Automatica 49 (2013) 1075–1082

Contents lists available at SciVerse ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Optimal linear state estimation over a packet-dropping network using linear temporal coding[☆]

Lidong He[a], Dongfang Han[b], Xiaofan Wang[a,1], Ling Shi[c]

[a] Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China

[b] School of Mathematics and Statistics, South-Central University for Nationalities, Wuhan 430074, China

[c] Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

ARTICLE INFO

Article history:
Received 23 February 2012
Received in revised form
25 September 2012
Accepted 29 November 2012
Available online 26 February 2013

Keywords:
Kalman filter
Correlated noise
Packet-dropping network
Linear temporal coding
Innovation sequence approach

ABSTRACT

We consider the problem of linear minimum mean square error estimation for a discrete-time system over a packet-dropping network. In order to improve the estimation performance, different from the standard approach of sending the current measurement data, we choose sending a linear combination of the current measurement and the measurement collected at the previous time, a method called *linear temporal coding*. We assume the packet arrival sequence is unknown and the noise contained in the packet may come from sensor or communication channel. In an effort to cope with colored noise caused by measurement combination, after comparing with the classic state augmentation approach and measurement differencing approach, we derive a recursive estimation algorithm by means of orthogonal projection principle and innovation sequence approach. Our algorithm consists of two parts: smooth and estimate. For large measurement noise case, numerical example shows the benefit of using linear temporal coding strategy, compared with directly sending the current data. On the contrary, when communication noise plays the dominating role, for scalar system we prove there is no benefit to choose this scheme.

1. Introduction

Thanks to the rapid development of wireless sensor and communication technology, recent years have witnessed a wealth of new applications of wireless sensor networks (WSNs) in many areas, including environmental monitoring and control, remote diagnostics and troubleshooting, factory automation and transportation (Mahalik, 2007). These systems are typically composed of spatially distributed low power sensors which are used to perform sensing, computing and communication. The information is exchanged among these spatially distributed devices through a shared communication network. This new area has attracted much research interest by different communities (Culler, Estrin, & Srivastava, 2004).

Compared with classic wired networks, WSNs provide many significant advantages, such as high flexibility, reduced wiring, low installation and maintenance costs, etc. At the same time, however, inserting a communication network introduces new research challenges, one of which is packet loss. The communication between sensors may be lost stochastically during the transmission, which is typically caused by transmission errors in physical network links, packets collisions or buffer overflows due to packets congestion (Hespanha, Naghshtabrizi, & Xu, 2007). Packet loss may severely degrade closed-loop system performance or even cause system instability. Thus, the effect of packet loss should not be neglected when designing the state estimator and controller. To compensate for packet loss, researchers have proposed many strategies.

Mesquita, Hespanha, and Nair (2009) suggested data retransmission which means a data packet is retransmitted until it is successfully received. This method is widely used in data network while the resulting delay may degrade performance of control system. At the same time, due to many practical reasons, network bandwidth is usually limited. Hence data retransmission is not applicable in networked control systems as in data network. Instead of retransmission, Mostofi and Murray (2009) designed another scheme to improve packet arrival rate, where a packet is used if the

[☆] The work by L.-D. He and X.-F. Wang is partially supported by the Major State Basic Research Development Program of China (973 Program) under Grant No. 2010CB731400 and Natural Science Foundation of China under Grant No. 61074125. The work by D. Han and L. Shi is supported by a HK RGC CRF grant under HKUST11/CRF/10. The material in this paper was partially presented at the 51st IEEE Conference on Decision and Control (CDC), December 10–13, 2012, Maui, Hawaii, USA. This paper was recommended for publication in revised form by Associate Editor Tongwen Chen under the direction of Editor Ian R. Petersen.

E-mail addresses: ldh1983@sjtu.edu.cn (L. He), eastuhan@gmail.com (D. Han), xfwang@sjtu.edu.cn (X. Wang), eesling@ust.hk (L. Shi).

[1] Tel.: +86 21 54747511; fax: +86 21 54747161.

0005-1098/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2013.01.043

corresponding signal-to-noise ratio of the channel is greater than a threshold and is discarded otherwise. In addition to these methods, one may also be interested in improving system performance while not taking extra communication resource. To facilitate the discussion, we first introduce the dynamical system considered in this paper

$$x_{k+1} = Ax_k + \omega_k, \tag{1}$$

$$y_k = Cx_k + v_k, \tag{2}$$

where measurement y_k from the sensor is sent over a packet-dropping network to a remote estimator. Let $\gamma_k = 1$ indicate that the packet successfully arrives at the estimator and $\gamma_k = 0$ otherwise. Denote z_k^E as the information received by the estimator.

Sun, Xie, Xiao, and Soh (2008) considered a compensation mechanism which was first proposed by Sahebsara, Chen, and Shah (2007) and can be described as

$$z_k^E = \gamma_k y_k + (1-\gamma_k) z_{k-1}^E.$$

Using state augmentation approach and innovation sequence approach they derived the optimal linear algorithm for prediction, estimation and smooth. Liang, Chen, and Pan (2010) extended the idea of Sun et al. (2008) to deal with the case when unreliable transmissions exist in both sensor and control channels. Different from Liang et al. (2010) and Sun et al. (2008), Zhang, Yu, and Feng (2011) considered the optimal linear estimation for system with spatially distributed sensors and each sensor i ($i \in \{1, \ldots, m\}$) sends the following data

$$z_k^{E,i} = \gamma_k^i y_k^i + (1-\gamma_k^i) r z_{k-1}^{E,i},$$

where $r \in [0, 1]$ is an adjustable weighting parameter.

