REFERENCES

[1] H. Al-Duwaish, M. N. Karim, and V. Chandrasekar, "Use of multilayer feedforward neural networks in identification and control of Wiener model," *Proc. Inst. Elect. Eng.*, vol. 143, no. 3, pp. 255–258, May 1996.
[2] J. S. Bendat, *Nonlinear System Analysis and Identification From Random Data*. New York: Wiley, 1990.
[3] S. A. Billings and S. Y. Fakhouri, "Theory of separable processes with applications to the identification of nonlinear systems," *Proc. Inst. Elect. Eng.*, vol. 125, no. 9, pp. 1051–1058, Oct. 1978.
[4] W. Greblicki, "Nonparametric identification of Wiener systems," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1487–1493, Sept. 1992.
[5] ——, "Nonparametric approach to Wiener system identification," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 538–545, June 1997.
[6] ——, "Recursive identification of Wiener systems," *Int. J. Appl. Math. Comput. Sci.*, vol. 11, no. 4, pp. 977–991, 2002.
[7] A. D. Kalafatis, L. Wang, and W. R. Cluett, "Identification of Wiener-type nonlinear systems in a noisy environment," *Int. J. Control*, vol. 66, no. 6, pp. 923–941, 1997.
[8] D. Westwick and M. Verhaegen, "Identifying MIMO Wiener systems using subspace model identification methods," *Signal Processing*, vol. 52, pp. 235–258, 1996.
[9] R. L. Wheeden and A. Zygmund, *Measure and Integral*. New York: Marcel Dekker, 1977.
[10] T. Wigren, "Convergence analysis of recursive identification algorithms based on nonlinear Wiener model," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 2191–2206, Nov. 1994.

# A Reorganized Innovation Approach to Linear Estimation

Huanshui Zhang, Lihua Xie, David Zhang, and Yeng Chai Soh

***Abstract*—This note will address a linear minimum variance estimation of discrete-time systems with instantaneous and delayed measurements. Although the problem may be approached via system augmentation and standard Kalman filtering, the computation of filter may be expensive when the dimension of the system is high and the measurement lag is significant. In this note, a new tool, termed as *reorganized innovation sequence*, is presented for deriving the optimal filter. The optimal filter is given by two Riccati difference equations (RDEs) with the same dimension as that of the original system. The approach is shown to induce saving of computational cost as compared to the system augmentation approach, especially when the delay is large. Further, it can be extended to solving the more complicated $H_\infty$ fixed-lag smoothing in Krein space.**

***Index Terms*—Delayed measurement, discrete-time systems, innovation analysis, optimal filtering, Riccati equations.**

## I. INTRODUCTION

Kalman filtering [2], [3] is concerned with the minimization of filtering error covariance (termed as $H_2$ estimation) and has become a major tool of state estimation since the 1960s; see, e.g., [6]–[13] and the references therein. It has been widely used in signal processing, communication and control applications. However, there are still some Kalman filtering problems which deserve further studies. One such problem is the optimal filtering of systems with various measurement delays. In the continuous-time context, the optimal estimation of time-delay systems has been well studied in the past decades; see [6], [7], and the references therein. The approaches in these works are usually related to solving a partial differential equation (PDE) which does not have an explicit solution in general. For the case of discrete-time systems, the problem has been investigated via system augmentation and standard Kalman filtering [8], [10], [16] or the polynomial approach [9], [11], [12]. Note that the augmented Kalman filtering approach is computationally expensive, especially when the dimension of the system is high and the measurement lags are large. On the other hand, the polynomial approach only addresses the steady-state filtering problem and it requires solving a much higher order of spectral factorization for systems with delays.

In this note, we will revisit the Kalman filtering problem for time-varying systems with measurement delays. We consider the system with instantaneous and delayed measurements which is described by

$$\mathbf{x}(t+1) = \Phi_t \mathbf{x}(t) + \Gamma_t \mathbf{u}(t) \tag{1.1}$$

$$\mathbf{y}(t) = H_t \mathbf{x}(t) + \mathbf{v}(t) \tag{1.2}$$

$$\mathbf{z}_{t-d}(t) = L_{t-d} \mathbf{x}(t-d) + \mathbf{v}_z(t) \tag{1.3}$$

where $\mathbf{x}(t) \in R^n$ is the state, $\mathbf{u}(t) \in R^r$ is the input noise, $\mathbf{y}(t) \in R^m$, and $\mathbf{z}_{t-d}(t) \in R^p$ are, respectively, the instantaneous and delayed measurements, $\mathbf{v}(t) \in R^m$ and $\mathbf{v}_z(t) \in R^p$ are the measurement noises. $d$ is the measurement delay which is an integer. The initial state $\mathbf{x}(0)$ and $\mathbf{u}(t)$, $\mathbf{v}(t)$ and $\mathbf{v}_z(t)$ are uncorrelated white noises with zero means and known covariance matrices $\mathcal{E}[\mathbf{x}(0)\mathbf{x}^T(0)] = P_0$, $\mathcal{E}[\mathbf{u}(i)\mathbf{u}^T(j)] = Q_u(i)\delta_{ij}$, $\mathcal{E}[\mathbf{v}(i)\mathbf{v}^T(j)] = Q_v(i)\delta_{ij}$ and $\mathcal{E}[\mathbf{v}_z(i)\mathbf{v}_z^T(j)] = Q_{v_z}(i)\delta_{ij}$, respectively.

Note that the previous estimation problem has important applications in many engineering problems such as in communications and sensor fusion [1] and networked control systems [15].

