

# The Discrete Kalman Filtering of a Class of Dynamic Multiscale Systems

Lei Zhang, Quan Pan, Paul Bao, and Hongcai Zhang

**Abstract**—This paper discusses the optimal estimation of a class of dynamic multiscale systems (DMS), which are observed by several sensors at different scales. The resolution and sampling frequencies of the sensors are supposed to decrease by a factor of two. By using the Haar wavelet transform to link the state nodes at each of the scales within a time block, we generalize the DMS into the standard state-space model, for which the Kalman filtering can be employed as the optimal estimation algorithm. The stochastic controllability and observability of time invariant DMS are analyzed and the stability of the Kalman filter is then discussed. Despite that the DMS model maybe become incompletely controllable and observable, it is proved that as long as the DMS is completely controllable and observable at the finest scale, the associated Kalman filter will be asymptotically stable. The scheme is illustrated with a two-scale Markov process.

**Index Terms**—Dynamic multiscale system (DMS), Kalman filtering, optimal estimation, wavelet transform.

## I. INTRODUCTION

IN THE LAST two decades, the multiscale autoregressive (MAR) framework [1]–[6] was developed to model a variety of random processes compactly and estimate them efficiently. The MAR was first motivated by Basseville *et al.* [1], and based on their works Chou *et al.* [2], [3] proposed the multiscale stochastic models and optimal estimation algorithms for a rich class of processes whitened by wavelet transforms (WT) [9]–[13]. Luetgten [5] and Frakt [6] contributed a lot to the stochastic realization of the MAR. MAR estimates the stationary processes from numerous measurements by multiscale techniques, which aim at saving much computation compared with the traditional linear minimum mean-square-error estimation (LMMSE).

In this paper, we aim at obtaining the real-time optimal estimation of a class of dynamic multiscale systems (DMS), which are observed by  $J$  sensors independently with different resolution. Then  $J$  sets of measurements are obtained. Since the observed target is the same, the measurements of the different sensors are correlated and they can be fused to estimate the state of the DMS optimally.

Manuscript received September 23, 2001; revised November 15, 2002. This paper was recommended by Associate Editor B. Baykal.

L. Zhang is with the Computing Department, The Hong Kong Polytechnic University, Hung Hum, Kowloon, Hong Kong. (E-mail: cslei@inet.polyu.edu.hk).

Q. Pan and H. Zhang are with the Automatic Control Department, Northwestern Polytechnic University, Xi'an 710072, China.

P. Bao is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, New Territory, Hong Kong.

Digital Object Identifier 10.1109/TCSII.2002.807573

Order the sensors by their resolution from 1 to  $J$  and suppose sensor 1 is of the highest resolution. We have the following equations to characterize the system:

$$\mathbf{x}_1(k_1) = A(k_1)\mathbf{x}_1(k_1) + B(k_1)\mathbf{w}(k_1) \quad (1-1)$$

$$\mathbf{z}_j(k_j) = C_j(k_j)\mathbf{x}_j(k_j) + \mathbf{v}_j(k_j), j = 1, 2, \dots, J \quad (1-2)$$

where  $\mathbf{x}_j(k_j) \in \mathfrak{R}^n \times 1$  is the state vector to be estimated and  $\mathbf{z}_j(k_j)$  is its measurement.  $k_j$  denotes the sampling time at scale  $j$ .  $A(k_1)$ ,  $B(k_1)$ , and  $C(k_j)$  are the system, input and measurement matrices.  $\mathbf{w}(k_1)$  and  $\mathbf{v}(k_j)$  are individually independent. They are Gaussian white processes with zero mean and covariance matrices  $Q(k_1)$  and  $R_j(k_j)$ , respectively. State  $\mathbf{x}_j$  belongs to a subspace of  $\mathbf{x}_1$ . It is determined by the resolution of sensor  $j$ . For the mentioned DMS, the well-known Kalman filtering [15], [16] could not be employed directly as the LMMSE algorithm.

An important case of the DMS is when the sampling frequencies of the sensors decrease by a factor of two from 1 to  $J$ . Fig. 1 illustrates the tree structure of the state nodes for such a DMS. In each time block  $\Delta T \cdot k$ , there are  $2^{J-1}$  state nodes at scale 1,  $2^{J-2}$  nodes at scale 2, ..., and sequentially only 1 node at scale  $J$ . The real-time optimal estimation of the state nodes should exploit all the measurements at each scale up to time  $\Delta T \cdot k$ .

Hong [7] presented a multiresolution-filtering scheme for such DMSs. WT is used to link the state nodes at different scales. See Fig. 2, node  $\mathbf{x}_j(n)$  is the low-pass output of  $\mathbf{x}_{j-1}(n)$  and the lost detail information is preserved in wavelet coefficient  $\mathbf{y}_j(n)$ . Hong first estimated  $\mathbf{x}_1(n)$  with the measurements at scale 1 within  $\Delta T \cdot k$ , then he wavelet decomposed the estimation to scale  $2 \sim J$  as the prediction of  $\mathbf{x}_j(n)$ . The updating was performed at each scale by the local measurements. At last, the local updated estimations were inversely transformed to scale 1 and fused together. It should be noted that in Hong's algorithm, the updating is imposed only on  $\mathbf{x}_j(n)$  but not on  $\mathbf{y}_j(n)$ , which is correlated with state  $\mathbf{x}_j(n)$  and measurement  $\mathbf{z}_j(n)$ .  $\mathbf{y}_j(n)$  should also be updated and it would make contributions to the estimation of  $\mathbf{x}_j(n)$  through the inverse WT. It can be seen that Hong's multiresolution-filtering is not an optimal estimation.

In [8], Hong *et al.* developed a multiscale Kalman filtering technique for standard state-space model, which is a special case of the DMS when only one sensor is working. Hong *et al.* decomposed the random signal at several scales and reformed the state-space model by state augmentation. The Kalman filtering of the new model is declared to yields better results than the standard Kalman filtering. But it should be noticed that the

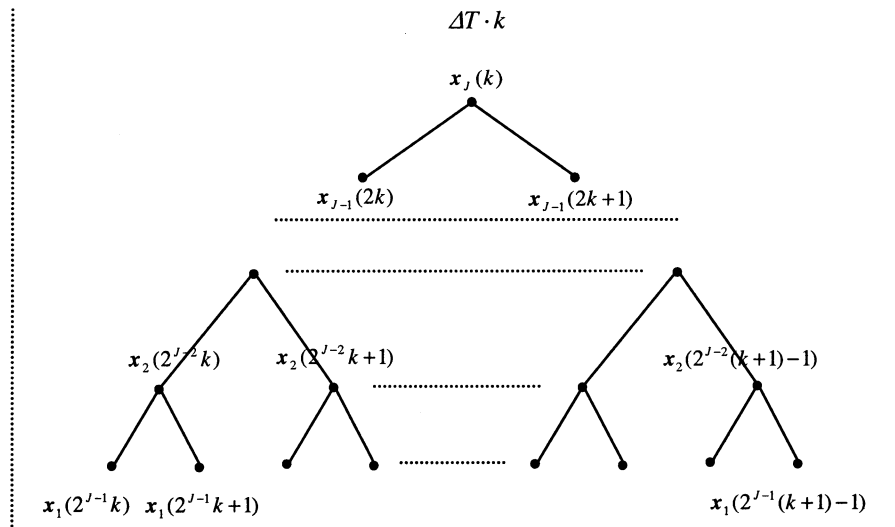
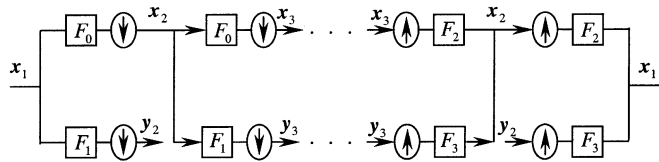


Fig. 1. Tree structure of the DMS state nodes within a time block.


