SRC TR 89-39

TECHNICAL RESEARCH REPORT

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

By

R. B. Sowers
&
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

RICHARD B. SOWERS[1] and ARMAND M MAKOWSKI[2]

Department of Electrical Engineering
University of Maryland
College Park, MD 20742

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 LMSE (or Kalman) estimates of the state given past observations. These formulae are obtained with the help of the Girsanov transformation for Gaussian white noise sequences, and display explicitly the dependence of the quantities of interest on the initial distribution. Applications of these results can be found in [5] and [6].

I. INTRODUCTION

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

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

defined on some probability triple $(\Omega, \mathcal{F}, P)$ which carries the $I\!R^n$-valued plant process $\{X_t,\ t = 0,1,\ldots\}$ and the $I\!R^k$-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

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

(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

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

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

When the plant and observation noises are *uncorrelated*, and the observation noise sequence $\{V_t,\ t = 0,1,\ldots\}$ is standard (i.e., $\Sigma^w_{t+1} = 0$ 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. We shall discover that the structure of the solution of the prediction problem with correlated noise is essentially the same as that for uncorrelated noise. Indeed, the only difference is in the propagation of a collection finite-dimensional sufficient statistics; the mapping from these statistics to the filter is the same as in [4]. In Section IV, we present, without proof, representations for the MMSE and LLSE estimates of X°_{t+1} on the basis of $\{Y_0, Y_1 \ldots Y_t\}$ for $t = 0,1,\ldots$. These representations are derived by using Theorem 1.

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 $T\phi : I\!R^n \times I\!R^n \times \mathcal{Q}_{2n} \to \mathcal{C}$ and $U\phi : I\!R^n \times I\!R^n \times \mathcal{Q}_n \times \mathcal{M}_{n\times n} \times \mathcal{Q}_{2n} \to \mathcal{C}$ by

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

and

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

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

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

---

[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.

Presented at the 1989 Conference on Information Sciences and Systems, Johns Hopkins University, March 22-2

II. THE FILTERING PROBLEM

II.1. The main results

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

$$\begin{aligned} P_{t+1} &= A_tP_tA_t' + \Sigma_{t+1}^w \\ &\quad - [A_tP_tH_t' + \Sigma_{t+1}^{wv}][H_tP_tH_t' + \Sigma_{t+1}^v]^{-1}[A_tP_tH_t' + \Sigma_{t+1}^{wv}]' \\ P_0 &= O_n \end{aligned} \qquad t = 0,1,\ldots \quad (2.1)$$

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

$$J_t := H_tP_tH_t' + \Sigma_{t+1}^v. \qquad t = 0,1,\ldots \quad (2.2)$$

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

$$\begin{aligned} Q_{t+1} &= A_tQ_t - [A_tP_tH_t' + \Sigma_{t+1}^{wv}]J_t^{-1}H_t(Q_t + \Psi(t,0)) \\ &\quad + \Sigma_{t+1}^{wv}(\Sigma_{t+1}^v)^{-1}H_t\Psi(t,0) \\ Q_0 &= O_n \end{aligned} \qquad t = 0,1,\ldots \quad (2.3)$$

and

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

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

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

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

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

and

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

Finally, define an auxiliary deterministic Q_n-valued sequence $\{M_t,\ t = 0,1,\ldots\}$ by

$$\begin{aligned} M_{t+1} &= M_t + \Psi(t,0)'H_t'(\Sigma_{t+1}^v)^{-1}H_t\Psi(t,0) \\ M_0 &= 0. \end{aligned} \qquad t = 0,1,\ldots \quad (2.8)$$

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.,

$$\mathcal{Y}_t := \sigma\{Y_0, Y_1, \ldots, Y_t\}. \qquad t = 0,1,\ldots \quad (2.9)$$

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

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

$$E[\phi(X_{t+1}^o)|\mathcal{Y}_t] = \frac{U\phi[\hat{X}_{t+1}, \hat{B}_{t+1}; M_{t+1}, \Psi(t+1,0); \Sigma_{t+1}]}{U\mathbb{1}[\hat{X}_{t+1}, \hat{B}_{t+1}; M_{t+1}, \Psi(t+1,0); \Sigma_{t+1}]} \quad (2.10)$$

holds true P-a.s.

