case, discrete-time or continuous-time, it is possible to define an output equation so that the overall system is all-pass. This is done next.

A. All-pass extension in discrete-time

Consider the discrete-time Lyapunov equation

$$P = APA' + BB'. \tag{6}$$

Since A has all eigenvalues inside the unit disc of the complex plane and (3) holds, (6) has as solution a matrix P which is positive definite. The state transformation

$$\xi = P^{-\frac{1}{2}}x, \tag{7}$$

and

$$F = P^{-\frac{1}{2}}AP^{\frac{1}{2}}, \ G = P^{-\frac{1}{2}}B, \tag{8}$$

brings (1) into

$$\xi(t+1) = F\xi(t) + Gw(t). \tag{9}$$

For this new system, the corresponding Lyapunov equation $X = FXF' + GG'$ has I_n as solution, where I_n denotes the $(n \times n)$ identity matrix. This fact, namely, that

$$I_n = FF' + GG' \tag{10}$$

implies that this $[F, G]$ can be embedded as part of an orthogonal matrix

$$U = \begin{bmatrix} F & G \\ H & J \end{bmatrix}, \tag{11}$$

i.e., a matrix such that $UU' = U'U = I_{n+p}$.

Define the transfer function

$$\mathbf{U}(z) := H(zI_n - F)^{-1}G + J \tag{12}$$

corresponding to

$$\xi(t+1) = F\xi(t) + Gw(t) \tag{13a}$$
$$\bar{w}(t) = H\xi(t) + Jw(t). \tag{13b}$$

This is also the transfer function of

$$x(t+1) = Ax(t) + Bw(t) \tag{14a}$$
$$\bar{w}(t) = \bar{B}'x(t) + Jw(t), \tag{14b}$$

where $\bar{B} := P^{-\frac{1}{2}}H'$, since the two systems are related by a similarity transformation. Hence,

$$\mathbf{U}(z) = \bar{B}'(zI_n - A)^{-1}B + J. \tag{15}$$

We claim that $\mathbf{U}(z)$ is a stable all-pass transfer function (with respect to the unit disc), i.e., that $\mathbf{U}(z)$ is a transfer function of a stable system and that

$$\mathbf{U}(z)\mathbf{U}(z^{-1})' = \mathbf{U}(z^{-1})'\mathbf{U}(z) = I_p. \tag{16}$$

The latter claim is immediate after we observe that, since $U'U = I_{n+p}$,

$$U' \left[\begin{array}{c} \xi(t+1) \\ \bar{w}(t) \end{array} \right] = \left[\begin{array}{c} \xi(t) \\ w(t) \end{array} \right],$$

and hence,

$$\xi(t) = F'\xi(t+1) + H'\bar{w}(t) \tag{17a}$$
$$w(t) = G'\xi(t+1) + J'\bar{w}(t) \tag{17b}$$

or, equivalently,

$$x(t) = PA'P^{-1}x(t+1) + P^{\frac{1}{2}}H'\bar{w}(t) \tag{18a}$$
$$w(t) = B'P^{-1}x(t+1) + J'\bar{w}(t). \tag{18b}$$

Setting

$$\bar{x}(t) := P^{-1}x(t+1), \tag{19}$$

(18) can be written

$$\bar{x}(t-1) = A'\bar{x}(t) + \bar{B}\bar{w}(t) \tag{20a}$$
$$w(t) = B'\bar{x}(t) + J'\bar{w}(t) \tag{20b}$$

with transfer function

$$\mathbf{U}(z)^* = B'(z^{-1}I_n - A')^{-1}\bar{B} + J'. \tag{21}$$

Either of the above systems inverts the dynamical relation $w \to \bar{w}$ (in (14) or (13)).
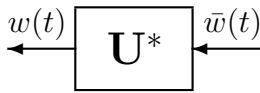

Fig. 1: Realization (14) in the forward time-direction.

Fig. 2: Realization (20) in the backward time-direction.

An algebraic proof of (16) is also quite immediate. In fact,

$$\mathbf{U}(z)\mathbf{U}(z^{-1})'$$
$$= \left[H(zI_n - F)^{-1}G + J\right]\left[H(z^{-1}I_n - F)^{-1}G + J\right]'$$
$$= H(zI_n - F)^{-1}GG'(z^{-1}I_n - F')^{-1}H' + JJ'$$
$$+ H(zI_n - F)^{-1}GJ' + JG'(z^{-1}I_n - F')^{-1}H$$

Now, using the identity

$$I_n - FF' = (zI_n - F)(z^{-1}I_n - F')$$
$$+ (zI_n - F)F' + F(z^{-1}I_n - F'),$$

(10) and $GJ' = -FH'$, obtained from $UU' = I_{n+p}$, this yields

$$\mathbf{U}(z)\mathbf{U}(z^{-1})' = HH' + JJ' = I_{n+p},$$

as claimed.

B. All-pass extension in continuous-time

Consider the continuous-time Lyapunov equation

$$AP + PA' + BB' = 0. \tag{22}$$

Since A has all its eigenvalues in the left half of the complex plane and since (3) holds, (22) has as solution a positive definite matrix P. Once again, applying (7-8), the system in (4) becomes

$$d\xi(t) = F\xi(t)dt + Gdw(t). \tag{23a}$$

We now seek a completion by adding an output equation

$$d\bar{w}(t) = H\xi(t)dt + Jdw(t) \tag{23b}$$

so that the transfer function

$$\mathbf{U}(s) := H(sI_n - F)^{-1}G + J \tag{24}$$

is all-pass (with respect to the imaginary axis), i.e.,

$$\mathbf{U}(s)\mathbf{U}(-s)' = \mathbf{U}(-s)'\mathbf{U}(s) = I_p. \tag{25}$$

For this new system, the corresponding Lyapunov equation has as solution the identity matrix and hence,

$$F + F' + GG' = 0. \tag{26}$$

Utilizing this relationship we note that

$$(sI_n - F)^{-1}GG'(-sI_n - F')^{-1}$$
$$= (sI_n - F)^{-1}(sI_n - F - sI_n - F')(-sI_n - F')^{-1}$$
$$= (sI_n - F)^{-1} + (-sI_n - F')^{-1},$$

and we calculate that

$$\mathbf{U}(s)\mathbf{U}(-s)'$$
$$= (H(sI_n - F)^{-1}G + J)(G'(-sI_n - F')^{-1}H' + J')$$
$$= JJ' + H(sI_n - F)^{-1}(GJ' + H')$$
$$\quad (JG' + H)(-sI_n - F')^{-1}H'.$$

For the product to equal the identity,

$$JJ' = I_p$$
$$H = -JG'.$$

Thus, we may take

$$J = I_p$$
$$H = -G',$$

and the forward dynamics

$$d\xi(t) = F\xi(t)dt + Gdw(t) \tag{27a}$$
$$d\bar{w}(t) = -G'\xi(t)dt + dw(t). \tag{27b}$$

Substituting $F = -F' - GG'$ from (26) into (27a) we obtain the reverse-time dynamics

$$d\xi(t) = -F'\xi(t)dt + Gd\bar{w}(t) \tag{28a}$$
$$dw(t) = G'\xi(t)dt + d\bar{w}(t). \tag{28b}$$

Now defining

$$\bar{x}(t) := P^{-1}x(t) \tag{29}$$

and using (7) and (8), (28) becomes

$$d\bar{x}(t) = -A'\bar{x}(t)dt + \bar{B}d\bar{w}(t) \tag{30a}$$
$$dw(t) = B'\bar{x}(t)dt + d\bar{w}(t), \tag{30b}$$

with transfer function

$$\mathbf{U}(s)^* = I_p + B'(sI_n + A')^{-1}\bar{B}, \tag{31}$$

where

$$\bar{B} := P^{-1}B. \tag{32}$$

Furthermore, the forward dynamics (27) can be expressed in the form

$$dx(t) = Ax(t)dt + Bdw(t) \tag{33a}$$
$$d\bar{w}(t) = \bar{B}'x(t)dt + dw(t) \tag{33b}$$

with transfer function

$$\mathbf{U}(s) = I_p - \bar{B}'(sI_n - A)^{-1}B. \tag{34}$$

III. Time-reversal of stationary linear stochastic systems

The development so far allows us to draw a connection between two linear stochastic systems having the same output and driven by a pair of arbitrary, but dual, stationary processes $w(t)$ and $\bar{w}(t)$, one evolving forward in time and one evolving backward in time. When one of these two processes is white noise (or, orthogonal increment process, in continuous-time), then so is the other. For this special case we recover results of [11] and [5], [6] in stochastic realization theory.

A. Time-reversal of discrete-time stochastic systems

Consider a stochastic linear system

$$x(t + 1) = Ax(t) + Bw(t) \tag{35a}$$
$$y(t) = Cx(t) + Dw(t) \tag{35b}$$

with an m-dimensional output process y, and x, u, A, B are defined as in Section II-A. All processes are stationary and the system can be thought as evolving forward in time from the remote past ($t = -\infty$).