 Fig. 2. Implementation of the wavelet transforms to the DMS state nodes. The downward arrows denote the subsampling and the upward arrows denote the upsampling.  $F_0$  and  $F_1$  are the analytic wavelet filters while  $F_2$  and  $F_3$  being the synthetic wavelet filters.

comparison is unfair. The Kalman filtering of the new model uses some measurements after time  $k$  to estimate the state  $\mathbf{x}_k$ . If we perform the associated Kalman smoothing with the standard state-space model, the same results will be gotten with much less computation than Hong's scheme.

In this paper, an optimal estimation algorithm for the mentioned DMS is presented. We employ Haar wavelet transform to represent the state projection across scales within a time block and generalize the DMS into the standard state-space model. Then the classical Kalman filtering would naturally be the LMMSE algorithm. The stochastic controllability and observability of the DMS and the stability of its associated Kalman filter are discussed in the time invariant case.

Section II focuses on the modeling of the DMS by Haar wavelet transform. The stochastic controllability and observability of time invariant DMS is analyzed in Section III and the stability of the Kalman filter is discussed in Section IV. Section V presents some illustrative examples and Section VI is the discussion and conclusion.

## II. DMS MODELING

Refer to Fig. 1, suppose state vector  $\mathbf{x} \in U_j$  and  $U_j \subset U_{j-1} \subset \dots \subset U_1$  is a closed subspace sequence. If  $\mathbf{x}_j$  is the linear projection of  $\mathbf{x}_{j-1}$  from space  $U_{j-1}$  to  $U_j$ :  $\mathbf{x}_j = P_{j-1,j} \cdot \mathbf{x}_{j-1}$ , where  $P_{j-1,j}$  is the linear projection operator, we have

$$\mathbf{x}_j = P_j \cdot \mathbf{x}_1 \quad (2-1)$$

where  $P_j = P_{j-1,j} \cdot P_{j-2,j-1} \cdot \dots \cdot P_{1,2}$ ,  $j = 2, 3, \dots, J$  and  $P_1$  is the identity operator  $I$ .

We want to implement the state space projection from the finest scale to other scales within a time block  $k \cdot \Delta T$ . The Haar wavelet whose lowpass filter has only two taps is a natural choice to approximate the linear state projection. Refer to Fig. 1, recursively there is

$$\mathbf{x}_j(k_j) = \left( \frac{\sqrt{2}}{2} \right)^{j-1} \sum_{i=0}^{2^{j-1}-1} \mathbf{x}_1(2^{j-1}k_j + i) \quad (2-2)$$

where  $k_j = 2^{J-j}k, 2^{J-j}k + 1, \dots, 2^{J-j}(k+1) - 1$ . At the coarsest scale  $J$ , the node  $\mathbf{x}_J(k)$  is the linear combination of all the nodes at scale 1 within time block  $\Delta T \cdot k$ . Define (2-3)–(2-4) at the bottom of the page where  $m_j = 0, 1, \dots, 2^{J-j} - 1$  and  $I_{n_{\mathbf{x}}}$  is the identity  $n_{\mathbf{x}} \times n_{\mathbf{x}}$  matrix. We have

$$\mathbf{x}_j(2^{J-j}k + m_j) = M_j(m_j) \cdot \bar{\mathbf{x}}(k). \quad (2-5)$$

From (1-2) there is

$$\mathbf{z}_j(2^{J-j}k + m_j) = C_j(2^{J-j}k + m_j) M_j(m_j) \bar{\mathbf{x}}(k) + \mathbf{v}_j(2^{J-j}k + m_j). \quad (2-6)$$

$$\bar{\mathbf{x}}(k) = \text{col} \{ \mathbf{x}_1(2^{J-1}k), \mathbf{x}_1(2^{J-1}k + 1), \dots, \mathbf{x}_1(2^{J-1}(k+1) - 1) \} \quad (2-3)$$

$$M_j(m_j) = \begin{bmatrix} \underbrace{0 \cdot I_{n_{\mathbf{x}}}, \dots, 0 \cdot I_{n_{\mathbf{x}}}}_{2^{j-1}m_j} & \underbrace{\left( \frac{\sqrt{2}}{2} \right)^{j-1} \cdot I_{n_{\mathbf{x}}}, \dots, \left( \frac{\sqrt{2}}{2} \right)^{j-1} \cdot I_{n_{\mathbf{x}}}}_{2^{j-1}} & \underbrace{0 \cdot I_{n_{\mathbf{x}}}, \dots, 0 \cdot I_{n_{\mathbf{x}}}}_{2^{j-1}(2^{J-j}-m_j-1)} \end{bmatrix}. \quad (2-4)$$

Denote

$$\bar{\mathbf{z}}_j(k) = \text{col} \{ \mathbf{z}_j(2^{J-j}k), \mathbf{z}_j(2^{J-j}k+1), \dots, \mathbf{z}_j(2^{J-j}(k+1)-1) \} \quad (2-7)$$

$$\bar{\mathbf{C}}_j(k) = \begin{bmatrix} C_j(2^{J-j}k)M_j(0) \\ C_j(2^{J-j}k+1)M_j(1) \\ \vdots \\ C_j(2^{J-j}(k+1)-1)M_j(2^{J-j}-1) \end{bmatrix} \quad (2-8)$$

$$\bar{\mathbf{v}}_j(k) = \text{col} \{ \mathbf{v}_j(2^{J-j}k), \mathbf{v}_j(2^{J-j}k+1), \dots, \mathbf{v}_j(2^{J-j}(k+1)-1) \}. \quad (2-9)$$

We have

$$\bar{\mathbf{z}}_j(k) = \bar{\mathbf{C}}_j(k)\bar{\mathbf{x}}(k) + \bar{\mathbf{v}}_j(k) \quad (2-10)$$

where the covariance of  $\bar{\mathbf{v}}_j(k)$  is shown in (2-11) at the bottom of the page. Denote

$$\bar{\mathbf{z}}(k) = \text{col} \{ \bar{\mathbf{z}}_J(k), \bar{\mathbf{z}}_{J-1}(k), \dots, \bar{\mathbf{z}}_1(k) \} \quad (2-12)$$

$$\bar{\mathbf{C}}(k) = \text{col} \{ \bar{\mathbf{C}}_J(k), \bar{\mathbf{C}}_{J-1}(k), \dots, \bar{\mathbf{C}}_1(k) \} \quad (2-13)$$

$$\bar{\mathbf{v}}(k) = \text{col} \{ \bar{\mathbf{v}}_J(k), \bar{\mathbf{v}}_{J-1}(k), \dots, \bar{\mathbf{v}}_1(k) \}. \quad (2-14)$$

Thus

$$\bar{\mathbf{z}}(k) = \bar{\mathbf{C}}(k)\bar{\mathbf{x}}(k) + \bar{\mathbf{v}}(k) \quad (2-15)$$

and the covariance of  $\bar{\mathbf{v}}(k)$  is

$$\bar{\mathbf{R}}(k) = \text{diag} [ \bar{\mathbf{R}}_J(k) \quad \bar{\mathbf{R}}_{J-1}(k) \quad \dots \quad \bar{\mathbf{R}}_1(k) ]. \quad (2-16)$$

Then (2-15) is the measurement equation of the DMS. The associated state-transition equation should be derived to complete the modeling. In time block  $\Delta T \cdot (k+1)$ , recursively by (1-1) shown in (2-17) at the bottom of the page. Denote (2-18)–(2-20) at the bottom of the page. Equation (2-17) can be rewritten as

$$\mathbf{x}_1(2^{J-1}(k+1) + m_1) = \bar{\mathbf{A}}(k, m_1)\mathbf{x}_1(2^{J-1}(k+1) - 1) + \bar{\mathbf{B}}(k, m_1)\bar{\mathbf{w}}(k). \quad (2-21)$$

Let

$$\bar{\mathbf{A}}(k) = \begin{bmatrix} O & \dots & O & \bar{\mathbf{A}}(k, 0) \\ O & \dots & O & \bar{\mathbf{A}}(k, 1) \\ \vdots & \ddots & \vdots & \vdots \\ O & \dots & O & \bar{\mathbf{A}}(k, 2^{J-1}-1) \end{bmatrix}$$

$$\bar{\mathbf{B}}(k) = \begin{bmatrix} \bar{\mathbf{B}}(k, 0) \\ \bar{\mathbf{B}}(k, 1) \\ \vdots \\ \bar{\mathbf{B}}(k, 2^{J-1}-1) \end{bmatrix}. \quad (2-22)$$