With the delayed measurement in (1.3), the system (1.1)–(1.3) is not in a standard form to which the standard Kalman filtering is applicable. Let $\mathbf{y}_s(t)$ denote the observation of the system (1.1)–(1.3) at time $t$, then

$$\mathbf{y}_s(t) = \begin{cases} \mathbf{y}(t), & 0 \le t < d \\ \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{z}_{t-d}(t) \end{bmatrix}, & t \ge d. \end{cases} \tag{1.4}$$

It follows that

$$\mathbf{y}_s(t) = \begin{cases} H_t \mathbf{x}(t) + \mathbf{v}_s(t), & 0 \le t < d \\ \begin{bmatrix} H_t & 0 \\ 0 & L_{t-d} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t-d) \end{bmatrix} + \mathbf{v}_s(t), & t \ge d \end{cases} \tag{1.5}$$

where

$$\mathbf{v}_s(t) = \begin{cases} \mathbf{v}(t), & 0 \le t < d \\ \begin{bmatrix} \mathbf{v}(t) \\ \mathbf{v}_z(t) \end{bmatrix}, & t \ge d \end{cases} \tag{1.6}$$

which is a white noise with zero mean and covariance matrix

$$Q_{v_s}(t) = \begin{cases} Q_v(t), & 0 \le t < d \\ \begin{bmatrix} Q_v(t) & 0 \\ 0 & Q_{v_z}(t) \end{bmatrix}, & t \ge d. \end{cases} \tag{1.7}$$

Manuscript received November 21, 2002; revised November 4, 2003. Recommended by Associate Editor R. Boel. The work of H. Zhang was supported by the National Nature Science Foundation of China under Grant 60174017, and was carried out in part at Nanyang Technological University, Singapore, and The Chinese University of Hong Kong, Hong Kong.

H. Zhang is with the Information and Control Research Center, Shenzhen Graduate School of Harbin Institute of Technology, HIT Campus Shenzhen University Town, Shenzhen 518055, P. R. China (e-mail: h_s_zhang@hit.edu.cn).
L. Xie and Y. C. Soh are with the BLK S2, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore.
D. Zhang is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong.

Digital Object Identifier 10.1109/TAC.2004.835599

The $H_2$ optimal estimation problem can be stated as: *Given the observation* $\{\{\mathbf{y}_s(i)\}_{i=0}^t\}$, *find a linear least mean square error estimator* $\hat{\mathbf{x}}(t|t)$ *of* $\mathbf{x}(t)$.

*Remark 1.1:* Traditionally, the aforementioned problem can be solved by transforming it into a standard Kalman filtering problem through system augmentation. By introducing an augmented state

$$\mathbf{x}_a^T(t) = \left[\mathbf{x}^T(t)\ \{L_{t-1}\mathbf{x}(t-1)\}^T\ \dots\ \{L_{t-d}\mathbf{x}(t-d)\}^T\right] \quad (1.8)$$

we obtain an augmentated state-space model

$$\mathbf{x}_a(t+1) = \Phi_a(t)\mathbf{x}_a(t) + \Gamma_a(t)\mathbf{u}(t), \quad (1.9)$$

$$\begin{bmatrix}\mathbf{y}(t)\\ \mathbf{z}_{t-d}(t)\end{bmatrix} = H_{L_a}(t)\mathbf{x}_a(t) + \mathbf{v}_s(t) \quad (1.10)$$

where

$$\Phi_a(t) = \begin{bmatrix}\Phi_t & & & & \\ L_t & & & & \\ & I_p & & & \\ & & \cdots & & \\ & & & I_p & 0\end{bmatrix} \quad \Gamma_a(t) = \begin{bmatrix}\Gamma_t\\ 0\\ \vdots\\ 0\end{bmatrix}$$

$$H_{L_a}(t) = \begin{bmatrix}H_a(t)\\ L_a(t)\end{bmatrix} \quad (1.11)$$

and

$$\begin{aligned} H_a(t) &= [H_t \quad 0 \quad \dots \quad 0]\\ L_a(t) &= [0 \quad \dots \quad 0 \quad I_p] \quad \text{for } t > d, \text{ and}\\ L_a(t) &= 0 \text{ for } t \le d. \end{aligned} \quad (1.12)$$

Then, the optimal estimate $\hat{\mathbf{x}}(t|t)$ is obtained by

$$\hat{\mathbf{x}}(t|t) = [I_n\ \dots\ 0]\hat{\mathbf{x}}_a(t|t) \quad (1.13)$$

where $\hat{\mathbf{x}}_a(t|t)$ is the optimal estimate of the previous augmented system which can be obtained by the standard Kalman filter. However, the augmentation leads to a much higher system dimension and, thus, a much higher computational cost.

*Remark 1.2:* It should be pointed out that special structure of $\Phi_a(t)$ in the aforementioned Kalman filtering formulation leads some computational simplification which to be discussed in Section IV, but nevertheless our approach to be presented in this note will be computationally more efficient.

In this note, we will propose a new method for the optimal filter design without resorting to system augmentation. Our approach is based on projection and a reorganized innovation sequence which is different from the standard Kalman innovation sequence. It is shown that the proposed approach is computationally attractive as compared with the augmentation approach. Furthermore, our approach can be extended to give a solution to the $H_\infty$ fixed-lag smoothing which has often been solved via a system augmentation approach [14].

The note is organized as follows. The reorganized innovation sequence and the associated RDEs are introduced in Section II. The optimal filter is derived in Section III which is given in terms of two RDEs having the same dimension as that of the original system. Some discussion and cost comparison with the existing system augmentation approach is given in Section IV. Conclusions are drawn in Section V.

