MARYLAND

TECHNICAL
RESEARCH
REPORT

SRC TR 88-85

Discrete-Time Filtering For Linear Systems In Correlated Noise With Non-Gaussian Initial Conditions

by

R.B. Sowers and A.M. Makowski

SYSTEMS RESEARCH CENTER

UNIVERSITY OF MARYLAND

COLLEGE PARK, MARYLAND 20742

DISCRETE-TIME FILTERING FOR LINEAR SYSTEMS IN CORRELATED NOISE WITH NON-GAUSSIAN INITIAL CONDITIONS

by

Richard B. Sowers[1] and Armand M. Makowski[2]
University of Maryland

ABSTRACT

We consider the one-step prediction problem for discrete-time linear systems in correlated plant and observation noises, and non-Gaussian initial conditions. Explicit representations are obtained for the MMSE and LMMSE (or Kalman) estimates of the state given past observations, as well as for the expected square of their difference. These formulae are obtained with the help of the Girsanov transformation for Gaussian white noise sequences, and display explicitly the dependency of the quantities of interest on the initial distribution. Applications of these results can be found in [5] and [6].

Key Words: Filtering, Linear systems, non-Gaussian initial conditions, correlated noises, Girsanov traansformation.

---

[1] Department of Electrical Engineering and Systems Research Center, University of Maryland, College Park, MD 20742. The work of this author was supported through an ONR Graduate Fellowship while completing the requirements towards the M.S. degree. Currently with the Department of Mathematics, University of Maryland, College Park, MD 20742.

[2] Electrical Engineering Department and Systems Research Center, University of Maryland, College Park, Maryland 20742. The work of this author was supported partially through the Engineering Research Centers Programs NSF CDR 88-03012 and partially through NSF grant ECS 83-51836.

I. INTRODUCTION

We consider the one-step prediction problem associated with the stochastic discrete-time linear dynamical system

$$\begin{aligned} X^\circ_{t+1} &= A_t X^\circ_t + W^\circ_{t+1} \\ X^\circ_0 &= \xi \\ Y_t &= H_t X^\circ_t + V^\circ_{t+1} \end{aligned} \qquad t = 0, 1, \ldots \quad (1.1)$$

defined on some probability triple $(\Omega, \mathcal{F}, P)$ which carries the $\mathbb{R}^n$-valued plant process $\{X_t,\ t = 0, 1, \ldots\}$ and the $\mathbb{R}^k$-valued observation process $\{Y_t,\ t = 0, 1, \ldots\}$. Here, for all $t = 0, 1, \ldots$, the matrices A_t and H_t are of dimension $n \times n$ and $n \times k$, respectively. Throughout we make the following assumptions (A.1)-(A.3), where

(A.1): The process $\{(W^\circ_{t+1}, V^\circ_{t+1}),\ t = 0, 1, \ldots\}$ is a zero-mean Gaussian White Noise (GWN) sequence with covariance structure $\{\Gamma_{t+1},\ t = 0, 1, \ldots\}$ given by

$$\Gamma_{t+1} := \mathrm{Cov}\begin{pmatrix} W^\circ_{t+1} \\ V^\circ_{t+1} \end{pmatrix} = \begin{pmatrix} \Sigma^w_{t+1} & \Sigma^{wv}_{t+1} \\ \Sigma^{vw}_{t+1} & \Sigma^v_{t+1} \end{pmatrix}, \qquad t = 0, 1, \ldots \quad (1.2)$$

(A.2): For all $t = 0, 1, \ldots$, the covariance matrix Σ^v_{t+1} is positive definite; and

(A.3): The initial condition ξ has distribution F with finite first and second moments μ and Δ, respectively, and is independent of the process $\{(W^\circ_{t+1}, V^\circ_{t+1}),\ t = 0, 1, \ldots\}$. No *a priori* assumptions, save those on the first two moments, are enforced on F.

The (one-step) prediction problem associated with (1.1) is defined as the problem of computing, for each $t = 0, 1, \ldots$, the conditional distribution of the state X°_{t+1} given the observations $\{Y_0, \ldots, Y_t\}$ or, equivalently, of evaluating the conditional expectation

$$E[\phi(X^\circ_{t+1}) |\ Y_0, \ldots, Y_t] \quad (1.3)$$

for all bounded Borel mappings $\phi : \mathbb{R}^n \to C$, with C denoting set of the complex numbers. In this paper, we solve the prediction problem (1.3) associated with (1.1)-(1.2).

For each $t = 0, 1, \ldots$, once the conditional distribution of X°_{t+1} given $\{Y_0, \ldots, Y_t\}$ is available, it is possible to construct the MMSE estimate $\hat{X}_{t+1} := E[X^\circ_{t+1} |\ Y_0, \ldots, Y_t]$ of X°_{t+1} on the basis of $\{Y_0, \ldots, Y_t\}$. In general, $\hat{X}_{t+1}$ is a *non*-linear function of $\{Y_0, \ldots, Y_t\}$, in contrast to the LMSE (or Kalman) estimate of X°_{t+1} on the basis of $\{Y_0, \ldots, Y_t\}$, which is by definition linear in $\{Y_0, \ldots, Y_t\}$, and which we denote by $\hat{X}^K_{t+1}$. We shall find representations for both $\{\hat{X}_t,\ t = 0, 1, \ldots\}$ and $\{\hat{X}^K_t,\ t = 0, 1, \ldots\}$, and then form the mean square error $\epsilon_t := E\|\hat{X}_t - \hat{X}^K_t\|^2]$ for $t = 1, 2, \ldots$. Simply stated, ϵ_t is a measure of the agreement between the MMSE and LMSE estimates of X°_t on the basis of $\{Y_0, \ldots, Y_{t-1}\}$, for $t = 1, 2, \ldots$.

When the plant and observation noises are *uncorrelated*, and the observation noise sequence $\{V_t,\ t = 0, 1, \ldots\}$ is standard (i.e., $\Sigma^{wv}_{t+1} = 0$ and $\Sigma^v_{t+1} = I_n$ for all $t = 0, 1, \ldots$), the prediction problem posed above is the discrete-time counterpart of the situation investigated in [4]. In Section II, we briefly outline in the discrete-time set-up the basic ingredients of the arguments developed in [4]. We then show in Section III how to modify these ideas in order to solve the prediction problem in the case of correlated noise. Once the solution to the prediction problem is available, we devote Section IV to the derivation of representations for $\{\hat{X}_t,\ t = 1, 2, \ldots\}$, $\{\hat{X}^K_t,\ t = 1, 2, \ldots\}$, and $\{\epsilon_t,\ t = 1, 2, \ldots\}$. A representation for $\{\hat{X}_t,\ t = 1, 2, \ldots\}$ is almost immediate, whereas a representation for $\{\hat{X}^K_t,\ t = 1, 2, \ldots\}$ is found from the former by taking the distribution F to be Gaussian. For each $t = 1, 2, \ldots$, a simple expression for ϵ_t is then found by evaluating the expectation $E[\|\hat{X}_t - \hat{X}^K_t\|^2]$.

The objectives of this paper are twofold: First we demonstrate that the technique of [4] carries over to the correlated noise situation without major difficulties. Moreover, we present expressions for the error terms $\{\epsilon_t,\ t = 1, 2, \ldots\}$ which explicitly display the dependence of the initial distribution F. These formulae form the basis for the large time asymptotic analysis carried out in [6] on the error terms $\{\epsilon_t,\ t = 1, 2, \ldots\}$. Many details have been omitted for the sake of brevity; additional information and material can be found in the thesis [5].

A word on the notation: For any positive integers n and m, we denote the space of $n \times m$ real matrices by $\mathcal{M}_{n\times m}$ and the cone of $n \times n$ symmetric positive-definite matrices by $\mathcal{Q}_n$. As in [4], for every Σ in $\mathcal{Q}_{2n}$, let X_Σ and B_Σ denote generic $I\!R^n$-valued random variables (RV's)such that (X_Σ, B_Σ) is a $I\!R^{2n}$-valued zero-mean Gaussian RV with covariance matrix Σ. For every bounded Borel mapping $\phi : I\!R^n \to \mathcal{C}$, we define the mappings $\mathcal{T}\phi : I\!R^n \times I\!R^n \times \mathcal{Q}_{2n} \to \mathcal{C}$ and $\mathcal{U}\phi : I\!R^n \times I\!R^n \times \mathcal{Q}_n \times \mathcal{M}_{n\times n} \times \mathcal{Q}_{2n} \to \mathcal{C}$ by

$$\mathcal{T}\phi[x, b; \Sigma] := \mathcal{E}\left[\phi(x + X_\Sigma) \exp[b' B_\Sigma]\right] \tag{1.4}$$

and

$$\mathcal{U}\phi[x, b; \Lambda, \Psi; \Sigma] := \mathcal{E}\left[\mathcal{T}\phi[x + \Psi\xi, \xi; \Sigma] \exp[b'\xi - \frac{1}{2}\xi'\Lambda\xi]\right] \tag{1.5}$$

with the understanding that $\mathcal{E}$ denotes integration with respect to the Gaussian distribution of the RV (X_Σ, B_Σ).