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 (2.10) 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 IR^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

$$E[\exp[i\theta'U_{t+1}]|\mathcal{F}_t] = \exp[-\frac{1}{2}\theta'\Lambda_{t+1}\theta] \qquad t = 0,1,\ldots \quad (2.11)$$

for every θ in IR^n. For any IR^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 IR^n and IR, respectively, by

$$\bar{U}_{t+1} := U_{t+1} - \Lambda_{t+1}\chi_t \qquad t = 0,1,\ldots \quad (2.12)$$

and

$$L_{t+1} := \prod_{s=0}^{t} \exp\left[\chi_s'U_{s+1} - \frac{1}{2}\chi_s'\Lambda_{s+1}\chi_s\right] \qquad t = 1,2,\ldots (2.13)$$

with $L_0 := 1$.

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

$$\bar{P}_{T+1}(A) := \int_A L_{T+1}dP, \quad A \text{ in } \mathcal{F}. \qquad (2.14)$$

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

$$\frac{d\bar{P}_{T+1}}{dP} = L_{T+1}; \qquad (2.15)$$

(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 (2.13) is simply

$$L_{t+1} := \prod_{s=0}^{t} \exp\left[\chi_s'\bar{U}_{s+1} + \frac{1}{2}\chi_s'\Lambda_{s+1}\chi_s\right]. \qquad t = 1,2,\ldots (2.16)$$

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 continu-ous-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 $\Sigma^{ww}_{t+1} = O_n$ and $\Sigma^{v}_{t+1} = I_n$ for $t = 0, 1, \ldots$, and fix a positive integer T. A careful 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*

$$X^\circ_t = X_t + Z_t \qquad t = 0,1,\ldots \quad (2.17)$$

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

$$\begin{aligned} X_{t+1} &= A_t X_t + W^\circ_{t+1} \\ X_0 &= 0 \end{aligned} \qquad t = 0,1,\ldots \quad (2.18)$$

and

$$\begin{aligned} Z_{t+1} &= A_t Z_t \\ Z_0 &= \xi, \end{aligned} \qquad t = 0,1,\ldots \quad (2.19)$$

in which case $Z_t = \Phi(t,0)\xi$ for $t = 0,1,\ldots$. However, for *any* decomposition of the form (2.17) we obtain

$$Y_t = H_t X_t + V_{t+1} \qquad t = 0,1,\ldots \quad (2.20)$$

where

$$V_{t+1} := V^\circ_{t+1} + H_t Z_t. \qquad t = 0,1,\ldots \quad (2.21)$$

If $\{(W^\circ_{t+1}, V_{t+1}),\ t = 0,1,\ldots,T\}$ were a GWN sequence under P, the prediction problem associated with (2.18)-(2.21) 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

$$\frac{d\bar{P}}{dP} := \exp\left[-\sum_{s=0}^{T}[H_s Z_s]' V^\circ_{s+1} - \frac{1}{2}\sum_{s=0}^{T}[H_s Z_s]'[H_s Z_s]\right]. \quad (2.22)$$

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

$$L_{t+1} := \exp\left[-\xi' \sum_{s=0}^{t}[H_s\Phi(s,0)]' V^\circ_{s+1} - \frac{1}{2}\xi' \sum_{s=0}^{t}[H_s\Phi(s,0)]'[H_s\Phi(s,0)]\xi\right] \qquad t = 0,1,\ldots \quad (2.23)$$

