Latin American Applied Research 35:121-127 (2005)

SAMPLED-DATA MINIMUM VARIANCE FILTERING

L. D. HERNANDEZ and R. H. MILOCCO

Grupo de Control Automático y Sistemas. Facultad de Ingeniería, U.N.Comahue
Buenos Aires 1400, 8300 Neuquén, Argentina
henaleo@hotmail.com, milocco@uncoma.edu.ar

Abstract— **This paper deals with the optimal solution to the sampled-data minimum variance filtering problem for linear systems with noise in the states and in the measurements. The solution is derived in the time-domain by using a fast sampling zero-order hold input discretization of the continuous time systems together with a lifting technique. The original sampled-data system is transformed into an equivalent LTI discrete-time system with infinite-dimensional input-output space. However, the designed filter is finite-dimensional. We derive both the existence conditions and the explicit expression of the desired filter and provide an illustrative numerical example.**

Keywords— **Filtering, lifting, T-periodic systems, sampled-data.**

I. INTRODUCTION

Minimum variance filtering problems have been extensively studied via both the state space (Kalman 1960; Anderson and Moore, 1975; Shaked 1976) and the polynomial system approach (Wiener 1950; Roberts and Newmann, 1987; Ahlén and Sternad, 1991; and Grimble 1995). The design techniques are based on continuous-time (ct) or discrete-time (dt) system descriptions. However, in most applications a more realistic situation is that in which a digital filter must be designed to interact with ct systems. In such cases, the estimated signal is formed by the output of the dt filter through a zero-order hold device. The goal is to match the piecewise estimations to the desired ct signal. In general, there are two classic approaches used to design the corresponding dt filter. The first approach consists of discretizing the ct system and designing a dt filter. In the second one, a digital implementation of the -optimal- filter obtained by a ct design is performed. Due to the intersample behaviour of the ct systems in both approaches, there is a serious performance degradation when the sampling is not fast enough. There is a third approach, called sampled-data design, in which the dt filter is designed taking into account the dynamics of the ct systems involved. The recent trends, such as techniques based on linear systems with jumps (Sun *et al.* 1993), and the lifting technique (Bamieh *et al.*, 1991 and Hara *et al.*, 1997), have been used to direct sampled-data design. Although these designs have been extensively used in feedback control systems, filters design received too little attention in spite of its importance in signal processing applications. Filtering sampled-data design has been investigated in the context of H_∞ in Sun *et al.* (1993) and Kabamba *et al.* (1993), and in the H_2 context in Milocco and De Doná (1996), Wang *et al.* (2001), and Milocco and Muravchik (2003). In Wang *et al.*(2001), a filtering sampled-data design based on the Error Covariance Assignment criterium is proposed, while in Milocco and De Doná (1996) and Milocco and Muravchik (2003), a frequency domain approach to MIMO linear filter design for sampled-data system is presented by using a polynomial approach. In this paper, we extend the results obtained in Milocco and De Doná (1996) to design MIMO sampled-data filters in the time-domain. The proposed solution allows us to obtain the sampled-data minimum variance estimation of the states as well as optimal solutions for minimum variance sampled-data filtering problems such as deconvolution, prediction and smoothing.

The paper is organized as follows: In section II, we provide a suitable description of the multivariable sampled-data system, *i.e.* the ct subsystem followed by a sampling stage at intervals of T sec., the dt subsystem or filter, and a holding device. Such systems can be represented with the help of T-periodic linear time-invariant systems. By means of a fast sampling zero-order hold input discretization of the ct systems together with a lifting technique, the original sampled-data system is transformed into an equivalent LTI dt system with infinite-dimensional input-output space. The cost to be minimized is defined as the averaged energy of the weighted output-error vector. In section III, the matrices of the dt filter state space representation are obtained such that the cost is minimized. In section IV, an example to illustrate the procedure is provided and finally, in section V, we present the conclusions.

II. PROBLEM FORMULATION

Consider the following ct time-invariant generalized system:

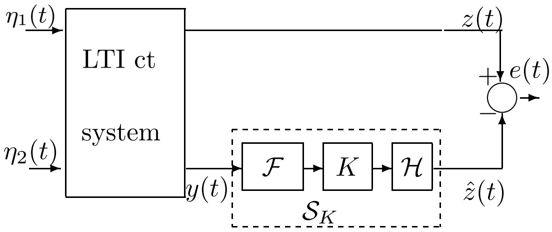

Figure 1: Filtering problem design.

$$\begin{aligned} S_c: \ \dot{x}(t) &= Ax(t) + B\eta_1(t), \\ y(t) &= C_1 x(t) + D\eta_2(t), \\ z(t) &= C_2 x(t), \end{aligned} \qquad (1)$$