Throughout, I_n denote the unit matrix in $\mathcal{M}_{n\times n}$, and let O_n denote the zero element in $\mathcal{M}_{n\times n}$, i.e., the $n \times n$ matrix whose elements are all zero. Elements of $I\!R^n$ are always interpreted as column vectors; transposition is denoted by $'$.

Let $\Phi(\cdot,\cdot)$ be the state transition matrix associated with $\{A_t, t = 0, 1, \ldots\}$, i.e.,

$$\begin{aligned} \Phi(t,t) &= I_n \\ \Phi(s+1,t) &= A_s \Phi(s,t), \qquad s = t, t+1, \ldots \end{aligned} \qquad t = 0, 1, \ldots \tag{1.6}$$

and let $\Psi(\cdot,\cdot)$ be the state transition matrix given by

$$\begin{aligned} \Psi(t,t) &= I_n \\ \Psi(s+1,t) &= [A_s - \Sigma^{wv}_{s+1}(\Sigma^v_{s+1})^{-1} H_s]\Psi(s,t). \qquad s = t, t+1, \ldots \end{aligned} \qquad t = 0, 1, \ldots \tag{1.7}$$

II. THE FILTERING PROBLEM

II.1. The main results

We define the $\mathcal{Q}_n$-valued sequence $\{P_t,\ t = 0, 1, \ldots\}$ by the recursions

$$\begin{aligned} P_{t+1} &= A_t P_t A_t' - [A_t P_t H_t' + \Sigma^{wv}_{t+1}][H_t P_t H_t' + \Sigma^v_{t+1}]^{-1}[A_t P_t H_t' + \Sigma^{wv}_{t+1}]' + \Sigma^w_{t+1} \\ P_0 &= O_n \end{aligned} \qquad t = 0, 1, \ldots \tag{2.1}$$

and, for convenience, we introduce the $\mathcal{Q}_k$-valued sequence $\{J_t,\ t = 0, 1, \ldots\}$, where

$$J_t := H_t P_t H_t' + \Sigma^v_{t+1}. \qquad t = 0, 1, \ldots \tag{2.2}$$

The deterministic sequences $\{Q_t,\ t = 0, 1, \ldots\}$ and $\{R_t,\ t = 0, 1, \ldots\}$ in $\mathcal{M}_{n\times n}$ and $\mathcal{Q}_n$, respectively, are now defined recursively by

$$\begin{aligned} Q_{t+1} &= A_t Q_t - [A_t P_t H_t' + \Sigma^{wv}_{t+1}] J_t^{-1} H_t\big(Q_t + \Psi(t,0)\big) + \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1} H_t \Psi(t,0) \\ Q_0 &= O_n \end{aligned} \qquad t = 0, 1, \ldots \tag{2.3}$$

and

$$\begin{aligned} R_{t+1} &= R_t - \big(Q_t + \Psi(t,0)\big)' H_t' J_t^{-1} H_t \big(Q_t + \Psi(t,0)\big) + \Psi'(t,0) H_t' H_t \Psi(t,0) \\ R_0 &= O_n. \end{aligned} \qquad t = 0, 1, \ldots \tag{2.4}$$

From these sequences, we form the $\mathcal{Q}_{2n}$-valued sequence $\{\Sigma_t,\ t = 0, 1, \ldots\}$ by setting

$$\Sigma_t = \begin{pmatrix} P_t & Q_t \\ Q_t' & R_t \end{pmatrix}. \qquad t = 0, 1, \ldots \tag{2.5}$$

We also generate the $\mathbb{R}^n$-valued processes $\{\bar{X}_t, t = 0,1,\ldots\}$ and $\{\bar{B}_t, t = 0,1,\ldots\}$ via the recursive relations

$$\begin{aligned} \bar{X}_{t+1} &= [A_t - [A_t P_t H_t' + \Sigma_{t+1}^{wv}] J_t^{-1} H_t] \bar{X}_t + [A_t P_t H_t' + \Sigma_{t+1}^{wv}] J_t^{-1} Y_t \\ \bar{X}_0 &= 0 \end{aligned} \qquad t = 0,1,\ldots \quad (2.6)$$

and

$$\begin{aligned} \bar{B}_{t+1} &= \bar{B}_t - (Q_t + \Psi(t,0))' H_t' J_t^{-1} H_t \bar{X}_t + (Q_t + \Psi(t,0))' H_t' J_t^{-1} Y_t \\ \bar{B}_0 &= 0. \end{aligned} \qquad t = 0,1,\ldots \quad (2.7)$$

The solution to the prediction problem associated with (1.1) can now be given. Define the filtration $\{\mathcal{Y}_t,\ t = 0,1,\ldots\}$ of $\mathcal{F}$ as the one generated by the observations $\{Y_t,\ t = 0,1,\ldots\}$, i.e.,

$$\mathcal{Y}_t := \sigma\{Y_0, Y_1, \ldots, Y_t\}. \qquad t = 0,1,\ldots \quad (2.8)$$

Moreover, let $\mathbb{I}$ denote the constant mapping $\mathbb{R}^n \to \mathbb{R} : x \to 1$.

Theorem 1. *For any bounded Borel mapping $\phi : \mathbb{R}^n \to \mathbb{R}$ and any $t = 0,1,\ldots$, the relationship*

$$E[\phi(X_{t+1}^{\circ})|\mathcal{Y}_t] = \frac{U\phi[\bar{X}_{t+1}, \bar{B}_{t+1}; M_{t+1}, \Psi(t+1,0); \Sigma_{t+1}]}{U\mathbb{I}[\bar{X}_{t+1}, \bar{B}_{t+1}; M_{t+1}, \Psi(t+1,0); \Sigma_{t+1}]} \quad P-a.s. \quad (2.9)$$

holds true.

Note that $\Psi(\cdot,\cdot) = \Phi(\cdot,\cdot)$ when $\Sigma_{t+1}^{wv} = O_n$ and $\Sigma_{t+1}^{v} = I_n$ for $t = 0,1,\ldots$, in which case (2.9) reduces to the discrete-time analog of the results of [4]. We readily see that the structure of the predictor in the general situation is not markedly different from what would have been obtained in the uncorrelated case. The noise correlation is encoded in the *universal* sufficient statistics [4] that parametrize the predictor, but does not affect the form of the statistics bearing functionals.

II.2. The discrete-time Girsanov transformation

The proof of these results hinges crucially on a discrete-time version of the Girsanov change of measure transformation [1], which is summarized here for easy reference. Let $\{\mathcal{F}_t,\ t = 0,1,\ldots\}$ be a filtration of $\mathcal{F}$, and let $\{U_{t+1},\ t = 0,1,\ldots\}$ be an $\mathbb{R}^n$-valued zero-mean $(\mathcal{F}_t, P)$ GWN sequence with correlation structure $\Lambda_{t+1} := E[U_{t+1}U_{t+1}']$ for $t = 0,1,\ldots$, i.e., for all $t = 0,1,\ldots$, the RV U_{t+1} is $\mathcal{F}_{t+1}$-measurable and

$$E[\exp[i\theta' U_{t+1}]\,|\mathcal{F}_t] = \exp[-\frac{1}{2}\theta'\Lambda_{t+1}\theta] \qquad t = 0,1,\ldots \quad (2.10)$$

for every θ in $\mathbb{R}^n$. For any $\mathbb{R}^n$-valued $\mathcal{F}_t$-adapted sequence $\{\chi_t,\ t = 0,1,\ldots\}$, we define the sequences $\{\bar{U}_{t+1},\ t = 0,1,\ldots\}$ and $\{L_t,\ t = 0,1,\ldots\}$ taking values in $\mathbb{R}^n$ and $\mathbb{R}$, respectively, by

$$\bar{U}_{t+1} := U_{t+1} - \Lambda_{t+1}\chi_t \qquad t = 0,1,\ldots \quad (2.11)$$

and

$$L_{t+1} := \prod_{s=0}^{t} \exp\left[\chi_s' U_{s+1} - \frac{1}{2}\chi_s'\Lambda_{s+1}\chi_s\right] \qquad t = 1,2,\ldots (2.12)$$

with $L_0 := 1$.

Fix a non-negative integer T, and define a measure $\bar{P}_{T+1}$ on $(\Omega, \mathcal{F})$ by

$$\bar{P}_{T+1}(A) := \int_A L_{T+1}\,dP, \quad A \text{ in } \mathcal{F}. \quad (2.13)$$

It is easy to see that

(a) The measure $\bar{P}_{T+1}$ is a probability measure which agrees with P on $\mathcal{F}_0$, and which is mutually absolutely continuous with P; in fact, its Radon-Nikodym derivative is given by

$$\frac{d\bar{P}_{T+1}}{dP} = L_{T+1}; \tag{2.14}$$

(b) The sequence $\{\bar{U}_{t+1},\ t = 0, 1, \ldots, T\}$ is a zero-mean $(\mathcal{F}_t, \bar{P}_{T+1})$ GWN process with $\bar{E}_{T+1}[\bar{U}_{t+1}\bar{U}'_{t+1}] = \Lambda_{t+1}$ for $t = 0, 1, \ldots, T$ (where $\bar{E}_{T+1}$ is the expectation operator associated with $\bar{P}_{T+1}$); and

(c) The process $\{L_t^{-1}, t = 0, 1, \ldots, T+1\}$ is an $(\mathcal{F}_t, \bar{P}_{T+1})$-martingale.