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]

$$E[\phi(X^\circ_{t+1})|\mathcal{Y}_t] = \frac{\bar{E}[\phi(X^\circ_{t+1})L^{-1}_{T+1}|\mathcal{Y}_t]}{\bar{E}[L^{-1}_{T+1}|\mathcal{Y}_t]} \quad P-a.s. \quad (2.24)$$

which holds for each bounded Borel mapping $\phi : IR^n \to \mathcal{C}$ 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

$$\begin{aligned} &\bar{E}\left[\phi(X^\circ_{t+1})\frac{d\bar{P}}{dP}\middle|\mathcal{Y}_t \vee \sigma\{\xi\}\right] \\ &\quad = \bar{E}\left[\bar{E}[\phi(X^\circ_{t+1})L^{-1}_{T+1}|\mathcal{F}_{t+1}]\middle|\mathcal{Y}_t \vee \sigma\{\xi\}\right] \\ &\quad = \bar{E}[\phi(X^\circ_{t+1})L^{-1}_{t+1}|\mathcal{Y}_t \vee \sigma\{\xi\}] \end{aligned} \quad (2.25)$$

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 IR^n-valued RV's $\{B_t,\ t = 0,1,\ldots\}$ and the Q_n-valued sequence $\{M_t,\ t = 0,1,\ldots\}$ by setting

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

and

$$M_{t+1} := \sum_{s=0}^{t}\Phi(s,0)' H_s' H_s \Phi(s,0) \qquad t = 0,1,\ldots \quad (2.27)$$

with $B_0 = 0$ and $M_0 = O_n$. From (2.21), (2.26) and (2.27), we observe that

$$L^{-1}_{t+1} = \exp\left[\xi' B_{t+1} - \frac{1}{2}\xi' M_{t+1}\xi\right] \qquad t = 0,1,\ldots \quad (2.28)$$

and readily conclude from (2.25) that

$$\begin{aligned} \bar{E}\left[\phi(X^\circ_{t+1})\frac{dP}{d\bar{P}}\middle|\mathcal{Y}_t \vee \sigma\{\xi\}\right] &= \exp\left[-\frac{1}{2}\xi' M_{t+1}\xi\right] \cdot \\ \bar{E}\left[\phi(X_{t+1} + \Phi(t+1,0)\xi)\exp[\xi' B_{t+1}]\middle|\mathcal{Y}_t \vee \sigma\{\xi\}\right]&. \end{aligned} \qquad t = 0,1,\ldots \quad (2.29)$$

By property **(B.2)**, we see from (2.18)-(2.20) 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

$$\hat{X}_{t+1} := \bar{E}[X_{t+1}|\mathcal{Y}_t] \quad \text{and} \quad \hat{B}_{t+1} := \bar{E}[B_{t+1}|\mathcal{Y}_t], \qquad t = 0,1,\ldots \quad (2.30)$$

with corresponding errors

$$\tilde{X}_{t+1} := X_{t+1} - \hat{X}_{t+1} \quad \text{and} \quad \tilde{B}_{t+1} := B_{t+1} - \hat{B}_{t+1}. \qquad t = 0,1,\ldots \quad (2.31)$$

As in [4], standard arguments [7, Sec. 2.7] imply that the RV's $\{\tilde{X}_{t+1}, \tilde{B}_{t+1}\}$ 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 IR^{2n}-valued RV $(\tilde{X}_{t+1}, \tilde{B}_{t+1})$ is a zero-mean *Gaussian* RV with covariance matrix Σ_{t+1} given by

$$\Sigma_{t+1} := \bar{E}\left[\begin{pmatrix}\tilde{X}_{t+1}\\ \tilde{B}_{t+1}\end{pmatrix}\begin{pmatrix}\tilde{X}_{t+1}\\ \tilde{B}_{t+1}\end{pmatrix}'\right] = \begin{pmatrix}P_{t+1} & Q_{t+1}\\ Q'_{t+1} & R_{t+1}\end{pmatrix}. \qquad t = 0,1,\ldots \quad (2.32)$$

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

$$P_{t+1} = \hat{E}\big[\tilde{X}_{t+1}\tilde{X}'_{t+1}\big], \quad Q_{t+1} = \hat{E}\big[\tilde{X}_{t+1}\tilde{B}'_{t+1}\big], \quad \text{and} \quad R_{t+1} = \hat{E}\big[\tilde{B}_{t+1}\tilde{B}'_{t+1}\big]. \tag{2.33}$$