A, B, C_1, C_2, and D are known constant matrices, and $\eta_1(t)$ and $\eta_2(t)$ are ct white noise input vectors -possibly correlated- affecting the states $(x(t))$ and the measured output $(y(t))$ respectively; $z(t)$ is the desired output. By sampling the signal $y(t)$, our aim is to design a dt filter to produce, through a zero-order hold, a piecewise-continuous signal $\hat{z}(t)$ to match the desired signal $z(t)$ in the sense of minimum variance. The flowgraph of the sampled-data filtering problem is depicted in Fig. 1. The signal $\hat{z}(t)$ is the zero-order hold output of the dt time-invariant filter $\mathcal{K}$ driven by samples $y_s[kT]$ of the signal $y(t)$ -with sampling period T. We call $\mathcal{F}$ an ideal sampler (with period T) and $\mathcal{H}$, a zero-order hold (with period T). The sampling operator maps the vector space of piecewise-continuous functions $\mathcal{W}$ to the space of the sequences defined on the set of integers, negative and positive $\mathcal{V}$, and is defined as

$$y = \mathcal{F}u \Leftrightarrow y[kT] = u[kT]. \qquad (2)$$

The hold operator maps $\mathcal{V}$ to $\mathcal{W}$ via

$$y = \mathcal{H}u \Leftrightarrow y(t) = u[kT] \quad (kT \le t < (k+1)T). \qquad (3)$$

$\mathcal{H}$ and $\mathcal{F}$ are synchronized and provide the interface between the digital and the analog parts of the system. We call $\mathcal{S}_{\mathcal{K}} = \mathcal{H}\mathcal{K}\mathcal{F}$ the sampled-data filter. Note that $\mathcal{S}_{\mathcal{K}}$ is a ct linear time-varying T-periodic operator that turns the complete filtering setup of Fig. 1 also into a linear time-varying T-periodic operator that we call $\mathcal{S}$. A T-periodic system is one whose response to an input delayed by exactly T-sec. is obtained by delaying the original output exactly by T-sec. If $\varphi(t,\tau)$ denotes the response of a time varying T-periodic linear system at time t when a vector of Dirac delta functions is applied to the input at time τ, then this impulse response satisfies $\varphi(t+T,\tau) = \varphi(t,\tau-T)$.

The time-varying T-periodic variance of the error $(e(t))$ between the desired $(z(t))$ and estimated $(\hat{z}(t))$ signals $(e(t) = z(t) - \hat{z}(t))$, when the input of the T-periodic operator $\mathcal{S}$ is driven by the input vector $\eta = [\eta_1^T \ \ \eta_2^T]^T$, is given by $\mathcal{E}\{e(t)e^T(t)\}$, where $\mathcal{E}$ means expectation. We define the cost function to be minimized with the sampled-data filter $\mathcal{S}_{\mathcal{K}}$ as the error variance averaged over the interval T as follows:

$$J(K) = \frac{1}{T}\int_0^T \mathcal{E}tr\{e(t)e(t)^T\}dt. \qquad (4)$$

Let $\varphi(t,\tau)$ denote the impulse response of $\mathcal{S}$. Thus, the error is given by

$$e(t) = \int_{-\infty}^{t} \varphi(t,\tau)\eta(\tau)d\tau \qquad (5)$$

and its covariance matrix becomes

$$\mathcal{E}\{e(t)e^T(t)\} = \\ = \int_{-\infty}^{t}\int_{-\infty}^{t} \varphi(t,\tau_1)\mathcal{E}\{\eta(\tau_1)\eta^T(\tau_2)\}\varphi^T(t,\tau_2)d\tau_1 d\tau_2. \qquad (6)$$

Since $\mathcal{E}\{\eta(\tau_1)\eta^T(\tau_2)\} = I\delta(\tau_1-\tau_2)$ the cost (4) can be written as

$$J(K) = \frac{1}{T}\int_0^T \mathrm{tr}\left\{\int_{-\infty}^{t}\varphi(t,\tau)\varphi^T(t,\tau)d\tau\right\}dt. \qquad (7)$$

Assuming the T-periodic operator $\mathcal{S}$ is band-limited or equivalently, that the power spectral density of the error decreases to zero as the frequency increases towards infinity, we show that the cost function can be approximated by a new cost associated to the zero-order hold discretized systems with small sampling period T_s. To this end, we follow the same line as in Hara *et al.* (1997). First, consider the small sampling period $T_s = T/\mathcal{N}$, where N is an integer. Then, the cost in (7) can be approximated by

$$J(K) \approx \frac{1}{T}\sum_{k=0}^{\mathcal{N}-1}\sum_{l=-\infty}^{k} \mathrm{tr}\left\{\varphi(kT_s,lT_s)\varphi^T(kT_s,lT_s)\right\}T_s^2. \qquad (8)$$

Let $\varphi_s(k,l)$ denote the impulse response of the fast sampling zero-order hold input discretization of $\mathcal{S}$. It is related to $\varphi(t,\tau)$ by

$$\varphi_s(k,l) = \int_{(l-1)T_s}^{lT_s} \varphi(kT_s,v)dv. \qquad (9)$$

In the limit, as T is fixed and $\mathcal{N}\to\infty$, the following approximation is valid:

$$\varphi_s(k,l) = T_s\varphi(kT_s,lT_s). \qquad (10)$$

By using $\varphi_s(k,l)$ in (8), we obtain

$$J(K) \approx \frac{1}{T}\sum_{k=0}^{\mathcal{N}-1}\sum_{l=-\infty}^{k} \mathrm{tr}\left\{\varphi_s(k,l)\varphi_s^T(k,l)\right\}. \qquad (11)$$

For a given sampling time T, equality holds for $\mathcal{N}\to\infty$.