An alternate expression for (2.12) is simply

$$L_{t+1} := \prod_{s=0}^{t} \exp\left[\chi'_s \bar{U}_{s+1} + \frac{1}{2}\chi'_s \Lambda_{s+1}\chi_s\right]. \qquad t = 1, 2, \ldots \tag{2.15}$$

II.3. The methodology for the uncorrelated case

As noted earlier, the solution to the filtering problem associated with the uncorrelated case can be found in [4] for the continuous-time version of (1.1). We briefly review the arguments of [4] in the discrete-time framework of this paper. Throughout the remainder of this section, we assume $\Sigma^{wv}_{t+1} = O_n$ and $\Sigma^{v}_{t+1} = I_n$ for $t = 0, 1, \ldots$, and fix a positive integer T. A carreful inspection of the solution of [4] reveals that it is articulated around the following two facts **(B.1)** and **(B.2)**, where

(B.1): *A decomposition of the RV's* $\{X^\circ_t,\ t = 0, 1, \ldots\}$ *of the form*

$$X^\circ_t = X_t + Z_t \qquad t = 0, 1, \ldots \tag{2.16}$$

with $\{X_t,\ t = 0, 1, \ldots\}$ *representing the effects of the plant noise process* $\{W^\circ_{t+1},\ t = 0, 1, \ldots\}$ *and* $\{Z_t,\ t = 0, 1, \ldots\}$ *representing the effects of the initial condition* ξ.

The most natural such decomposition is described by the recursions

$$\begin{aligned} X_{t+1} &= A_t X_t + W^\circ_{t+1} \\ X_0 &= 0 \end{aligned} \qquad t = 0, 1, \ldots \tag{2.17}$$

and

$$\begin{aligned} Z_{t+1} &= A_t Z_t \\ Z_0 &= \xi, \end{aligned} \qquad t = 0, 1, \ldots \tag{2.18}$$

in which case $Z_t = \Phi(t, 0)\xi$ for $t = 0, 1, \ldots$. However, for *any* decomposition of the form (2.16) we obtain

$$Y_t = H_t X_t + V_{t+1} \qquad t = 0, 1, \ldots \tag{2.19}$$

where

$$V_{t+1} := V^\circ_{t+1} + H_t Z_t. \qquad t = 0, 1, \ldots \tag{2.20}$$

If $\{(W^\circ_{t+1}, V_{t+1}),\ t = 0, 1, \ldots, T\}$ were a GWN sequence under P, the prediction problem associated with (2.17)-(2.20) would fall within the purview of Kalman filtering. With this in mind, we now use the Girsanov transformation to find a new measure under which to carry out the calculations.

(B.2): *A probability measure* $\bar{P}$ *on* $(\Omega, \mathcal{F})$, *which is mutually absolutely continuous with* P *and which agrees with* P *on* $\sigma\{\xi\}$, *such that under* $\bar{P}$, $\{(W^\circ_{t+1}, V_{t+1}),\ t = 0, 1, \ldots, T\}$ *is a GWN sequence independent of the RV* ξ.

This probability measure $\bar{P}$ is defined by the Radon-Nikodym derivative

$$\frac{d\bar{P}}{dP} := \exp\left[-\sum_{s=0}^{T}[H_s Z_s]' V^\circ_{s+1} - \frac{1}{2}\sum_{s=0}^{T}[H_s Z_s]'[H_s Z_s]\right]. \tag{2.21}$$

In view of this last relation, we define the $I\!R$-valued RV's $\{L_t,\ t=0,1,\ldots\}$ by

$$L_{t+1} := \exp\left[-\xi' \sum_{s=0}^{t} [H_s\Phi(s,0)]' V^{\circ}_{s+1} - \frac{1}{2}\xi' \sum_{s=0}^{t} [H_s\Phi(s,0)]'[H_s\Phi(s,0)]\xi\right] \qquad t=0,1,\ldots \quad (2.22)$$

with $L_0 = 1$, and observe that $d\bar{P}/dP = L_{T+1}$. We may use this probability measure $\bar{P}$ to solve our original filtering problem through the well-known relationship [3, Sec. 27.4]

$$E[\phi(X^{\circ}_{t+1})|\mathcal{Y}_t] = \frac{\bar{E}[\phi(X^{\circ}_{t+1})L^{-1}_{T+1}|\mathcal{Y}_t]}{\bar{E}[L^{-1}_{T+1}|\mathcal{Y}_t]} \quad P-a.s. \qquad (2.23)$$

which holds for each bounded Borel mapping $\phi : I\!R^n \to C$ and $t = 0,1,\ldots,T$. Here $\bar{E}$ denotes the expectation operator associated with $\bar{P}$.

We recall that $\{L^{-1}_t,\ t=0,1,\ldots,T+1\}$ is an $(\mathcal{F}_t,\bar{P})$-martingale by virtue of the Girsanov transformation. Thus, fixing ϕ and $t=0,1,\ldots,T$, we see from the law of iterated conditioning that

$$\begin{aligned}\bar{E}\left[\phi(X^{\circ}_{t+1})\frac{dP}{d\bar{P}}\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\right] &= \bar{E}\left[\bar{E}[\phi(X^{\circ}_{t+1})L^{-1}_{T+1}|\mathcal{F}_{t+1}]\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\right] \\ &= \bar{E}\left[\phi(X^{\circ}_{t+1})L^{-1}_{t+1}|\mathcal{Y}_t \vee \sigma\{\xi\}\right]\end{aligned} \qquad (2.24)$$

since X°_{t+1} is clearly $\mathcal{F}_{t+1}$-measurable and $\mathcal{Y}_t \subset \mathcal{F}_{t+1}$.

To pursue the discussion, we introduce the $I\!R^n$-valued RV's $\{B_t,\ t=0,1,\ldots\}$ and the $\mathcal{Q}_n$-valued sequence $\{M_t,\ t=0,1,\ldots\}$ by setting

$$B_{t+1} := \sum_{s=0}^{t} \Phi(s,0)' H'_s V_{s+1} \qquad t=0,1,\ldots \quad (2.25)$$

and

$$M_{t+1} := \sum_{s=0}^{t} \Phi(s,0)' H'_s H_s \Phi(s,0) \qquad t=0,1,\ldots \quad (2.26)$$

with $B_0 = 0$ and $M_0 = O_n$. From (2.20), (2.25) and (2.26), we observe that

$$L^{-1}_{t+1} = \exp\left[\xi' B_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi\right] \qquad t=0,1,\ldots \quad (2.27)$$

and readily conclude from (2.24) that

$$\begin{aligned}&\bar{E}\left[\phi(X^{\circ}_{t+1})\frac{dP}{d\bar{P}}\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\right] \\ &= \exp\left[-\frac{1}{2}\xi' M_{t+1}\xi\right]\bar{E}\left[\phi(X_{t+1}+\Phi(t+1,0)\xi)\exp[\xi' B_{t+1}]\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\right].\end{aligned} \qquad t=0,1,\ldots \quad (2.28)$$

By property **(B.2)**, we see from (2.17)-(2.19) that under $\bar{P}$, the RV's $\{X_{t+1}, B_{t+1}\}$ and $\{Y_0, Y_1, \ldots, Y_t\}$ are *jointly* Gaussian (and *independent* of the σ-field $\sigma\{\xi\}$). Motivated by standard facts for Gaussian RV's [7, Sec. 2.7], we thus define the MMSE sequences $\{\hat{X}_{t+1},\ t=0,1,\ldots\}$ and $\{\hat{B}_{t+1},\ t=0,1,\ldots\}$ by

$$\hat{X}_{t+1} := \bar{E}[X_{t+1}|\mathcal{Y}_t] \quad \text{and} \quad \hat{B}_{t+1} := \bar{E}[B_{t+1}|\mathcal{Y}_t], \qquad t=0,1,\ldots \quad (2.29)$$

with corresponding errors

$$\tilde{X}_{t+1} := X_{t+1} - \hat{X}_{t+1} \quad \text{and} \quad \tilde{B}_{t+1} := B_{t+1} - \hat{B}_{t+1}. \qquad t=0,1,\ldots \quad (2.30)$$

As in [4], standard arguments [7, Sec. 2.7] imply that the RV's $\{\bar{X}_{t+1}, \bar{B}_{t+1}\}$ are $\bar{P}$-independent of $\mathcal{Y}_t$, whence also $\bar{P}$-independent of the σ-field $\mathcal{Y}_t \vee \sigma\{\xi\}$ since the RV's $\{X_{t+1}, B_{t+1}\}$ and $\{Y_0, Y_1, \ldots, Y_t\}$ are $\bar{P}$-independent of the σ-field $\sigma\{\xi\}$. Moreover, under $\bar{P}$, the $I\!R^{2n}$-valued RV $(\bar{X}_{t+1}, \bar{B}_{t+1})$ is a zero-mean *Gaussian* RV with covariance matrix Σ_{t+1} given by

$$\Sigma_{t+1} := \bar{E}\left[\begin{pmatrix}\bar{X}_{t+1}\\ \bar{B}_{t+1}\end{pmatrix}\begin{pmatrix}\bar{X}_{t+1}\\ \bar{B}_{t+1}\end{pmatrix}'\right] = \begin{pmatrix}P_{t+1} & Q_{t+1}\\ Q'_{t+1} & R_{t+1}\end{pmatrix}. \qquad t = 0, 1, \ldots \quad (2.31)$$