## II. Reorganized Innovation Sequence

In this section, we will present a solution to the $H_2$ estimation of the system (1.1)–(1.3) involving delayed measurements using the projection in Hilbert space. The key to our discussion in this section is to reorganize the instantaneous and delayed measurements and introduce an associated innovation sequence.

As is well known, given the measurement sequence $\{\mathbf{y}_s(i)\}_{i=0}^t$, the optimal state estimator $\hat{\mathbf{x}}(t|t)$ is the projection of $\mathbf{x}(t)$ onto the linear space spanned by the measurement sequence, denoted by $\mathcal{L}\{\{\mathbf{y}_s(i)\}_{i=0}^t\}$ [2], [16].

First, observe from (1.4) that for $d > t \ge 0$

$$\mathcal{L}\left\{\{\mathbf{y}_s(i)\}_{i=0}^t\right\} = \mathcal{L}\left\{\{\mathbf{y}(i)\}_{i=0}^t\right\} \quad (2.1)$$

and the estimator $\hat{\mathbf{x}}(t|t)$ is a standard $H_2$ estimator associated with (1.1)–(1.2). When $t \ge d$, it is easy to know that the linear space $\mathcal{L}\{\{\mathbf{y}_s(i)\}_{i=0}^t\}$ is equivalent to

$$\mathcal{L}\left\{\begin{bmatrix}\mathbf{y}(0)\\ \mathbf{z}_0(d)\end{bmatrix}, \begin{bmatrix}\mathbf{y}(1)\\ \mathbf{z}_1(1+d)\end{bmatrix}, \dots, \begin{bmatrix}\mathbf{y}(t-d)\\ \mathbf{z}_{t-d}(t)\end{bmatrix}\right. \\ \left.\mathbf{y}(t-d+1), \dots, \mathbf{y}(t)\right\}. \quad (2.2)$$

Denote

$$\mathbf{y}_f(i) \triangleq \begin{bmatrix}\mathbf{y}(i)\\ \mathbf{z}_i(i+d)\end{bmatrix}, \qquad i = 0, 1, \dots, t-d. \quad (2.3)$$

It is easy to know that $\mathbf{y}_f(i)$ satisfies

$$\mathbf{y}_f(i) = \begin{bmatrix}H_i\\ L_i\end{bmatrix}\mathbf{x}(i) + \mathbf{v}_f(i), \qquad i = 0, 1, \dots, t-d \quad (2.4)$$

with

$$\mathbf{v}_f(i) = \begin{bmatrix}\mathbf{v}(i)\\ \mathbf{v}_z(i+d)\end{bmatrix} \quad (2.5)$$

being a white noise of zero mean and covariance matrix $Q_{v_f}(i) = \begin{bmatrix}Q_v(i) & 0\\ 0 & Q_{v_z}(i+d)\end{bmatrix}$. It should be noted that $\mathbf{y}_f(i)$ contains measurements of the state $\mathbf{x}(i)$ at time instants $i$ and $i+d$. Observe that (1.1) and (2.4) give a standard state-space representation.

The following notations will be used throughout the note:

| | |
|---|---|
| $t_d$ | $t-d$; |
| $\hat{\xi}(j\|t)$ | optimal estimate of $\xi(j)$ given $\{\mathbf{y}_s(0), \dots, \mathbf{y}_s(t)\}$; |
| $\hat{\xi}(j\|t+i, t)$ | the estimate of $\xi(j)$ given $\{\mathbf{y}_f(0), \dots \mathbf{y}_f(t); \mathbf{y}(t+1), \dots \mathbf{y}(t+i), i \ge 0\}$. |

It is obvious that $\hat{\xi}(j|t,t)$ is the standard Kalman estimator for system (1.1) and (2.4), and the estimator $\hat{\mathbf{x}}(t|t)$ to be sought can be redenoted as $\hat{\mathbf{x}}(t|t, t_d)$. In other words, the optimal estimation problem is equivalent to finding the optimal estimate $\hat{\mathbf{x}}(t|t, t_d)$ of $\mathbf{x}(t)$ which will be discussed later in Section III.

### A. Reorganized Innovation Sequence

To define the reorganized innovation and the associated Riccati equation, we introduce the following stochastic sequence:

$$\mathbf{w}(t+i, t) \triangleq \mathbf{y}(t+i) - \hat{\mathbf{y}}(t+i|t+i-1, t), \ \ i > 0 \quad (2.6)$$

$$\begin{aligned}\mathbf{w}(t, t) &\triangleq \mathbf{y}_f(t) - \hat{\mathbf{y}}_f(t|t-1, t-1)\\ \hat{\mathbf{y}}_f(0|-1, -1) &= 0\end{aligned} \quad (2.7)$$

where $\hat{\mathbf{y}}(t+i|t+i-1, t)$ is the projection of $\mathbf{y}(t+i)$ onto the linear space of $\{\mathbf{y}_f(0), \dots, \mathbf{y}_f(t); \mathbf{y}(t+1), \dots, \mathbf{y}(t+i-1)\}$ and $\hat{\mathbf{y}}_f(t|t-1, t-1)$ is the projection of $\mathbf{y}_f(t)$ onto the linear space of $\{\mathbf{y}_f(0), \dots \mathbf{y}_f(t-1)\}$. It is clear that $\mathbf{w}(t,t)$ is the standard Kalman filtering innovation sequence for the system (1.1) and (2.4). We then have the following relationships:

$$\mathbf{w}(t+i,t) = H_{t+i}\mathbf{e}(t+i,t) + \mathbf{v}(t+i),\ i>0 \tag{2.8}$$

$$\mathbf{w}(t,t) = \begin{bmatrix} H_t \\ L_t \end{bmatrix}\mathbf{e}(t,t) + \mathbf{v}_f(t) \tag{2.9}$$

where

$$\mathbf{e}(t+i,t) \triangleq \mathbf{x}(t+i) - \hat{\mathbf{x}}(t+i|t+i-1,t),\ i>0 \tag{2.10}$$

$$\mathbf{e}(t,t) \triangleq \mathbf{x}(t) - \hat{\mathbf{x}}(t|t-1,t-1). \tag{2.11}$$

It is clear that $\mathbf{e}(t+1,t) = \mathbf{e}(t+1,t+1)$. The following lemma shows that $\{\mathbf{w}(\cdot,\cdot)\}$ is in fact the innovation sequence for the model of (1.1), (2.4), and (2.5).