In order to have a working expression of the approximated cost (11), we are interested in transforming the fast zero-order hold input discretization of the T-periodic system into an equivalent LTI dt system. To this end, note that the following equality holds for the impulse response of the fast sampled approximation $\varphi_s(k,l)$:

$$\varphi_s(k\mathcal{N}+i, l\mathcal{N}+j) = \varphi_s((k-l)\mathcal{N}+i, j); \quad \forall\, i,j = 0,1,\cdots,\mathcal{N}-1 \text{ and } l \leq k.$$

Then, the impulse response $\varphi_s(k,l)$ of the time-varying system can be grouped into blocks forming the impulse response of an augmented dimension time invariant-system. The impulse response $\tilde{\varphi}(h)$ for $h = k-l$ of such time-invariant system is given by

$$\tilde{\varphi}(h) = \begin{bmatrix} \varphi_s(h\mathcal{N},0) & \cdots & \varphi_s(h\mathcal{N},\mathcal{N}-1) \\ \vdots & \ddots & \vdots \\ \varphi_s(h\mathcal{N}+\mathcal{N}-1,0) & \cdots & \varphi_s(h\mathcal{N}+\mathcal{N}-1,\mathcal{N}-1) \end{bmatrix},$$

where each element $\tilde{\varphi}_{r,s}(h)$ of $\tilde{\varphi}(h)$ represents the *r-th* output at time kT, due to an impulse at time $hT=0$ in the *s-th* input component of the $\mathcal{N}\times\mathcal{N}$ multivariable operator $\tilde{S}$. Notice that $\tilde{S}$ is the lifted construction of the fast sampled dt hold input approximation of the periodical system $\mathcal{S}$. By using the impulse response of the lifted system, the following relationship is obtained:

$$\lim_{\mathcal{N}\to\infty} \sum_{k=0}^{\mathcal{N}-1} \sum_{l=-\infty}^{k} \operatorname{tr}\left\{\varphi_s(k,l)\varphi_s^T(k,l)\right\} = \sum_{h=0}^{\infty} \operatorname{tr}\left\{\tilde{\varphi}(h)\tilde{\varphi}^T(h)\right\} = J_N(K).$$

From the equality above, it is clear that the following holds:

$$J(K) = \lim_{\mathcal{N}\to\infty} \frac{1}{T} J_N(K) \tag{12}$$

Note that an equivalent interpretation of $J_N(K)$ is given by

$$J_N(K) = \mathcal{E}\operatorname{tr}\left\{\tilde{e}(h)\tilde{e}^T(h)\right\}, \tag{13}$$

where $\tilde{e}(h)$ is the output of the fast sampled zero-order hold and lifted system $\tilde{S}$ driven by a white noise sequence with identity covariance matrix, and $J_N(K)$ is the sum of the variance of each of the elements in $\tilde{e}(h)$. Having obtained a working expression of the cost we minimize $J_N(K)$ instead of $J(K)$. Thus, in order to obtain $\tilde{e}(h)$, we need the lifting version of the fast sampled dt zero-order hold input approximation of the periodical system $\mathcal{S}$.

Assume the zero-order hold input discretization with sampling time T_s of the ct system (1), together with the low-pass antialiasing filter, is given by

$$S_d: \begin{aligned} x[(k+1)T_s] &= \Phi x[kT_s] + \Gamma\eta[kT_s], \\ y[kT_s] &= L_1 x[kT_s], \\ z[kT_s] &= L_2 x[kT_s], \end{aligned} \tag{14}$$

where $\eta[k]$ is a white noise sequence vector with covariance matrix $\mathcal{E}\{\eta[k]\eta^T[k]\} = I$ and the constant matrices Φ, Γ, L_1, and L_2 are known. Our purpose is to lift the system from sampling time T_s to T. This lifting operation can be formulated by means of a block implementation as used, for instance, in Keller and Anderson (1992). The transformed signal $\tilde{\eta}(k\mathcal{N}T_s)$ is defined by

$$\begin{aligned} \tilde{\eta}[k\mathcal{N}T_s)] = {} & [\eta^T[k\mathcal{N}T_s] \;\; \eta^T[(k\mathcal{N}+1)T_s]\cdots \\ & \cdots\cdots\cdots\, \eta^T[(k\mathcal{N}+\mathcal{N}-1)T_s]]^T. \end{aligned}$$