Robinson and Kumar (2007) proposed a scheme, called *linear temporal coding* in which

$$z_k^E = \gamma_k(\alpha y_k + \beta y_{k-1}),$$

where α, β are two weighting parameters. Assume $\{\gamma_k\}$ is available, they derived the optimal ratio between α and β to minimize the estimation error covariance at time $k+t$ when $(\gamma_k, \gamma_{k+1}, \ldots, \gamma_{k+t}) = (1, 0, \ldots, 0, 1)$ and demonstrated its advantages when compared with $z_k^E = \gamma_k y_k$.

In Robinson and Kumar (2007), the estimator proposed has a mechanism to detect whether $\gamma_k = 1$ or 0. In some applications, however, such γ_k may not be available. For example, when there exists communication noise in the channel (Nahi, 1969), it is difficult for the estimator to infer whether packet dropping occurs. In this paper, we consider the following problem: when $\{\gamma_k\}$ is not available to the estimator, can we still improve the estimation performance using linear temporal coding strategy? If so, how does the packet arrival rate affect the optimal weighting parameter? Intuitively the more unreliable the network is, the more advantageous is using linear temporal coding since the estimator can partially recover the lost information y_{k-1} at time k. In order to answer these questions, a first step is to derive a linear optimal estimation algorithm. To derive such a state estimation algorithm, most of the works (Liang et al., 2010; Robinson & Kumar, 2007; Sun et al., 2008; Zhang et al., 2011) chose *state augmentation approach*, which is a standard method in linear filtering. On the basis of the state augmentation form (6) and taking γ_k as a system mode, Fletcher, Rangan, and Goyal (2004) gave another minimum mean square error estimator using *Markov Jump Linear System* (MJLS) theory when $\{\gamma_k\}$ is available. In the case when $\{\gamma_k\}$ is unknown, Costa (1994) derived the linear minimum mean square error (LMMSE) estimator for MJLS. More general cases were considered in their book (Costa, Fragoso, & Marques, 2005). The resulting estimation expressions via state augmentation approach, however, are often complicated, from which it is not easy to study the effect of weighting parameters on estimation performance.

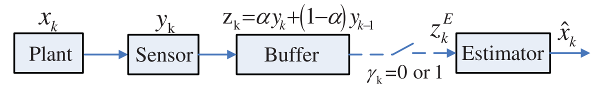

Fig. 1. System block diagram.

The main contributions of this paper are summarized as follows:

(1) We consider LMMSE estimation for a discrete-time system using *linear temporal coding* strategy when the packet arrival sequence $\{\gamma_k\}$ is unknown under two noise assumptions: large measurement noise covariance and large communication noise covariance, shown in (4) and (5), respectively. To the best of our knowledge, this formulation is novel.
(2) After comparing with two classic algorithms, i.e., state augmentation approach and measurement differencing approach, we derive a recursive estimation algorithm using the innovation sequence approach and orthogonal projection principle. The algorithm is composed of two parts: smooth and estimate.
(3) Simulation shows that choosing proper ratio of y_k and y_{k-1} can improve the estimation performance for large measurement noise case. By contrast, when communication noise plays the dominating role, for scalar systems we prove that it is no beneficial to choose this strategy.

The remainder of this paper is organized as follows. In Section 2, the mathematical model is given. In Section 3 we discuss two well-known filtering methods. In Section 4 we propose a linear optimal estimation algorithm using innovation sequence approach. Examples are given in Section 5 to demonstrate the developed theory. Some concluding remarks and future work are given at the end.

2. Problem setup

Consider the discrete linear time-invariant system (1)–(2) (Fig. 1), where $x_k \in \mathbb{R}^n$ is the state vector, $y_k \in \mathbb{R}^m$ is the sensor measurement, $\omega_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are zero-mean random vectors with $\mathbb{E}[\omega_i \omega_j'] = \delta_{ij} Q (Q > 0)$, $\mathbb{E}[v_i v_j'] = \delta_{ij} R_v (R_v > 0)$ and $\mathbb{E}[\omega_i v_j'] = 0$, $\forall i, j$. Here δ is the Kronecker delta function, i.e., $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ otherwise. The initial state is also assumed to be zero-mean with covariance Π_0. We assume the pair (A, C) is detectable and $(A, \sqrt{Q})$ is stabilizable.

Assume measurements collected by the sensor are sent over a packet-dropping network to the estimator. Recall that γ_k indicates whether the packet sent at time k successfully arrives at the estimator or not, i.e., $\gamma_k = 1$ means the packet arrives and $\gamma_k = 0$ otherwise. Furthermore the packet arrival process $\{\gamma_k\}$ is assumed to be independent, identically distributed with $\mathbb{E}[\gamma_k] = \lambda$. In this paper, we assume there is no acknowledgment to indicate whether the packet arrives or not.

Unlike the standard method of transmitting y_k, we consider *linear temporal coding* and send z_k to the remote estimator, where z_k is defined as

$$z_k = \alpha y_k + (1-\alpha) y_{k-1} \tag{3}$$

for a properly chosen $\alpha \in [0, 1]$. In this paper, we also consider communication noise n_k (Dey, Leong, & Evans, 2009; Mostofi & Murray, 2009), which is zero-mean white with covariance R_n and is independent of $\{\omega_k\}$ and $\{v_k\}$. Hence the information received by the estimator is $z_k^E = \gamma_k z_k + n_k$. We consider two cases in this paper:

T1: Large measurement noise covariance $R_v \gg R_n$ where the effect of n_k can be ignored. Thus

$$z_k^E = \gamma_k\{[\alpha C x_k + (1-\alpha) C x_{k-1}] + [\alpha v_k + (1-\alpha) v_{k-1}]\}. \tag{4}$$

T2: Large communication noise covariance $R_v \ll R_n$ where the effect of v_k can be ignored. Thus

$$z_k^E = \gamma_k[\alpha C x_k + (1-\alpha)C x_{k-1}] + n_k. \tag{5}$$

The goal of this paper is to derive a LMMSE estimator under these two noise assumptions with unknown $\{\gamma_k\}$. On that basis, we will also evaluate the influence of α on the estimation performance.