Then the state-transition equation of the DMS is

$$\bar{\mathbf{x}}(k+1) = \bar{\mathbf{A}}(k)\bar{\mathbf{x}}(k) + \bar{\mathbf{B}}(k)\bar{\mathbf{w}}(k) \quad (2-23)$$

where  $\bar{\mathbf{w}}(k)$  is a Gaussian white process uncorrelated with  $\bar{\mathbf{v}}(k)$  and its covariance matrix is shown in (2-24) at the bottom of the next page.

$$\bar{\mathbf{R}}_j(k) = \text{diag} [ R_j(2^{J-j}k) \quad R_j(2^{J-j}k+1) \quad \dots \quad R_j(2^{J-j}(k+1)-1) ]. \quad (2-11)$$

$$\mathbf{x}_1(2^{J-1}(k+1) + m_1) = \prod_{n=1}^{m_1+1} A(2^{J-1}(k+1) + m_1 - n) \bullet \mathbf{x}_1(2^{J-1}(k+1) - 1)$$

$$+ \sum_{i=-1}^{m_1-1} \prod_{n=1}^{m_1-i-1} A(2^{J-1}(k+1) + m_1 - n) \bullet B(2^{J-1}(k+1) + i) \mathbf{w}(2^{J-1}(k+1) + i). \quad (2-17)$$

$$\bar{\mathbf{w}}(k) = \text{col} \{ \mathbf{w}(2^{J-1}(k+1) - 1), \mathbf{w}(2^{J-1}(k+1)), \dots, \mathbf{w}(2^{J-1}(k+2) - 2) \} \quad (2-18)$$

$$\bar{\mathbf{A}}(k, m_1) = \prod_{n=1}^{m_1+1} A(2^{J-1}(k+1) + m_1 - n) \quad (2-19)$$

$$\bar{\mathbf{B}}(k, m_1) = \begin{bmatrix} \left( \prod_{n=1}^{m_1} A(2^{J-1}(k+1) + m_1 - n) \bullet B(2^{J-1}(k+1) - 1) \right)^T \\ \left( \prod_{n=1}^{m_1-1} A(2^{J-1}(k+1) + m_1 - n) \bullet B(2^{J-1}(k+1)) \right)^T \\ \vdots \\ (B(2^{J-1}(k+1) + m_1 - 1))^T \\ O \\ \vdots \\ O \end{bmatrix}^T. \quad (2-20)$$

Equations (2-23) and (2-15) make up the state-space model of the DMS

$$\begin{cases} \bar{\mathbf{x}}(k+1) = \bar{A}(k)\bar{\mathbf{x}}(k) + \bar{B}(k)\bar{\mathbf{w}}(k) \\ \bar{\mathbf{z}}(k) = \bar{C}(k)\bar{\mathbf{x}}(k) + \bar{\mathbf{v}}(k) \end{cases} \quad (2-25)$$

Obviously, the model meets the requirements of the standard Kalman filtering, which is then the LMMSE estimation algorithm of the DMS. Denote by  $\hat{\bar{\mathbf{x}}}(k)$  the Kalman filtering result of model (2-25).  $\hat{\bar{\mathbf{x}}}(k)$  consists of the LMMSE of those nodes at the finest scale. The LMMSE of the nodes at coarser scales can be directly obtained from  $\hat{\bar{\mathbf{x}}}(k)$ . We have the following theorem.

*Theorem 2-1:* Suppose  $\hat{\bar{\mathbf{x}}}(k)$  is the LMMSE of  $\bar{\mathbf{x}}(k)$ , then the LMMSE of node  $\mathbf{x}_j(2^{J-j}k + m_j)$  is  $M_j(m_j) \cdot \hat{\bar{\mathbf{x}}}(k)$ . ■

*Proof:* Denote  $\vec{\mathbf{z}}_k = \text{col}\{\bar{\mathbf{z}}(1), \bar{\mathbf{z}}(2), \dots, \bar{\mathbf{z}}(k)\}$ . Since  $\bar{\mathbf{x}}(k)$  and  $\vec{\mathbf{z}}_k$  are jointly Gaussian distributed, the LMMSE of  $\bar{\mathbf{x}}(k)$  conditioned on  $\vec{\mathbf{z}}_k$  can be represented as

$$\hat{\bar{\mathbf{x}}}(k) = E\left[\bar{\mathbf{x}}(k) \middle| \vec{\mathbf{z}}_k\right] = L \cdot \vec{\mathbf{z}}_k + \mathbf{b} \quad (2-26)$$

where  $L$  is a matrix and  $\mathbf{b}$  is a vector. Denote by  $\tilde{\bar{\mathbf{x}}}(k) = \bar{\mathbf{x}}(k) - \hat{\bar{\mathbf{x}}}(k)$  the estimation error. Since LMMSE is unbiased [15, p. 94, theorem 2.3] and according to the orthogonal projection theorem [15, p. 95, theorem 2.5], there are

$$E[\tilde{\bar{\mathbf{x}}}(k)] = \mathbf{0} \text{ and } E[\tilde{\bar{\mathbf{x}}}(k) \cdot \vec{\mathbf{z}}_k^T] = \mathbf{0}. \quad (2-27)$$

Let  $\hat{\mathbf{x}}_j(2^{J-j}k + m_j) = M_j(m_j) \cdot \hat{\bar{\mathbf{x}}}(k)$ , it is easy to obtain

$$\begin{aligned} \tilde{\mathbf{x}}_j(2^{J-j}k + m_j) &= \mathbf{x}_j(2^{J-j}k + m_j) - \hat{\mathbf{x}}_j(2^{J-j}k + m_j) \\ &= M_j(m_j) \cdot \tilde{\bar{\mathbf{x}}}(k). \end{aligned} \quad (2-28)$$

So, from (2-27) we have

$$E[\tilde{\mathbf{x}}_j(2^{J-j}k + m_j)] = M_j(m_j) \cdot E[\tilde{\bar{\mathbf{x}}}(k)] = \mathbf{0} \quad (2-29)$$

$$E[\tilde{\mathbf{x}}_j(2^{J-j}k + m_j) \cdot \vec{\mathbf{z}}_k^T] = M_j(m_j) \cdot E[\tilde{\bar{\mathbf{x}}}(k) \cdot \vec{\mathbf{z}}_k^T] = \mathbf{0}. \quad (2-30)$$

Thus, according to the orthogonal projection theorem, the LMMSE of  $\mathbf{x}_j(2^{J-j}k + m_j)$  is

$$E\left[\mathbf{x}_j(2^{J-j}k + m_j) \middle| \vec{\mathbf{z}}_k\right] = \hat{\mathbf{x}}_j(2^{J-j}k + m_j) M_j(m_j) \cdot \hat{\bar{\mathbf{x}}}(k) \quad (2-31)$$

In fact  $\hat{\mathbf{x}}_j(2^{J-j}k + m_j) = L' \cdot \vec{\mathbf{z}}_k + \mathbf{b}'$  where  $L' = M_j(m_j) \cdot L$  and  $\mathbf{b}' = M_j(m_j) \cdot \mathbf{b}$ . End of proof. ■