Clearly, the matrices P_{t+1}, Q_{t+1} and R_{t+1} are elements of $\mathcal{Q}_n$, $\mathcal{M}_{n\times n}$ and $\mathcal{Q}_n$, respectively, with the interpretation that

$$P_{t+1} = \bar{E}[\bar{X}_{t+1}\bar{X}'_{t+1}], \; Q_{t+1} = \bar{E}[\bar{X}_{t+1}\bar{B}'_{t+1}] \quad \text{and} \quad R_{t+1} = \bar{E}[\bar{B}_{t+1}\bar{B}'_{t+1}] \quad (2.32)$$

The RV's $\bar{X}_{t+1} + \Phi(t+1, 0)\xi$, $\bar{B}_{t+1}$ and ξ are all $\mathcal{Y}_t \vee \sigma\{\xi\}$-measurable, and from the remarks made earlier, we conclude [2, Prop. 6.1.15] through (2.28) that

$$\begin{aligned}&\bar{E}\left[\phi(X^\circ_{t+1})\frac{dP}{d\bar{P}}\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\right]\\ &= \exp\left[-\frac{1}{2}\xi' M_{t+1}\xi + \xi'\bar{B}_{t+1}\right] \mathcal{T}\phi\left[\bar{X}_{t+1} + \Phi(t+1, 0)\xi, \xi; \Sigma_{t+1}\right]\end{aligned} \qquad t = 0, 1, \ldots \quad (2.33)$$

where the mapping $\mathcal{T}\phi$ is defined by (1.4).

From (2.33), we now readily obtain by the law of iterated conditioning that

$$\begin{aligned}&\bar{E}\left[\phi(X^\circ_{t+1})\frac{dP}{d\bar{P}}\Big|\mathcal{Y}_t\right]\\ &= \bar{E}\left[\exp\left[-\frac{1}{2}\xi' M_{t+1}\xi + \xi'\bar{B}_{t+1}\right] \mathcal{T}\phi\left[\bar{X}_{t+1} + \Phi(t+1, 0)\xi, \xi; \Sigma_{t+1}\right]\Big|\mathcal{Y}_t\right]\\ &= \mathcal{U}\phi\left[\bar{X}_{t+1}, \bar{B}_{t+1}; M_{t+1}, \Phi(t+1, 0); \Sigma_{t+1}\right]\end{aligned} \qquad t = 0, 1, \ldots \quad (2.34)$$

where the mapping $\mathcal{U}\phi$ is defined by (1.5). We have used the fact that the RV's $\{\bar{X}_{t+1}\bar{B}_{t+1}\}$ are $\mathcal{Y}_t$-measurable and therefore $\bar{P}$-independent of $\sigma\{\xi\}$. The reader will readily check that substitution of (2.34) (with arbitrary ϕ and with $\phi = \mathbb{1}$) results in (2.9) since $\Psi(\cdot,\cdot) = \Phi(\cdot,\cdot)$ under the assumptions $\Sigma^{wv}_{t+1} = O_n$ and $\Sigma^v_{t+1} = I_n$ for $t = 0, 1, \ldots$, made here. □

III. THE CORRELATED CASE

We now show how the arguments outlined in the preceding section for the uncorrelated case need to be modified so as to handle the correlated case as well. Let T be a fixed non-negative integer, and consider a decomposition of $\{X^\circ_t,\ t = 0, 1, \ldots\}$ of the form (2.16), and define $\{V_{t+1},\ t = 0, 1, \ldots\}$ by (2.20). If $\Sigma^v_{t+1} = O_n$ for $t = 0, 1, \ldots$, we would arrive at the probability measure $\bar{P}$ characterized by property **(B.2)** as follows: Define the filtration $\{\mathcal{F}_t,\ t = 0, 1, \ldots\}$ by

$$\mathcal{F}_{t+1} := \mathcal{F}_0 \vee \sigma\{V^\circ_{s+1},\ s = 0, 1, \ldots, t\} \qquad t = 0, 1, \ldots (3.1)$$

with $\mathcal{F}_0 := \sigma\{\xi, W^\circ_{s+1},\ s = 0, 1, \ldots\}$, and observe that the sequence $\{V^\circ_{t+1},\ t = 0, 1, \ldots\}$ is an $(\mathcal{F}_t, P)$ zero-mean GWN sequence. The Girsanov transformation implies that $\bar{P}$ as defined in (2.21) enjoys property **(B.2)**. However, if $\Sigma^{wv}_{t+1} \neq O_n$ for $t = 0, 1, \ldots$, then the sequence $\{V^\circ_{t+1},\ t = 0, 1, \ldots\}$ is *not* necessarily an $(\mathcal{F}_t, P)$ zero-mean GWN sequence because now the sequence $\{V^\circ_{t+1},\ t = 0, 1, \ldots\}$ may *not* be independent of $\mathcal{F}_0$, in which case $\bar{P}$ given by (2.21) need *not* enjoy property **(B.2)**.

We may overcome this difficulty when the plant and observation noise sequences have an *arbitrary* covariance structure by performing a Girsanov transformation on the *joint* $I\!R^{n+k}$-valued sequence $\{(W^\circ_{t+1}, V^\circ_{t+1}),\ t = 0, 1, \ldots\}$. With this in mind, we change the definition (3.1) to read instead

$$\mathcal{F}_{t+1} := \mathcal{F}_0 \vee \sigma\{W^\circ_{t+1}, V^\circ_{s+1},\ s = 0, 1, \ldots t\} \qquad t = 0, 1, \ldots \quad (3.2)$$

with $\mathcal{F}_0 := \sigma\{\xi\}$. We now define the $I\!R^{n+k}$-valued sequence $\{(W_{t+1}, V_{t+1}),\ t = 0, 1, \ldots\}$ by

$$\begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix} = \begin{pmatrix} W^{\circ}_{t+1} \\ V^{\circ}_{t+1} \end{pmatrix} - \begin{pmatrix} \Sigma^{w}_{t+1} & \Sigma^{wv}_{t+1} \\ \Sigma^{vw}_{t+1} & \Sigma^{v}_{t+1} \end{pmatrix} \begin{pmatrix} \varphi^{w}_t \\ \varphi^{v}_t \end{pmatrix} \qquad t = 0, 1, \ldots \quad (3.3)$$

where $\{\varphi^w_t,\ t = 0, 1, \ldots\}$ and $\{\varphi^v_t,\ t = 0, 1, \ldots\}$ are $\mathcal{F}_t$-adapted sequences taking values in $I\!R^n$ and $I\!R^k$, respectively, which we yet have to specify. Reviewing the Girsanov transformation, we see that for *any* two such sequences $\{\varphi^w_t,\ t = 0, 1, \ldots\}$ and $\{\varphi^v_t,\ t = 0, 1, \ldots\}$ if we define $\{(W_{t+1}, V_{t+1}),\ t = 0, 1, \ldots\}$ by (3.3), we can find a probability measure $\bar{P}$ on $(\Omega, \mathcal{F})$ satisfying (B.3) where

(B.3): *The probability measure $\bar{P}$ is mutually absolutely continuous with P and agrees with P on $\mathcal{F}_0$. Furthermore, $\{(W_{t+1}, V_{t+1}),\ t = 0, 1, \ldots, T\}$ is a zero-mean $(\mathcal{F}_t, \bar{P})$ GWN sequence with the same covariance structure under $\bar{P}$ as the covariance structure under P of the original noise sequence $\{(W^{\circ}_{t+1}, V^{\circ}_{t+1}),\ t = 0, 1, \ldots, T\}$.*

Now if we impose the constraints (2.20), the sequences $\{\varphi^w_t,\ t = 0, 1, \ldots\}$ and $\{\varphi^v_t,\ t = 0, 1, \ldots\}$ in (3.3) must necessarily have the form

$$\varphi^w_t = \varphi_t \quad \text{and} \quad \varphi^v_t = -(\Sigma^v_{t+1})^{-1}[\Sigma^{vw}_{t+1}\varphi_t + H_t Z_t] \qquad t = 0, 1, \ldots \quad (3.4)$$

for some unspecified $\mathcal{F}_t$-adapted sequence $\{\varphi_t,\ t = 0, 1, \ldots\}$ taking values in $I\!R^n$. Injecting (3.4) into (3.3), we obtain

$$W_{t+1} = W^{\circ}_{t+1} + \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}H_t Z_t [\Sigma^w_{t+1} - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}\Sigma^{vw}_{t+1}]\varphi_t \qquad t = 0, 1, \ldots \quad (3.5)$$

and the appropriate probability measure $\bar{P}$ given by the Girsanov theorem and satisfying (**B.3**) is then defined by

$$\begin{aligned} \frac{d\bar{P}}{dP} = \exp\Bigg[&\sum_{s=0}^{T}\Big[\varphi'_s[W^{\circ}_{s+1} - \Sigma^{wv}_{s+1}(\Sigma^v_{s+1})^{-1}V^{\circ}_{s+1}] - Z'_s H'_s(\Sigma^v_{s+1})^{-1}V^{\circ}_{s+1}\Big] \\ &+ \frac{1}{2}\sum_{s=0}^{T}\Big[\varphi'_s[\Sigma^w_{s+1} - \Sigma^{wv}_{s+1}(\Sigma^v_{s+1})^{-1}\Sigma^{vw}_{s+1}]\varphi_s + Z'_s H'_s(\Sigma^v_{s+1})^{-1}H_s Z_s\Big]\Bigg]. \end{aligned} \quad (3.6)$$