*Lemma 2.1:* $\{\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d); \mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t,t_d)\}$ is the innovation sequence which spans the same linear space as

$$\mathcal{L}\{\mathbf{y}_f(0),\dots,\mathbf{y}_f(t_d);\ \mathbf{y}(t_d+1),\dots,\mathbf{y}(t)\}$$

or, equivalently, $\mathcal{L}\{\mathbf{y}_s(0),\dots,\mathbf{y}_s(t)\}$.

*Proof:* First, it is readily seen from (2.6) and (2.7) that $\mathbf{w}(i,i)$, $i \le t_d$ (or $\mathbf{w}(t_d+i,t_d)$, $d \ge i > 0$) is a linear combination of the observations $\mathbf{y}_f(0),\dots,\mathbf{y}_f(i)$ (or $\mathbf{y}_f(0),\dots,\mathbf{y}_f(t_d)$; $\mathbf{y}(t_d+1),\dots,\mathbf{y}(t_d+i)$). Conversely, $\mathbf{y}_f(i)$, $t_d \ge i \ge 0$, (or $\mathbf{y}(t_d+i)$, $d \ge i > 0$) can be given in terms of a linear combination of $\mathbf{w}(0,0),\dots,\mathbf{w}(i,i)$ (or $\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d)$; $\mathbf{w}(t_d+1,t_d),\dots\mathbf{w}(t_d+i,t_d)$). Thus, $\{\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d);\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t,t_d)\}$ spans the same linear space as $\mathcal{L}\{\mathbf{y}_f(0),\dots,\mathbf{y}_f(t_d);\mathbf{y}(t_d+1),\dots,\mathbf{y}(t)\}$ or equivalently $\mathcal{L}\{\mathbf{y}_s(0),\dots\mathbf{y}_s(t)\}$. Next, we show that $\mathbf{w}(\cdot,\cdot)$ is an uncorrelated sequence. In fact, for any $d \ge i > 0$ and $t_d \ge j \ge 0$, from (2.8) we have

$$\mathcal{E}\left[\mathbf{w}(t_d+i,t_d)\mathbf{w}^T(j,j)\right] = \mathcal{E}\left[H_{t_d+i}\mathbf{e}(t_d+i,t_d)\mathbf{w}^T(j,j)\right] + \mathcal{E}\left[\mathbf{v}_f(t_d+i)\mathbf{w}^T(j,j)\right]. \tag{2.12}$$

Note that $\mathcal{E}[\mathbf{v}_f(t_d+i)\mathbf{w}^T(j,j)] = 0$. Since $\mathbf{e}(t_d+i,t_d)$ is the state prediction error, it follows that $\mathcal{E}[\mathbf{e}(t_d+i,t_d)\mathbf{w}^T(j,j)] = 0$ and, thus, $\mathcal{E}[\mathbf{w}(t_d+i,t_d)\mathbf{w}^T(j,j)] = 0$, which implies that $\mathbf{w}(j,j)$ $(t_d \ge j \ge 0)$ is uncorrelated with $\mathbf{w}(t_d+i,t_d)$ $(d \ge i > 0)$. Similarly, it can be verified that $\mathbf{w}(i,i)$ is uncorrelated with $\mathbf{w}(j,j)$ for $i \ne j$ and $\mathbf{w}(t_d+i',t_d)$ is uncorrelated with $\mathbf{w}(t_d+j',t_d)$ for $i' \ne j'$, where $t_d \ge i,j \ge 0$ and $d \ge i',j' > 0$. Hence, $\{\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d);\mathbf{w}(t_d+1,t_d),\dots\mathbf{w}(t,t_d)\}$ is an innovation sequence. This completes the proof of the lemma.

The white noise sequence, $\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d),\mathbf{w}(t_d+1,t_d),\dots\mathbf{w}(t,t_d)$, calculated from the reorganized observations $\mathbf{y}_f(0),\dots,\mathbf{y}_f(t_d),\mathbf{y}(t_d+1),\dots,\mathbf{y}(t)$, is termed as the *reorganized innovation sequence*. Similarly, for any $s > t_d$, the sequence $\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d);\ \mathbf{w}(t_d+1,t_d),\dots\mathbf{w}(s,t_d)$ is also termed as the *reorganized innovation sequence*, and spans the same linear space as $\mathcal{L}\{\mathbf{y}_f(0),\dots,\mathbf{y}_f(t_d);\mathbf{y}(t_d+1),\dots,\mathbf{y}(s)\}$. The reorganized innovation sequence will play a key role in deriving the optimal estimator in this note.

### B. Riccati Equation

*Definition 2.1:*

$$\mathcal{P}^{t_d+i}_{t_d+i,t_d} \triangleq \mathcal{E}\left[\mathbf{e}(t_d+i,t_d)\mathbf{e}^T(t_d+i,t_d)\right];\ i>0 \tag{2.13}$$

is termed as the covariance matrix of the estimation error $\mathbf{e}(t_d+i,t_d)$.