### III. STOCHASTIC CONTROLLABILITY AND OBSERVABILITY OF TIME INVARIANT DMS

The matrices  $\bar{A}(k)$ ,  $\bar{B}(k)$ , and  $\bar{C}(k)$  in model (2-25) contain many zero elements. In general, the stochastic controllability and observability of the DMS model would be lost. Here we discussed them for the time invariant DMS, where matrices  $A(\bullet)$ ,  $B(\bullet)$ , and  $C_j(\bullet)$  are constant and the covariance matrices of  $\mathbf{w}(\bullet)$  and  $\mathbf{v}_j(\bullet)$  are constants  $Q$  and  $R_j$ . The model of the time invariant DMS is

$$\begin{cases} \bar{\mathbf{x}}(k+1) = \bar{A} \cdot \bar{\mathbf{x}}(k) + \bar{B} \cdot \bar{\mathbf{w}}(k) \\ \bar{\mathbf{z}}(k) = \bar{C} \cdot \bar{\mathbf{x}}(k) + \bar{\mathbf{v}}(k) \end{cases} \quad (3-1)$$

with

$$\begin{aligned} \bar{A} &= \begin{bmatrix} O & \dots & O & A \\ O & \dots & O & A^2 \\ \vdots & \ddots & \vdots & \vdots \\ O & \dots & O & A^{2^{j-1}} \end{bmatrix} \\ \bar{B} &= \begin{bmatrix} B & O & \dots & O \\ AB & B & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ A^{2^{j-1}-1}B & A^{2^{j-1}-2}B & \dots & B \end{bmatrix} \\ \bar{C} &= \text{col}\{\bar{C}_J, \bar{C}_{J-1}, \dots, \bar{C}_1\}. \end{aligned} \quad (3-2)$$

The covariance matrices of  $\bar{\mathbf{w}}(\bullet)$  and  $\bar{\mathbf{v}}(\bullet)$  are constants  $\bar{Q}$  and  $\bar{R}$ .

#### A. Stochastic Controllability

Since  $Q$  is positive, there exists a matrix  $Q_1$  such that  $Q_1 \cdot Q_1^T = Q$ . So we have  $\bar{Q}_1 \cdot \bar{Q}_1^T = \bar{Q}$ , where  $\bar{Q}_1 = \text{diag}[Q_1 \ \dots \ Q_1]$ . Let  $\bar{B}_q = \bar{B} \cdot \bar{Q}_1$ , the controllability matrix of pair  $(\bar{A}, \bar{B}_q)$  is

$$\Omega = [\bar{B}, \bar{A}\bar{B}_q, \dots, \bar{A}^{n_{\bar{\mathbf{x}}}-1}\bar{B}_q]. \quad (3-3)$$

where  $n_{\bar{\mathbf{x}}} = 2^{J-1}n_{\mathbf{x}}$ . Pair  $(\bar{A}, \bar{B}_q)$  is completely controllable if and only if  $\text{rank}(\Omega) = n_{\bar{\mathbf{x}}}$  [16].

*Theorem 3-1:* Suppose at the finest scale pair  $(A, B_q)$  is completely controllable, where  $B_q = B \cdot Q_1$ . Then pair  $\bar{A}, \bar{B}_q$  is completely controllable if and only if matrix  $B$  is of full row rank. ■

*Proof:* From the special structure of  $\bar{A}$  we have (3-4) at the bottom of the page. It is well-known that  $(\bar{A}, \bar{B}_q)$  is completely

$$\bar{Q}(k) = \text{diag}[Q(2^{J-1}(k+1) - 1) \quad Q(2^{J-1}(k+1)) \quad \dots \quad Q(2^{J-1}(k+2) - 2)]. \quad (2-24)$$

$$\bar{A}^n \bar{B}_q = \begin{bmatrix} A^{n2^{j-1}}B_q & A^{n2^{j-1}-1}B_q & \dots & A^{(n-1)2^{j-1}+1}B_q \\ A^{n2^{j-1}+1}B_q & A^{n2^{j-1}}B_q & \dots & A^{(n-1)2^{j-1}+2}B_q \\ \vdots & \vdots & \ddots & \vdots \\ A^{(n+1)2^{j-1}-1}B_q & A^{(n+1)2^{j-1}-2}B_q & \dots & A^{n2^{j-1}}B_q \end{bmatrix}. \quad (3-4)$$

controllable if and only if  $\Omega$  is of full row rank. Let

$$P_1 = \begin{bmatrix} I_{n\mathbf{x}} & O & \cdots & O \\ -A & I_{n\mathbf{x}} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & -A & I_{n\mathbf{x}} \end{bmatrix}. \quad (3-5)$$

Obviously,  $P_1$  is nonsingular. Transforming  $\Omega$  by  $P_1$  shown in (3-6) at the bottom of the page. Since pair  $(A, B_q)$  is completely controllable, its controllability matrix  $[B, AB_q, \dots, A^{n_{\mathbf{x}}-1}B_q]$  is of full row rank. So,  $\Omega_1^u$  must have full row rank.  $\Omega_1^u$  and  $\Omega_1^d$  are linearly independent in row. It can be observed that  $\Omega_1$  is of full row rank if and only if  $B_q$  has full row rank. Because  $Q_1$  is nonsingular, it is equivalent that  $B$  is of full row rank. End of proof. ■

### B. Stochastic Observability

Since  $\bar{R}$  is positive, pair  $(\bar{C}, \bar{A})$  is completely observable if and only if its observability matrix

$$\Lambda = [\bar{C}^T, (\bar{C}A)^T, \dots, (\bar{C}A^{n_{\mathbf{x}}-1})^T]^T \quad (3-7)$$

is of full column rank  $\text{rank}(\Lambda) = n_{\mathbf{x}}$ [16]. We have the following theorem.

**Theorem 3-2:** Suppose at the finest scale pair  $(C_1, A)$  is completely observable, and then pair  $(\bar{C}, \bar{A})$  is completely observable if and only if

$$\begin{cases} \text{rank} \left( [C_2^T \ C_1^T]^T \right) = n_{\mathbf{x}}, & J = 2 \\ \text{rank} (C_1) = n_{\mathbf{x}}, & J > 2. \end{cases} \quad (3-8)$$

*Proof:* Denote  $\bar{A}^n = [O \ O \ \cdots \ O \ \bar{A}_{r,n}]$ , where  $\bar{A}_{r,n} = \text{col} \{ A^{(n-1)2^{J-1}+1}, A^{(n-1)2^{J-2}+1}, \dots, A^{n2^{J-1}} \}$ . We have  $\bar{C}\bar{A}^n = [O \ \cdots \ O \ \bar{C}\bar{A}_{r,n}]$ . Denote  $\bar{C} = \begin{bmatrix} \bar{C}^{lu} & \bar{C}^{ru} \\ O & C_1 \end{bmatrix}$ , there will be

$$\Lambda = \begin{bmatrix} \bar{C}^{lu} & \bar{C}^{ru} \\ O & C_1 \\ O & \bar{C}\bar{A}_{r,1} \\ \vdots & \vdots \\ O & \bar{C}\bar{A}_{r,n_{\mathbf{x}}-1} \end{bmatrix} = \begin{bmatrix} \bar{C}^{lu} & \bar{C}^{ru} \\ O & \Lambda^{rd} \end{bmatrix}. \quad (3-9)$$

Pair  $(\bar{C}, \bar{A})$  is completely observable if and only if  $\Lambda$  is of full column rank. Since  $(C_1, A)$  is completely observable, its observability matrix  $[C_1^T, (C_1A)^T, \dots, (C_1A^{n_{\mathbf{x}}-1})^T]^T$  has full

column rank. Each row of the matrix is included in  $\Lambda^{rd}$ , so  $\Lambda^{rd}$  is of full column rank if and only if  $\bar{C}^{lu}$  has full column rank.

If  $J = 2$ ,  $\bar{C}^{lu} = [\sqrt{2}/2C_2^T \ C_1^T]^T$  and it has full column rank when  $\text{rank} \left( [C_2^T \ C_1^T]^T \right) = n_{\mathbf{x}}$ . If  $J > 2$ ,  $\bar{C}^{lu}$  has  $(2^{J-1} - 1)n_{\mathbf{x}}$  columns in total. Divide it into  $(2^{J-1} - 1)$  blocks from left to right and each block has  $n_{\mathbf{x}}$  columns. Subtract the first block from the second block, and with some linear row transforming, the second block will be  $[O^T \ \cdots \ O^T \ O^T \ C_1^T \ O^T \ \cdots \ O^T]^T$ . To make  $\bar{C}^{lu}$  be of full rank,  $C_1$  must have full column rank. Since each block of  $\bar{C}^{lu}$  contains a  $C_1$  unit and those  $C_1$ 's are not overlapped in row,  $\bar{C}^{lu}$  has full column rank if and only if  $\text{rank}(C_1) = n_{\mathbf{x}}$ . End of proof. ■

### IV. STABILITY OF THE KALMAN FILTER

Although pair  $(\bar{C}, \bar{A}, \bar{B}_q)$  would be incompletely controllable and observable in general, it will be shown that as long as at the finest scale  $(C_1, A, B_q)$  is completely controllable and observable, the Kalman filter of pair  $(\bar{C}, \bar{A}, \bar{B}_q)$  will be asymptotically stable. Let us review a lemma and two definitions first [15].