Note that the new vector signal $\tilde{\eta}(kT)$ corresponds to a compound signal with sampling period T. The fast system S_d can be lifted to the system $\tilde{S_d}$, which maps input blocks $\tilde{\eta}(kT)$ to output blocks $\tilde{y}(kT)$ and $\tilde{z}(kT)$. This lifted system $\tilde{S_d}$ is associated with the sampling time T and is given by

$$\tilde{S}_d: \begin{aligned} x[(k+1)\mathcal{N}]T_s] &= \tilde{\Phi}x[k\mathcal{N}T_s)] + \tilde{\Gamma}\tilde{\eta}[k\mathcal{N}T_s] \\ \tilde{y}[k\mathcal{N}T_s] &= \tilde{C}_1 x[k\mathcal{N}T_s] + \tilde{D}_1\tilde{\eta}[k\mathcal{N}T_s] \\ \tilde{z}[k\mathcal{N}T_s] &= \tilde{C}_2 x[k\mathcal{N}T_s] + \tilde{D}_2\tilde{\eta}[k\mathcal{N}T_s] \end{aligned} \tag{15}$$

The structure of the matrices is as follows:

$$\tilde{\Phi} = \Phi^{\mathcal{N}}; \quad \tilde{\Gamma} = [\Phi^{\mathcal{N}-1}\Gamma \quad \Phi^{\mathcal{N}-2}\Gamma \quad \ldots \quad \Gamma];$$

$$\tilde{C}_i = \begin{pmatrix} L_i \\ L_i\Phi \\ \cdot \\ \cdot \\ \cdot \\ L_i\Phi^{\mathcal{N}-1} \end{pmatrix};$$

$$\tilde{D}_i = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 \\ L_i\Gamma & 0 & 0 & \ldots & 0 \\ L_i\Phi\Gamma & L_i\Gamma & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ L_i\Phi^{\mathcal{N}-2}\Gamma & L_i\Phi^{\mathcal{N}-3}\Gamma & \ldots & L_i\Gamma & 0 \end{pmatrix}. \tag{16}$$

Considering the dt lifted version of the system, the sampled output is given by $y_s[kT)] = \mathcal{F}\tilde{y}[kT)]$, which corresponds to the first n elements of the vectorized signal $\tilde{y}[k\mathcal{N}T_s]$, where n is the number of ct measured signals. Then, the sampler $\mathcal{F}$ is represented by the $n\times n\mathcal{N}$ matrix $\mathcal{F} = [I_n\ 0 \cdots 0]$. The discretized hold element $\mathcal{H}$ produces a sequence of $\mathcal{N}$ pulses equal to each filter output. Thus, it can be represented by a matrix of dimension $p\mathcal{N}\times p$ given by $\mathcal{H} = [I_p \cdots I_p]^T$, where p is the number of the filter outputs. Finally, taking into account that $\mathcal{F}\tilde{D}_1 = 0$, the problem design can be stated as follows: Given the fast sampled and lifted system

$$\begin{aligned} x[(k+1)T] &= \tilde{\Phi}x[kT] + \tilde{\Gamma}\tilde{\eta}[kT], \\ y_s[kT] &= L_1 x[kT], \\ \tilde{z}[kT] &= \tilde{C}_2 x[kT] + \tilde{D}_2\tilde{\eta}[kT], \end{aligned} \tag{17}$$

where $L_1 = \mathcal{F}\tilde{C}_1$, find the stationary dt filter $\mathcal{K}$ given by

$$\mathcal{K}: \begin{aligned} \hat{x}[kT] &= \Phi_f\hat{x}[(k-1)T] + \Gamma_f y_s[kT], \\ \hat{z}[kT] &= \mathcal{H}L_f\hat{x}[kT] + \mathcal{H}D_f y_s[kT], \end{aligned} \tag{18}$$

such that the cost $J_N(K)$ given by

$$J_N(K) = \mathcal{E}\mathrm{tr}\left\{\tilde{e}(kT)\tilde{e}^T(kT)\right\}, \quad (19)$$

where $\tilde{e} = \hat{z}[kT] - \tilde{z}[kT]$ is minimized. In the next section we derive the optimal constant matrices Φ_f, Γ_f, L_f, and D_f of K.

III. MINIMUM VARIANCE FILTER DESIGN

Using the state space representation of the dt lifted system in (17) and the dt filter K in (18), the state space equations for the estimation error $\tilde{e}[k]$ are

$$\begin{aligned}\mathcal{X}[k+1] &= \mathcal{A}\mathcal{X}[k] + \mathcal{B}\tilde{\eta}[k]\\ \tilde{e}[k] &= \mathcal{C}\mathcal{X}[k] + \mathcal{D}\tilde{\eta}[k],\end{aligned} \quad (20)$$

where the sampling time T is dropped from now on and $\mathcal{X}$, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ are given by

$$\mathcal{X}[k] = \begin{pmatrix} x[k] \\ \hat{x}[k] \end{pmatrix}; \ \mathcal{A} = \begin{pmatrix} \tilde{\Phi} & 0 \\ \Gamma_f L_1 \tilde{\Phi} & \Phi_f \end{pmatrix};$$

$$\mathcal{B} = \begin{pmatrix} \Gamma \\ \Gamma_f L_1 \Gamma \end{pmatrix}; \ \mathcal{C} = \begin{pmatrix} \tilde{C}_2 - \mathcal{H}D_f L_1 & -\mathcal{H}L_f \end{pmatrix};$$

$$\text{and} \quad \mathcal{D} = \tilde{D}_2. \quad (21)$$

Then, the stationary covariance matrix of $\mathcal{X}$ is

$$P = \mathcal{E}\left\{\mathcal{X}[k]\mathcal{X}^T[k]\right\} = \begin{pmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{pmatrix}. \quad (22)$$