The RV's $\tilde{X}_{t+1} + \Phi(t+1,0)\xi$, $\tilde{B}_{t+1}$ 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 (2.29) that

$$\hat{E}\Big[\phi(X^\circ_{t+1})\frac{dP}{d\hat{P}}\Big|\mathcal{Y}_t \vee \sigma\{\xi\}\Big] = \exp\Big[-\frac{1}{2}\xi' M_{t+1}\xi + \xi'\tilde{B}_{t+1}\Big]\cdot \mathcal{T}\phi\big[\tilde{X}_{t+1} + \Phi(t+1,0)\xi, \xi; \Sigma_{t+1}\big] \qquad t = 0,1,\ldots \tag{2.34}$$

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

From (2.34), we now readily obtain by the law of iterated conditioning that

$$\begin{aligned}\hat{E}\Big[\phi(X^\circ_{t+1})\frac{dP}{d\hat{P}}\Big|\mathcal{Y}_t\Big] &= \hat{E}\Big[\exp\Big[-\frac{1}{2}\xi' M_{t+1}\xi + \xi'\tilde{B}_{t+1}\Big]\cdot \mathcal{T}\phi\big[\tilde{X}_{t+1} + \Phi(t+1,0)\xi, \xi; \Sigma_{t+1}\big]\Big|\mathcal{Y}_t\Big] \\ &= \mathcal{U}\phi\big[\tilde{X}_{t+1}, \tilde{B}_{t+1}; M_{t+1}, \Phi(t+1,0); \Sigma_{t+1}\big] \qquad t = 0,1,\ldots\end{aligned} \tag{2.35}$$

where the mapping $\mathcal{U}\phi$ is defined by (1.5). We have used the fact that the RV's $\{\tilde{X}_{t+1}\tilde{B}_{t+1}\}$ are $\mathcal{Y}_t$-measurable and therefore $\hat{P}$-independent of $\sigma\{\xi\}$. The reader will readily check that substitution of (2.35) (with arbitrary ϕ and with $\phi = \mathbb{1}$) results in (2.10) 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 (2.17), and define $\{V_{t+1},\ t = 0,1,\ldots\}$ by (2.21). If $\Sigma^{wv}_{t+1} = O_n$ for $t = 0,1,\ldots$, we would arrive at the probability measure $\hat{P}$ characterized by property **(B.2)** as follows: Define the filtration $\{\mathcal{F}_t,\ t = 0,1,\ldots\}$ by

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

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 $\hat{P}$ as defined in (2.22) 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 $\hat{P}$ given by (2.22) 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* IR^{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

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

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

$$\begin{pmatrix} W_{t+1} \\ V_{t+1}\end{pmatrix} = \begin{pmatrix} W^\circ_{t+1} \\ V^\circ_{t+1}\end{pmatrix} - \begin{pmatrix} \Sigma^{w}_{t+1} & \Sigma^{wv}_{t+1} \\ \Sigma^{vw}_{t+1} & \Sigma^{v}_{t+1}\end{pmatrix}\begin{pmatrix}\varphi^w_t \\ \varphi^v_t\end{pmatrix} \qquad t = 0,1,\ldots \tag{3.3}$$

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 IR^n and IR^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 $\hat{P}$ on $(\Omega, \mathcal{F})$ satisfying (B.3) where

(B.3): *The probability measure $\hat{P}$ 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 $(\mathcal{F}_t, \hat{P})$ GWN sequence with the same covariance structure under $\hat{P}$ 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 (2.21), 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