**Lemma 4-1:** Suppose time invariant linear system is

$$\begin{cases} \mathbf{x}(k+1) = F \cdot \mathbf{x}(k) + \Gamma \cdot \mathbf{w}(k) \\ \mathbf{z}(k) = H \cdot \mathbf{x}(k) + \mathbf{v}(k) \end{cases} \quad (4-1)$$

where  $\mathbf{w}(\bullet)$  is Gaussian white process with zero mean and covariance  $Q$ . If pair  $(F, H)$  is completely detectable and pair  $(F, \Gamma \cdot Q_1)$  is completely stabilizable for any  $Q_1$  with  $Q_1 \cdot Q_1^T = Q$ , then the system's Kalman filter is asymptotically stable. ■

**Definition 4-1:** Pair  $(F, \Gamma)$  is completely stabilizable if there exists a nonsingular matrix  $T$  such that

$$TFT^{-1} = \begin{bmatrix} F_{11} & F_{12} \\ O & F_{22} \end{bmatrix}, \quad \Gamma T = \begin{bmatrix} \Gamma_1 \\ O \end{bmatrix} \quad (4-2)$$

with  $(F_{11}, \Gamma_1)$  is completely controllable and  $|\lambda_i(F_{22})| < 1$ . ■

**Definition 5-2:** Pair  $(F, H)$  is completely detectable if there exists a nonsingular matrix  $T$  such that

$$TFT^{-1} = \begin{bmatrix} F_{11} & O \\ F_{21} & F_{22} \end{bmatrix}, \quad HT^{-1} = [H_1 \ O] \quad (4-3)$$

with  $(F_{11}, H_1)$  is completely observable and  $|\lambda_i(F_{22})| < 1$ . ■

Based on the above lemma and definitions, we have the following theorem.

**Theorem 4-1:** If pair  $(C_1, A, B_q)$  is completely controllable and observable, then the Kalman filter of pair  $(\bar{C}, \bar{A}, \bar{B}_q)$  is asymptotically stable. ■

$$\Omega_1 = P_1 \Omega = \begin{bmatrix} B_q & O & \cdots & O & A^{2^{J-1}}B_q & \cdots & AB_q & \cdots & A^{(n_{\mathbf{x}}-1)2^{J-1}}B_q & \cdots & A^{(n_{\mathbf{x}}-2)2^{J-1}+1}B_q \\ O & B_q & \cdots & O & O & \cdots & O & \cdots & O & \cdots & O \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & B_q & O & \cdots & O & \cdots & O & \cdots & O \end{bmatrix} = \begin{bmatrix} \Omega_1^u \\ \Omega_1^d \end{bmatrix} \quad (3-6)$$

*Proof:*

$(\bar{A}, \bar{B}_q)$  is *Completely Stabilizable*: For  $n$ -dimensional pair  $(F, \Gamma)$ , there exists nonsingular matrix  $T$  such that  $T \cdot \Omega_F = \text{col}\{\Omega_c, O\}$  with  $\Omega_c$  having full row rank, where controllability matrix  $\Omega_F = [\Gamma \quad F\Gamma \quad \dots \quad F^{n-1}\Gamma]$ . Then  $T \cdot \mathbf{x} = \text{col}\{\mathbf{x}_c, \mathbf{x}_{\bar{c}}\}$  with  $\mathbf{x}_c$  being completely controllable and  $\mathbf{x}_{\bar{c}}$  being uncontrollable [16]. This gives a way to find  $T$ . It has been shown in the proof of theorem 3-1 that  $\Omega_1 = P_1\Omega = \text{col}\{\Omega_1^u, \Omega_1^d\}$ , where  $\Omega_1^u$  has full row rank and linearly independent with  $\Omega_1^d$  in row. So

$$\begin{aligned} \bar{A}_1 = P_1 \bar{A} P_1^{-1} &= \left[ \begin{array}{c|ccc} A^{2^{j-1}} & A^{2^{j-1}-1} & \dots & A^2 & A \\ \hline O & O & \dots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \dots & O & O \end{array} \right] \\ &= \left[ \begin{array}{c|c} A^{2^{j-1}} & \bar{A}_1^{ru} \\ \hline O & O \end{array} \right]. \end{aligned} \quad (4-4)$$

There exists a nonsingular matrix  $P_2$  such that

$$\Omega_2 = \begin{bmatrix} I & O \\ O & P_2 \end{bmatrix} \Omega_1 = \begin{bmatrix} \Omega_1^u \\ P_2 \Omega_1^d \end{bmatrix} = \begin{bmatrix} \Omega_1^u \\ \Omega_{1,P}^d \end{bmatrix} = \begin{bmatrix} \Omega_2^u \\ O \end{bmatrix} \quad (4-5)$$

where  $\Omega_{1,P}^d$  is of full row rank. Then  $\Omega_2^u$  has full rank due to  $\Omega_{1,P}^d$  and  $\Omega_1^u$  being linearly independent in row. Now we have

$$\begin{aligned} \bar{A}_2 &= \begin{bmatrix} I & O \\ O & P_2 \end{bmatrix} \bar{A}_1 \begin{bmatrix} I & O \\ O & P_2 \end{bmatrix}^{-1} = \begin{bmatrix} A^{2^{j-1}} & \bar{A}_1^{ru} P_2^{-1} \\ O & O \end{bmatrix} \\ &= \left[ \begin{array}{c|cc} A^{2^{j-1}} & \bar{A}_{1,l}^{ru} & \bar{A}_{1,r}^{ru} \\ \hline O & O & O \\ O & O & O \end{array} \right] = \left[ \begin{array}{c|c} \bar{A}_2^{lu} & \bar{A}_2^{ru} \\ \hline O & O \end{array} \right]. \end{aligned} \quad (4-6)$$

$\bar{A}_2$  is the canonical controllable decomposition of  $\bar{A}$ . The subsystem matrix of the controllable elements is  $\bar{A}_2^{lu}$  and that of the uncontrollable elements is  $O$ , whose eigenvalues are all zeros. According to the definition 4-1,  $(\bar{A}, \bar{B}_q)$  is completely stabilizable. ■

$(\bar{C}, \bar{A})$  is *Completely Detectable*: For  $n$ -dimensional pair  $(F, H)$ , there exists nonsingular matrix  $T$  such that  $\Lambda_F \cdot T^{-1} = [\Lambda_o \quad O]$  with  $\Lambda_o$  being of full column rank, where observability matrix  $\Lambda_F = \text{col}\{H, HF, \dots, HF^{n-1}\}$ . Then  $T \cdot \mathbf{x} = \text{col}\{\mathbf{x}_o, \mathbf{x}_{\bar{o}}\}$ , where  $\mathbf{x}_o$  being completely observable and  $\mathbf{x}_{\bar{o}}$  being unobservable. This gives a way to find  $T$ . In the proof of theorem 3-2 we have

$$\Lambda = \left[ \begin{array}{c|c} \Lambda^{lu} & \Lambda^{ru} \\ \hline O & \Lambda^{rd} \end{array} \right] = [\Lambda^l \mid \Lambda^r] \quad (4-7)$$

where  $\Lambda^r$  has full rank and independent with  $\Lambda^l$  in column. There exists nonsingular matrix  $P_1$  such that

$$\Lambda_1 = \Lambda \cdot \begin{bmatrix} P_1 & O \\ O & I \end{bmatrix} = [\Lambda^l P_1 \quad \Lambda^r] = [O \mid \Lambda_P^l \quad \Lambda^r] = [O \mid \Lambda_1^r] \quad (4-8)$$

where  $\Lambda_1^r$  has full column rank. Thus

$$\begin{aligned} \bar{A}_1 &= \begin{bmatrix} P_1 & O \\ O & I \end{bmatrix}^{-1} \bar{A} \begin{bmatrix} P_1 & O \\ O & I \end{bmatrix} = \begin{bmatrix} O & P^{-1} \bar{A}^{ru} \\ O & A^{2^{j-1}} \end{bmatrix} \\ &= \left[ \begin{array}{c|c} O & \bar{A}_P^{ru} \\ \hline O & \bar{A}_P^{rd} \end{array} \right] = \left[ \begin{array}{c|c} O & \bar{A}_1^{ru} \\ \hline O & \bar{A}_1^{rd} \end{array} \right]. \end{aligned} \quad (4-9)$$