From (2.8) and (2.9), the innovation covariance matrix

$$Q_w(t_d+i,t_d) \triangleq \mathcal{E}\left[\mathbf{w}(t_d+i,t_d)\mathbf{w}^T(t_d+i,t_d)\right],\ i \ge 0$$

is given by

$$Q_w(t_d+i,t_d) = \begin{cases} H_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}H^T_{t_d+i} + Q_v(t_d+i), & i>0 \\ \begin{bmatrix} H_{t_d} \\ L_{t_d}\end{bmatrix}\mathcal{P}^{t_d}_{t_d,t_d-1}\begin{bmatrix} H_{t_d} \\ L_{t_d}\end{bmatrix}^T + Q_{v_f}(t_d), & i=0. \end{cases} \tag{2.14}$$

We have the following results.

*Theorem 2.1:* The cross-covariance matrix $\mathcal{P}^{t_d+i}_{t_d+i,t_d}$ can be calculated as follows.

- For $i=1$, $\mathcal{P}^{t_d+1}_{t_d+1,t_d}$ is given by the following standard RDE:

$$\begin{aligned}\mathcal{P}^{t_d+1}_{t_d+1,t_d} &= \Phi_{t_d}\mathcal{P}^{t_d}_{t_d,t_d-1}\Phi^T_{t_d} - \Phi_{t_d}\mathcal{P}^{t_d}_{t_d,t_d-1}\begin{bmatrix} H_{t_d} \\ L_{t_d}\end{bmatrix}^T \\ &\quad \times Q_w^{-1}(t_d,t_d)\begin{bmatrix} H_{t_d} \\ L_{t_d}\end{bmatrix}\mathcal{P}^{t_d}_{t_d,t_d-1}\Phi^T_{t_d} + \Gamma_{t_d}Q_u(t_d)\Gamma^T_{t_d} \\ \mathcal{P}^0_{0,-1} &= P_0 \end{aligned} \tag{2.15}$$

where

$$Q_w(t_d,t_d) = \begin{bmatrix} H_{t_d} \\ L_{t_d}\end{bmatrix}\mathcal{P}^{t_d}_{t_d,t_d-1}\begin{bmatrix} H_{t_d} \\ L_{t_d}\end{bmatrix}^T + Q_{v_f}(t_d). \tag{2.16}$$

- For $i>1$, $\mathcal{P}^{t_d+i}_{t_d+i,t_d}$ is given by

$$\begin{aligned}\mathcal{P}^{t_d+i+1}_{t_d+i+1,t_d} &= \Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}\Phi^T_{t_d+i} - \Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d} \\ &\quad \times H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)H_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}\Phi^T_{t_d+i} \\ &\quad + \Gamma_{t_d+i}Q_u(t_d+i)\Gamma^T_{t_d+i}\end{aligned} \tag{2.17}$$

where the initial condition $\mathcal{P}^{t_d+1}_{t_d+1,t_d}$ is from (2.15) and

$$Q_w(t_d+i,t_d) = H_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}H^T_{t_d+i} + Q_v(t_d+i). \tag{2.18}$$

*Proof:* For $i=1$, it is obvious that $\mathcal{P}^{t_d+1}_{t_d+1,t_d}$ is the covariance matrix of the one-step-ahead prediction error of the state $\mathbf{x}(t_d+1)$ associated with (1.1) and (2.4). Thus, following the standard Kalman filtering theory, $\mathcal{P}^{t_d+1}_{t_d+1,t_d}$ satisfies the Riccati equation (2.15).

For $i>1$, note that $\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d)$ is the projection of the state $\mathbf{x}(t_d+i+1)$ onto the linear space spanned by $\{\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d),\dots\mathbf{w}(t_d+1,t_d),\dots\mathbf{w}(t_d+i,t_d)\}$. Since $\mathbf{w}(\cdot,\cdot)$ is a white noise, the estimator $\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d)$ is calculated by using the projection formula as

$$\begin{aligned}&\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d) \\ &\quad = \mathrm{Proj}\{\mathbf{x}(t_d+i+1|\mathbf{w}(0,0),\dots\mathbf{w}(t_d,t_d), \\ &\qquad\qquad \mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t_d+i-1,t_d)\} \\ &\qquad + \mathrm{Proj}\{\mathbf{x}(t_d+i+1|\mathbf{w}(t_d+i,t_d)\} \\ &\quad = \Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d) \\ &\qquad + \Phi_{t_d+i}\mathcal{E}\left[\mathbf{x}(t_d+i)\mathbf{e}^T(t_d+i,t_d)\right] \\ &\qquad \times H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)\mathbf{w}(t_d+i,t_d) \\ &\quad = \Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d) + \Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d} \\ &\qquad \times H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)\mathbf{w}(t_d+i,t_d).\end{aligned} \tag{2.19}$$

It is readily obtained from (1.1) and (2.19) that

$$\begin{aligned}\mathbf{e}(t_d+i+1,t_d) &= \mathbf{x}(t_d+i+1)-\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d)\\ &= \Phi_{t_d+i}\mathbf{e}(t_d+i,t_d)+\Gamma_{t_d+i}\mathbf{u}(t_d+i)-\Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}\\ &\quad\times H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)\mathbf{w}(t_d+i,t_d). \end{aligned}\tag{2.20}$$

Since $\mathbf{e}(t_d+i+1,t_d)$ is uncorrelated with $\mathbf{w}(t_d+i,t_d)$ and so is $\mathbf{e}(t_d+i,t_d)$ with $\mathbf{u}(t_d+i)$, it follows from the aforementioned equation that

$$\begin{aligned}&\mathcal{P}^{t_d+i+1}_{t_d+i+1,t_d}+\Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)\\ &\quad\times H_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}\Phi^T_{t_d+i}\\ &\qquad=\Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}\Phi^T_{t_d+i}+\Gamma_{t_d+i}Q_u(t_d+i)\Gamma^T_{t_d+i}\end{aligned}\tag{2.21}$$

which is (2.17).