To formalize this, we introduce some notation. Let $\mathbf{H}$ be the Hilbert space spanned by $\{w_k(t); t \in \mathbb{Z}, k = 1, 2, \ldots, n\}$, endowed with the inner product $\langle \lambda, \mu \rangle = \mathrm{E}\{\lambda\mu\}$, and let $\mathbf{H}_t^-(w)$ and $\mathbf{H}_t^+(w)$ be the (closed) subspaces spanned by $\{w_k(s); s \leq t - 1, k = 1, \ldots, m\}$ and $\{w_k(s); s \geq t, k = 1, \ldots, m\}$, respectively. Define $\mathbf{H}_t^-(y)$ and $\mathbf{H}_t^+(y)$ accordingly in terms of the output process process y. Then the stochastic system (35) evolves forward in time in the sense that

$$\mathbf{H}_t^-(z) \subset \mathbf{H}_t^-(w) \perp \mathbf{H}_t^+(w), \tag{36}$$

where $\mathbf{A} \perp \mathbf{B}$ means that elements of the subspaces $\mathbf{A}$ and $\mathbf{B}$ are mutually orthogonal, and where $\mathbf{H}_t^-(z)$ is formed as above in terms of

$$z(t) = \begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix};$$

see [18, Chapter 6] for more details.

Next we construct a stochastic system

$$\bar{x}(t - 1) = A'\bar{x}(t) + \bar{B}\bar{w}(t) \tag{37a}$$
$$y(t) = \bar{C}\bar{x}(t) + \bar{D}\bar{w}(t), \tag{37b}$$

which evolves backward in time from the remote future ($t = \infty$) in the sense that the processes $\bar{x}, x, \bar{w}, w$ relate as in the previous section. More specifically, as shown in Section II-A, $\mathbf{H}^-(\bar{w}) \subset \mathbf{H}^-(w)$ and $\mathbf{H}^+(w) \subset \mathbf{H}^+(\bar{w})$ for all t, as examplified in Figures 1 and 2.

In fact, the all-pass extension (14) of (35a) yields

$$\bar{w}(t) = \bar{B}'x(t) + Jw(t) \tag{38}$$

It follows from (20b) that (38) can be inverted to yield

$$w(t) = B'\bar{x}(t) + J'\bar{w}(t), \tag{39}$$

where $\bar{x}(t) = P^{-1}x(t+1)$, and that we have the reverse-time recursion

$$\bar{x}(t - 1) = A'\bar{x}(t) + \bar{B}\bar{w}(t). \tag{40a}$$

Then inserting (39) and

$$x(t) = P\bar{x}(t - 1) = PA'\bar{x}(t) + P\bar{B}\bar{w}(t)$$

into (35b), we obtain

$$y(t) = \bar{C}\bar{x}(t) + \bar{D}\bar{w}(t), \tag{40b}$$

where $\bar{D} := CP\bar{B} + DJ'$ and

$$\bar{C} := CPA' + DB'. \tag{41}$$

Then, (40) is precisely what we wanted to establish.

The white noise w is normalized in the sense of (2). Since $\mathbf{U}$, given by (15), is all-pass, $\bar{w}$ is also a normalized white noise process, i.e.,

$$\mathrm{E}\{\bar{w}(t)\bar{w}(s)'\} = I_p\delta_{t-s}.$$

From the reverse-time recursion (37a)

$$\bar{x}(t) = \sum_{k=t+1}^{\infty} (A')^{k-(t+1)}\bar{B}\bar{w}(k).$$

Since, $\bar{w}$ is a white noise process, $\mathrm{E}\{\bar{x}(t)\bar{w}(s)'\} = 0$ for all $s \leq t$. Consequently, (37) is a backward stochastic realization in the sense defined above.

Moreover, the transfer functions

$$\mathbf{W}(z) = C(zI_n - A)^{-1}B + D \tag{42}$$

of (35) and

$$\bar{\mathbf{W}}(z) = \bar{C}(z^{-1}I_n - A')^{-1}\bar{B} + \bar{D} \tag{43}$$

of (37) satisfy

$$\mathbf{W}(z) = \bar{\mathbf{W}}(z)\mathbf{U}(z). \tag{44}$$

In the context of stochastic realization theory, $\mathbf{U}(z)$ is called *structural function* ([13], [14]).

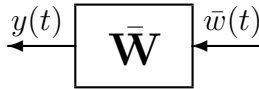Fig. 3: The forward stochastic system (35).

Fig. 4: The backward stochastic system (37)

B. Time-reversal of continuous-time stochastic systems

We now turn to the continuous-time case. Let

$$dx = Ax\,dt + B\,dw \tag{45a}$$
$$dy = Cx\,dt + D\,dw \tag{45b}$$

be a stochastic system with x, w, A, B as in Section II-B, evolving forward in time from the remote past ($t = -\infty$). Now let $\mathbf{H}$ be the Hilbert space spanned by the *increments* of the components of w on the real line $\mathbb{R}$, endowed with the

same inner product as above, and let $\mathbf{H}_t^-(du)$ and $\mathbf{H}_t^+(du)$ be the (closed) subspaces spanned by the increments of the components of U on $(-\infty, t]$ and $[t, \infty)$, respectively. Define $\mathbf{H}_t^-(dy)$ and $\mathbf{H}_t^+(dy)$ accordingly in terms of the output process y. All processes have stationary increments and the stochastic system (45) evolves forward in time in the sense that

$$\mathbf{H}_t^-(dz) \subset \mathbf{H}_t^-(dw) \perp \mathbf{H}_t^+(dw), \tag{46}$$

where $\mathbf{H}_t^-(dz)$ is formed in terms of

$$z(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}. \tag{47}$$

The all-pass extension of Section II-B yields

$$d\bar{w} = dw - \bar{B}'x\,dt \tag{48}$$

as well as the reverse-time relation

$$d\bar{x} = -A'\bar{x}\,dt + \bar{B}\,d\bar{w} \tag{49a}$$
$$dw = B'\bar{x}\,dt + d\bar{w}, \tag{49b}$$

where $\bar{x}(t) = P^{-1}x(t)$. Inserting (49b) into

$$dy = CP\bar{x}\,dt + D\,dw$$

yields

$$dy = \bar{C}\bar{x}\,dt + D\,d\bar{w},$$

where

$$\bar{C} = CP + DB'. \tag{50}$$

Thus, the reverse-time system is

$$d\bar{x} = -A'\bar{x}\,dt + \bar{B}\,d\bar{w} \tag{51a}$$
$$dy = \bar{C}\bar{x}\,dt + D\,d\bar{w}. \tag{51b}$$

From this, we deduce that the system (45) has the backward property

$$\mathbf{H}_t^+(d\bar{z}) \subset \mathbf{H}_t^+(d\bar{w}) \perp \mathbf{H}_t^-(d\bar{w}), \tag{52}$$

where $\mathbf{H}_t^+(d\bar{z})$ is formed as above in terms of

$$\bar{z}(t) = \begin{bmatrix} \bar{x}(t) \\ y(t) \end{bmatrix}.$$

We also note that the transfer function

$$\mathbf{W}(s) = C(sI_n - A)^{-1}B + D$$

of (45) and the transfer function

$$\bar{\mathbf{W}}(s) = \bar{C}(sI_n + A')^{-1}\bar{B} + D$$

of (51) also satisfy

$$\mathbf{W}(s) = \bar{\mathbf{W}}(s)\mathbf{U}(s)$$

as in discrete-time.

Note that the orthogonal-increment process w is normalized in the sense of (5). Since $\mathbf{U}(s)$ is all-pass,

$$d\bar{w} = du - \bar{B}'x\,dt \tag{53}$$

also defines a stationary orthogonal-increment process $\bar{w}$ such that

$$\{d\bar{w}(t)d\bar{w}(t)'\} = I_p dt.$$

It remains to show that (51) is a backward stochastic realization, that is, at each time t the past increments of $\bar{w}$ are orthogonal to $\bar{x}(t)$. But this follows from the fact that

$$\bar{x}(t) = \int_t^\infty e^{-A'(t-s)} \bar{B} d\bar{w}(s)$$

and $\bar{w}$ has orthogonal increments.

IV. TIME REVERSAL OF NON-STATIONARY STOCHASTIC SYSTEMS

In a similar manner non-stationary stochastic systems admit unitary extensions which in turn allows us to construct dual time-reversed stochastic models that share the same state process. The case of discrete-time dynamics is documented in the appendix, whereas the continuous-time counterpart is explained next as prelude to smoothing and interpolation that will follow.

A. Unitary extension

The covariance matrix function $P(t) := \mathrm{E}\{x(t)x(t)'\}$ of the time-varying state representation

$$dx = A(t)x(t)dt + B(t)dw, \quad x(0) = x_0 \quad (54)$$

with x_0 a zero-mean stochastic vector with covariance matrix $P_0 = \mathrm{E}\{x_0 x_0'\}$, satisfies the matrix-valued differential equation

$$\dot{P}(t) = A(t)P(t) + P(t)A(t)' + B(t)B(t)' \quad (55)$$

with $P(0) = P_0$. Throughout we assume total reachability [18, Section 15.2], and therefore $P(t) > 0$ for all $t > 0$.