3. Classic filtering algorithms for system with colored noise

For the large measurement noise case, one notes that z_k depends on both v_k and v_{k-1}. Hence the noise term in z_k is no longer white. In this section, we state two well-known methods, i.e., *state augmentation approach* and *measurement differencing approach*, to tackle the correlated noise.

3.1. State augmentation approach (Anderson & Moore, 1979)

For system (1), (4), one can augment the state as follows

$$X_{k+1} = \bar{A}X_k + W_k,$$
$$z_k^E = \gamma_k F(\bar{C}X_k + \xi_k),$$

where $X_k = [x_k; x_{k-1}]$, $W_k = [\omega_k; 0]$, $\xi_k = [v_k; v_{k-1}]$, $\bar{A} = [A\ 0; I\ 0]$, $\bar{C} = [C\ 0; 0\ C]$, $F = [\alpha I\ (1-\alpha)I]$. By introducing a dummy variable ξ_k, we can further rewrite it as

$$\begin{aligned} X_{k+1} &= \bar{A}X_k + W_k,\\ \xi_{k+1} &= G\xi_k + Hv_{k+1}, \\ z_k^E &= \gamma_k F(\bar{C}X_k + \xi_k), \end{aligned} \tag{6}$$

where $G = [0\ 0; I\ 0]$, $H = [I; 0]$ and $Q_W = [Q\ 0; 0\ 0]$.

Filtering algorithm for system (6) with no packet loss was discussed in Anderson and Moore (1979). In principle, there is no difficulty in dealing with system (6) with colored noise and pack loss. One can use classical Kalman filtering theory to design the LMMSE estimator for $[X_k; \xi_k]$ as in Anderson and Moore (1979), and then combine that with the idea in Nahi (1969) to deal with unknown packet-dropping sequences. Taking γ_k as a mode, one can also use MJLS theory (Costa, 1994; Costa et al., 2005) to derive a LMMSE estimator for $[X_k; \xi_k]$, which has been discussed by Fletcher et al. (2004) when $\{\gamma_k\}$ is available. The resulting filter expressions via the state augmentation approach, however, often quite complex and not easy to analyze due to the high dimension of the augmented system.

3.2. Measurement differencing approach (Bryson & Henrikson, 1968; Henrikson, 1968)

Bryson and Henrikson (1968) proposed another scheme, the measurement differencing approach.

(1) Single step correlated measurement noise: Suppose the colored noise and the measurement equation of the system (1) can be written as

$$\varepsilon_{k+1} = \Psi\varepsilon_k + v_k,$$
$$y_k = Cx_k + \varepsilon_k,$$

where v_k is white noise. Define the measurement difference term as

$$\begin{aligned} \widehat{y}_k &\triangleq y_k - \Psi y_{k-1} \\ &= \widehat{C}x_{k-1} + (C\omega_{k-1} + v_{k-1}), \end{aligned}$$

where $\widehat{C} \triangleq CA - \Psi C$. Now the new measurement noise $C\omega_{k-1} + v_{k-1}$ only correlates with the process noise, which can be easily dealt with. Henrikson extended the idea to more general systems in his Ph.D. Thesis (Henrikson, 1968).

(2) Two steps correlated measurement noise (Henrikson, 1968): Assume the measurement noise can be described as

$$\begin{bmatrix}\varepsilon_{k+1}^1 \\ \varepsilon_{k+1}^2\end{bmatrix} = \begin{bmatrix}\Psi_{11} & I \\ 0 & \Psi_{22}\end{bmatrix}\begin{bmatrix}\varepsilon_k^1 \\ \varepsilon_k^2\end{bmatrix} + \begin{bmatrix}0 \\ 1\end{bmatrix} v_k, \tag{7}$$

$$y_k = Cx_k + \varepsilon_k^1.$$

Define

$$\begin{aligned} \widehat{y}_k &\triangleq y_k - (\Psi_{11} + \Psi_{22})y_{k-1} + \Psi_{22}\Psi_{11}y_{k-2} \\ &= \widehat{M}_k + v_{k-2}, \end{aligned}$$

where

$$\widehat{M}_k \triangleq Cx_k - (\Psi_{11} + \Psi_{22})Cx_{k-1} + \Psi_{22}\Psi_{11}Cx_{k-2},$$

which does not depend on the measurement noise. Hence the colored noise term is eliminated.

In our problem setup, the measurement equation can be written as

$$z_k = F\bar{C}X_k + [\alpha v_k + (1-\alpha)v_{k-1}].$$

Here we consider a perfect network with all γ_k's being 1 for illustrating the main idea. Given any ψ, calculate the measurement difference term as follows:

$$\begin{aligned} \widehat{z}_k &\triangleq z_k - \psi z_{k-1} \\ &= (F\bar{C}X_k - \psi F\bar{C}X_{k-1}) - (1-\alpha)\psi v_{k-2} \\ &\quad + [(1-\alpha)I - \alpha\psi]v_{k-1} + \alpha v_k. \end{aligned}$$

The colored noise term, however, cannot be removed using the measurement differencing approach.

4. Optimal linear estimator using innovation sequence approach

Unlike the two classic methods, in this section we derive a LMMSE estimator for system (1)–(3) using orthogonal projection principle and innovation sequence approach. Two important lemmas are introduced first, based on which lie our LMMSE estimation algorithm. In this paper, we define Proj.$[X|Y]$ as the LMMSE estimation of X given Y, i.e., the projection of X on the space spanned by Y. $\mathbb{E}[X]$ stands for mathematical expectation of X.