*Remark 2.1:* Observe that (2.15) is the standard RDE associated with the Kalman filtering for (1.1) and (2.4), and (2.17) is the RDE for (1.1) and (1.2).

## III. Optimal Estimate $\hat{\mathbf{x}}(t|t)$

In this section, we will give the solution to the optimal filtering problem.

Based on the discussion in the previous section, the following results are obtained by applying the reorganized innovation sequence.

*Theorem 3.1:* Consider the system (1.1)–(1.3). Given $d>0$, the optimal filter $\hat{\mathbf{x}}(t|t)=\hat{\mathbf{x}}(t|t,t_d)$ is given by

$$\hat{\mathbf{x}}(t|t,t_d)=\hat{\mathbf{x}}(t|t-1,t_d)+\mathcal{P}^t_{t,t_d}H^T_tQ_w^{-1}\times(t,t_d)\left[y(t)-H_t\hat{\mathbf{x}}(t|t-1,t_d)\right]\tag{3.1}$$

where $\hat{\mathbf{x}}(t|t-1,t_d)$ is calculated recursively as

$$\begin{aligned}\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d) &=\Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d)\\ &\quad+\Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)\\ &\quad\times\left[\mathbf{y}(t_d+i)-H_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d)\right],\\ &\qquad i=1,\dots,d-1\end{aligned}\tag{3.2}$$

while $Q_w(t_d+i,t_d)=H_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}H^T_{t_d+i}+Q_v(t_d+i)$ and $\mathcal{P}^{t_d+i}_{t_d+i,t_d}(i=2,\dots,d)$ is computed by (2.17). The initial value $\hat{\mathbf{x}}(t_d+1|t_d,t_d)$ in (3.2) can be computed from the recursion

$$\begin{aligned}\hat{\mathbf{x}}(t_d+1|t_d,t_d) &=\Phi_{t_d}\hat{\mathbf{x}}(t_d|t_d-1,t_d-1)\\ &\quad+\Phi_{t_d}\mathcal{P}^{t_d}_{t_d,t_d-1}\begin{bmatrix}H_{t_d}\\ L_{t_d}\end{bmatrix}^TQ_w^{-1}(t_d,t_d)\\ &\quad\times\left[\mathbf{y}_f(t_d)-\begin{bmatrix}H_{t_d}\\ L_{t_d}\end{bmatrix}\hat{\mathbf{x}}(t_d|t_d-1,t_d-1)\right]\\ \hat{\mathbf{x}}(0|-1,-1)&=0\end{aligned}\tag{3.3}$$

where $Q_w(t_d,t_d)=\left[\begin{smallmatrix}H_{t_d}\\ L_{t_d}\end{smallmatrix}\right]\mathcal{P}^{t_d}_{t_d,t_d-1}\left[\begin{smallmatrix}H_{t_d}\\ L_{t_d}\end{smallmatrix}\right]^T+Q_{v_f}(t_d)$ and $\mathcal{P}^{t_d}_{t_d,t_d-1}$ is as in (2.15).

*Proof:* By applying Lemma 2.1, $\hat{\mathbf{x}}(t|t)=\hat{\mathbf{x}}(t|t,t_d)$ is the projection of the state $\mathbf{x}(t)$ onto the linear space spanned by $\{\mathbf{w}(0,0),\dots\mathbf{w}(t_d,t_d),\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t,t_d)\}$. Since $\mathbf{w}(\cdot,\cdot)$ is a white noise, the filter $\hat{\mathbf{x}}(t|t,t_d)$ is calculated by using the projection formula as

$$\begin{aligned}\hat{\mathbf{x}}(t|t,t_d) &=\mathrm{Proj}\left\{\mathbf{x}(t)|\mathbf{w}(0,0),\dots\mathbf{w}(t_d,t_d),\right.\\ &\qquad\left.\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t-1,t_d)\right\}\\ &\quad+\mathrm{Proj}\left\{\mathbf{x}(t)|\mathbf{w}(t,t_d)\right\}\\ &=\hat{\mathbf{x}}(t|t-1,t_d)+\mathcal{E}\left[\mathbf{x}(t)\mathbf{w}^T(t,t_d)\right]\\ &\quad\times Q_w^{-1}(t,t_d)\mathbf{w}(t,t_d)\\ &=\hat{\mathbf{x}}(t|t-1,t_d)+\mathcal{P}^t_{t,t_d}H^T_tQ_w^{-1}(t,t_d)\\ &\quad\times\left[y(t)-H_t\hat{\mathbf{x}}(t|t-1,t_d)\right]\end{aligned}\tag{3.4}$$

which is (3.1). Similarly, from Lemma 2.1, $\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d)$ $(i>0)$ is the projection of the state $\mathbf{x}(t_d+i+1)$ onto the linear space spanned by the innonvation $\{\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d),\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t_d+i,t_d)\}$, it follows from the projection formula that

$$\begin{aligned}&\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d)\\ &\quad=\mathrm{Proj}\left\{\mathbf{x}(t_d+i+1)|\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d),\right.\\ &\qquad\quad\left.\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t_d+i,t_d)\right\}\\ &\quad=\Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i)|t_d+i,t_d)+\Gamma_{t_d+i}\\ &\qquad\times\mathrm{Proj}\left\{\mathbf{u}(t_d+i)|\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d),\right.\\ &\qquad\quad\left.\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t_d+i,t_d)\right\}.\end{aligned}\tag{3.5}$$