In order to complete the specification of the decomposition (2.16) and of the probability measure (3.6), we must specify $\{X_t,\ t = 0, 1, \ldots\}$, $\{Z_t,\ t = 0, 1, \ldots\}$, and $\{\varphi_t,\ t = 0, 1, \ldots\}$. To that end we rewrite the evolution of $\{X^{\circ}_t,\ t = 0, 1, \ldots\}$ in terms of $\{X_t,\ t = 0, 1, \ldots\}$, $\{Z_t,\ t = 0, 1, \ldots\}$ and $\{W_{t+1},\ t = 0, 1, \ldots\}$. Since we wish to use the properties of $\bar{P}$, it is more natural to write this evolution in terms of $\{W_{t+1},\ t = 0, 1, \ldots\}$ rather than in terms of $\{W^{\circ}_{t+1},\ t = 0, 1, \ldots\}$, and this leads to

$$\begin{aligned} X_{t+1} + Z_{t+1} &= X^{\circ}_{t+1} \\ &= A_t X^{\circ}_t + W^{\circ}_{t+1} \\ &= A_t(X_t + Z_t) + W_{t+1} - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}H_t Z_t \\ &\quad + [\Sigma^w_{t+1} - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}\Sigma^{vw}_{t+1}]\varphi_t \\ &= A_t X_t + [A_t - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}H_t]Z_t + W_{t+1} \\ &\quad + [\Sigma^w_{t+1} - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}\Sigma^{vw}_{t+1}]\varphi_t. \end{aligned} \qquad t = 0, 1, \ldots \quad (3.7)$$

This suggests a separation of the dynamics in the form

$$\begin{aligned} X_{t+1} &= A_t X_t + W_{t+1} + [\Sigma^w_{t+1} - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}\Sigma^{vw}_{t+1}]\varphi_t - \pi_t \\ X_0 &= \zeta \end{aligned} \qquad t = 0, 1, \ldots \quad (3.8)$$

and

$$\begin{aligned} Z_{t+1} &= [A_t - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}H_t]Z_t + \pi_t \\ Z_0 &= \xi - \zeta \end{aligned} \qquad t = 0, 1, \ldots \quad (3.9)$$

where ζ and $\{\pi_t,\ t=0,1,\ldots\}$ are $I\!R^n$-valued RV's yet to be specified. We shall simply assume that

$$\varphi_t = 0, \quad \pi_t = 0 \quad \text{and} \quad \zeta = 0. \qquad t=0,1,\ldots \quad (3.10)$$

At this point, a summary of the relevant quantities is in order under the constraints (3.10).

- **The effect of the initial condition**

$$\begin{aligned} Z_{t+1} &= [A_t - \Sigma^{wv}_{t+1}(\Sigma^v_{t+1})^{-1}H_t]Z_t \\ Z_0 &= \xi, \end{aligned} \qquad t=0,1,\ldots \quad (3.11)$$

which may also be written as $Z_t = \Psi(t,0)\xi$ for $t=0,1,\ldots$.

- **The noise processes**

$$\begin{aligned} \begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix} &= \begin{pmatrix} W^\circ_{t+1} \\ V^\circ_{t+1} \end{pmatrix} - \begin{pmatrix} \Sigma^w_{t+1} & \Sigma^{wv}_{t+1} \\ \Sigma^{vw}_{t+1} & \Sigma^v_{t+1} \end{pmatrix} \begin{pmatrix} 0 \\ -(\Sigma^v_{t+1})^{-1}H_tZ_t \end{pmatrix} \\ &= \begin{pmatrix} W^\circ_{t+1} + \Sigma^{vw}_{t+1}(\Sigma^v_{t+1})^{-1}H_tZ_t \\ V^\circ_{t+1} + H_tZ_t \end{pmatrix}. \end{aligned} \qquad t=0,1,\ldots \quad (3.12)$$

- **The auxiliary system**

$$\begin{aligned} X_{t+1} &= A_tX_t + W_{t+1} \\ X_0 &= 0 \\ Y_t &= H_tX_t + V_{t+1}. \end{aligned} \qquad t=0,1,\ldots \quad (3.13)$$

- **The change of measure**

$$\frac{d\bar{P}}{dP} = \exp\Big[-\sum_{s=0}^{T} Z'_sH'_s(\Sigma^v_{s+1})^{-1}V^\circ_{s+1} + \frac{1}{2}\sum_{s=0}^{T} Z'_sH'_s(\Sigma^v_{s+1})^{-1}H_sZ_s\Big]. \quad (3.14)$$

The properties of our decomposition and change of measure are summarized in

Proposition 1. *Let the filtration $\{\mathcal{F}_t,\ t=0,1,\ldots\}$ be given by (3.2). If the sequences $\{X_t,\ t=0,1,\ldots\}$, $\{Z_t,\ t=0,1,\ldots\}$ and $\{(W_{t+1},V_{t+1}),\ t=0,1,\ldots\}$ are defined by (3.11)-(3.13) and if the probability measure $\bar{P}$ is defined by (3.14), then $\bar{P}$ and P are mutually absolutely continuous and the process $\{(W_{t+1},V_{t+1}),\ t=0,1,\ldots,T\}$ is a zero-mean $(\mathcal{F}_t,\bar{P})$ GWN sequence with covariance structure structure $\{\Gamma_{t+1},\ t=0,1,\ldots,T\}$ under $\bar{P}$.*

Motivated by the form of (3.14), we define the $I\!R$-valued sequence $\{L_t,\ t=0,1,\ldots\}$ by

$$L_{t+1} := \exp\Big[-\sum_{s=0}^{t} Z'_sH'_s(\Sigma^v_{s+1})^{-1}V^\circ_{s+1} + \frac{1}{2}\sum_{s=0}^{t} Z'_sH'_s(\Sigma^v_{s+1})^{-1}H_sZ_s\Big] \qquad t=0,1,\ldots (3.15)$$

with $L_0 = 1$, and observe that $d\bar{P}/dP = L_{T+1}$. The Girsanov transformation now implies that $\{L_t^{-1},\ t=0,1,\ldots,T+1\}$ is an $(\mathcal{F}_t,\bar{P})$-martingale, and by the same arguments as the ones leading to (2.24) we conclude that

$$\bar{E}\Big[\phi(X^\circ_{t+1})\frac{d\bar{P}}{dP}\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\Big] = \bar{E}\big[\phi(X^\circ_{t+1})L^{-1}_{t+1}\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big]. \qquad t=0,1,\ldots \quad (3.16)$$

Since

$$L^{-1}_{t+1} = \exp\Big[\sum_{s=0}^{t} Z'_sH'_s(\Sigma^v_{s+1})^{-1}V^\circ_{s+1} - \frac{1}{2}\sum_{s=0}^{t} Z'_sH'_s(\Sigma^v_{s+1})^{-1}H_sZ_s\Big], \qquad t=0,1,\ldots (3.17)$$

we see from (2.20) that

$$L^{-1}_{t+1} = \exp\Big[\xi'B_{t+1} - \frac{1}{2}\xi'M_{t+1}\xi\Big] \qquad t=0,1,\ldots \quad (3.18)$$

where

$$B_{t+1} := \sum_{s=0}^{t} \Psi(s,0)' H_s' (\Sigma^v_{s+1})^{-1} V_{s+1} \qquad t = 0,1,\ldots \quad (3.19)$$

and

$$M_{t+1} := \sum_{s=0}^{t} \Psi(s,0)' H_s' (\Sigma^v_{s+1})^{-1} H_s \Psi(s,0) \qquad t = 0,1,\ldots \quad (3.20)$$

with $B_0 = 0$ and $M_0 = O_n$.