A unitary extension of (54) is somewhat more complicated than in the discrete time case. In fact, differentiating

$$\xi(t) = P(t)^{-\frac{1}{2}} x(t) \quad (56)$$

we obtain

$$d\xi = F(t)\xi(t)dt + G(t)dw, \quad (57)$$

where

$$F(t) = P(t)^{-\frac{1}{2}} A(t)P(t)^{\frac{1}{2}} + R(t), \quad (58a)$$
$$G(t) = P(t)^{-\frac{1}{2}} B(t) \quad (58b)$$

with

$$R(t) = \left[\frac{d}{dt}P(t)^{-\frac{1}{2}}\right] P(t)^{\frac{1}{2}}. \quad (59)$$

In fact,

$$d\xi = P(t)^{-\frac{1}{2}} dx + R(t)\xi(t)dt. \quad (60)$$

Differentiating $P(t)^{-\frac{1}{2}} P(t)P(t)^{-\frac{1}{2}} = I_n$, we obtain

$$P(t)^{-\frac{1}{2}} \dot{P} P(t)^{-\frac{1}{2}} = -R(t) - R(t)',$$

and hence the (55) yields

$$F(t) + F(t)' + G(t)G(t)' = 0. \quad (61)$$

Using (61) to eliminate F in (57), we obtain

$$d\xi = -F(t)'\xi(t)dt + G(t)d\bar{w}, \quad (62)$$

where

$$d\bar{w} = dw - G(t)'\xi(t)dt, \quad (63)$$

which can also be written

$$d\bar{w} = dw - \bar{B}(t)'x(t)dt, \quad (64)$$

where $\bar{B}(t) := P(t)^{-1} B(t)$.

Proposition 1: A process $\bar{w}$ satisfying (63) has orthogonal increments with the normalized property (5). Moreover,

$$\mathrm{E}\{[\bar{w}(t) - \bar{w}(s)]\xi(t)'\} = 0 \quad (65)$$

for all $s \le t$.

Proof: As is well-known, the solution of (57) can be written in the form

$$\xi(t) = \Phi(t, s)\xi(s) + \int_s^t \Phi(t, \tau)G(\tau)dw, \quad (66)$$

where $\Phi(t, s)$ is the transition matrix with the property

$$\frac{\partial \Phi}{\partial t}(t, s) = F(t)\Phi(t, s), \quad \Phi(s, s) = I_n \quad (67a)$$
$$\frac{\partial \Phi}{\partial s}(t, s) = -\Phi(t, s)F(s), \quad \Phi(t, t) = I_n \quad (67b)$$

Let $s \le t$. Then, in view of (63), a straight-forward calculation yields

$$\bar{w}(t) - \bar{w}(s) = w(t) - w(s)$$
$$- M(t, s)\xi(s) - \int_s^t M(t, \tau)G(\tau)dw, \quad (68)$$

where

$$M(t, s) = \int_s^t G(\tau)'\Phi(\tau, s)d\tau. \quad (69)$$

Therefore,

$$\mathrm{E}\{[\bar{w}(t) - \bar{w}(s)][\bar{w}(t) - \bar{w}(s))'\} = I_p(t - s) + \Delta(t, s),$$

where

$$\Delta(t, s) = M(t, s)M(t, s)' + \int_s^t M(t, \tau)G(\tau)G(\tau)'M(t, \tau)'d\tau$$
$$- \int_s^t [M(t, \tau)G(\tau) + G(\tau)'M(t, \tau)'] d\tau.$$

However, $\Delta(t, s)$ is identically zero. To see this, first note that

$$\frac{\partial M}{\partial s}(t, s) = -M(t, s)F(s) - G(s)'. \quad (70)$$

Then, in view of (61), a simple calculation shows that

$$\frac{\partial \Delta}{\partial s}(t, s) \equiv 0.$$

Since $\Delta(t, t) = 0$, the assertion follows. Hence the incremental covariance is normalized.

Next, we show that $\bar{w}(t)$ has orthogonal increments. To this end, choose arbitrary times $s \leq t \leq a \leq b$ on the interval $[0, T]$, where we choose a and b fixed, and show that

$$Q(t, s) := \mathrm{E}\{[\bar{w}(b) - \bar{w}(a)][\bar{w}(t) - \bar{w}(s))']\}$$

is identically zero for all $s \leq t$. Using (68) and

$$\bar{w}(b) - \bar{w}(a) = w(b) - w(s) - M(b, a)\Phi(a, s)\xi(s)$$
$$- M(b, a)\int_s^b \Phi(a, \tau)G(\tau)dw - \int_a^b M(b, \tau)dw$$

computed analogously, we obtain

$$Q(t, s) = M(b, a)\left[\Phi(a, s)M(t, s)' - \int_s^b \Phi(a, \tau)G(\tau)d\tau\right.$$
$$\left. + \int_s^b \Phi(a, \tau)G(\tau)G(\tau)'M(t, \tau)d\tau\right].$$

Then, again using (61), we see that

$$\frac{\partial M}{\partial s}(t, s) \equiv 0,$$

so, since $Q(t, t) = 0$, we see that $Q(t, s)$ is identically zero, establishing that $\bar{w}(t)$ has orthogonal increments.

Finally, we use the same trick to show (65). In fact, for $s \leq t$, (66) and (68) yield

$$\mathrm{E}\{[\bar{w}(t) - \bar{w}(s))\xi(t)'\} = -M(t, s)\Phi(t, s)'$$
$$+ \int_s^t G(\tau)'\Phi(t, \tau)'d\tau - \int_s^t M(t, \tau)G(\tau)G(\tau)')\Phi(t, \tau)'d\tau,$$

the partial derivative of which with respect to s is identical zero; this is seen by again using (61). Therefore, since (65) is zero for $s = t$, it is identical zero for all $s \leq t$, as claimed. This concludes the proof of Proposition 1. ∎

Consequently, (57) and (64) form a forward unitary system

$$dx = A(t)x(t)dt + B(t)dw \tag{71a}$$
$$d\bar{w} = dw - \bar{B}(t)'x(t)dt, \tag{71b}$$

The corresponding backward unitary system is obtained through the transformation

$$\bar{x}(t) = P(t)^{\frac{1}{2}}\xi(t), \tag{72}$$

which yields

$$d\bar{x} = P(t)^{-\frac{1}{2}}d\xi + R(t)\xi(t)dt. \tag{73}$$

This together with (62) and (63) yields

$$d\bar{x} = -A(t)'\bar{x}(t)dt + \bar{B}(t)d\bar{w} \tag{74a}$$
$$dw = B(t)'\bar{x}(t)dt + d\bar{w}, \tag{74b}$$

B. Time reversal in continuous-time systems

Next we derive the backward stochastic system corresponding to the non-stationary forward stochastic system

$$dx = A(t)x(t)dt + B(t)dw, \quad x(0) = x_0 \tag{75a}$$
$$dy = C(t)x(t)dt + D(t)dw, \quad y(0) = 0 \tag{75b}$$

defined on the finite interval $[0, T]$, where x_0 (with covariance P_0) and the normalized Wiener process w are uncorrelated. To this end, apply the transformation

$$\bar{x}(t) = P(t)^{-1}x(t) \tag{76}$$

together with (74b) to (75b) to obtain

$$dy = \bar{C}(t)\bar{x}(t) + D(t)d\bar{w},$$

where

$$\bar{C}(t) = C(t)P(t) + D(t)B(t). \tag{77}$$

This together with (74a) yields the the backward system corresponding to (75), namely

$$d\bar{x} = -A(t)'\bar{x}(t)dt + \bar{B}(t)d\bar{w} \tag{78a}$$
$$dy = \bar{C}(t)\bar{x}(t)dt + D(t)d\bar{w}. \tag{78b}$$

with end-point condition $\bar{x}(T) = P(T)^{-1}x(T)$ uncorelated to the Wiener process $\bar{w}$.

The backward realization (78) was derived in [3], but in cumbersome way, requiring the proof that $\bar{w}(t)$ is a normalized process with orthogonal increments to be suppressed. What is new here is imposing the unitary map between w and $\bar{w}$, making the analysis much simpler and more natural.

V. KALMAN FILTERING WITH MISSING OBSERVATIONS

We consider the linear stochastic system (75) which does not have a purely deterministic component that enables exact estimation of components of x from y, an assumption that we retain in the rest of the paper. In the engineering literature is often the case that the stochastic system (75) represented as

$$\dot{x}(t) = A(t)x(t) + B(t)\dot{w}(t), \quad x(0) = x_0 \tag{79a}$$
$$\dot{y}(t) = C(t)x(t) + D(t)\dot{w}(t) \tag{79b}$$

where the formal "derivative" $\dot{w}$ is white noise, i.e., $\mathrm{E}\{\dot{w}(t)\dot{w}(s)'\} = I\delta(t - s)$ with $\delta(t - s)$ being the Dirac "function". Of course $\dot{x}$, $\dot{y}$ and $\dot{w}$ are to be interpreted as generalized stochastic processes. From a mathematically rigorous point of view, observing $\dot{y}$ makes little sense since, for any fixed t, $\dot{y}(t)$ has infinite variance and contains no information about the state process x. However, observations of $\dot{y}$ could be interpreted as observations of the increments dy of y in a precise meaning to be defined next. On the other hand, one can think of (75) as a system of type

$$dz = M(t)z(t)dt + N(t)dw(t), \quad \text{where } z(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix},$$

and one would like to determine the optimal linear least-squares estimate of $x(t)$ given past observed values of y.

Generally this distinction between observing y or dy is not important. However, when there is loss of information over an interval (t_1, t_2), there are two different information patterns depending on whether dy or y is observed. The difference consists in whether $\Delta y := y(t_2) - y(t_1)$ is part of the observation record or not. These two cases will be dealt with separately in subsections below. In fact, the former, which is common in engineering applications, is provided as a simplified preliminary, whereas our main interest is in the latter. To this end, we first introduce some notation.