Permute  $\Lambda_1$  in column as  $\Lambda_2 = \Lambda_1 \begin{bmatrix} O & I \\ I & O \end{bmatrix} = [\Lambda_1^r \quad O]$ , then

$$\bar{A}_2 = \begin{bmatrix} O & I \\ I & O \end{bmatrix}^{-1} \left[ \begin{array}{c|c} O & \bar{A}_1^{ru} \\ \hline O & \bar{A}_2^{rd} \end{array} \right] \begin{bmatrix} O & I \\ I & O \end{bmatrix} = \left[ \begin{array}{c|c} \bar{A}_1^{rd} & O \\ \hline \bar{A}_1^{ru} & O \end{array} \right]. \quad (4-10)$$

$\bar{A}_2$  is the canonical observable decomposition of  $\bar{A}$ . The subsystem matrix of the observable elements is  $\bar{A}_1^{rd}$  and that of the unobservable elements is  $O$ , whose eigenvalues are all zeros. According to definition 4-2,  $(\bar{C}, \bar{A})$  is completely detectable. ■

Now that  $(\bar{C}, \bar{A}, \bar{B}_q)$  is completely detectable and stabilizable. According to the lemma 4-1, the corresponding Kalman filter will be asymptotically stable. End of Proof. ■

## V. EXAMPLES

We take scalar Markov processes as illustrative examples, whose system equation at the finest scale is

$$x_1(k_1 + 1) = ax_1(k_1) + bw(k_1) \quad (5-1)$$

where  $w(\bullet)$  is Gaussian white noise with zero mean and variance  $q = 1$ . Suppose the DMS has two scales. We discuss it in two cases. One is that the measurements are available only at the finest scale and the other is that the measurements are available at both the scales.

### A. Measurements are Available Only at the Finest Scale

Suppose that only the measurements at scale 1 are available, i.e.

$$z_1(k_1) = c_1 x_1(k_1) + v_1(k_1) \quad (5-2)$$

where  $v_1(\bullet)$  is Gaussian white noise with zero-mean and variance  $r_1$ . It is individually uncorrelated with  $w(\bullet)$ . Naturally the standard Kalman filtering could be employed to obtain the real time LMMSE of  $x_1(\bullet)$ . Now we assume that the system has two scales but there is no measurement at the second scale. According to Section II, the elements of the time invariant DMS model (3-1) are

$$\bar{\mathbf{x}}(k) = \text{col}\{x_1(2k), x_1(2k+1)\}$$

$$\bar{\mathbf{z}}(k) = \text{col}\{z_1(2k), z_1(2k+1)\}$$

$$\bar{\mathbf{w}}(k) = \text{col}\{w(2k+1), w(2k+2)\}$$

$$\bar{\mathbf{v}}(k) = \text{col}\{v_1(2k), v_1(2k+1)\}$$

$$\bar{A} = \begin{bmatrix} 0 & a \\ 0 & a^2 \end{bmatrix}, \bar{B} = \begin{bmatrix} b & 0 \\ ab & b \end{bmatrix}, \bar{C} = \begin{bmatrix} c_1 & 0 \\ 0 & c_1 \end{bmatrix}$$

$$\bar{Q} = \text{Cov}(\bar{\mathbf{w}}) = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix}, \bar{R} = \text{Cov}(\bar{\mathbf{v}}) = \begin{bmatrix} r_1 & 0 \\ 0 & r_1 \end{bmatrix}.$$

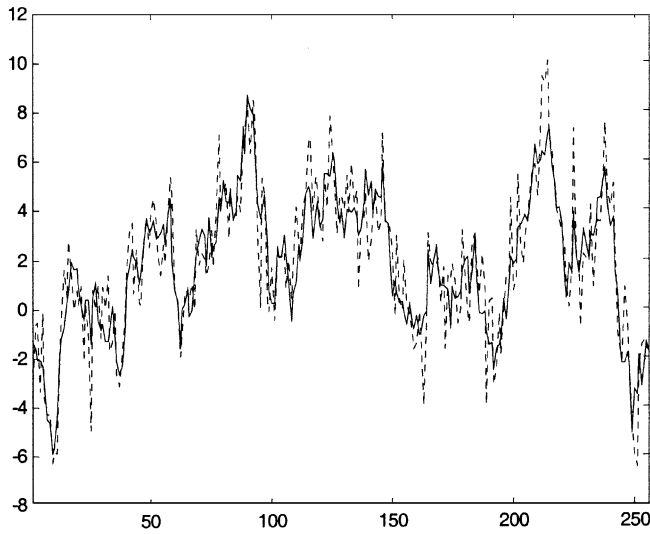


Fig. 3. True-state sequence  $x_1(k_1)$  (solid) and its measurement  $z_1(k_1)$  (dotted).

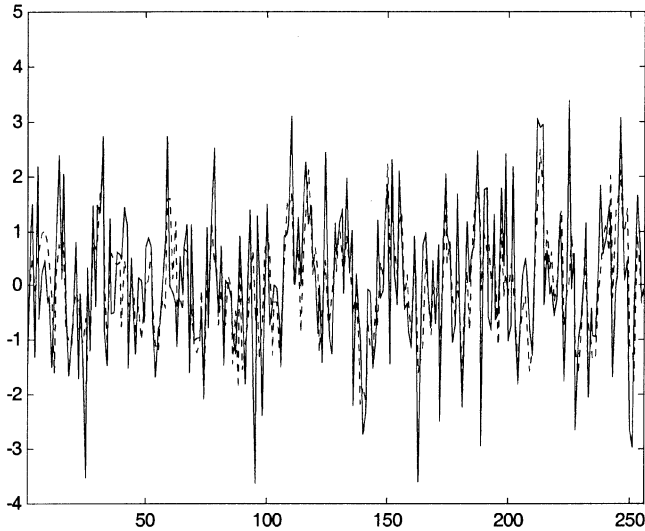


Fig. 4. Measurement noise  $v_1(k_1)$  (solid) and estimation error  $\tilde{x}_1(k_1)$  (dotted). The noise compression ratio is  $c_r = 1.47$ . (No Measurements available at the second scale.)

Actually, in this case the Kalman filtering of model (3-1) is equivalent to the “half Kalman filtering + half Kalman smoothing” of pair  $(c_1, a, b)$ , the subsystem at scale 1. That is to say, once the Kalman filtering of  $x_1(2k)$  and  $x_1(2k + 1)$  is finished, the filtering result of  $x_1(2k)$  is then smoothed by measurement  $z_1(2k + 1)$ . The following experiments validate this observation.

Let  $a = 0.94$ ,  $b = 1$ ,  $c_1 = 1$ , and  $r_1 = 1.69$ . Fig. 3 shows a sequence of true state  $x_1(k_1)$  and its measurement  $z_1(k_1)$ . The estimation result of  $x_1(k_1)$  by the presented scheme is denoted as  $\hat{x}_1(k_1)$ . Fig. 4 compares the measurement noise  $v_1(k_1)$  with the estimation error  $\tilde{x}_1(k_1) = x_1(k_1) - \hat{x}_1(k_1)$ . The noise compression ratio  $c_r$ , which is defined as the ratio of the norm of  $v_1$  to that of  $\tilde{x}_1$ , is 1.47. Fig. 5 plots the “half Kalman filtering + half Kalman smoothing” result of pair  $(c_1, a, b)$  as well as  $\hat{x}_1(k_1)$ . Obviously the two curves are the same.