As before, we define the sequences of conditional means $\{\bar{X}_{t+1},\ t = 0,1,\ldots\}$ and $\{\bar{B}_{t+1},\ t = 0,1,\ldots\}$ by (2.29), with corresponding errors $\{\tilde{X}_{t+1},\ t = 0,1,\ldots\}$ and $\{\tilde{B}_{t+1},\ t = 0,1,\ldots\}$ given by (2.30). The RV's X_{t+1}, B_{t+1}, and $\{Y_0,\ldots,Y_t\}$ may all be represented as linear combinations of $\{(W_{t+1}, V_{t+1}),\ t = 0,1,\ldots,T\}$, and are thus jointly Gaussian and independent of $\sigma\{\xi\}$ under $\bar{P}$. As argued in the uncorrelated case, under $\bar{P}$, the $I\!R^{2n}$-valued RV $(\tilde{X}_{t+1}, \tilde{B}_{t+1})$ is a zero-mean *Gaussian* RV with covariance matrix Σ_{t+1} which is is $\bar{P}$*-independent* of the σ-field $\mathcal{Y}_t \vee \sigma\{\xi\}$. Hence standard results on conditional expectations [2, Prop. 6.1.15] validates the following chain of equalities

$$\bar{E}\left[\phi(X^o_{t+1}) \exp[\xi' B_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi] | \mathcal{Y}_t \vee \sigma\{\xi\}\right]$$

$$\begin{aligned} &= \bar{E}\left[\phi(\tilde{X}_{t+1} + \bar{X}_{t+1} + \Psi(t+1,0)\xi) \exp[\xi'\tilde{B}_{t+1} + \xi'\bar{B}_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi] | \mathcal{Y}_t \vee \sigma\{\xi\}\right] \\ &= \exp[-\frac{1}{2}\xi' M_{t+1}\xi] \bar{E}\left[\phi(\tilde{X}_{t+1} + x)\exp[b'\tilde{B}_{t+1}]\right]_{x=\bar{X}_{t+1}+\Psi(t+1,0)\xi,\, b=\bar{B}_{t+1}} \\ &= \exp[-\frac{1}{2}\xi' M_{t+1}\xi] \mathcal{T}\phi[\bar{X}_{t+1} + \Psi(t+1,0)\xi, \xi; \Sigma_{t+1}] \end{aligned} \quad (3.21)$$

where Σ_{t+1} has the decomposition (2.31)-(2.32). Removing the conditioning upon $\sigma\{\xi\}$, we find

$$\begin{aligned} \bar{E}[\phi(X^o_{t+1})\frac{dP}{d\bar{P}}|\mathcal{Y}_t] &= \bar{E}[\mathcal{T}\phi[\bar{X}_{t+1} + \Psi(t+1,0)\xi, \xi; \Sigma_{t+1}] \exp[\xi'\bar{B}_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi]|\mathcal{Y}_t] \\ &= \mathcal{U}\phi[\bar{X}_{t+1}, \bar{B}_{t+1}; M_{t+1}, \Psi(t+1,0); \Sigma_{t+1}] \end{aligned} \quad (3.22)$$

since $(\bar{X}_{t+1}, \bar{B}_{t+1})$ is $\mathcal{Y}_t$-measurable and therefore $\bar{P}$-independent of $\sigma\{\xi\}$.

At this point, we have solved the prediction problem over the finite horizon $t = 0,1,\ldots,T$. Indeed we readily obtain (2.9) by injecting (3.22) (for arbitrary ϕ and for $\phi = \mathbb{1}$) into (2.23). The only remaining problem is to calculate $\{(\bar{X}_t, \bar{B}_t),\ t = 0,1,\ldots,T+1\}$ and $\{\Sigma_{t+1},\ t = 0,1,\ldots,T+1\}$. We combine (3.13) and (3.19) to rewrite the dynamics of $\{(X_t, B_t),\ t = 0,1,\ldots,T+1\}$ and $\{Y_t,\ t = 0,1,\ldots,T\}$ by

$$\begin{aligned} \begin{pmatrix} X_{t+1} \\ B_{t+1} \end{pmatrix} &= \begin{pmatrix} A_t & 0 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} X_t \\ B_t \end{pmatrix} + \begin{pmatrix} I_n & 0 \\ 0 & \Psi'(t,0)H_t'(\Sigma^v_{t+1})^{-1} \end{pmatrix} \begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix} \\ \begin{pmatrix} X_0 \\ B_0 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad\qquad t = 0,1,\ldots \quad (3.23) \\ Y_t &= \begin{pmatrix} H_t & 0 \end{pmatrix} \begin{pmatrix} X_t \\ B_t \end{pmatrix} + \begin{pmatrix} 0 & I_k \end{pmatrix} \begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix}. \end{aligned}$$

By applying the Kalman filtering equations to this system (under $\bar{P}$), after appropriate identification, we easily arrive at the equations (2.1)-(2.7) satisfied by the sequence of $I\!R^{2n}$-valued RV's $\{(\bar{X}_t, \bar{B}_t),\ t = 0,1,\ldots,T+1\}$ and the $\mathcal{M}_{2n\times 2n}$-valued sequence $\{\Sigma_t,\ t = 0,1,\ldots,T+1\}$. The calculations are tedious, and the details are left to the interested reader [5].

The final step now consists in extending these results from the finite horizon $t = 0,1,\ldots,T$ to the infinite horizon $t = 0,1,\ldots$. To that end, note the following: The dynamics of the sequences $\{(\bar{X}_t, \bar{B}_t),\ t = 0,1,\ldots,T+1\}$ and $\{\Sigma_t,\ t = 0,1,\ldots,T+1\}$ are *independent* of T. Moreover, although the transformed measure $\bar{P}$ used in the derivation depends *a priori* on T, the definitions of the mappings $\mathcal{T}\phi$ and $\mathcal{U}\phi$ are independent of T. These remarks are sufficient to yield Theorem 1 from the finite-horizon results of this section. □

Following on the comments made at the end of the proof, we could have displayed explicitly the dependence of the transformed measure $\bar{P}$ on the parameter T, say through the notation $\bar{P}_{T+1}$. Although $\bar{P}_{T+1} = \bar{P}_T$ on the σ-field $\mathcal{F}_T$ for all $T = 0, 1, \ldots$, and the probability measure $\bar{P}_{T+1}$ is mutually absolutely continuous with respect to P, it is *not* true in general [5] that the *projective system* $\{\bar{P}_T,\ T = 0, 1, \ldots\}$ has a limit $\bar{P}$ which is absolutely continuous with respect to P on the σ-field $\vee_T \mathcal{F}_T$, i.e., there does not exist necessarily a probability measure $\bar{P}$ on $\vee_T \mathcal{F}_T$ such that $\bar{P}$ is absolutely continuous with respect to P, and $\bar{P}_T = \bar{P}$ on the σ-field $\mathcal{F}_T$ for all $T = 0, 1, \ldots$. Although this could *a priori* complicate matters for the infinite-horizon situation, we shall not concern ourselves with this difficulty in what follows. Indeed, in the remainder of this paper, only statements for finite t will be made and the notation $\bar{P}$ (and $\bar{E}$) will be used throughout with the understanding that $\bar{P} = \bar{P}_{T+1}$ for some $t < T$. As should be clear from earlier comments, the exact choice of T is irrelevant.

IV. REPRESENTATIONS FOR $\{\hat{X}_t,\ t = 0, 1, \ldots\}$, $\{\hat{X}_t^K,\ t = 0, 1, \ldots\}$ AND $\{\epsilon_t,\ t = 0, 1, \ldots\}$.

Using Theorem 1, we now develop formulae for $\{\hat{X}_t,\ t = 0, 1, \ldots\}$, $\{\hat{X}_t^K,\ t = 0, 1, \ldots\}$ and $\{\epsilon_t,\ t = 0, 1, \ldots\}$. We do this under the additional assumption (A.4), where

(A.4): The covariance matrix Δ is positive-definite.

To state these representation results, we find it convenient to introduce the auxiliary quantities $\{Q_t^*,\ t = 0, 1, \ldots\}$ and $\{R_t^*,\ t = 0, 1, \ldots\}$ in $\mathcal{M}_{n\times n}$ and $\mathcal{Q}_n$, respectively, by setting

$$Q_t^* := Q_t + \Psi(t, 0) \quad \text{and} \quad R_t^* := M_t - R_t. \qquad t = 0, 1, \ldots \tag{4.1}$$

With this notation, we then have

Theorem 2. *For all $t = 0, 1, \ldots$, the representations*

$$\hat{X}_{t+1} = \bar{X}_{t+1} + Q_{t+1}^* \frac{\int_{\mathbb{R}^n} z \exp\left[z'\bar{B}_{t+1} - \frac{1}{2} z' R_{t+1}^* z\right] dF(z)}{\int_{\mathbb{R}^n} \exp\left[z'\bar{B}_{t+1} - \frac{1}{2} z' R_{t+1}^* z\right] dF(z)} \tag{4.2}$$

and

$$\hat{X}_{t+1}^K = \bar{X}_{t+1} + Q_{t+1}^* \left[R_{t+1}^* + \Delta^{-1}\right]^{-1} \left[\bar{B}_{t+1} + \Delta^{-1}\mu\right] \tag{4.3}$$

hold P-a.s.

Before giving a proof of this result, several points are in order:

(i): The estimates $\hat{X}_{t+1}$ and $\hat{X}_{t+1}^K$ are well defined for $t = 0, 1, \ldots$, for we may write

$$X_{t+1}^\circ = \Phi(t+1, 0)\xi + \sum_{s=0}^{t} \Phi(t, s) W_{s+1}^\circ. \qquad t = 0, 1, \ldots \tag{4.4}$$

Since the RV's ξ and $\{(W_{t+1}^\circ, V_{t+1}^\circ),\ t = 0, 1, \ldots\}$ are all P-square-integrable, so are the RV's $\{X_{t+1}^\circ, Y_t,\ t = 0, 1, \ldots\}$, and the claim follows.