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

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

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

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

$$\frac{d\hat{P}}{dP} = \exp\Big[\sum_{s=0}^{T}\big[\varphi'_s[W^\circ_{s+1} - \Sigma^{wv}_{s+1}(\Sigma^v_{s+1})^{-1}V^\circ_{s+1}] - Z'_s H'_s(\Sigma^v_{s+1})^{-1}V^\circ_{s+1}\big] + \frac{1}{2}\sum_{s=0}^{T}\big[\varphi'_s[\Sigma^w_{s+1} - \Sigma^{wv}_{s+1}(\Sigma^v_{s+1})^{-1}\Sigma^{vw}_{s+1}]\varphi_s + Z'_s H'_s(\Sigma^v_{s+1})^{-1}H_s Z_s\big]\Big]. \tag{3.6}$$

In order to complete the specification of the decomposition (2.17) 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 $\hat{P}$, 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

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

This suggests a separation of the dynamics in the form

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

and

$$Z_{t+1} = [A_t - \Sigma^{wv}_{t+1}(\Sigma^{v}_{t+1})^{-1}H_t]Z_t + \pi_t \qquad t = 0,1,\dots \quad (3.9)$$
$$Z_0 = \xi - \zeta$$

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

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

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

- The effect of the initial condition

$$Z_{t+1} = [A_t - \Sigma^{wv}_{t+1}(\Sigma^{v}_{t+1})^{-1}H_t]Z_t \qquad t = 0,1,\dots \quad (3.11)$$
$$Z_0 = \xi,$$

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

- The noise processes

$$\begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix} = \begin{pmatrix} W^o_{t+1} \\ V^o_{t+1} \end{pmatrix} - \begin{pmatrix} \Sigma^{w}_{t+1} & \Sigma^{wv}_{t+1} \\ \Sigma^{vw}_{t+1} & \Sigma^{v}_{t+1} \end{pmatrix} \begin{pmatrix} 0 \\ -(\Sigma^{v}_{t+1})^{-1}H_tZ_t \end{pmatrix}$$
$$= \begin{pmatrix} W^o_{t+1} + \Sigma^{vw}_{t+1}(\Sigma^{v}_{t+1})^{-1}H_tZ_t \\ V^o_{t+1} + H_tZ_t \end{pmatrix}. \qquad t = 0,1,\dots \quad (3.12)$$

- The auxiliary system

$$X_{t+1} = A_tX_t + W_{t+1}$$
$$X_0 = 0 \qquad t = 0,1,\dots \quad (3.13)$$
$$Y_t = H_tX_t + V_{t+1}.$$

- The change of measure

$$\frac{d\bar{P}}{dP} = \exp\Big[-\sum_{s=0}^{T} Z_s'H_s'(\Sigma^{v}_{s+1})^{-1}V^o_{s+1} + \frac{1}{2}\sum_{s=0}^{T} Z_s'H_s'(\Sigma^{v}_{s+1})^{-1}H_sZ_s\Big]. \quad (3.14)$$

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

Proposition 1. *Let the filtration $\{\mathcal{F}_t,\ t = 0,1,\dots\}$ be given by (3.2). If the sequences $\{X_t,\ t = 0,1,\dots\}$, $\{Z_t,\ t = 0,1,\dots\}$ and $\{(W_{t+1}, V_{t+1}),\ t = 0,1,\dots\}$ 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 agree on $\mathcal{F}_0$, and the process $\{(W_{t+1}, V_{t+1}),\ t = 0,1,\dots,T\}$ is a zero-mean $(\mathcal{F}_t, P)$ GWN sequence with covariance structure structure $\{\Gamma_{t+1},\ t = 0,1,\dots,T\}$ under $\bar{P}$.*

Motivated by the form of (3.14), we define the IR-valued sequence $\{L_t,\ t = 0,1,\dots\}$ by

$$L_{t+1} := \exp\Big[-\sum_{s=0}^{t} Z_s'H_s'(\Sigma^{v}_{s+1})^{-1}V^o_{s+1} + \frac{1}{2}\sum_{s=0}^{t} Z_s'H_s'(\Sigma^{v}_{s+1})^{-1}H_sZ_s\Big] \qquad t = 0,1,\dots (3.15)$$

with $L_0 = 1$, and observe that $d\bar{P}/dP = L_{T+1}$. The Girsanov transformation now implies that $\{L_t^{-1},\ t = 0,1,\dots,T+1\}$ is an $(\mathcal{F}_t, \bar{P})$-martingale, and by the same arguments as the ones leading to (2.25) we conclude that