Consider the stochastic system (75) on a finite interval $[0, T]$. As before, let $\mathbf{H}$ be the Hilbert space spanned by $\{w_k(t) - w_k(s); s, t \in [0, T], k = 1, 2, \ldots, m\}$, endowed with the inner product $\langle \lambda, \mu \rangle = \mathrm{E}\{\lambda \mu\}$. For any $\lambda \in \mathbf{H}$ and any subspace $\mathbf{A}$, let $\mathrm{E}^{\mathbf{A}}$ denote the orthogonal projection of λ onto $\mathbf{A}$. We denote by $\mathbf{H}_{[t_1, t_2]}(dy)$ the (closed) subspace generated by the components of the increments of the observation process y over the window $[t_1, t_2]$. In particular, we shall also use the notations $\mathbf{H}_t^-(dy) := \mathbf{H}_{[0,t]}(dy)$ and $\mathbf{H}_t^+(dy) := \mathbf{H}_{[t,T]}(dy)$.

Suppose that the output process or its increments are available for observation only on some subintervals of $[0, T]$, namely $\mathcal{I}_k$, $k = 1, 2, \ldots, \nu$. Next we want to define $\overset{\circ}{\mathbf{H}}$ as the proper subspace of $\mathbf{H}_{[0,T]}(dy)$ spanned by the observed data. In the case that only the increments dy or, equivalently, the "derivative" $\dot{y}$ is observed, we simply define

$$\overset{\circ}{\mathbf{H}} := \mathbf{H}_{\mathcal{I}_1}(dy) \vee \mathbf{H}_{\mathcal{I}_2}(dy) \vee \cdots \vee \mathbf{H}_{\mathcal{I}_\nu}(dy),$$

In the case that the process y is observed, we need to expand $\overset{\circ}{\mathbf{H}}$ by adding the subspaces spanned by the increments Δy over the complementary intervals without observation. In either case, we define

$$\overset{\circ}{\mathbf{H}}_t^- := \overset{\circ}{\mathbf{H}} \cap \mathbf{H}_t^-(dy) \quad \text{and} \quad \overset{\circ}{\mathbf{H}}_t^+ := \overset{\circ}{\mathbf{H}} \cap \mathbf{H}_t^+(dy). \quad (80)$$

Then Kalman filtering with missing observations amounts to determining a recursion for x_- where

$$a'x_-(t) = \mathrm{E}^{\overset{\circ}{\mathbf{H}}_t^-} a'x(t), \quad \text{for all } a \in \mathbb{R}^n. \quad (81)$$

A. Observing dy only

When observations are available on the interval $[0, t_1]$, the Kalman filter on that interval is given by

$$dx_- = A(t)x_-(t)dt + K_-(t)(dy(t) - C(t)x_-(t)dt) \quad (82a)$$

$$K_- = (Q_-C' + BD')R^{-1} \quad (82b)$$

$$\dot{Q}_-(t) = AQ_- + Q_-A' - K_-RK_-' + BB' \quad (82c)$$

with $R(t) = D(t)D(t)'$ and initial conditions $x_-(0) = 0$ and $Q(0) = P_0$. Here $Q_-(t)$ is the error covariance

$$Q_-(t) := \mathrm{E}\{[x(t) - x_-(t)][x(t) - x_-(t)]'\}, \quad (83)$$

which, by the nondeterministic assumption, is positive definite for all t.

Next suppose the observation process becomes unavailable over the interval $[t_1, t_2] \subset [0, T]$. Then the Kalman filter needs to be modified accordingly. In fact, for any $t \in [t_1, t_2)$, (81) holds with the space of observations $\overset{\circ}{\mathbf{H}}_t^- := \mathbf{H}_{t_1}^-(dy)$, and consequently

$$a'x_-(t) = \mathrm{E}^{\mathbf{H}_{t_1}^-(dy)} a'x(t) = a'\Phi(t, t_1)x_-(t_1).$$

This corresponds to setting $K_-(t) = 0$ in (82) on the interval $[t_1, t)$ so that

$$dx_- = A(t)x_-(t)dt \quad (84a)$$

with initial condition $x_-(t_1)$ given by (82a). The error covariance Q_- is then given by the Lyapunov equation

$$\dot{Q}_-(t) = AQ_- + Q_-A' + BB' \quad (84b)$$

with initial the condition $Q_-(t_1)$ given by the value produced in the previous interval.

Then suppose observations of dy become available again on the interval $[t_2, t_3]$. Then, for any $t \in [t_2, t_3)$, we have

$$\overset{\circ}{\mathbf{H}}_t^+ = \mathbf{H}_{[0,t_1]} \vee \mathbf{H}_{[t_2,t]},$$

so the Kalman estimate is generated by (82) but now with initial conditions $x_-(t_2)$ and $Q_-(t_2)$ being those computed in the previous step without observation. In the case there are more intervals, one proceeds similarly by alternating between filters (82) and (84) depending on whether increments dy are available or not.

In an identical manner, a cascade of backward Kalman filters generates a process $\bar{x}_+(t)$ based on the backward stochastic realization (78) and the observation windows $[t, T]$. Assuming that there are observations in a final interval ending at $t = T$, on that interval the Kalman estimate

$$a'\bar{x}_+(t) = \mathrm{E}^{\overset{\circ}{\mathbf{H}}_t^+} a'\bar{x}(t), \quad (85)$$

with initial observation space $\overset{\circ}{\mathbf{H}}_t^+ := \mathbf{H}_{[t,T]}$, is generated by the backward Kalman filter

$$d\bar{x}_+ = -A(t)'\bar{x}_+(t)dt$$
$$+ \bar{K}_+(t)(dy(t) - \bar{C}(t)\bar{x}_+(t)dt) \quad (86a)$$

$$\bar{K}_+ = -(\bar{Q}_+\bar{C}' - \bar{B}D')R^{-1} \quad (86b)$$

$$\dot{\bar{Q}}_+ = -A'\bar{Q}_+ - \bar{Q}_+A + \bar{K}_+R(t)\bar{K}_+' - \bar{B}\bar{B}' \quad (86c)$$

and initial conditions $\bar{x}_+(T) = 0$ and $\bar{Q}_+(T) = \bar{P}(T)$ for $\bar{x}_+$ and the error covariance

$$\bar{Q}_+(t) := \mathrm{E}\{[\bar{x}(t) - \bar{x}_+(t)][\bar{x}(t) - \bar{x}_+(t)]'\}, \quad (87)$$

which like $Q_-(t)$ is positive definite for all t. During periods of no observations of dy, we then set the gain $\bar{K}_+ = 0$. This update is obtained from the backward time stochastic model (74) in an identical manner to that of (84).

Consequently, both the underlying process as well as the filter can run in either time-direction. This duality becomes essential in subsequent sections where we will be concerned with smoothing and interpolation.

B. Observing y

Now consider the case that y, and note merely dy, is available for observation on all intervals $\mathcal{I}_k$, $k = 1, 2, \ldots, \nu$. Under this scenario and with a continuous-time process the dynamics of Kalman filtering become hybrid, requiring both continuous-time filtering when data is available as well as a discrete-time update across intervals where measurements are not available.

Then on the first interval $[0, t_1]$ the Kalman estimate (82) will still be valid. However, when t reaches the endpoint t_2 of the interval of no information and an observation of y is obtained again, the subspace of observed data becomes

$$\mathring{\mathbf{H}}_{t_2}^- = \mathbf{H}_{t_1}^- \vee \mathbf{H}(\Delta y),$$

where $\Delta y := y(t_2) - y(t_1)$. Computing $x(t_2)$ across the window $(t_1, t_2]$ as a function of $x(t_1)$ and the noise components we have that

$$x(t_2) = \underbrace{\Phi(t_2, t_1)}_{A_d} x(t_1) + \underbrace{\int_{t_1}^{t_2} \Phi(t_2, s) B \, dw(s)}_{u_1(t_1)}$$

while

$$y(t_2) = y(t_1) + \int_{t_1}^{t_2} C(t) x(t) dt + \int_{t_1}^{t_2} D(t) dw(t).$$

Therefore,

$$\Delta y = \underbrace{\int_{t_1}^{t_2} C(t) \Phi(t, t_1) dt}_{C_d} x(t_1) + u_2(t_1)$$

where

$$u_2(t_1) = \int_{t_1}^{t_2} C(t) \int_{t}^{t_1} \Phi(t, s) B(s) dw(s) dt$$
$$+ \int_{t_1}^{t_2} D(s) dw(s)$$
$$= \int_{t_1}^{t_2} \underbrace{\left(\int_{t}^{t_2} C(t) \Phi(t, s) dt B(s) + D(s) \right)}_{M(s)} dw(s).$$

Thus, we obtain the discrete-time update

$$x(t_2) = A_d x(t_1) + B_d v(t_1) \tag{88a}$$
$$\Delta y = C_d x(t_1) + D_d v(t_1) \tag{88b}$$

where

$$u(t_1) = \begin{pmatrix} u_1(t_1) \\ u_2(t_1) \end{pmatrix} = \begin{pmatrix} B_d \\ D_d \end{pmatrix} v(t_1)$$

and B_d and D_d are chosen so that

$$\begin{pmatrix} B_d \\ D_d \end{pmatrix} (B_d', \ D_d') = \int_{t_1}^{t_2} \begin{pmatrix} \Phi(t_2, s) BB' \Phi(t, s) & \Phi(t_2, s) BM(s)' \\ M(s) B' \Phi(t_2, s)' & M(s) M(s)' \end{pmatrix} ds$$

while $E\{v(t_1) v(t_1)'\} = I$.