Lemma 4.1 (Anderson & Moore, 1979)**.** *Suppose the random variable $[X; Y]$ has mean $[m_x; m_y]$ and covariance $[\Sigma_{xx}, \Sigma_{xy}; \Sigma_{yx}, \Sigma_{yy}]$, respectively. Then*

$$\text{Proj.}[X|Y] = m_x + \Sigma_{xy}\Sigma_{yy}^{-1}(Y - m_y), \tag{8}$$

with replacement of the inverse by a pseudo-inverse if necessary.

Lemma 4.2 (Anderson & Moore, 1979)**.** *Suppose that $X, Y_1, \ldots, Y_k$ are jointly distributed, with $Y_1, \ldots, Y_k$ mutually uncorrelated, i.e., $\Sigma_{y_iy_j} = 0$ $(i \neq j)$. Then*

$$\text{Proj.}[X|Y_1, \ldots, Y_k] = \sum_{i=1}^{k} \text{Proj.}[X|Y_i] - (k-1)m_x. \tag{9}$$

We also define some symbols as follows: $A \sim B$ implies that the variable A depends on B; Z_1^k is the space spanned by the vectors $z_1^E, \ldots, z_k^E$; $\hat{z}_k^- \triangleq$ Proj.$[z_k^E|Z_1^{k-1}]$ is the LMMSE estimation of $z_k^E = \gamma_k z_k + n_k$ given Z_1^{k-1}; $\tilde{z}_k \triangleq z_k^E - \hat{z}_k^-$ is the innovation contained in z_k^E; $\tilde{Z}_1^k$ is the space spanned by $\tilde{z}_1, \ldots, \tilde{z}_k$; $\hat{x}_k^- \triangleq$ Proj.$[x_k|Z_1^{k-1}]$, $\hat{x}_k \triangleq$ Proj.$[x_k|Z_1^k]$, $\hat{x}_k^+ \triangleq$ Proj.$[x_k|Z_1^{k+1}]$ stand for prediction, estimation and smooth for state x_k, respectively. Their corresponding errors are written as $e_k^- \triangleq x_k - \hat{x}_k^-$, $e_k \triangleq x_k - \hat{x}_k$, $e_k^+ \triangleq x_k - \hat{x}_k^+$. We also define their error covariances as $P_k^- \triangleq \mathbb{E}[e_k^-(e_k^-)']$, $P_k \triangleq \mathbb{E}[e_k e_k']$, $P_k^+ \triangleq \mathbb{E}[e_k^+(e_k^+)']$. $J_{k-1} \triangleq \text{Cov}(x_{k-1}, \tilde{z}_k) = \mathbb{E}[(x_{k-1} - \text{Proj.}[x_{k-1}|Z_1^{k-1}])(\tilde{z}_k - \text{Proj.}[\tilde{z}_k|Z_1^{k-1}])']$ is the covariance of x_{k-1} and $\tilde{z}_k$

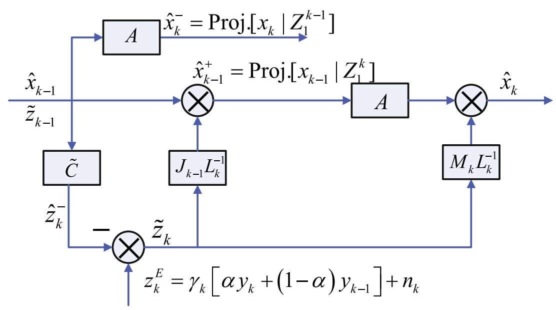

Fig. 2. Filter with linear temporal coding.

given Z_1^{k-1}.[2] Similarly define $L_k \triangleq \mathrm{Cov}(\tilde{z}_k, \tilde{z}_k)$ and $M_k \triangleq \mathrm{Cov}(\omega_{k-1}, \tilde{z}_k)$. Finally denote $S_k \triangleq \mathbb{E}[x_k x_k']$ and xx' is abbreviated as $(x)(\bullet)'$ when the expression of x is lengthy to state.

The central idea of our estimation algorithm is shown graphically in Fig. 2. Instead of computing $\hat{x}_k$ directly, one first gets the smooth of x_{k-1}, i.e., $\hat{x}_{k-1}^+ \sim (\hat{x}_{k-1}, \tilde{z}_k)$. To derive $\hat{x}_k$, it is then better to use $\hat{x}_{k-1}^+ = \mathrm{Proj.}[x_{k-1}|Z_1^k]$ than $\hat{x}_k^- = \mathrm{Proj.}[x_k|Z_1^{k-1}] = A\mathrm{Proj.}[x_{k-1}|Z_1^{k-1}]$, since the former one has access to more information. Building on the above idea, we have the following result.

Theorem 4.3. *A LMMSE estimation algorithm for system (1)–(3) with unknown $\{\gamma_k\}$ series consists of two steps:*

(1) *Smooth:*

$$\begin{aligned} \hat{x}_{k-1}^+ &= \hat{x}_{k-1} + J_{k-1}L_k^{-1}\tilde{z}_k, \\ P_{k-1}^+ &= P_{k-1} - J_{k-1}L_k^{-1}J_{k-1}', \end{aligned} \tag{10}$$

(2) *Estimate:*

$$\begin{aligned} \hat{x}_k &= A\hat{x}_{k-1}^+ + M_k L_k^{-1}\tilde{z}_k, \\ P_k &= AP_{k-1}A' + Q - (AJ_{k-1} + M_k)L_k^{-1}(AJ_{k-1} + M_k)'. \end{aligned} \tag{11}$$

Proof. We first list some equations obtained by orthogonal projection principle (Anderson & Moore, 1979), which will be frequently used.

$$\mathbb{E}[e_k \hat{x}_k'] = 0, \qquad \mathbb{E}[e_k \tilde{z}_k'] = 0,$$
$$\mathrm{Proj.}[\tilde{z}_k|Z_1^{k-1}] = \mathrm{Proj.}[\tilde{z}_k|\tilde{Z}_1^{k-1}] = 0.$$