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

Since

$$L^{-1}_{t+1} = \exp\Big[\sum_{s=0}^{t} Z_s'H_s'(\Sigma^{v}_{s+1})^{-1}V^o_{s+1} - \frac{1}{2}\sum_{s=0}^{t} Z_s'H_s'(\Sigma^{v}_{s+1})^{-1}H_sZ_s\Big], \qquad t = 0,1,\dots (3.17)$$

we see from (2.21) that

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

where

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

and

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

with $B_0 = 0$ and $M_0 = O_n$. We may verify that the definition of $\{M_t, t = 0,1,\dots\}$ of (3.20) is equivalent to that of (2.8).

As before, we define the sequences of conditional means $\{\hat{X}_{t+1},\ t = 0,1,\dots\}$ and $\{\hat{B}_{t+1},\ t = 0,1,\dots\}$ by (2.30), with corresponding errors $\{\tilde{X}_{t+1},\ t = 0,1,\dots\}$ and $\{\tilde{B}_{t+1},\ t = 0,1,\dots\}$ given by (2.31). The RV's X_{t+1}, B_{t+1}, and $\{Y_0,\dots,Y_t\}$ may all be represented as linear combinations of $\{(W_{t+1}, V_{t+1}),\ t = 0,1,\dots,T\}$, and are thus jointly Gaussian and independent of $\sigma\{\xi\}$ under $\bar{P}$. As argued in the uncorrelated case, under $\bar{P}$, the IR^{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

$$\begin{aligned}
&\bar{E}\big[\phi(X^o_{t+1})\exp[\xi'B_{t+1} - \tfrac{1}{2}\xi'M_{t+1}\xi]\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big] \\
&= \bar{E}\big[\phi(\tilde{X}_{t+1} + \hat{X}_{t+1} + \Psi(t+1,0)\xi)\cdot \\
&\qquad \exp[\xi'\hat{B}_{t+1} + \xi'\tilde{B}_{t+1} - \tfrac{1}{2}\xi'M_{t+1}\xi]\big|\mathcal{Y}_t \vee \sigma\{\xi\}\big] \\
&= \exp[-\tfrac{1}{2}\xi'M_{t+1}\xi]\bar{E}\big[\phi(\tilde{X}_{t+1} + x)\exp[b'\tilde{B}_{t+1}]\big]_{\substack{x=\hat{X}_{t+1}+\Psi(t+1,0)\xi \\ b=B_{t+1}}} \\
&= \exp[-\tfrac{1}{2}\xi'M_{t+1}\xi]\mathcal{T}\phi[\hat{X}_{t+1} + \Psi(t+1,0)\xi, \xi; \Sigma_{t+1}] \qquad (3.21)
\end{aligned}$$

where Σ_{t+1} has the decomposition (2.32)-(2.33). Removing the conditioning upon $\sigma\{\xi\}$, we find

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

since $(\hat{X}_{t+1}, \hat{B}_{t+1})$ 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 $t = 0,1,\dots,T$. Indeed we readily obtain (2.10) by

injecting (3.22) (for arbitrary ϕ and for $\phi = \mathbb{I}$) into (2.24). 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

$$\begin{pmatrix} X_{t+1} \\ B_{t+1} \end{pmatrix} = \begin{pmatrix} A_t & 0 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} X_t \\ B_t \end{pmatrix} + \begin{pmatrix} I_n & 0 \\ 0 & \Psi'(t,0)H_t'(\Sigma_{t+1}^v)^{-1} \end{pmatrix} \begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix}$$

$$\begin{pmatrix} X_0 \\ B_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$Y_t = \begin{pmatrix} H_t & 0 \end{pmatrix} \begin{pmatrix} X_t \\ B_t \end{pmatrix} + \begin{pmatrix} 0 & I_k \end{pmatrix} \begin{pmatrix} W_{t+1} \\ V_{t+1} \end{pmatrix}. \qquad t = 0,1,\ldots \quad (3.23)$$