Hence, across the window of missing data the Kalman state estimate x_- is now generated by a discrete-time Kalman-filter step

$$x_-(t_2) = A_d x_-(t_1) + K_d(\Delta y - C_d x_-(t_1)) \tag{89a}$$
$$K_d = (A_d Q(t_1) C_d' + B_d D_d')$$
$$\times (C_d Q(t_1) C_d' + D_d D_d')^{-1} \tag{89b}$$

with initial conditions $x_-(t_1)$ and $Q(t_1)$ given by (82) and the error covariance at t_2 by

$$Q(t_2) = A_d Q(t_1) A_d' - K_d(C_d Q(t_1) C_d'$$
$$+ D_d D_d') K_d' + B_d B_d'. \tag{89c}$$

In the next interval $[t_2, t_3]$, where observations of y are available, the new Kalman estimate (81) with

$$\mathring{\mathbf{H}}_t^+ = \mathbf{H}_{[0, t_1]} \vee \mathbf{H}(\Delta y) \vee \mathbf{H}_{[t_2, t]}$$

is again generated by the continuous-time Kalman filter (82) starting from $x_-(t_2)$ and $Q(t_2)$ given by (89).

Again given an observation pattern, where intermittently y becomes unavailable for observation, the Kalman estimate (81) can be generated in precisely this manner by a cascade of continuous and discrete-time Kalman filters.

Remark 2: The observation pattern of a continuous-time stochastic model, where y becomes unavailable over particular time-windows, is closely related to hybrid stochastic models where continuous-time diffusion is punctuated by discrete-time transitions. Indeed, unless interpolation of the statistics within windows of unavailable data is the goal, the end points of such intervals can be identified and the same hybrid model utilized to capture the dynamics.

Remark 3: A common engineering scenario is the case where the signal is lost while the observation noise is still present. This amounts to having $C \equiv 0$ over the corresponding window, and the Kalman estimates are obtained by merely running the filters (82) and (86) in the two time directions with the modified condition on C. This situation does not cover the information patterns discussed above since, whenever $BD' \neq 0$, the Kalman gains do not vanish and information about the state process is available even when C is zero.

C. Smoothing

Given these intermittent forward and backward Kalman estimates, we shall derive a formula for the smoothing estimate

$$a'\hat{x}(t) := \mathrm{E}^{\overset{\circ}{\mathbf{H}}} a'x(t), \quad a \in \mathbb{R}^n, \tag{90}$$

valid for both the cases discussed above, where

$$\overset{\circ}{\mathbf{H}} := \overset{\circ}{\mathbf{H}}_t^- \vee \overset{\circ}{\mathbf{H}}_t^+ \subset \mathbf{H}_{[0,T]}(dy) \tag{91}$$

is the complete subspace of observations. This is discussed next.

VI. GEOMETRY OF FUSION

Consider the system (75), and let $\mathbf{X}(t)$ be the (finite-dimensional) subspace in $\mathbf{H}$ spanned by the components of the stochastic state vector $x(t)$. Then it can be shown [18, Chapter 7] that $\mathbf{H}_{[0,t]}(dy) \perp \mathbf{H}_{[t,T]}(dy) \mid \mathbf{X}_t$, where $\mathbf{A} \perp \mathbf{B} \mid \mathbf{X}$ denotes the conditional orthogonality

$$\langle \alpha - \mathrm{E}^{\mathbf{X}} \alpha, \beta - \mathrm{E}^{\mathbf{X}} \beta \rangle = 0 \quad \text{for all } \alpha \in \mathbf{A}, \beta \in \mathbf{B}. \tag{92}$$

Next, let $\mathbf{X}_-(t)$ and $\mathbf{X}_+(t)$ be the subspaces spanned by the components of the (intermittent) Kalman estimates $x_-(t)$ and $\bar{x}_+(t)$, respectively. Then since $\mathbf{X}_-(t) \subset \overset{\circ}{\mathbf{H}}_t^- \subset \mathbf{H}_{[0,t]}(dy)$ and $\mathbf{X}_+(t) \subset \overset{\circ}{\mathbf{H}}_t^+ \subset \mathbf{H}_{[t,T]}(dy)$, we have

$$\mathbf{X}_-(t) \perp \mathbf{X}_+(t) \mid \mathbf{X}(t),$$

which is equivalent to

$$\mathrm{E}^{\mathbf{X}_+(t)} a'x_-(t) = \mathrm{E}^{\mathbf{X}_+(t)} \mathrm{E}^{\mathbf{X}(t)} a'x_-(t), \quad a \in \mathbb{R}^n \tag{93a}$$

[18, Proposition 2.4.2]. Therefore the diagram

$$\begin{array}{ccc}
\mathbf{X}_- & \overset{\mathrm{E}^{\mathbf{X}_+}|_{\mathbf{X}_-}}{\longrightarrow} & \mathbf{X}_+ \\
{\scriptstyle \mathrm{E}^{\mathbf{X}}|_{\mathbf{X}_-}} \searrow & & \nearrow {\scriptstyle \mathrm{E}^{\mathbf{X}_+}|_{\mathbf{X}}} \\
& \mathbf{X} &
\end{array} \tag{93b}$$

commutes, where the argument t has been suppressed.

Lemma 4: Let $x(t)$, $\bar{x}(t)$, $x_-(t)$ and $\bar{x}_+(t)$ be defined as above. Then, for each $t \in [0, T]$,

(i) $\mathrm{E}\{x(t)x_-(t)'\} = P_-(t)$
(ii) $\mathrm{E}\{\bar{x}(t)\bar{x}_+(t)'\} = \bar{P}_+(t)$
(iii) $\mathrm{E}\{\bar{x}_+(t)x_-(t)'\} = \bar{P}_+(t)P_-(t)$,

where $P_-(t) := \mathrm{E}\{x_-(t)x_-(t)'\}$ is the state covariance of the Kalman estimate $x_-(t)$ and $\bar{P}_+(t) := \mathrm{E}\{\bar{x}_+(t)\bar{x}_+(t)'\}$ is the covariance of the backward Kalman estimate $\bar{x}_+(t)$.

Proof: By the definition of the Kalman filter, (81) holds, and consequently the components of the estimation error $x(t) - x_-(t)$ are orthogonal to $\mathbf{H}_t^-$ and hence to the components of $x_-(t)$. Therefore,

$$\mathrm{E}\{x(t)x_-(t)'\} = \mathrm{E}\{x_-(t)x_-(t)'\} = P_-(t),$$

proving condition (i). Condition (ii) follows from a symmetric argument. To prove (iii) we use condition (93). To this end, first note that, by the usual projection formula,

$$\begin{aligned}
\mathrm{E}^{\mathbf{X}_+(t)} a'x_-(t) &= \mathrm{E}\{a'x_-(t)\bar{x}_+(t)\}\bar{P}_+(t)^{-1}\bar{x}_+(t) \\
&= a' \mathrm{E}\{x_-(t)\bar{x}_+(t)'\}x_+(t),
\end{aligned} \tag{94}$$

where $x_+(t) := \bar{P}_+(t)^{-1}\bar{x}_+(t)$ is the dual basis in $\mathbf{X}_+(t)$ such that $\mathrm{E}\{x_+(t)\bar{x}_+(t)'\} = I$. Moreover,

$$\begin{aligned}
\mathrm{E}^{\mathbf{X}(t)} a'x_-(t) &= E\{a'x_-(t)x(t)'\}P(t)^{-1}x(t) \\
&= a' \mathrm{E}\{x_-(t)x(t)'\}\bar{x}(t) = a'P_-(t)\bar{x}(t),
\end{aligned}$$

where we have used condition (i) and (76). Next, set $b := P_-a$ and form

$$\begin{aligned}
\mathrm{E}^{\mathbf{X}_+(t)} b'\bar{x}(t) &= \mathrm{E}\{b'\bar{x}(t)\bar{x}_+(t)\}\bar{P}_+(t)^{-1}\bar{x}_+(t) \\
&= b' \mathrm{E}\{\bar{x}(t)\bar{x}_+(t)'\}x_+(t) \\
&= b'\bar{P}_+(t)x_+(t),
\end{aligned}$$

by condition (ii), and consequently

$$\mathrm{E}^{\mathbf{X}_+(t)} E^{\mathbf{X}(t)}a'x_-(t) = a'P_-(t)\bar{P}_+(t)x_+(t). \tag{95}$$

Then condition (iii) follows from (93a), (94) and (95). ∎

Remark 5: The proof of condition (iii) in Lemma 4 could be simplified if $\bar{x}_+$ were a regular backward Kalman estimate without intermittent loss of information. In this case, $x_+ = \bar{P}_+^{-1}\bar{x}_+$ would be generated by a forward stochastic realization belonging to the same class as (75) and $\mathrm{E}\{\bar{x}_+(t)x_-(t)'\} = \bar{P}_+(t) \mathrm{E}\{x_+(t)x_-(t)\} = \bar{P}_+(t) \mathrm{E}\{x_-(t)x_-(t)\}$.