By replacing $\mathcal{X}[k]$ of equation (20) in equation (22) and (19), it is easy to see that the stationary covariance matrix and the cost $J_N(K)$ fulfil

$$P = \mathcal{A}\ P\ \mathcal{A}^T + \mathcal{B}\mathcal{B}^T \quad (23)$$

$$J_N(\mathcal{K}) = tr\{\mathcal{C}\ P\ \mathcal{C}^T + \mathcal{D}\mathcal{D}^T\}. \quad (24)$$

Note that the covariance matrix P does not depend on L_f nor on D_f. Then, in order to obtain the optimal values of Φ_f, Γ_f, L_f, and D_f, we start by minimizing the cost $J_N(\mathcal{K})$ with respect to L_f. To this end, let L_f be all possible matrices of constants parameterized as $L_f = L_{fo} + \epsilon\Upsilon$, where ϵ is a small scalar; Υ, an arbitrary constant matrix; and L_{fo}, the optimal matrix which minimizes the cost $J_N(K)$ in (24). No matter which Υ is chosen, the minimum of the cost function is obtained for the choice of the scalar value $\epsilon = 0$. The necessary condition for the minimum is given by $\frac{\partial J_N(K)}{\partial\epsilon}|_{\epsilon=0} = 0$ and the sufficient condition is $\frac{\partial^2 J_c}{\partial\epsilon^2}|_{\epsilon=0} > 0$. The last condition is always satisfied (due to the convexity of $J_N(K)$). By replacing $\mathcal{C}$ given by (21) in (24) and by using $L_f = L_{fo} + \epsilon\Upsilon$, the following optimal value of L_f is obtained by deriving the cost with respect to ϵ and equating the result to zero:

$$L_{fo} = \frac{\mathcal{H}^T(\tilde{C}_2 - \mathcal{H}D_f L_1)P_{12}P_2^{-1}}{\mathcal{N}}. \quad (25)$$

By using (25), we can rewrite $\mathcal{C}$ as

$$\mathcal{C} = \begin{pmatrix} a & -\mathcal{H}\mathcal{H}^T a \frac{P_{12}P_2^{-1}}{\mathcal{N}} \end{pmatrix}, \quad (26)$$

where

$$a = \tilde{C}_2 - \mathcal{H}D_f L_1. \quad (27)$$

Then, the cost $J_N(K)$ in (24) can be written as

$$J_N(\mathcal{K}) = tr\left\{\begin{pmatrix} a & \frac{-\mathcal{H}\mathcal{H}^T}{\mathcal{N}}a \end{pmatrix} \begin{pmatrix} P_1 & L \\ L & L \end{pmatrix}\begin{pmatrix} a^T \\ -a^T\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} \end{pmatrix} + \tilde{D}_2\tilde{D}_2^T\right\}, \quad (28)$$

where

$$L = P_{12}P_2^{-1}P_{12}^T. \quad (29)$$

We can rewrite the cost in a much more suitable form as

$$J_N(\mathcal{K}) = tr\left\{\begin{pmatrix} \frac{\mathcal{H}\mathcal{H}^T}{N}a & \frac{-\mathcal{H}\mathcal{H}^T}{\mathcal{N}}a \end{pmatrix}\begin{pmatrix} P_1 & L \\ L & L \end{pmatrix}\begin{pmatrix} a^T\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} \\ -a^T\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} \end{pmatrix} - \frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}}aP_1a^T\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} + aP_1a^T + \tilde{D}_2\tilde{D}_2^T\right\}. \quad (30)$$

Since P_1, a, and $\tilde{D}_2$ do not depend on Φ_f and Γ_f, we are able to find the minimum of the cost with respect to Φ_f and Γ_f by just minimizing the first term. Thus, we rewrite the first term as

$$tr\left\{\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}}a\begin{pmatrix} I & -I \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & P_{12}P_2^{-1} \end{pmatrix}\begin{pmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & P_2^{-1}P_{12}^T \end{pmatrix}\begin{pmatrix} I \\ -I \end{pmatrix}a^T\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}}\right\}. \quad (31)$$