(ii): The expression (4.3) provides a non-standard representation for the Kalman filter associated with system (1.1). This representation is notable in that it explicitly displays the effects of the mean μ and covariance Δ of the initial condition ξ; the only dependence of the filtering formulae on μ and Δ is through the affine mapping $x \mapsto [R_{t+1}^* + \Delta^{-1}]^{-1}[x + \Delta^{-1}\mu]$.

(iii): Observing from (3.20) that

$$M_{t+1} = M_t + \Psi(t, 0)' H_t' (\Sigma_{t+1})^{-1} H_t \Psi(t, 0), \qquad t = 0, 1, \ldots \tag{4.5}$$

we readily see from (1.6), (2.3) and (2.4) that

$$Q^*_{t+1} = \left[A_t - [A_t P_t H'_t + \Sigma^{wv}_{t+1}]J_t^{-1}H_t\right]Q^*_t \qquad t = 0, 1, \ldots \quad (4.6)$$
$$Q^*_0 = I_n,$$

and

$$R^*_{t+1} = R^*_t + Q^{*\prime}_t H'_t J_t^{-1} H_t Q^*_t \qquad t = 0, 1, \ldots \quad (4.7)$$
$$R^*_0 = O_n.$$

Note also that the dynamics (2.7) then simplifies into

$$\bar{B}_{t+1} = \bar{B}_t - Q^{*\prime}_t H'_t J_t^{-1} H_t \hat{X}_t + Q^{*\prime}_t H'_t J_t^{-1} Y_t \qquad t = 0, 1, \ldots \quad (4.8)$$
$$\bar{B}_0 = 0.$$

The following two technical lemmas will be useful in the forthcoming discussion.

Lemma 1. *For $t = 0, 1, \ldots$, $\bar{B}_{t+1}$ is a zero-mean Gaussian RV with covariance R^*_{t+1} under $\bar{P}$.*

Proof. Fix $t = 0, 1, \ldots$. The RV $\bar{B}_{t+1}$ is normally distributed under $\bar{P}$. Moreover, we note from the Orthogonality Principle that

$$\bar{E}[B_{t+1}] = \bar{E}[\bar{B}_{t+1}] \quad (4.9)$$

and

$$\bar{E}[B_{t+1}B'_{t+1}] = \bar{E}[\bar{B}_{t+1}\bar{B}'_{t+1}] + \bar{E}[\tilde{B}_{t+1}\tilde{B}'_{t+1}] \quad (4.10)$$

since $\bar{E}[\tilde{B}_{t+1}] = 0$ and $\bar{E}[\bar{B}_{t+1}\tilde{B}'_{t+1}] = 0$.

The definition (3.19) implies the relation

$$B_{t+1} = B_t + \Psi'(t,0)H'_t(\Sigma^v_{t+1})^{-1}V_{t+1}, \quad (4.11)$$

so that

$$\bar{E}[B_{t+1}] = \bar{E}[B_t] + \Psi'(t,0)H'_t(\Sigma^v_{t+1})^{-1}\bar{E}[V_{t+1}] = \bar{E}[B_t] \quad (4.12)$$

and

$$\bar{E}[B_{t+1}B'_{t+1}] = \bar{E}[B_tB'_t] + \Psi'(t,0)H'_t(\Sigma^v_{t+1})^{-1}H_t\Psi(t,0) \quad (4.13)$$

since the RV B_t is $\bar{P}$-independent of V_{t+1}. From the fact $\bar{B}_0 = 0$, we readily conclude that $\bar{E}[B_t] = 0$ for *all* $t = 0, 1, \ldots$, while by comparison of (4.5) with (4.13) we see that necessarily $\bar{E}[B_tB'_t] = M_t$ for *all* $t = 0, 1, \ldots$. Finally, from the definitions (2.29)-(2.30), we see that

$$\bar{E}[\bar{B}_{t+1}] = \bar{E}[B_{t+1}] = 0 \qquad t = 0, 1, \ldots \quad (4.14)$$

and

$$\bar{E}[\bar{B}_{t+1}\bar{B}'_{t+1}] = \bar{E}[B_{t+1}B'_{t+1}] - \bar{E}[\tilde{B}_{t+1}\tilde{B}'_{t+1}] \qquad t = 0, 1, \ldots \quad (4.15)$$
$$= M_{t+1} - R_{t+1} = R^*_{t+1}.$$

□

Lemma 2. *For any $t = 0, 1, \ldots$ and any $I\!R$-valued, nonnegative $\mathcal{Y}_t \vee \sigma\{\xi\}$-measurable RV X, the relation*

$$E[X] = \bar{E}\left[X \exp[\xi'\bar{B}_{t+1} - \frac{1}{2}\xi' R^*_{t+1}\xi]\right] \quad (4.16)$$

holds true.

Proof. We readily see that

$$\begin{aligned}E[X] &= \bar{E}[XL_{t+1}^{-1}] \\ &= \bar{E}\big[X\exp[\xi' B_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi]\big] = \bar{E}\big[X\frac{d\bar{P}}{dP}\big] \\ &= \bar{E}\Big[\bar{E}\big[X\exp[\xi' B_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi]\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big]\Big] \\ &= \bar{E}\Big[X\exp[-\frac{1}{2}\xi' M_{t+1}\xi]\bar{E}\Big[\exp[\xi' B_{t+1}]\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\Big],\end{aligned} \tag{4.17}$$

where we have used the law of iterated conditioning and the measurability of $X\exp[-\frac{1}{2}\xi' M_{t+1}\xi]$ with respect to the σ-field $\mathcal{Y}_t \vee \sigma\{\xi\}$. But

$$\bar{E}\big[\exp[\xi' B_{t+1}]\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big] = \bar{E}\big[\exp[\xi'\tilde{B}_{t+1}]\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big]\exp[\xi'\bar{B}_{t+1}], \tag{4.18}$$

since the RV $\bar{B}_{t+1}$ is $\mathcal{Y}_t \vee \sigma\{\xi\}$-measurable. Under $\bar{P}$, $\tilde{B}_{t+1}$ is a zero-mean Gaussian RV with covariance R_{t+1}, which is independent of the σ-field $\mathcal{Y}_t \vee \sigma\{\xi\}$. Consequently, we get

$$\bar{E}\big[\exp[\xi'\tilde{B}_{t+1}]\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big] = \bar{E}\big[\exp[x'\tilde{B}_{t+1}]\big]\Big|_{x=\xi} = \exp\big[\frac{1}{2}\xi' R_{t+1}\xi\big] \tag{4.19}$$

and (4.16) follows from (4.1), (4.17) and (4.19). □

A proof of Theorem 2. Our first step consists in finding a representation for the conditional characteristic function $E\big[\exp[i\theta' X^\circ_{t+1}]\big|\mathcal{Y}_t\big]$. Under the enforced moment assumptions on ξ, we then recover an expression for the conditional mean by differentiating with respect to θ. Finally, by substituting a Gaussian distribution for F in this representation for $\hat{X}_{t+1}$, we obtain a formula for $\hat{X}^K_{t+1}$:

For θ in $\mathbb{R}^n$, we define the mapping $\psi_\theta : \mathbb{R}^n \to \mathcal{C}$ by

$$\psi_\theta(x) := \exp[i\theta' x], \quad x \text{ in } \mathbb{R}^n \tag{4.20}$$

and introduce the $\mathcal{C}$-valued RV $q_{t+1}(\theta)$ by setting

$$q_{t+1}(\theta) := \mathcal{U}\psi_\theta[\bar{X}_{t+1}, \bar{B}_{t+1}; M_{t+1}, \Psi(t+1,0); \Sigma_{t+1}]. \qquad t = 0,1,\ldots \tag{4.21}$$

Since $\psi_0 = \mathbb{1}$, we obtain from Theorem 1 that

$$E\big[\exp[i\theta' X^\circ_{t+1}]\big|\mathcal{Y}_t\big] = \frac{q_{t+1}(\theta)}{q_{t+1}(0)}, \qquad t = 0,1,\ldots \tag{4.22}$$

whence

$$\hat{X}_{t+1} = -i\nabla_\theta E\big[\exp[i\theta' X^\circ_{t+1}]\big|\mathcal{Y}_t\big]_{\theta=0} = -i\frac{\nabla_\theta q_{t+1}(\theta)\big|_{\theta=0}}{q_{t+1}(0)} \qquad t = 0,1,\ldots \tag{4.23}$$

since the RV X°_{t+1} is P-integrable.