Lemma 6: For each $t \in [0, T]$, the smoothing estimate $\hat{x}(t)$, defined by (90), is given by

$$a'\hat{x}(t) = E^{\mathbf{H}_t^{\square}} a'x(t), \quad a \in \mathbb{R}^n, \tag{96}$$

where $\mathbf{H}_t^{\square}$ is the subspace

$$\mathbf{H}_t^{\square} = \mathbf{X}_-(t) \vee \mathbf{X}_+(t). \tag{97}$$

Proof: Following [14], [3], [18], define $\mathbf{N}^-(t) := \overset{\circ}{\mathbf{H}}_t^- \ominus \mathbf{X}_-(t)$ and $\mathbf{N}^+(t) := \overset{\circ}{\mathbf{H}}_t^+ \ominus \mathbf{X}_+(t)$. Then

$$\overset{\circ}{\mathbf{H}} = \mathbf{N}^-(t) \oplus \mathbf{H}_t^{\square} \oplus \mathbf{N}^+(t).$$

Now, $a'(x(t) - x_-(t))$ is orthogonal to $\overset{\circ}{\mathbf{H}}_t^-$ and hence to $\mathbf{N}^-(t)$. Also $a'x_-(t) \perp \mathbf{N}^-(t)$. Hence $a'x(t) \perp \mathbf{N}^-(t)$ as well. In the same way we see that $a'x(t) \perp \mathbf{N}^+(t)$. Therefore (96) follows. ∎

Consequently, the information from the two Kalman filters can be fused into the smoothing estimate

$$\hat{x}(t) = L_-(t)x_-(t) + \bar{L}_+(t)\bar{x}_+(t) \tag{98}$$

for some matrix functions L_- and $\bar{L}_+$.

VII. UNIVERSAL TWO-FILTER FORMULA

To obtain a robust and particularly simple smoothing formula that works also with an intermittent observation pattern, we assume that the stochastic system (75) has already been transformed via (58) so that, for all $t \in [0, T]$,

$$x(t) = \bar{x}(t) \tag{99}$$

and therefore

$$P(t) = \mathrm{E}\{x(t)x(t)'\} = I = \bar{P}(t). \tag{100}$$

Then the error covariances in the filtering formulas of Section V are

$$Q_- = I - P_- \quad \text{and} \quad \bar{Q}_+ = I - \bar{P}_+. \tag{101}$$

Consequently, $x(t)$, $\bar{x}(t)$, $P_-(t)$ and $\bar{P}_+(t)$ are all bounded in norm by one for all $t \in [0, T]$.

Theorem 7: Suppose that (99) holds. For every $t \in [0, T]$, we have the formula

$$\hat{x}(t) = Q(t)\left(Q_-(t)^{-1}x_-(t) + \bar{Q}_+(t)^{-1}\bar{x}_+(t)\right) \tag{102}$$

for the smoothing estimate (90), where the estimation error

$$Q(t) := \mathrm{E}\left\{(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))'\right\} \tag{103}$$

is given by

$$Q(t)^{-1} = Q_-(t)^{-1} + \bar{Q}_+(t)^{-1} - I, \tag{104}$$

and where x_-, $\bar{x}_+$, Q_- and Q_+ are given by (82) and (86) with boundary conditions $x_-(0) = \bar{x}_+(T) = 0$ and $Q_-(0) = Q_+(T) = I$.

Proof: Clearly the matrix functions L_- and $\bar{L}_+$ in (98) can be determined from the orthogonality relations

$$\mathrm{E}\{[x(t) - \hat{x}(t)]x_-(t)'\} = 0 \tag{105a}$$

and

$$\mathrm{E}\{[x(t) - \hat{x}(t)]\bar{x}_+(t)'\} = 0. \tag{105b}$$

By Lemma 4, (105) yields

$$P_- - L_-P_- - \bar{L}_+\bar{P}_+P_- = 0$$
$$\bar{P}_+ - L_-P_-\bar{P}_+ - \bar{L}_+\bar{P}_+ = 0,$$

which, in view of the fact that P_- and $\bar{P}_+$ are positive definite, yields

$$L_- + \bar{L}_+\bar{P}_+ = I \tag{106a}$$

$$L_-P_- + \bar{L}_+ = I \tag{106b}$$

Again by orthogonality and Lemma 4,

$$Q = \mathrm{E}\left\{(x - \hat{x})x'\right\} = I - L_-P_- - \bar{L}_+\bar{P}_+,$$

which, in view of (106) and the relations (101), yields

$$L_- = QQ_-^{-1} \quad \text{and} \quad \bar{L}_+ = Q\bar{Q}_+^{-1}. \tag{107}$$

Then (102) follows from (98) and (107). To prove (104) eliminate $\bar{L}_+$ in (106) to obtain

$$L_-(I - P_-\bar{P}_+) = \bar{Q}_+,$$

which together with (107) yields

$$Q^{-1} = Q_-^{-1}(I - P_-\bar{P}_+)\bar{Q}_+^{-1}.$$

However,

$$I - P_-\bar{P}_+ = \bar{Q}_+ + Q_- - Q_-\bar{Q}_+,$$

and hence (104) follows. ∎

In the special case with no loss of observation this is a normalized version of the Mayne-Frazer two-filter formula [1], [2], which however in [1], [2] was formulated in terms of x_- and x_+ rather than $\bar{x}_+$, where x_+ is the state process of the forward stochastic system of the backward Kalman filter. (For the corresponding formula in terms of x_- and $\bar{x}_+$, see [3], [18]; also cf. [21], where an independent derivation was given.) With a single interval of loss of observation the formula (102) reduces to a version of the interpolation formulas in [6]. The remarkable fact, discovered here, is that the same formula (102) holds for any intermittent observations structure and by a cascade of continuous and discrete-time forward and backward Kalman filters, as needed depending on the assumed information pattern.

VIII. RECAP OF COMPUTATIONAL STEPS

Given a system (75) with state covariance (55), make the normalizing substitution

$$\begin{aligned}
A(t) &\leftarrow P(t)^{-\frac{1}{2}}A(t)P(t)^{\frac{1}{2}} + R(t) \\
B(t) &\leftarrow P(t)^{-\frac{1}{2}}B(t) \\
C(t) &\leftarrow C(t)P(t)^{\frac{1}{2}}
\end{aligned} \tag{108}$$

with $R(t) = \left[\frac{d}{dt}P(t)^{-\frac{1}{2}}\right]P(t)^{\frac{1}{2}}$. Next, we compute the intermittent forward and backward Kalman filter estimates x_- and $\bar{x}_+$, respectively, along the lines of Section V, where, due to the normalization, $Q_-(0) = \bar{Q}_+(T) = I_n$. Then the smoothing estimate is given by

$$\hat{x}(t) = Q(t)\left(Q_-(t)^{-1}x_-(t) + \bar{Q}_+(t)^{-1}\bar{x}_+(t)\right),$$

where

$$Q(t) = \left(Q_-(t)^{-1} + \bar{Q}_+(t)^{-1} - I\right)^{-1}.$$

IX. AN EXAMPLE

We now illustrate the results of the paper on a specific numerical example. We consider the continuous-time diffusion process

$$\begin{aligned}
dx_1(t) &= x_2(t)dt \\
dx_2(t) &= -0.3x_1(t)dt - 0.7x_2(t)dt + dw(t) \\
dy(t) &= x_1(t)dt + dv(t)
\end{aligned}$$

where w and v are thought to be independent standard Wiener processes. Here, x_1 is thought of as position and x_2 as velocity of a particle that is steered by stochastic excitation in dw, in the presence of a restoring force $0.3x_1$ and frictional force $0.7x_2$. Then dy/dt represents measurement of the position and dv/dt represents measurement noise (white).
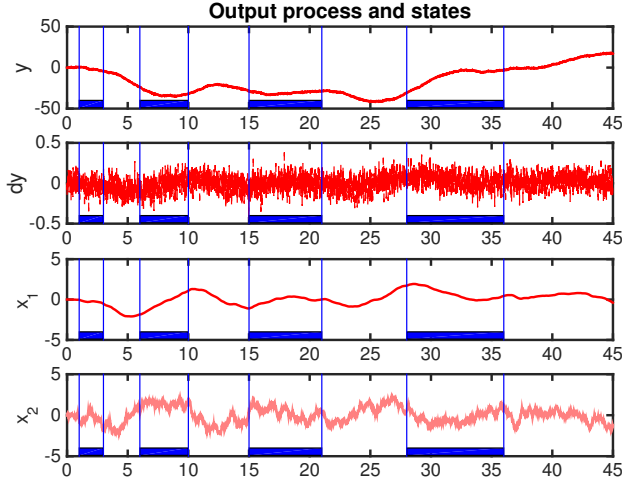

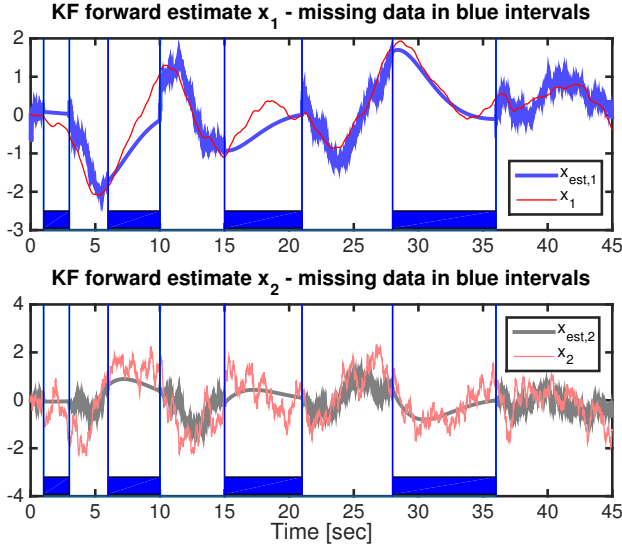Fig. 5: Sample paths of output process, increment, and state processes