The last equality above holds true for the innovation sequence is derived via the so-called Gram–Schmidt procedure. Then calculate L_k as

$$\begin{aligned} L_k &= \mathrm{Cov}(\tilde{z}_k, \tilde{z}_k) \\ &= \mathbb{E}[(\tilde{z}_k - \mathrm{Proj.}[\tilde{z}_k|\tilde{Z}_1^{k-1}])(\tilde{z}_k - \mathrm{Proj.}[\tilde{z}_k|\tilde{Z}_1^{k-1}])'] \\ &= \mathbb{E}[\tilde{z}_k \tilde{z}_k'], \end{aligned} \tag{12}$$

where the third equality is due to $\mathrm{Proj.}[\tilde{z}_k|\tilde{Z}_1^{k-1}] = 0$. Similarly,

$$J_{k-1} = \mathrm{Cov}(x_{k-1}, \tilde{z}_k) = \mathbb{E}[e_{k-1}\tilde{z}_k'], \tag{13}$$

$$M_k = \mathrm{Cov}(\omega_{k-1}, \tilde{z}_k) = \mathbb{E}[\omega_{k-1}\tilde{z}_k']. \tag{14}$$

From Lemma 4.2, we have

$$\begin{aligned} \hat{x}_{k-1}^+ &= \mathrm{Proj.}[x_{k-1}|Z_1^k] = \mathrm{Proj.}[x_{k-1}|\tilde{Z}_1^k] \\ &= \mathrm{Proj.}[x_{k-1}|\tilde{Z}_1^{k-1}] + \mathrm{Proj.}[x_{k-1}|\tilde{z}_k] - \mathbb{E}[x_{k-1}], \end{aligned}$$

where $\mathrm{Proj.}[x_{k-1}|\tilde{z}_k]$ can be computed via Lemma 4.1 as follows

$$\mathrm{Proj.}[x_{k-1}|\tilde{z}_k] = \mathbb{E}[x_{k-1}] + \mathrm{Cov}(x_{k-1}, \tilde{z}_k)[\mathrm{Cov}(\tilde{z}_k, \tilde{z}_k)]^{-1}\tilde{z}_k.$$

Recall the definition of J_{k-1}, L_k and $\hat{x}_{k-1}$, we have

$$\hat{x}_{k-1}^+ = \hat{x}_{k-1} + J_{k-1}L_k^{-1}\tilde{z}_k.$$

Hence

$$\begin{aligned} e_{k-1}^+ &= x_{k-1} - \hat{x}_{k-1}^+ = e_{k-1} - J_{k-1}L_k^{-1}\tilde{z}_k, \\ P_{k-1}^+ &= \mathbb{E}[e_{k-1}^+ e_{k-1}'^+] = P_{k-1} - J_{k-1}L_k^{-1}J_{k-1}'. \end{aligned} \tag{15}$$

Combining (12) and (13) leads to the last equality above. Using again Lemmas 4.1 and 4.2 with the fact $\mathrm{Proj.}[\omega_{k-1}|\tilde{Z}_1^{k-1}] = 0$ yields

$$\begin{aligned} \hat{x}_k &= \mathrm{Proj.}[x_k|Z_1^k] = \mathrm{Proj.}[(Ax_{k-1} + \omega_{k-1})|\tilde{Z}_1^k] \\ &= A\hat{x}_{k-1}^+ + \mathrm{Proj.}[\omega_{k-1}|\tilde{Z}_1^{k-1}] \\ &\quad + \mathrm{Cov}(\omega_{k-1}, \tilde{z}_k)[\mathrm{Cov}(\tilde{z}_k, \tilde{z}_k)]^{-1}\tilde{z}_k \\ &= A\hat{x}_{k-1}^+ + M_k L_k^{-1}\tilde{z}_k \\ &= A\hat{x}_{k-1} + (AJ_{k-1} + M_k)L_k^{-1}\tilde{z}_k, \end{aligned} \tag{16}$$

$$\begin{aligned} e_k &= x_k - \hat{x}_k \\ &= Ae_{k-1} + \omega_{k-1} - (AJ_{k-1} + M_k)L_k^{-1}\tilde{z}_k. \end{aligned} \tag{17}$$

The last equality of (16) follows on using expression (10). From (12)–(14), together with (17) and $\mathbb{E}[e_{k-1}\omega_{k-1}'] = 0$, we get

$$\begin{aligned} P_k &= \mathbb{E}[e_k e_k'] \\ &= AP_{k-1}A' + Q - (AJ_{k-1} + M_k)L_k^{-1}(AJ_{k-1} + M_k)'. \quad \square \end{aligned}$$

Remark 4.4. The estimation algorithm of Theorem 4.3 is relatively general, namely, it does not rely on whether the noise is white or colored. Calculation of J_{k-1}, M_k, L_k and $\tilde{z}_k$ depends on the different models of z_k^E, which will be presented in the next few subsections.

Remark 4.5. When there only exists correlated noise, innovation sequence approach has been considered by Jiang, Zhou, and Zhu (2010). Different from their work, our linear temporal coding strategy is novel in that we consider both smooth and estimate, and we consider estimation over an unreliable network.