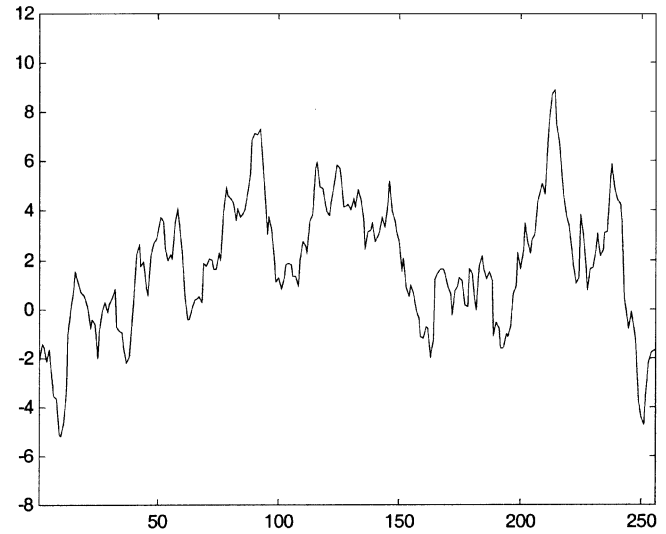


Fig. 5. Estimation  $\hat{x}_1(k_1)$  by the proposed scheme and the estimation by the “half Kalman filtering + half Kalman smoothing” processing. The two curves are overlapped because they are the same.

### B. Measurements are Available at Both the Scales

Now suppose the measurements are also available at the second scale, i.e.

$$\begin{cases} z_1(k_1) = c_1 x_1(k_1) + v_1(k_1) \\ z_2(k_2) = c_2 x_2(k_2) + v_2(k_2) \end{cases} \quad (5-3)$$

where Gaussian white noise  $v_2(\bullet)$  is with zero-mean and variances  $r_2$ . It is uncorrelated with  $v_1(\bullet)$  and  $w(\bullet)$ . Now in model (3-1), there are

$$\bar{\mathbf{x}}(k) = \text{col} \{x_1(2k), x_1(2k + 1)\}$$

$$\bar{\mathbf{z}}(k) = \text{col} \{z_2(k), z_1(2k), z_1(2k + 1)\}$$

$$\bar{\mathbf{w}}(k) = \text{col} \{w(2k + 1), w(2k + 2)\}$$

$$\bar{\mathbf{v}}(k) = \text{col} \{v_2(k), v_1(2k), v_1(2k + 1)\}$$

$$\bar{\mathbf{A}} = \begin{bmatrix} 0 & a \\ 0 & a^2 \end{bmatrix}, \bar{\mathbf{B}} = \begin{bmatrix} b & 0 \\ ab & b \end{bmatrix}, \bar{\mathbf{C}} = \begin{bmatrix} \frac{\sqrt{2}}{2}c_2 & \frac{\sqrt{2}}{2}c_2 \\ c_1 & 0 \\ 0 & c_1 \end{bmatrix}$$

$$\bar{\mathbf{Q}} = \text{Cov}(\bar{\mathbf{w}}) = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix}, \bar{\mathbf{R}} = \text{Cov}(\bar{\mathbf{v}}) = \begin{bmatrix} r_2 & 0 & 0 \\ 0 & r_1 & 0 \\ 0 & 0 & r_1 \end{bmatrix}.$$

Let  $c_2 = 1$  and  $r_2 = 1.44$ . Fig. 6 shows the true-state sequence  $x_2(k_2)$  and its measurement  $z_2(k_2)$ .  $x_2(k_2)$  is taken as the lowpass output of  $x_1(k_1)$ 's Haar wavelet transform. Its estimation by the presented scheme is denoted as  $\hat{x}_2(k_2)$ . Fig. 7 compares  $v_1(k_1)$  with  $\tilde{x}_1(k_1)$ . The noise compression ratio  $c_r$  is 1.85. It is higher than that in the last section because the measurements at the second scale improve the estimation accuracy. Fig. 8 compares noise  $v_2(k_2)$  with estimation error  $\tilde{x}_2(k_2) = x_2(k_2) - \hat{x}_2(k_2)$ . The noise compression ratio  $c_r$  is 1.50.

## VI. CONCLUSION AND DISCUSSION

The modeling and optimal estimation of a class of DMSs that observed independently by several sensors at different scales, to which the traditional Kalman filtering could not be applied

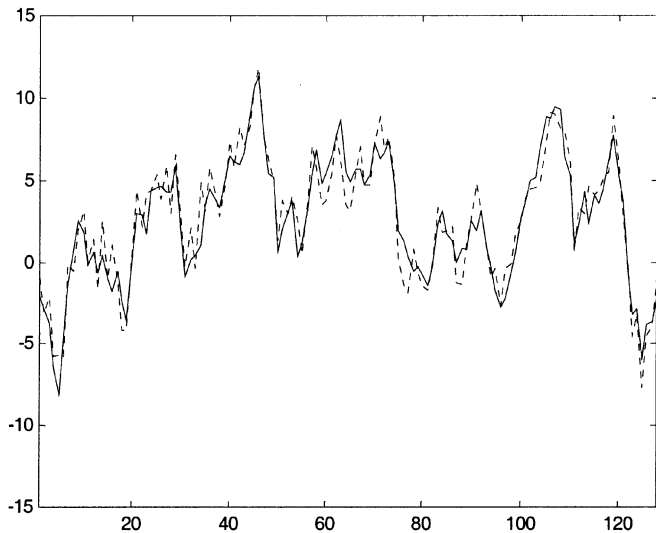


Fig. 6. True-state sequence  $x_2(k_2)$  (solid) and its measurement  $z_2(k_2)$  (dotted).

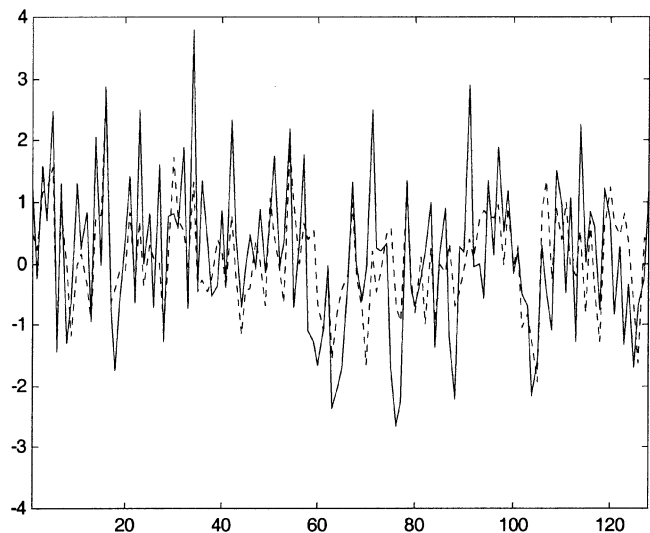


Fig. 8. Measurement noise  $v_2(k_2)$  (solid) and estimation error  $\bar{x}_2(k_2)$  (dotted). The noise compression ratio is  $c_r = 1.50$ . (Measurements are available at both the scales.)