Fix $t = 0,1,\ldots$ and θ in $\mathbb{R}^n$. For all u and v in $\mathbb{R}^n$, we easily check (with the notation of Section 1) that

$$\begin{aligned}\mathcal{T}\psi_\theta\big[u, v; \Sigma_{t+1}\big] &= \mathcal{E}\Big[\exp[i\theta' u]\exp\Big[\begin{pmatrix} i\theta \\ v\end{pmatrix}'\begin{pmatrix} X_{\Sigma_{t+1}} \\ B_{\Sigma_{t+1}}\end{pmatrix}\Big]\Big] \\ &= \exp[i\theta' u]\exp\big[-\frac{1}{2}\theta' P_{t+1}\theta + i\theta' Q_{t+1}v + \frac{1}{2}v' R_{t+1}v\big],\end{aligned} \tag{4.24}$$

whence

$$\mathcal{T}\psi_\theta\big[x + \Psi(t+1,0)\xi, \xi; \Sigma_{t+1}\big] = \exp\big[ix'\theta - \frac{1}{2}\theta' P_{t+1}\theta\big]\exp\big[-i\theta'(Q_{t+1} - \Psi(t+1,0))\xi + \frac{1}{2}\xi' R_{t+1}\xi\big]$$

for all x in $I\!R^n$. From (1.5) we have

$$\begin{aligned} q_{t+1}(\theta) &= \mathcal{E}\left[T\psi_\theta\left[x+\Psi(t+1,0)\xi,\xi;\Sigma_{t+1}\right]\exp\left[b'\xi-\frac{1}{2}\xi' M_{t+1}\xi\right]\right]_{x=\bar{X}_{t+1},b=\bar{B}_{t+1}} \\ &= \exp\left[i\theta'\bar{X}_{t+1}-\frac{1}{2}\theta' P_{t+1}\theta\right]\mathcal{E}\left[\exp\left[i\theta' Q^*_{t+1}\xi+b'\xi-\frac{1}{2}\xi' R^*_{t+1}\xi\right]\right]_{x=\bar{X}_{t+1},b=\bar{B}_{t+1}} \\ &= \exp[i\theta'\bar{X}_{t+1}-\frac{1}{2}\theta' P_{t+1}\theta]\int_{I\!R^n}\exp\left[z'(iQ^{*\prime}_{t+1}\theta+\bar{B}_{t+1})-\frac{1}{2}z' R^*_{t+1}z\right]dF(z) \end{aligned} \quad (4.25)$$

and (4.22) now takes the form

$$\begin{aligned} &E\left[\exp[i\theta' X^\circ_{t+1}]\middle|\mathcal{Y}_t\right] \\ &=\exp\left[i\theta'\bar{X}_{t+1}-\frac{1}{2}\theta' P_{t+1}\theta\right]\frac{\int_{I\!R^n}\exp\left[z'\left[iQ^{*\prime}_{t+1}\theta+\bar{B}_{t+1}\right]-\frac{1}{2}z' R^*_{t+1}z\right]dF(z)}{\int_{I\!R^n}\exp\left[z'\bar{B}_{t+1}-\frac{1}{2}z' R^*_{t+1}z\right]dF(z)}\ . \end{aligned} \quad (4.26)$$

It is now a simple exercise to verify that under the enforced integrability assumptions, the equality

$$\begin{aligned} &\left[\nabla_\theta\int_{I\!R^n}\exp\left[z'(iQ^{*\prime}_{t+1}\theta+\bar{B}_{t+1})-\frac{1}{2}z' R^*_{t+1}z\right]dF(z)\right]_{\theta=0} \\ &= iQ^*_{t+1}\int_{I\!R^n} z\exp[z'\bar{B}_{t+1}-\frac{1}{2}z' R^*_{t+1}z]dF(z) \end{aligned} \quad (4.27)$$

holds P-a.s.; the proof of (4.27) is omitted fror the sake of brevity as the details are available in [5] for the interested reader. The result (4.2) now follows upon combining (4.23), (4.26) and (4.27).

Now let G be a Gaussian distribution with mean μ and covariance Δ. A representation for $\hat{X}^K_{t+1}$ for $t=0,1,\ldots$ may be found by simply substituting G to F in (4.2). This is so because when $F=G$, the MMSE estimate of X°_{t+1} given the observations $\{Y_0,\ldots,Y_t\}$ coincides with the estimate generated by the Kalman filter, whose dependence on the initial condition is only through its second order properties. The claim then follows from the observation that F and G have the same mean and covariance.

Recalling the assumption (**A.4**) that Δ be positive-definite, we find by direct evaluation that the expressions

$$\begin{aligned} &\int_{I\!R^n} z\exp[z'\bar{B}_{t+1}-\frac{1}{2}z' R^*_{t+1}z]dG(z) \\ &=\frac{\left[R^*_{t+1}+\Delta^{-1}\right]^{-1}\left[\bar{B}_{t+1}+\Delta^{-1}\mu\right]}{\sqrt{\det[\Delta R^*_{t+1}+I_n]}}\cdot\exp\left[-\frac{1}{2}\mu'\Delta^{-1}\mu+\frac{1}{2}\left[\bar{B}_{t+1}+\Delta^{-1}\mu\right]'\left[R^*_{t+1}+\Delta^{-1}\right]^{-1}\left[\bar{B}_{t+1}+\Delta^{-1}\mu\right]\right] \end{aligned} \quad (4.28)$$

and

$$\begin{aligned} &\int_{I\!R^n}\exp[z'\bar{B}_{t+1}-\frac{1}{2}z' R^*_{t+1}z]dG(z) \\ &=\frac{1}{\sqrt{\det[\Delta R^*_{t+1}+I_n]}}\cdot\exp\left[-\frac{1}{2}\mu'\Delta^{-1}\mu+\frac{1}{2}\left[\bar{B}_{t+1}+\Delta^{-1}\mu\right]'\left[R^*_{t+1}+\Delta^{-1}\right]^{-1}\left[\bar{B}_{t+1}+\Delta^{-1}\mu\right]\right] \end{aligned} \quad (4.29)$$

hold P-a.s. for $t=0,1,\ldots$. The conclusion (4.3) follows as we use (4.28) and (4.29) in (4.2). □

Theorem 2 now leads us to a simple representation of the errors $\{\epsilon_{t+1},\ t=0,1,\ldots\}$. In what follows, for each Λ in $\mathcal{Q}_n$, G_Λ denotes a normal distribution with zero mean and covariance Λ.

Theorem 3. *The representation*

$$\epsilon_{t+1}=\int_{I\!R^n}\frac{\left\|Q^*_{t+1}\int_{I\!R^n}\{z-[R^*_{t+1}+\Delta^{-1}]^{-1}[b+\Delta^{-1}\mu]\}\exp[z'b-\frac{1}{2}z' R^*_{t+1}z]dF(z)\right\|^2}{\int_{I\!R^n}\exp[z'b-\frac{1}{2}z' R^*_{t+1}z]dF(z)}dG_{R^*_{t+1}}(b)$$

$$t = 0, 1, \ldots \quad (4.30)$$

holds true.

Proof. We observe directly from Theorem 2 that

$$\hat{X}_{t+1} - \hat{X}^K_{t+1} = Q^*_{t+1} \cdot \frac{\int_{\mathbb{R}^n} \{z - [R^*_{t+1} + \Delta^{-1}]^{-1}[\bar{B}_{t+1} + \Delta^{-1}\mu]\} \exp[z'\bar{B}_{t+1} - \frac{1}{2}z'R^*_{t+1}z]dF(z)}{\int_{\mathbb{R}^n} \exp[z'\bar{B}_{t+1} - \frac{1}{2}z'R^*_{t+1}z]dF(z)} \quad (4.31)$$

for all $t = 0, 1, \ldots$, whence

$$\epsilon_{t+1} = E\left[\left\| Q^*_{t+1} \frac{\int_{\mathbb{R}^n} \{z - [R^*_{t+1} + \Delta^{-1}]^{-1}[\bar{B}_{t+1} + \Delta^{-1}\mu]\} \exp[z'\bar{B}_{t+1} - \frac{1}{2}z'R^*_{t+1}z]dF(z)}{\int_{\mathbb{R}^n} \exp[z'\bar{B}_{t+1} - \frac{1}{2}z'R^*_{t+1}z]dF(z)} \right\|^2\right]$$

$$= \bar{E}\left[\frac{\left\| Q^*_{t+1} \int_{\mathbb{R}^n} \{z - [R^*_{t+1} + \Delta^{-1}]^{-1}[\bar{B}_{t+1} + \Delta^{-1}\mu]\} \exp[z'\bar{B}_{t+1} - \frac{1}{2}z'R^*_{t+1}z]dF(z) \right\|^2}{\int_{\mathbb{R}^n} \exp[z'\bar{B}_{t+1} - \frac{1}{2}z'R^*_{t+1}z]dF(z)}\right]$$

upon using Lemma 2. We now obtain (4.30) by a simple application of Lemma 1 on this last relation. □

IV. REFERENCES

[1] G. DiMasi and W. Runggaldier, "On measure transformations for combined filtering and parameter estimation in discrete time," *Systems & Control Letters* **2**, pp. 57-62 (1982).

[2] R. Laha and V. Rohatgi, *Probability Theory*, John Wiley & Sons, New York (NY) (1979).

[3] M. Loéve, *Probability Theory II*, Fourth Edition, Springer-Verlag, New York (NY) (1978).

[4] A. M. Makowski, "Filtering formulae for partially observed linear systems with non-Gaussian initial conditions," *Stochastics* **16**, pp. 1-24 (1986).

[5] R. B. Sowers, *New Discrete-Time Filtering Results*, M.S. Thesis, Electrical Engineering Department, University of Maryland, College Park (MD), August 1988.

[6] R. B. Sowers and A.M. Makowski, "Filtering for discrete-time linear systems with non-Gaussian initial conditions: Asymptotic behavior of the difference between the MMSE and LMSE estimates," in preparation (1988).

[7] E. Wong, *Introduction to Random Processes*, Springer-Verlag, New York (NY) (1983).