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 IR^{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 $T\phi$ and $U\phi$ 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 THE MMSE AND LLSE FILTERS

Using Theorem 1, we may develop formulae for the MMSE and LLSE estimates of X_{t+1}° on the basis of $\{Y_0, Y_1, \ldots, Y_t\}$ for any $t = 0,1,\ldots$. In the interest of brevity, we shall omit the proofs. The reader is referred to [5] for a full exposition.

In addition to assumptions (A.1) to (A.3), we shall need an additional assumption (A.4), where

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

We shall also find it convenient to introduce the auxiliary quantities $\{Q_t^\circ,\ t = 0,1,\ldots\}$ and $\{R_t^\circ,\ t = 0,1,\ldots\}$ in $\mathcal{M}_{n\times n}$ and $\mathcal{Q}_n$, respectively, given by

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

Algebraic manipulation then reveals that these quantities propagate according to

$$\begin{aligned} Q_{t+1}^\circ &= \left[A_t - [A_t P_t H_t' + \Sigma_{t+1}^{wv}]J_t^{-1}H_t\right]Q_t^\circ \\ Q_0^\circ &= I_n, \end{aligned} \qquad t = 0,1,\ldots \quad (4.2)$$

and

$$\begin{aligned} R_{t+1}^\circ &= R_t^\circ + Q_t^{\circ\prime}H_t'J_t^{-1}H_tQ_t^\circ \\ R_0^\circ &= O_n. \end{aligned} \qquad t = 0,1,\ldots \quad (4.3)$$

The dynamics (2.7) then also simplifies into

$$\begin{aligned} \bar{B}_{t+1} &= \bar{B}_t - Q_t^{\circ\prime}H_t'J_t^{-1}H_t\bar{X}_t + Q_t^{\circ\prime}H_t'J_t^{-1}Y_t \\ \bar{B}_0 &= 0. \end{aligned} \qquad t = 0,1,\ldots \quad (4.4)$$

We then have:

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

$$E[X_{t+1}^\circ|\mathcal{Y}_t] = \bar{X}_{t+1} + Q_{t+1}^\circ \frac{\int_{IR^n} z \exp\left[z'\bar{B}_{t+1} - \frac{1}{2}z'R_{t+1}^\circ z\right]dF(z)}{\int_{IR^n} \exp\left[z'\bar{B}_{t+1} - \frac{1}{2}z'R_{t+1}^\circ z\right]dF(z)} \quad (4.5)$$

P-a.s., and if we denote by $\hat{X}_{t+1}^K$ the LLSE estimate of X_{t+1}° on the basis of $\{Y_0, Y_1, \ldots, Y_t\}$, then

$$\hat{X}_{t+1}^K = \bar{X}_{t+1} + Q_{t+1}^\circ\left[R_{t+1}^\circ + \Delta^{-1}\right]^{-1}\left[\bar{B}_{t+1} + \Delta^{-1}\mu\right] \quad (4.6)$$

P-a.s. for all $t = 0,1,\ldots$.

Outline of Proof. Apply Theorem 1 to find the conditional characteristic function; i.e., let $\phi_\theta(z) := \exp[i\theta'z]$. We then differentiate the resulting expression with respect to θ to arrive at (4.5). Since the LLSE or Kalman filter depends solely on the first and second moments, and since the MMSE and LLSE filters coincide if ξ has a Gaussian distribution, we arrive at (4.6) by replacing F in (4.5) by a Gaussian distribution with mean μ and covariance Δ. This representation for the LLSE filter 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 $x \mapsto [R_{t+1}^\circ + \Delta^{-1}]^{-1}[x + \Delta^{-1}\mu]$.

V. 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 (1989).

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