Taking into account that the covariance matrix of the error in the states ($\tilde{x}$) is

$$\tilde{P} = \mathcal{E}(\tilde{x}[k]\tilde{x}^T[k]) = \begin{pmatrix} I & -I \end{pmatrix} P \begin{pmatrix} I \\ -I \end{pmatrix}, \quad (32)$$

where $\tilde{x} = x - \hat{x}$, we will show that the minimum of (31) is reached by minimizing $\tilde{P}$ in the positive definite sense. In the minimum, the estimation error $\tilde{x}$ is uncorrelated with the estimation $\hat{x}$ of x. This is the optimal error that can be achieved for the covariance matrix $\tilde{P}$ since otherwise, if the estimation error $\tilde{x}$ was correlated with $\hat{x}$, a further decrease in $\tilde{P}$ could be achieved. Then, $\mathcal{E}(\tilde{x}\hat{x}^T) = 0$ holds, which leads to the following equality: $P_{12} = \mathcal{E}(x\hat{x}^T) = \mathcal{E}(\tilde{x}\hat{x}^T) + \mathcal{E}(\hat{x}\hat{x}^T) \Rightarrow P_{12} = P_2$. By using this equality, equation (31) is written as

$$tr\left\{\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}}a\tilde{P}a^T\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}}\right\}. \quad (33)$$

Since the term in (33) is positive definite, its optimum value is given by the minimum of $\tilde{P}$. Then, we need to find the minimum of $\tilde{P}$ with respect to Φ_f and Γ_f, subjected to the bias constraint $\mathcal{E}(\tilde{x}) = 0$. The optimal stationary solution to this problem is given by the stationary dt Kalman filter (Anderson and Moore, 1975) according to the following equations:

$$\begin{aligned} \Psi &= (\tilde{\Phi}\tilde{P}\tilde{\Phi}^T + \tilde{\Gamma}\tilde{\Gamma}^T) \\ \Gamma_f &= \Psi L_1^T (L_1 \Psi L_1^T)^{-1} \\ \Phi_f &= \tilde{\Phi} - \Gamma_f L_1 \tilde{\Phi}. \end{aligned} \quad (34)$$

The optimal value of D_f still remains to be obtained. To this end, we replace (33) in (30), which gives

$$\begin{aligned} J(\mathcal{K}) &= tr\{\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} a(\tilde{P} - P_1)a^T \frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} + aP_1a^T + \\ &\quad +\tilde{D}_2\tilde{D}_2^T\} \\ &= tr\{\frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} a\tilde{P}a^T \frac{\mathcal{H}\mathcal{H}^T}{\mathcal{N}} + \tilde{D}_2\tilde{D}_2^T\}. \end{aligned} \quad (35)$$

In the same way as with L_{fo}, we obtain the optimum D_{fo} by using the parametrization $D_f = D_{fo} + \epsilon\Upsilon$. By deriving the cost with respect to ε, and equating to zero, the optimal value is

$$D_f = \frac{\mathcal{H}\tilde{C}_2\tilde{P}L_1^T}{\mathcal{N}}(L_1\tilde{P}L_1^T)^{-1}. \quad (36)$$

The optimal stationary dt filer K is given by (18), where Φ_f, Γ_f, and D_f are given by (34) and (36). L_f from (25) is given by

$$L_f = \frac{\mathcal{H}^T(\tilde{C}_2 - \mathcal{H}D_fL_1)}{\mathcal{N}}. \quad (37)$$

The expression of $\tilde{P}$ still remains to be derived. This can be made by using (32) with P defined by (23). By means of the expressions of Φ_f and Γ_f, the following dt algebraic Riccati equation (DARE) is achieved:

$$\tilde{P} = \Psi - \Psi L_1^T(L_1\Psi L_1^T)^{-1}L_1\Psi. \quad (38)$$

Then, the minimum of the cost function is given by (35) together with the solution of (38).

IV. EXAMPLE

Consider the following scalar prediction problem depicted in Fig. 2. The prediction of the continuous time signal $z(t)$ at time $t+1$ *sec.* is desired. The signal $z(t)$ is the output of the ct system S_1. The measured output $y_s(t)$ is the signal $z(t)$ corrupted by additive noise $v(t)$. $z(t)$ is the output of the ct system S_2. Both systems are excited by the inputs $\eta_1(t)$ and $\eta_2(t)$, which are uncorrelated white noise processes with spectrum amplitudes α^2 and β^2 respectively. The predictor is a digital ensemble that operates at a sampling rate of $T = 1sec.$. In Fig. 2, $\hat{z}(t+T|t)$ means the prediction of $z(t)$ at time $t+1$, taking into account measurements until time t. The same consideration is made for the discrete signal $\hat{z}(k+1|k)$. This prediction problem design can be formulated as in Fig. 3, which is similar to the assemble in Fig. 1.

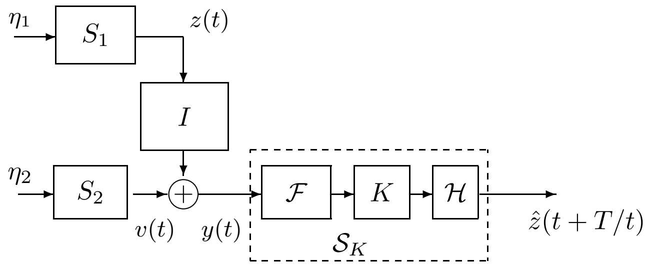

Figure 2: Prediction problem set-up.