Noting that $\mathbf{u}(t_d+i)$ is uncorrelated with the innovation $\mathbf{w}(0,0),\dots,\mathbf{w}(t_d,t_d),\mathbf{w}(t_d+1,t_d),\dots,\mathbf{w}(t_d+i,t_d)$, we have

$$\begin{aligned}&\hat{\mathbf{x}}(t_d+i+1|t_d+i,t_d)\\ &\quad=\Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i,t_d)\\ &\quad=\Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d)\\ &\qquad+\Phi_{t_d+i}\mathcal{E}\left[\mathbf{x}(t_d+i)\mathbf{w}^T(t_d+i,t_d)\right]\\ &\qquad\times Q_w^{-1}(t_d+i,t_d)\mathbf{w}(t_d+i,t_d)\\ &\quad=\Phi_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d)\\ &\qquad+\Phi_{t_d+i}\mathcal{P}^{t_d+i}_{t_d+i,t_d}H^T_{t_d+i}Q_w^{-1}(t_d+i,t_d)\\ &\qquad\times\left[\mathbf{y}(t_d+i)-H_{t_d+i}\hat{\mathbf{x}}(t_d+i|t_d+i-1,t_d)\right]\end{aligned}\tag{3.6}$$

which is (3.2). Furthermore, $\hat{\mathbf{x}}(t_d+1|t_d,t_d)$ is the standard Kalman filter of (1.1) and (2.4), which is obviously given by (3.3).

*Remark 3.1:* The Kalman filtering solution for (1.1)–(1.3) with delayed measurement has been given by applying the reorganized innovation analysis. Different from the standard Kalman filtering approach, our approach consists of two parts. The first is (3.2) and (2.17), which is the Kalman formulation for the system (1.1)–(1.2). The second part is (3.3) and (2.15), which is the Kalman formulation for system (1.1) and (2.4). Observe that the solution only relies on two Riccati recursions of dimension $n\times n$. This is in comparison with the traditional augmentation method where one Riccati equation of dimension $(n+d\times p)\times(n+d\times p)$ is involved. In the following section, we will demonstrate that the proposed method indeed possesses computational advantages over the latter.

*Remark 3.2:* The aforementioned reorganized innovation analysis in Hilbert space can be extended to Krein space to address the more complicated $H_\infty$ fixed-lag smoothing [4] and $H_\infty$ estimation problem for time-delay systems [5].

## IV. Discussions and Comparison

The purpose of the section is to compare the computational cost of the presented approach and the traditional augmentation method. As additions are much faster than multiplications and divisions, it is the number of multiplications and divisions that is used as the operation count. Let $MD$ denote the number of multiplications and divisions.

First, note that the algorithm by Theorem 3.1 can be summarized as follows:

i) compute matrix $\mathcal{P}_{t_d+1,t_d}^{t_d+1}$ using the RDE (2.15);
ii) compute $\mathcal{P}_{t_d+i,t_d}^{t_d+i}$, $i = 2, \ldots, d$ using (2.17);
iii) compute $\hat{\mathbf{x}}(t|t, t_d)$ using (3.1)–(3.3).

It is easy to know that the total $MD$ number of obtaining $\hat{\mathbf{x}}(t|t, t_d)$ for one step, denoted as $MD_{\text{new}}$, is given by

$$\begin{aligned} MD_{\text{new}} = & \left[3n^3+(3m+r)n^2+2m^2n+m^3\right]d+(6p+1)n^2 \\ & +\left[2(m+p)^2-2m^2+4mp+2p^2+m+p\right]n \\ & +2(m+p)^3+(m+p)^2-2m^3. \end{aligned} \tag{4.1}$$

On the other hand, recall the Kalman filtering for the augmented state-space model (1.8)–(1.12). The optimal filter $\hat{\mathbf{x}}_a(t|t)$ is computed by

$$\begin{aligned} \hat{\mathbf{x}}_a(t+1|t+1) = & \Phi_a(t)\hat{\mathbf{x}}_a(t|t) + P_a(t)H_{L_a}^T(t)Q_w^{-1}(t) \\ & \times \left\{ \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{z}_{t-d}(t) \end{bmatrix} - H_{L_a}(t+1)\Phi_a(t)\hat{\mathbf{x}}_a(t|t) \right\} \end{aligned} \tag{4.2}$$

where the matrix $P_a(t)$ satisfies the following RDE:

$$\begin{aligned} P_a(t+1) = & \Phi_a(t)P_a(t)\Phi_a^T(t) - \Phi_a(t)P_a(t)H_{L_a}^T(t)Q_w^{-1}(t) \\ & \times H_{L_a}(t)P_a(t)\Phi_a^T(t) + \Gamma_a(t)Q_u(t)\Gamma_a^T(t) \end{aligned} \tag{4.3}$$

with $Q_w(t) = H_{L_a}(t)P_a(t)H_{L_a}^T(t) + Q_{v_s}(t)$. In view of the special structure of the matrices $\Phi_a(t)$, $\Gamma_a(t)$, $H_a(t)$ and $L_a(t)$, the calculation burden for the RDE (4.3) can be reduced. We partition $P_a(t)$ as

$$P_a(t) = \{P_{a,ij}(t), 1 \le i \le d+1, 1 \le j \le d+1\}$$

where the dimension of $P_{a,11}(t)$ is $n \times n$, the dimension of $P_{a,ii}(t)$, $i > 1$, is $p \times p$, and the other blocks are of corresponding dimensions. Also, introduce a similar partition for $\Pi_a(t)$. Then, the RDE (4.3) is simplified as [14]