Fig. 6: Kalman estimates in the forward time direction

Numerical simulation over $[0, T]$ with $T = 45$ (units of time) produces a time-function $y(t)$ which is sampled with integer multiples of $\Delta t = 0.01$ (units). The interval $[0, T]$ is partitioned into

$$[0, T] = \cup_{i=1}^{9}[t_{i-1}, t_i]$$

where $t_0 = 0$ and $t_i - t_{i-1} = i$ (units). Measurements of y are made available for purposes of state estimation over the intervals $[t_{i-1}, t_i]$ for $i = 1, 3, 5, 9$. Over the complement set

of intervals, data are not made available for state estimation; these intervals where data are not to be used are marked by a thick blue baseline in the figures. In Figure 5 we display sample paths of the output process y, increments dy, and state-processes x_1 and x_2.
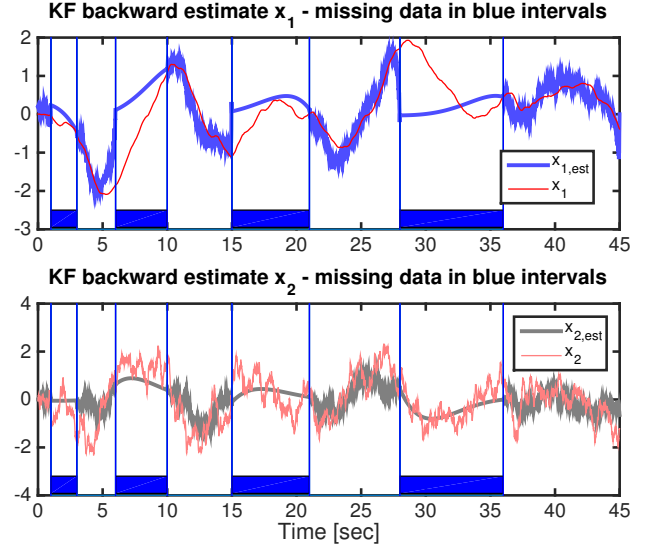

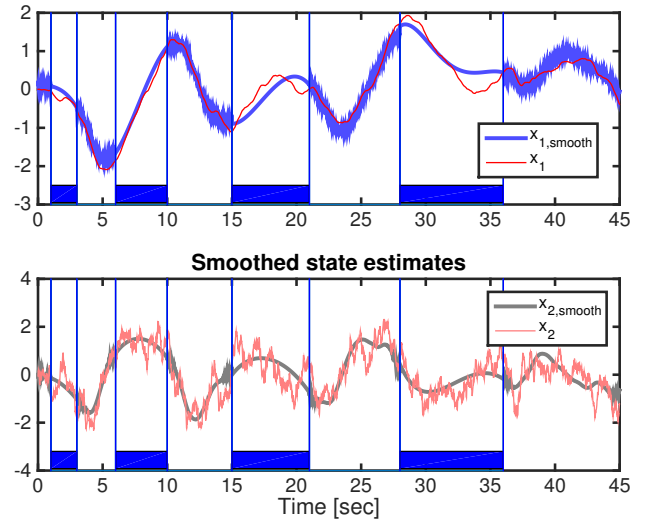Fig. 7: Kalman estimates in the backward time direction

Fig. 8: Interpolation/smoothed estimates by fusion of Kalman forward and backward estimates

The process increments dy over $[t_{i-1}, t_i]$ for $i = 1, 3, 5, 9$ as well as the increments Δy across the $[t_{i-1}, t_i]$ for $i = 2, 4, 6, 8$ are used in the two-filter formula for the purpose of smoothing. The Kalman estimates for the states in the forward and backwards in time directions, $x_-(t)$ and $\bar{x}_+(t)$ are shown in Figures 6 and 7, respectively. The fusion of the two using (102) is shown in Figure 8. It is worth observing the nature and fidelity of the estimates. In the forward direction, across intervals where data is not available, x_- becomes

increasing more unreliable whereas the opposite is true for $\bar{x}_+$, as expected. The smoothing estimate is generally an improvement to those of the two Kalman filters as seen in Figure 8. In particular, it is worth noting x_2 (in subplot 2), where, over windows of available observations, estimates have considerably less variance in the middle of the interval where the weights $(Q(t)Q_-(t)^{-1}$ and $Q(t)\bar{Q}_+(t)^{-1})$ in (102) are equalized, whereas sample paths become increasing rugged at the two ends where one of the two Kalman estimates has significantly higher variance, and the corresponding mixing coefficient becomes relatively smaller.

X. Concluding remarks

Historically the problem of interpolation has been considered from the beginning of the study of stochastic processes [22], [23]. Early accounts and treatments were cumbersome and non-explicit as the problem was considered difficult [7], [8], [9], [10]. In a manner that echoes the development of Kalman filtering, the problem became transparent and computable for ouput processes of linear stochastic systems [5], [6], [18].

This paper builds on developments in stochastic realization theory [11], [24] and presents a unified and generalized two-filter formula for smoothing and interpolation in continuous time for the case of intermittent availability of data over an operating window. The analysis considers two alternative information patterns where increments of the output process or the output process itself is recorded when information becomes available. The second information pattern appears most natural to us in this continuous-time setting, and this is our main problem. Nevertheless, in either case, two Kalman filters run in opposite time-directions, designed on the basis of a forward and a backward model for the process, respectively. Fusion of the respective estimates is effected via linear mixing in a manner similar to the Mayne-Fraser formula and applies to both smoothing and interpolation intermixed. In earlier works, smoothing and interpolation have been considered separate problems [18, Chapter 15]. The balancing normalization also simplifies the mixing formula and makes it completely time symmetric.

The theory relies on time-reversal of stochastic models. We provide a new derivation of such a reversal which has the convenient property of being balanced. It is based on lossless imbedding of linear systems and effects the time reversal through a unitary transformation. Interestingly, time symmetry in statistical and physical laws have occupied some of the most prominent minds in science and mathematics. In particular, closer to our immediate interests, dual time-reversed models have been employed to model, in different time-directions, Brownian or Schrödinger bridges [25], [26], a subject which is related to reciprocal processes [27], [28], [29], [30]. A natural extension of the present work in fact is in the direction of general reciprocal dynamics [28], [29] and the question of whether similar two-filter formula are possible.

Appendix: Time reversal of non-stationary discrete-time systems

Next, instead of (1), consider the non-stationary state dynamics

$$x(t+1) = A(t)x(t) + B(t)w(t), \quad x(0) = x_0, \quad (109)$$

on a finite time-window $[0, T]$, where, for simplicity we now assume that the covariance matrix $P_0 := P(0)$ of the zero-mean stochastic vector x_0 is positive definite, i.e., $P_0 = \mathrm{E}\{x_0 x_0'\} > 0$. Then the state covariance matrix $P(t) := \mathrm{E}\{x(t)x(t)'\}$ will satisfy the Lyapunov difference equation

$$P(t+1) = A(t)P(t)A(t)' + B(t)B(t)'. \quad (110)$$

The state transformation

$$\xi(t) = P(t)^{-\frac{1}{2}}x(t) \quad (111)$$

brings the system (109) into the form

$$\xi(t+1) = F(t)\xi(t) + G(t)w(t), \quad (112)$$

where now $\mathrm{E}\{\xi(t)\xi(t)'\} = I_n$ for all t and

$$F(t) = P(t+1)^{-\frac{1}{2}}A(t)P(t)^{\frac{1}{2}}, \quad (113a)$$

$$G(t) = P(t+1)^{-\frac{1}{2}}B. \quad (113b)$$

The Lyapunov difference equation then reduces to

$$I_n = F(t)F(t)' + G(t)G(t)' \quad (114)$$

allowing us to embed $[F, G]$ as part of a time-varying orthogonal matrix

$$U(t) = \begin{bmatrix} F(t) & G(t) \\ H(t) & J(t) \end{bmatrix}. \quad (115)$$

This amounts to extending (112) to

$$\xi(t+1) = F(t)\xi(t) + G(t)w(t) \quad (116a)$$

$$\bar{w}(t) = H(t)\xi(t) + J(t)w(t), \quad (116b)$$

or, in the equivalent form

$$\begin{bmatrix} \xi(t+1) \\ \bar{w}(t) \end{bmatrix} = U(t) \begin{bmatrix} \xi(t) \\ w(t) \end{bmatrix}. \quad (117)$$