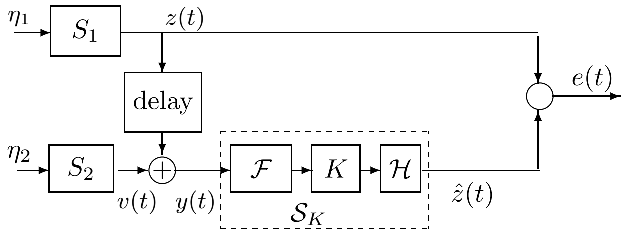

Figure 3: Equivalent representation of the prediction problem set-up.

In the example, systems S_1 and S_2 are given by the transfer functions

$$\begin{aligned} S1(s) &= \frac{\alpha((1/1.41\omega_o)s + 1)}{(1/\omega_0)s^2 + (1.41/\omega_0)s + 1} \\ S2(s) &= \frac{\beta}{(1/\omega_0)s^2 + (1.41/\omega_0)s + 1}. \end{aligned} \quad (39)$$

The optimal sampled-data predictor K, which minimizes the predictor error variance, is obtained by performing the following steps:

1. Obtain the zero-order hold discretization of ct systems S_1 and S_2 at sampling time $T_s = T/\mathcal{N}$ with $T = 1sec$ and an arbitrary value of N.

2. Build the state equations of both systems including the delay and express it as in equation (14). Note that the number of states increases from four to five.

3. Lift the dt system and apply the operator $\mathcal{F}$ to obtain the state space representation as in (17).

4. Obtain the value of $\tilde{P}$ using the DARE equation (38).

5. Compute the cost $J_N(K)$ of equation (35).

6. Compare the cost obtained in the former step with respect to the previous iteration. If there exists no significant difference, then the optimal filter approximation has been obtained. Otherwise, increase the value of $\mathcal{N}$ and return to step 1.

For different values of ω_o (the natural frequency of the systems), the optimal discrete predictor was designed by following the steps previously described. Optimal dt filters of first order-degree are obtained for every ω_o -keeping the sampling rate $T = 1sec$ constant. The cost $J_N(K)$ tends towards the minimum as $\mathcal{N}$ increases. In our example, the cost obtained with $\mathcal{N} = 20$ constitutes, in fact, an approximation to the optimal solution with an error lower than 1%. In Figs. 4 and 5, the minimum of the cost obtained versus the natural frequency ω_o is shown.

In order to asses the performance of the proposed design method, the cost without predictor ($K = 0$), the cost with $\mathcal{N} = 1$ -which corresponds to the standard discrete time design-, and the cost obtained by means of the optimal ct design are depicted. The ct filter design was achieved by using a 10th Padé delay approximation. By increasing the order of the Padé approximation, no significant improvements were made. The different costs were obtained in the case where the noise amplitudes are $\alpha = 1$ and $\beta = 0.01$, and they are depicted in Fig. 4. In Fig. 5, the same costs are shown when both α and β are equal to 1.

The example shows that considerable improvements can be made with respect to the classical discrete time design ($\mathcal{N} = 1$) by using the proposed sampled-data design method. The differences in cost among the different designs become less significant as the natural frequency decreases. This is obvious since the minimum of the cost obtained by means of the optimal ct filter is equivalent to the discrete one as $T\omega_o$ tends to zero. Fewer improvements are obtained in the case of higher noise power (Fig. 5). The filter parameters in the case of Fig. 4 when $\omega_n = 0.6$ are the following:

$$\mathcal{K}: \quad \hat{x}[kT] = -0.2684\hat{x}[(k-1)T] + y_s[kT], \\ \hat{z}[kT] = 3.847\hat{x}[kT] - 2.063y_s[kT]. \qquad (40)$$

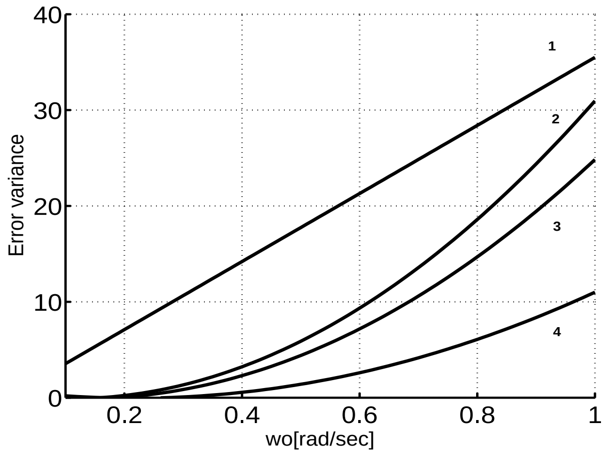

Figure 4: Cost versus ω_o with $\alpha = 1$ and $\beta = 0.01$ (1)-Cost without predictor K=0; (2)-Cost using the dt solution; (3)-Cost using the sampled-data design; and (4)- Cost using the ct design.