4.1. Large measurement noise covariance: $R_v \gg R_n$

We rewrite the measurement Eq. (4) as

$$z_k^E = \gamma_k(\tilde{C}x_{k-1} + \tilde{v}_k), \tag{18}$$

where

$$\tilde{C} \triangleq C[\alpha A + (1-\alpha)I], \tag{19}$$

$$\tilde{v}_k \triangleq \alpha C\omega_{k-1} + \alpha v_k + (1-\alpha)v_{k-1}. \tag{20}$$

Note that $\{\omega_k\}$ and $\{\tilde{v}_k\}$ are still zero-mean noise, but the newly defined $\{\tilde{v}_k\}$ is a colored noise sequence, i.e., $\mathbb{E}[\tilde{v}_k\tilde{v}_{k-1}'] \neq 0$. We can write them in a compact form as

$$\mathbb{E}\left[\begin{bmatrix} \omega_{k-1} \\ \tilde{v}_{k-1} \\ \tilde{v}_k \end{bmatrix}\begin{bmatrix} \omega_{k-1}' & \tilde{v}_{k-1}' & \tilde{v}_k' \end{bmatrix}\right] = \begin{bmatrix} Q & 0 & \Omega \\ 0 & \tilde{R}_0 & \tilde{R}_1 \\ \Omega' & \tilde{R}_1 & \tilde{R}_0 \end{bmatrix}, \tag{21}$$

where

$$\Omega \triangleq \alpha QC', \qquad \tilde{R}_1 \triangleq \alpha(1-\alpha)R_v,$$
$$\tilde{R}_0 \triangleq \alpha^2 CQC' + [\alpha^2 + (1-\alpha)^2]R_v.$$

Now we derive explicit expressions for J_{k-1}, M_k and L_k.

[2] The covariance here cannot be deemed as the conditional covariance since we use LMMSE estimation $\mathrm{Proj.}[x_{k-1}|Z_1^{k-1}]$ instead of conditional expectation $\mathbb{E}[x_{k-1}|Z_1^{k-1}]$.

For the unknown Bernoulli random variable γ_k, we have

$$\mathbb{E}[(\gamma_k-\lambda)^2]=\lambda(1-\lambda),$$
$$\mathbb{E}[\gamma_k]=\mathbb{E}[\gamma_k^2]=\lambda,\qquad \mathbb{E}[\gamma_i\gamma_j]=\lambda^2\quad(i\neq j).$$

Compute $\hat{z}_k^-$ as

$$\begin{aligned}\hat{z}_k^- &= \mathrm{Proj}.[z_k^E|Z_1^{k-1}]=\mathrm{Proj}.[\gamma_k(\tilde{C}x_{k-1}+\tilde{v}_k)|\tilde{Z}_1^{k-1}]\\ &=\lambda\{\tilde{C}\hat{x}_{k-1}+\mathrm{Proj}.[\tilde{v}_k|\tilde{Z}_1^{k-1}]\}.\end{aligned}\tag{22}$$

Since

$$\tilde{Z}_1^{k-1}\sim(x_0;\omega_0,\ldots,\omega_{k-2};v_0,\ldots,v_{k-1};\gamma_1,\ldots,\gamma_{k-1}),\tag{23}$$

we have $\mathrm{Proj}.[\omega_{k-1}|\tilde{Z}_1^{k-1}]=0$, $\mathrm{Proj}.[v_k|\tilde{Z}_1^{k-1}]=0$ and $\mathbb{E}[v_{k-1}(\hat{z}_{k-1}^-)']=0$. By Lemma 4.1,

$$\begin{aligned}\mathrm{Proj}.[\tilde{v}_k|\tilde{Z}_1^{k-1}] &= (1-\alpha)\{\mathrm{Proj}.[v_{k-1}|\tilde{Z}_1^{k-2}]\\ &\qquad+\mathrm{Cov}(v_{k-1},\tilde{z}_{k-1})L_{k-1}^{-1}\tilde{z}_{k-1}\}\\ &=(1-\alpha)\mathbb{E}[v_{k-1}\tilde{z}_{k-1}']L_{k-1}^{-1}\tilde{z}_{k-1}.\end{aligned}\tag{24}$$

Since

$$\begin{aligned}\mathbb{E}[v_{k-1}\tilde{z}_{k-1}'] &= \mathbb{E}[\gamma_{k-1}v_{k-1}z_{k-1}']=\mathbb{E}[\gamma_{k-1}v_{k-1}\tilde{v}_{k-1}']\\ &=\mathbb{E}[\gamma_{k-1}v_{k-1}\alpha v_{k-1}']=\lambda\alpha R_v.\end{aligned}\tag{25}$$

Substituting (24)–(25) into (22) yields

$$\hat{z}_k^- = \lambda[\tilde{C}\hat{x}_{k-1}+\lambda\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1}],\tag{26}$$

$$\begin{aligned}\tilde{z}_k &= z_k^E-\hat{z}_k^-\\ &=\gamma_k\tilde{C}e_{k-1}+(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}+\gamma_k\tilde{v}_k-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1}.\end{aligned}\tag{27}$$

As shown in (10), in order to derive $\hat{x}_{k-1}^+$, we calculate J_{k-1} in (13) as

$$J_{k-1}=\mathbb{E}[e_{k-1}\tilde{z}_k']=\lambda\{P_{k-1}\tilde{C}'+\mathbb{E}[e_{k-1}\tilde{v}_k']\},\tag{28}$$

where we use the orthogonality condition $\mathbb{E}[e_{k-1}\hat{x}_{k-1}']=0$ and $\mathbb{E}[e_{k-1}\tilde{z}_{k-1}']=0$. Since

$$\mathbb{E}[e_{k-1}\tilde{v}_k']=-(1-\alpha)\mathbb{E}[\hat{x}_{k-1}v_{k-1}'],$$

we derive the expression for $\hat{x}_{k-1}$ (or $\hat{x}_k$) first. As Eq. (16) shows, we need to calculate M_k in (14) as

$$M_k=\mathbb{E}[\omega_{k-1}\tilde{z}_k']=\mathbb{E}[\omega_{k-1}(\gamma_k\tilde{v}_k)']=\lambda\Omega.\tag{29}$$