$$\begin{aligned} \Pi_a(t) = & P_a(t) - P_{a,1}(t)H_t^T\left[Q_v(t)+H_tP_{a,11}(t)H_t^T\right]^{-1} \\ & \times H_tP_{a,1}^T(t) \end{aligned} \tag{4.4}$$

$$\begin{aligned} \Sigma_a(t) = & \Pi_a(t) - \Pi_{a,d+1}(t)\left[Q_{v_z}(t)+\Pi_{a,(d+1)(d+1)}(t)\right]^{-1} \\ & \times \Pi_{a,d+1}^T(t) \end{aligned} \tag{4.5}$$

$$P_a(t+1) = \Gamma_a(t)Q_u(t)\Gamma_a^T(t)+\Phi_a(t)\Sigma_a(t)\Phi_a^T(t) \tag{4.6}$$

where $P_{a,i}(t)$ and $\Pi_{a,i}(t)$ represent the $i$th column blocks of $P_a(t)$ and $\Pi_a(t)$, respectively. Suppose that $\Sigma_a(t)$ is partitioned similarly to $P_a(t)$ and $\Pi_a(t)$. By taking into account the structure of the matrices $\Phi_a(t)$ and $\Gamma_a(t)$, the operation number of calculating $\hat{\mathbf{x}}_a(t+1|t+1)$ by the previous ormula, which is denoted as $MD_{\text{aug}}$, is given as

$$\begin{aligned} MD_{\text{aug}} = & p^2(n+p)d^2 + \left[4pn^2 + (4p^2+3mp)n \right. \\ & \left. + (m^2+m+p^2+p)p\right]d \\ & + 4n^3 + (4m+3p+1)n^2 \\ & + (2m^2+2m-p^2+2p)n \\ & + m^3 + m^2 + p^3. \end{aligned} \tag{4.7}$$

From (4.1) and (4.7), it is clear that $MD_{\text{aug}}$ is of magnitude $O(d^2)$ whereas $MD_{\text{new}}$ is linear in $d$. Thus, when the delay $d$ is sufficiently large, it is easy to know that $MD_{\text{aug}} > MD_{\text{new}}$. Moreover, the larger the $d$, the larger the ratio $MD_{\text{aug}}/MD_{\text{new}}$. To see this, we consider one example.

*Example 5.1:* Consider the system (1.1)–(1.3), with $n = 3, m = 1$, $r = 1$, and $p = 3$. The $MD$ numbers of the proposed approach and the system augmentation approach are compared in the following table for various values of $d$:

| $d$ | 1 | 2 | 3 | 6 | 12 |
|---|---|---|---|---|---|
| $MD_{\text{new}}$ | 629 | 753 | 877 | 1249 | 1993 |
| $MD_{\text{aug}}$ | 605 | 1052 | 1607 | 3920 | 11 462 |
| $\frac{MD_{\text{aug}}}{MD_{\text{new}}}$ | 0.9618 | 1.3971 | 1.8324 | 3.1385 | 5.7511. |

## V. Conclusion

In this note, we have revisited the $H_2$ estimation problem for discrete-time systems with instantaneous and delayed measurements by a reorganized innovation analysis. Our contributions are twofold. First, the presented approach simplifies the calculation of the estimator as compared with the traditional system augmentation approach. Second, the new concept of reorganized innovation can be extended to solving the $H_\infty$ estimation, particularly the long standing $H_\infty$ fixed-lag smoothing and the $H_\infty$ estimation for both the discrete and continuous time-delay systems [4], [5].

## References


[1] L. A. Klein, *Sensor and Data Fusion Concepts and Applications*. Bellingham, WA: SPIE Press, 1999.
[2] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 1999.
[3] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng., Trans. ASME-D*, vol. 82, no. 1, pp. 35–45, 1960.
[4] H. Zhang, L. Xie, and Y. C. Soh, "A unified approach to linear estimation for discrete-time systems—Part II: $\boldsymbol{H}_\infty$ estimation," in *Proc. IEEE Conf. Decision Control*, Dec. 2001, pp. 2923–2928.
[5] H. Zhang, D. Zhang, and L. Xie, "$\boldsymbol{H}_\infty$ fixed-lag smoothing and prediction for linear continous-time systems," in *Proc. 2003 Amer. Control Conf.*, pp. 4201–4206.
[6] H. Kwakernaak, "Optimal filtering in linear systems with time delays," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 169–173, Feb. 1967.
[7] M. Briggs, "Linear filtering for time-delay systems," *IMA J. Math. Control Inform.*, vol. 6, pp. 167–178, 1989.
[8] J. S. Meditch, *Stochastic Optimal Linear Estimation and Control*. New York: McGraw-Hill, 1969.
[9] L. Chisci and E. Mosca, "Polynomial equations for the linear MMSE state estimation," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 623–626, May 1992.
[10] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. Upper Saddle River, NJ: Prentice-Hall, 1984.
[11] V. Kucera, *Discrete Linear Control. The Polynomial Equation Approach*. New York: Wiley, 1979.
[12] ——, "New results in state estimation and regulation," *Automatica*, vol. 17, pp. 745–748, Sept. 1981.
[13] J. M. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach*. New York: Academic, 1983.
[14] Y. Theodor and U. Shaked, "Game theory approach to $\boldsymbol{H}_\infty$ optimal discrete-time fixed-point and fixed-lag smoothing," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1944–1948, Sept. 1994.
[15] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of networked control systems," *IEEE Contr. Syst. Mag.*, pp. 84–99, Feb. 2001.
[16] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Upper Saddle River, NJ: Prentice-Hall, 1979.