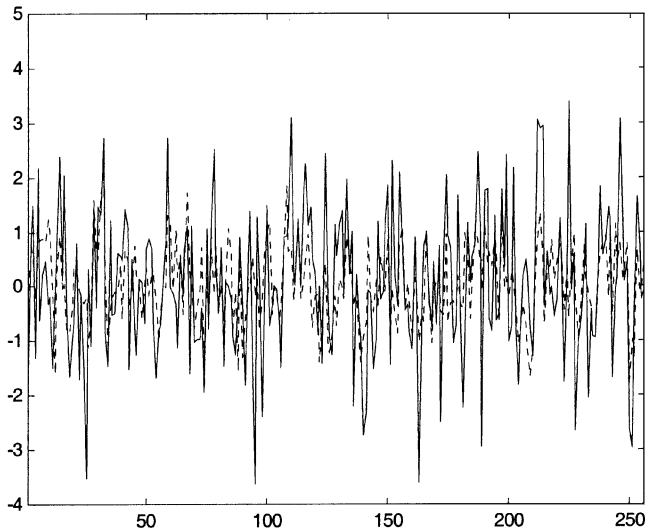


Fig. 7. Measurement noise  $v_1(k_1)$  (solid) and estimation error  $\bar{x}_1(k_1)$  (dotted). The noise compression ratio is  $c_r = 1.85$ . (Measurements are available at both the scales.)

directly, were proposed here. Using Haar wavelet transform to approximate the state projection between scales, we generalized the DMS into the standard state-space model and then the Kalman filtering is employed as the LMMSE algorithm. In the time invariant case, we prove that as long as the DMS is stochastically completely controllable and observable at the finest scale, its associated Kalman filter will be asymptotically stable. Examples were presented with a two-scale Markov process to illustrate the proposed scheme and its relationship with the traditional Kalman filtering and smoothing.

The computation will increase rapidly with the increasing of scale number  $J$ . If the dimension of state  $\mathbf{x}_1$  is  $n_x$ , the dimension of augmented state vector  $\bar{\mathbf{x}}$  will be  $2^{J-1}n_x$ . Suppose the dimension of measurement matrix  $C_j$  is  $n_z^j \times n_x$ , and then the dimension of measurement matrix  $\bar{C}$  will be  $n_z \times 2^{J-1}n_x$ , where  $n_z = \sum_{j=1}^J 2^{j-1}n_z^j$ . A matrix inverse operation of di-

mension  $n_z \times n_z$  will occur when calculating the gain matrix  $K$  in Kalman filtering [14], [15], which needs very heavy computation  $O(n_z^3)$ .

Fortunately, a fast algorithm of the DMS estimation was proposed by Zhang [17]. With observation that the measurements  $z_j(k_j)$  are independent of each other inter-scale and intra-scale, Zhang employed the sequential Kalman filtering of  $\bar{\mathbf{x}}(k)$ , i.e., in each time block  $\Delta T \cdot k$ , the estimation of  $\bar{\mathbf{x}}(k)$  is updated by measurement  $z_j(k_j)$  one by one. So the inverse computation of  $n_z \times n_z$  matrix is replaced by the inverse of those small matrices  $n_z^j \times n_z^j$ . Furthermore, Zhang pointed out that the computation could still be reduced much at the finest scale because the associated state transition (1-1) is available for this scale. The sequential Kalman filtering with all the measurements at scale 1 is equivalent to the fixed-interval Kalman smoothing of the subsystem at scale 1. Thus, for the  $2^{J-1}$  measurements at scale 1 (about half number of the  $2^J - 1$  measurements within a time block), the  $2^{J-1}n_x$  dimensional DMS is reduced to the  $n_x$  dimensional dynamic system.

## REFERENCES

- [1] M. Basseville, A. Benveniste, K. Chou, S. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory*, vol. 38, pp. 766–784, Mar. 1992.
- [2] K. Chou, A. S. Willsky, and A. Benveniste, "Multiscale recursive estimation, data fusion, and regularization," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 464–478, Mar. 1994.
- [3] K. Chou, A. S. Willsky, and R. Nikoukhah, "Multiscale systems, Kalman filters, and Riccati equations," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 479–492, Mar. 1994.
- [4] K. Daoudi, A. Frakt, and A. S. Willsky, "Multiscale autoregressive models and wavelets," *IEEE Trans. Inform. Theory*, vol. 45, pp. 828–845, Apr. 1999.
- [5] M. Luetgen, W. Karl, A. S. Willsky, and R. Tenney, "Multiscale representations of Markov random fields," *IEEE Trans. Signal Processing*, vol. 41, pp. 3377–3396, Dec. 1993.
- [6] A. Frakt, "Internal multiscale autoregressive processes, stochastic realization, and covariance extension," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, 1999.
- [7] L. Hong, "Multiresolutional distributed filtering," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 853–856, Apr. 1994.



- [8] L. Hong, G. Chen, and C. K. Chui, "A filter-bank-based Kalman filtering technique for wavelet estimation and decomposition of random signals," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 237–241, Feb. 1998.
- [9] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992, CBMS-NSF Series in Appl. Math..
- [10] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, pp. 674–693, Nov. 1989.
- [11] M. Vetterli and C. Herley, "Wavelet and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 2207–2232, Sept. 1992.
- [12] Jawerth and W. Sweldens, "An overview of wavelet based multiresolution analyzes," *SIAM Review*, vol. 36, pp. 377–412, 1994.
- [13] G. Strang and T. Nguyen, *Wavelet and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.
- [14] J. M. Mendal, *Lessons in Digital Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [15] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1979.
- [16] C.-T. Chen, *Linear System Theory and Design*. New York: Holt, Rinehart and Winston, 1970.
- [17] L. Zhang, "The optimal estimation of a class of dynamic systems," Ph.D. dissertation, Northwestern Polytechnic Univ., Xi'an, China, 2001.



**Lei Zhang** was born in China, in 1974. He received the B.S. degree from Shenyang Institute of Aeronautic Engineering, Shenyang, China, in 1995 and the M.S. and Ph.D degrees in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 1998 and 2001, respectively.

Currently, he is a Research Associate in the Department of Computing, at The Hong Kong Polytechnic University, Kowloon, Hong Kong. His research interests include digital signal and image processing, optimal estimation theory, information

fusion, and wavelet transform.

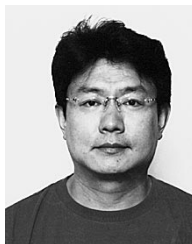


**Quan Pan** received the B.S. degree in automatic control from Huazhong Institute of Technology, Huazhong, China, in 1982 and the M.S. and Ph.D. degrees, both in control theory and control engineering, from Northwestern Polytechnical University, Xi'an, China, in 1991 and 1997, respectively.

In 1991, he joined Northwestern Polytechnical University, Xi'an, China, where from 1998 to 2002, he was the Dean of the Graduate School and is currently a Professor and the Director of the Information Fusion Institute. His research interests

include information fusion, target tracking, automatic target recognition, multiscale system theory, and signal processing. He has published more than 40 journal papers and 20 conference papers.

Dr. Pan received the National Youth Award of Outstanding Contribution to Science and Technology in China, in 1998. He is a member of the International Society of Information Fusion.



**Paul Bao** received the Ph.D. degree from the University of Calgary, Calgary, ON, Canada, in 1988.

From 1988 to 1990, he served on the faculty of the Computer Science Department, University of Calgary. From 1990 to 1995, he was a Staff Analyst at IBM, Canada. From 1995 to 2002, he was an Associate Professor in the Computing Department, at The Hong Kong Polytechnic University. Currently, he is a Professor in the Information Engineering Department, at The Chinese University of Hong Kong. His current research interests include computer graphics and image processing, image-based rendering and virtual reality, digital signal processing, and multimedia systems for mobile computing.



**Hongcai Zhang** received the B.S. degree in gyroscope engineering and the M.S. degree in nonlinear control theory and its applications, from Northwestern Polytechnical University, Xi'an, China, in 1961 and 1964, respectively.

In 1964, he joined the Department of Automatic Control, Northwestern Polytechnical University, where from 1987 to 1992, he was a Vice Dean of the Department of Automatic Control, and, since 1988, he has been a Professor in the same department.

From 1996 to 1998, he was a Curator of the Library at Northwestern Polytechnical University. Currently, he is the Director of the Control Engineering Institute, Northwestern Polytechnical University. He has published more than 45 journal papers and 25 conference papers. His research interests include information fusion, target tracking, multiscale system theory, and signal processing.