Hence, since $\mathrm{E}\{\xi(t)\xi(t)'\} = I_n$ and $\mathrm{E}\{w(t)w(t)'\} = I_p$, and assuming that $\mathrm{E}\{\xi(t)w(t)'\} = 0$,

$$\mathrm{E}\left\{ \begin{bmatrix} \xi(t+1) \\ \bar{w}(t) \end{bmatrix} \begin{bmatrix} \xi(t+1) \\ \bar{w}(t) \end{bmatrix}' \right\} = U(t)U(t)' = I_{n+p}, \quad (118)$$

which yields

$$\mathrm{E}\{\xi(t+1)\bar{w}(t)'\} = 0, \quad (119a)$$

$$\mathrm{E}\{\bar{w}(t)\bar{u}(t)'\} = I_p. \quad (119b)$$

Moreover, from (116) we have

$$\bar{u}(t+k) = H(t+k)\Phi(t+k,t)\xi(t)$$
$$+ \sum_{j=t}^{t+k-1} H(t+k)\Phi(t+k,j+1)G(j)w(j) + J(t)w(t)$$

for $k > 0$, where

$$\Phi(s,t) = \begin{cases} F(s-1)F(s-2)\cdots F(t) & \text{for } s > t \\ I_n & \text{for } s = t. \end{cases}$$

Therefore, since $F(t)H(t)' + G(t)J(t)' = 0$ by the unitarity of $U(t)$,

$$\mathrm{E}\{\bar{u}(t+k)\bar{u}(t)'\}$$
$$= H(t+k)\Phi(t+k,t+1)[F(t)H(t)' + G(t)J(t)'] = 0.$$

Consequently, $\bar{u}$ is a white noise process. Finally, premultiplying (117) by $U(t)'$, we then obtain

$$\xi(t) = F(t)'\xi(t+1) + H(t)'\bar{w}(t) \tag{120a}$$
$$w(t) = G(t)'\xi(t+1) + J(t)'\bar{w}(t), \tag{120b}$$

which, in view of (119), is a backward stochastic system.

Using the transformation (111), (116) yields the forward representation

$$x(t+1) = A(t)x(t) + B(t)w(t) \tag{121a}$$
$$\bar{w}(t) = \bar{B}(t)'x(t) + J(t)w(t), \tag{121b}$$

where $\bar{B}(t) := P(t)^{-\frac{1}{2}}H(t)'$. Likewise (120) and

$$\bar{x}(t) = P(t+1)^{-1}x(t+1), \tag{122}$$

yields the backward representation

$$\bar{x}(t-1) = A(t)'\bar{x}(t) + \bar{B}(t)\bar{w}(t) \tag{123a}$$
$$w(t) = B(t)'\bar{x}(t) + J(t)'\bar{w}(t). \tag{123b}$$

Remark 8: When considered on the doubly infinite time axis, equation (117) defines an isometry. Indeed, assuming that the input is squarely summable, the fact that $U(t)$ is unitary for all t directly implies that

$$\sum_{-\infty}^{N} \|\bar{w}\|^2 + \|\xi(t+1)\|^2 = \sum_{-\infty}^{N} \|w(t)\|^2.$$

Then, $\xi(t) \to 0$ as $t \to \infty$, provided $\Phi(t,s) \to 0$ as $s \to -\infty$. It follows that

$$\sum_{t=-\infty}^{\infty} \|\bar{w}(t)\|^2 = \sum_{t=-\infty}^{\infty} \|w(t)\|^2.$$

We are now in a position to derive a backward version of a non-stationary stochastic system

$$x(t+1) = A(t)x(t) + B(t)w(t), \quad x(0) = x_0 \tag{124a}$$
$$y(t) = C(t)x(t) + D(t)w(t) \tag{124b}$$

where x_0 and the normalized white-noise process w are uncorrelated and $\mathrm{E}\{x_0 x_0'\} = P_0$. In fact, inserting the transformations (122) and (123a) into (124b) yields

$$y(t) = \bar{C}\bar{x}(t) + \bar{D}\bar{w}(t),$$

where

$$\bar{C} = C(t)P(t)A(t)' + D(t)B(t)' \tag{125}$$
$$\bar{D} = C(t)P(t)\bar{B}(t) + D(t)J(t)' \tag{126}$$

From that we have the backward system

$$\bar{x}(t-1) = A(t)'\bar{x}(t) + \bar{B}(t)\bar{w}(t) \tag{127a}$$
$$y(t) = \bar{C}(t)\bar{x}(t) + \bar{D}(t)\bar{w}(t) \tag{127b}$$

with the boundary condition $\bar{x}(T-1) = P(T)^{-1}x(T)$ being uncorrelated to the white-noise process $\bar{w}$.

REFERENCES

[1] D. Q. Mayne, "A solution of the smoothing problem for linear dynamic systems," *Automatica*, vol. 4, pp. 73–92, 1966.

[2] D. Fraser and J. Potter, "The optimum linear smoother as a combination of two optimum linear filters," *Automatic Control, IEEE Transactions on*, vol. 14, no. 4, pp. 387–390, 1969.

[3] F. A. Badawi, A. Lindquist, and M. Pavon, "A stochastic realization approach to the smoothing problem," *IEEE Trans. Automat. Control*, vol. 24, no. 6, pp. 878–888, 1979.

[4] F. Badawi, A. Lindquist, and M. Pavon, "On the Mayne-Fraser smoothing formula and stochastic realization theory for nonstationary linear stochastic systems," in *Decision and Control including the Symposium on Adaptive Processes, 1979 18th IEEE Conference on*, vol. 18. IEEE, 1979, pp. 505–510.

[5] M. Pavon, "New results on the interpolation problem for continuous-time stationary increments processes," *SIAM journal on Control and Optimization*, vol. 22, no. 1, pp. 133–142, 1984.

[6] ——, "Optimal interpolation for linear stochastic systems," *SIAM journal on Control and Optimization*, vol. 22, no. 4, pp. 618–629, 1984.

[7] K. Karhunen, *Zur Interpolation von stationären zufälligen Funktionen*. Suomalainen tiedeakatemia, 1952.

[8] Y. Rozanov, *Stationary random processes*. Holden-Day, San Francisco, 1967.

[9] P. Masani, "Review: Yu.A. Rozanov, stationary random processes," *The Annals of Mathematical Statistics*, vol. 42, no. 4, pp. 1463–1467, 1971.

[10] H. Dym and H. P. McKean, *Gaussian processes, function theory, and the inverse spectral problem*. Courier Dover Publications, 2008.

[11] A. Lindquist and G. Picci, "On the stochastic realization problem," *SIAM J. Control Optim.*, vol. 17, no. 3, pp. 365–389, 1979.

[12] ——, "Forward and backward semimartingale models for Gaussian processes with stationary increments," *Stochastics*, vol. 15, no. 1, pp. 1–50, 1985.

[13] ——, "Realization theory for multivariate stationary Gaussian processes," *SIAM J. Control Optim.*, vol. 23, no. 6, pp. 809–857, 1985.

[14] ——, "A geometric approach to modelling and estimation of linear stochastic systems," *J. Math. Systems Estim. Control*, vol. 1, no. 3, pp. 241–333, 1991.

[15] A. Lindquist and M. Pavon, "On the structure of state-space models for discrete-time stochastic vector processes," *IEEE Trans. Automat. Control*, vol. 29, no. 5, pp. 418–432, 1984.

[16] G. Michaletzky, J. Bokor, and P. Várlaki, *Representability of stochastic systems*. Budapest: Akadémiai Kiadó, 1998.

[17] G. Michaletzky and A. Ferrante, "Splitting subspaces and acausal spectral factors," *J. Math. Systems Estim. Control*, vol. 5, no. 3, pp. 1–26, 1995.

[18] A. Lindquist and G. Picci, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*. Springer-Verlag, Berlin Heidelberg, 2015.

[19] T. T. Georgiou, "The Carathéodory–Fejér–Pisarenko decomposition and its multivariable counterpart," *Automatic Control, IEEE Transactions on*, vol. 52, no. 2, pp. 212–228, 2007.

[20] T. Georgiou and A. Lindquist, "On time-reversibility of linear stochastic models," *arXiv preprint arXiv:1309.0165*, 2013.

[21] J. E. Wall Jr, A. S. Willsky, and N. R. Sandell Jr, "On the fixed-interval smoothing problem," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 5, no. 1-2, pp. 1–41, 1981.

[22] A. N. Kolmogorov, *Stationary sequences in Hilbert space*. John Crerar Library National Translations Center, 1978.

[23] A. M. Yaglom, "On problems about the linear interpolation of stationary random sequences and processes," *Uspekhi Matematicheskikh Nauk*, vol. 4, no. 4, pp. 173–178, 1949.

[24] M. Pavon, "Stochastic realization and invariant directions of the matrix Riccati equation," *SIAM Journal on Control and Optimization*, vol. 18, no. 2, pp. 155–180, 1980.

[25] M. Pavon and A. Wakolbinger, "On free energy, stochastic control, and Schrödinger processes," in *Modeling, Estimation and Control of Systems with Uncertainty*. Springer, 1991, pp. 334–348.

[26] P. Dai Pra and M. Pavon, "On the Markov processes of Schrödinger, the Feynman-Kac formula and stochastic control," in *Realization and Modelling in System Theory*. Springer, 1990, pp. 497–504.

[27] B. Jamison, "Reciprocal processes," *Probability Theory and Related Fields*, vol. 30, no. 1, pp. 65–86, 1974.

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

[29] B. C. Levy, R. Frezza, and A. J. Krener, "Modeling and estimation of discrete-time Gaussian reciprocal processes," *Automatic Control, IEEE Transactions on*, vol. 35, no. 9, pp. 1013–1023, 1990.

[30] P. Dai Pra, "A stochastic control approach to reciprocal diffusion processes," *Applied mathematics and Optimization*, vol. 23, no. 1, pp. 313–329, 1991.