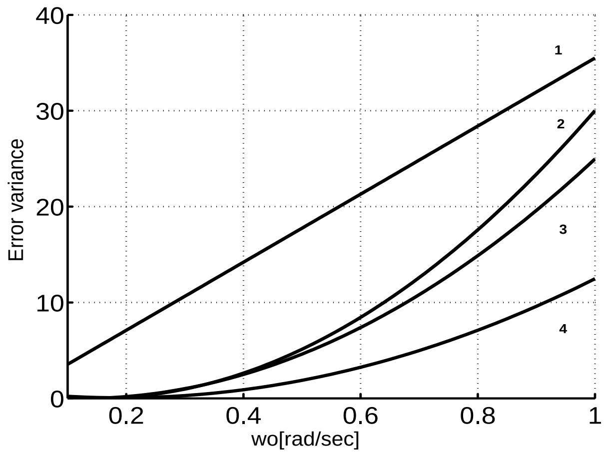

Figure 5: Cost versus ω_o with $\alpha = 1$ and $\beta = 1$ (1)-Cost without predictor K=0; (2)-Cost using the dt solution; (3)-Cost using the sampled-data design, and (4)-Cost using the ct design.

V. CONCLUSIONS

A sampled-data minimum variance filter design problem enclosing deconvolution, prediction, and smoothing has been presented. The optimal solution can be found by an approximation method based on lifted systems. The optimal filter is designed by using the solution of the algebraic Riccati equation. Although the original sampled-data system is transformed into an equivalent LTI dt system with infinite-dimensional input-output space, the Riccati equation and the designed filter are finite dimensional.

Our design requires the solution of a dt algebraic Riccati equation instead of the two diophantine equations and the spectral factorization used in the polynomial approach design as in Milocco (1996). There exist standard algorithms in Arnold and Laub (1984) to solve the DARE efficiently.

An example of sampled data prediction has been presented to show the design procedure. In the example, under certain sampling conditions, improvements up to 20% with respect to the classical discrete time solutions were obtained without increasing the filter structure complexity.

The filter obtained is optimal when the ct systems are linear and time invariant. However, the design can be extended to deal with time-variant system. In such case, instead of using the expressions of the linear time invariant lifted system given in (16), equivalent expressions for linear time-variant lifted systems should be used. The approach can also be extended to deal with uncertain systems as in Milocco and Muravchik (2003).

Acknowledgments

This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), and Universidad Nacional del Comahue (UNC), both

of República Argentina.

REFERENCES

Ahlén, A. and M. Sternad, " Wiener filter design using polynomial equations," *IEEE Trans. Signal Proc.,* **39**, no. 11, 2387-2399, (1991).

Anderson, B.D.O. and J.B. Moore, *Optimal Filtering,* Prentice-Hall, Englehood Cliff, New Jersey, (1975).

Arnold, W.F. , and A.J. Laub, "Generalized Eigenproblem Algorithms and Software for Algebraic Riccati Equations," *Proc. IEEE,* **72**, 1746-1754, (1984).

Bamieh, B. J. Pearson, B. Francis, and A. Tannenbaum,"A lifting technique for linear periodic system with application to sampled-data control," *Syst. Contr. Lett.,* **17**, 78-88, (1991).

Grimble, M.J. "Polynomial systems approach to optimal linear filtering and prediction," *Int. J. Control,* **41**, 1545-1564, (1995).

Hara, S. , B.D.O. Anderson and H. Fujioka,"Relating H_2 and H_∞ norms bounds for sampled-data systems," *IEEE Trans. Automat. Contr.,* **42**, 858-863, (1997).

Kabamba, P.T. and S. Hara, "Worst case analysis of sampled-data control systems," *IEEE Trans. Automat. Contr.,* **38**, no. 9, 1337-1357, (1993).

Kalman, R.E. "A new approach to linear filtering and prediction problems," *ASME Trans. Journal of Basic Engineering,* **82D**, 35-45, (1960).

Keller, J. and B. Anderson, "A new approach to the discretization of continuous-time controller," *IEEE Trans. Automat. Contr.,* **37**, 214-223, (1992).

Milocco, R. and J. DeDoná,"Optimal parallel model solutions using the polynomial approach," *Int. J. Control,* **65**, no. 4, 681-697, (1996).

Milocco, R. and C. Muravchik, "H2 Optimal linear robust sampled-data filtering design using polynomial approach" *IEEE Trans. Signal Proc.,* **51**, 1816-1824, (2003).

Roberts, A. and M. Newmann, "Polynomial approach to Wiener filtering," *Int. J. Contr.,* **45**, 1953-1959, (1987).

Shaked, U. "A generalized transfer function approach to linear stationary filtering and steady-state optimal control problems," *Int. J. Control,* **29**, 741-770, (1976).

Sun, W. , K. Nagpal and P. Khargonekar, "H_∞ control and filtering for sampled-data systems," *IEEE Trans. Automat. Contr.,* **38**, no. 8, 1162-1175, (1993).

Wang, Z., Huang B. and P. Huo , "Sampled-data filtering with error covariance assignment" *IEEE Trans. Signal Proc.,* **49**, no. 3, 666-670, (2001).

Wiener, N. *Extrapolation, Interpolation and Smoothing of Stationary Time Series,* The Technology Press and Wiley, New York, (1950).

Received: October 8, 2003.
Accepted for publication: October 27, 2004.
Recommended by Guest Editors C. E. D´Attellis and S. Ponce.