Then

$$\begin{aligned}\mathbb{E}[e_{k-1}\tilde{v}_k'] &= -(1-\alpha)\mathbb{E}[\hat{x}_{k-1}v_{k-1}']\\ &=-(1-\alpha)(AJ_{k-2}+M_{k-1})L_{k-1}^{-1}\mathbb{E}[\tilde{z}_{k-1}v_{k-1}']\\ &=-\lambda(AJ_{k-2}+M_{k-1})L_{k-1}^{-1}\tilde{R}_1,\end{aligned}\tag{30}$$

where the second equality uses (16) and the last equality follows by (25). Combining (28) with (30) yields

$$J_{k-1}=\lambda P_{k-1}\tilde{C}'-\lambda^2(AJ_{k-2}+M_{k-1})L_{k-1}^{-1}\tilde{R}_1.\tag{31}$$

Now let us turn back to calculate L_k. Employing again the orthogonal condition $\mathbb{E}[e_{k-1}\hat{x}_{k-1}']=0$, we have

$$S_k=\mathbb{E}[x_kx_k']=AS_{k-1}A'+Q,\tag{32}$$

$$P_{k-1}=\mathbb{E}[e_{k-1}e_{k-1}']=\mathbb{E}[e_{k-1}x_{k-1}'],\tag{33}$$

$$\mathbb{E}[\hat{x}_{k-1}(\hat{x}_{k-1})']=\mathbb{E}[x_{k-1}(\hat{x}_{k-1})']=S_{k-1}-P_{k-1}.\tag{34}$$

As

$$\tilde{z}_k=\gamma_k\tilde{C}e_{k-1}+(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}+\gamma_k\tilde{v}_k-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1},$$

one can easily verify that

$$\mathbb{E}[(\gamma_k\tilde{C}e_{k-1})(\bullet)']=\lambda\tilde{C}P_{k-1}\tilde{C}',$$
$$\mathbb{E}[\{(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}\}\{\bullet\}']=\lambda(1-\lambda)\tilde{C}(S_{k-1}-P_{k-1})\tilde{C}',$$
$$\mathbb{E}[(\gamma_k\tilde{v}_k)(\bullet)']=\lambda\tilde{R}_0,$$
$$\mathbb{E}[(-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1})(\bullet)']=\lambda^4\tilde{R}_1L_{k-1}^{-1}\tilde{R}_1.$$

From the orthogonal principle, we have

$$\mathbb{E}[(\gamma_k\tilde{C}e_{k-1})\{(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}\}']=0,$$
$$\mathbb{E}[(\gamma_k\tilde{C}e_{k-1})(-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1})']=0.$$

Since neither $\hat{x}_{k-1}$ nor $\tilde{z}_{k-1}$ relies on γ_k, and $\mathbb{E}[\gamma_k-\lambda]=0$, we obtain

$$\mathbb{E}[\{(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}\}(-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1})']=0.$$

In view of (20), (23) and (25), we have

$$\mathbb{E}[\tilde{v}_k\tilde{z}_{k-1}']=(1-\alpha)\mathbb{E}[v_{k-1}\tilde{z}_{k-1}']=\lambda\tilde{R}_1.$$

Hence

$$\begin{aligned}&\mathbb{E}[(\gamma_k\tilde{v}_k)(-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1})'+(-\lambda^2\tilde{R}_1L_{k-1}^{-1}\tilde{z}_{k-1})(\gamma_k\tilde{v}_k)']\\ &\quad=-2\lambda^4\tilde{R}_1L_{k-1}^{-1}\tilde{R}_1.\end{aligned}$$

From (30), one has

$$\mathbb{E}[e_{k-1}\tilde{v}_k']=-\mathbb{E}[\hat{x}_{k-1}\tilde{v}_k']=-\lambda(AJ_{k-2}+M_{k-1})L_{k-1}^{-1}\tilde{R}_1,$$

where λ comes from γ_{k-1} and does not relate to γ_k. Therefore

$$\begin{aligned}&\mathbb{E}[(\gamma_k\tilde{C}e_{k-1})(\gamma_k\tilde{v}_k)'+\{(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}\}(\gamma_k\tilde{v}_k)'\\ &\qquad+(\gamma_k\tilde{v}_k)(\gamma_k\tilde{C}e_{k-1})'+(\gamma_k\tilde{v}_k)\{(\gamma_k-\lambda)\tilde{C}\hat{x}_{k-1}\}']\\ &\quad=-\lambda^3[\tilde{C}(AJ_{k-2}+M_{k-1})L_{k-1}^{-1}\tilde{R}_1+\tilde{R}_1L_{k-1}^{-1}(AJ_{k-2}+M_{k-1})'\tilde{C}'].\end{aligned}$$

Combining all the results above yields

$$\begin{aligned}L_k &= \mathbb{E}[\tilde{z}_k\tilde{z}_k']\\ &=\lambda\{\tilde{C}P_{k-1}\tilde{C}'+(1-\lambda)\tilde{C}(S_{k-1}-P_{k-1})\tilde{C}'+\tilde{R}_0\\ &\quad-\lambda^3\tilde{R}_1L_{k-1}^{-1}\tilde{R}_1-\lambda^2\tilde{C}(AJ_{k-2}+M_{k-1})L_{k-1}^{-1}\tilde{R}_1\\ &\quad-\lambda^2\tilde{R}_1L_{k-1}^{-1}(AJ_{k-2}+M_{k-1})'\tilde{C}'\}.\end{aligned}$$

Up to this point, we have obtained all the parameters for the filter expression (10)–(11) when $R_v\gg R_n$. For the initial conditions, one can make $\hat{x}_0=0$ and $P_0=\Pi_0$. Suppose we do not send packet at time $k=1$, which means $\tilde{z}_1=0$, $\hat{x}_1=A\hat{x}_0$, $P_1=AP_0A'+Q$, then $\tilde{z}_2=\gamma_2z_2-\lambda\tilde{C}\hat{x}_1$, $J_1=\lambda P_1\tilde{C}'$ and $L_2=\lambda[\tilde{C}P_1\tilde{C}'+(1-\lambda)\tilde{C}(S_1-P_1)\tilde{C}'+\tilde{R}_0]$. The complete estimation algorithm is given by Algorithm 1 and denoted as

$$(\hat{x}_k,P_k)=\Phi(\hat{x}_{k-1},P_{k-1},S_{k-1},\lambda,\Omega,\tilde{R}_0,\tilde{R}_1).$$

4.2. Large communication noise covariance: $R_v\ll R_n$

Rewrite Eq. (5) as

$$z_k^E=\gamma_k\tilde{C}x_{k-1}+\bar{n}_k,\tag{35}$$

where

$$\bar{n}_k\triangleq\gamma_k\alpha C\omega_{k-1}+n_k,\tag{36}$$

and $\tilde{C}$ is same as in (19). Clearly $\{\omega_{k-1}\}$ and $\{\bar{n}_k\}$ are zero-mean white with

$$\begin{aligned}&\mathbb{E}\left[\begin{bmatrix}\omega_{k-1}\\ \bar{n}_k\end{bmatrix}\begin{bmatrix}\omega_{k-1}' & \bar{n}_k'\end{bmatrix}\right]\\ &\quad=\begin{bmatrix}Q & \lambda\alpha QC'\\ \lambda\alpha CQ & \lambda\alpha^2CQC'+R_n\end{bmatrix}\triangleq\begin{bmatrix}Q & \bar{\Omega}\\ \bar{\Omega}' & \bar{R}_n\end{bmatrix}.\end{aligned}\tag{37}$$

Algorithm 1 Estimation algorithm for $R_v \gg R_n$

1: Initial conditions: $\tilde{R}_0 \triangleq \alpha^2 CQC' + [\alpha^2 + (1-\alpha)^2]R_v$, $\tilde{R}_1 \triangleq \alpha(1-\alpha)R_v$, $\Omega \triangleq \alpha QC'$, $\tilde{C} \triangleq C[\alpha A + (1-\alpha)I]$, $x_0 \sim \mathcal{N}(0, \Pi_0)$, $S_0 = \Pi_0$, $\hat{x}_0 = 0$, $P_0 = \Pi_0$, $\tilde{z}_1 = 0$, $\hat{x}_1 = A\hat{x}_0$, $P_1 = AP_0A' + Q$, $J_1 = \lambda P_1 \tilde{C}'$ and $L_2 = \lambda[\tilde{C}P_1\tilde{C}' + (1-\lambda)\tilde{C}(S_1 - P_1)\tilde{C}' + \tilde{R}_0]$, $\hat{x}_2, P_2, S_1, S_2$ are calculated from step 5–8.
When $k \geq 3$,

2: LMMSE estimation of $z_k^E = \gamma_k z_k$ given Z_1^{k-1}
$$\hat{z}_k^- = \lambda[\tilde{C}\hat{x}_{k-1} + \lambda \tilde{R}_1 L_{k-1}^{-1}(\gamma_{k-1} z_{k-1} - \hat{z}_{k-1}^-)],$$

3: Covariance of x_{k-1} and $\tilde{z}_k$ given Z_1^{k-1}
$$J_{k-1} = \lambda P_{k-1}\tilde{C}' - \lambda^2 (AJ_{k-2} + M_{k-1}) L_{k-1}^{-1}\tilde{R}_1,$$

4: Variance of $\tilde{z}_k$ and $\tilde{z}_k$ given Z_1^{k-1}
$$\begin{aligned} L_k = \lambda\{&\tilde{C}P_{k-1}\tilde{C}' + (1-\lambda)\tilde{C}(S_{k-1} - P_{k-1})\tilde{C}' + \tilde{R}_0 \\ &- \lambda^3 \tilde{R}_1 L_{k-1}^{-1}\tilde{R}_1 - \lambda^2 \tilde{C}(AJ_{k-2} + M_{k-1})L_{k-1}^{-1}\tilde{R}_1 \\ &- \lambda^2 \tilde{R}_1 L_{k-1}^{-1}(AJ_{k-2} + M_{k-1})'\tilde{C}'\}, \end{aligned}$$

5: Smooth step
$$\hat{x}_{k-1}^+ = \hat{x}_{k-1} + J_{k-1}L_k^{-1}(\gamma_k z_k - \hat{z}_k^-),$$
$$P_{k-1}^+ = P_{k-1} - J_{k-1}L_k^{-1}J_{k-1}',$$

6: Covariance of ω_{k-1} and $\tilde{z}_k$ given Z_1^{k-1}
$$M_k = \lambda\Omega,$$

7: Estimate step
$$\hat{x}_k = A\hat{x}_{k-1}^+ + M_k L_k^{-1}(\gamma_k z_k - \hat{z}_k^-),$$
$$P_k = AP_{k-1}A' + Q - (AJ_{k-1} + M_k)L_k^{-1}(AJ_{k-1} + M_k)',$$

8: Recursion for $S_k \triangleq \mathbb{E}[x_k x_k']$
$$S_k = AS_{k-1}A' + Q.$$

We follow a similar approach as in the $R_v \gg R_n$ case to derive the expression of the parameter J_{k-1}, M_k, L_k and innovation term $\tilde{z}_k$. First recall that when $R_v \gg R_n$,

$$\hat{z}_k^- = \lambda\{\tilde{C}\hat{x}_{k-1} + \text{Proj.}[\tilde{v}_k | \tilde{Z}_1^{k-1}]\}.$$

Along the same line, when $R_v \ll R_n$ we need to calculate

$$\hat{z}_k^- = \lambda\tilde{C}\hat{x}_{k-1} + \text{Proj.}[\overline{n}_k | \tilde{Z}_1^{k-1}].$$

In this situation,

$$\tilde{Z}_1^{k-1} \sim (x_0; \omega_0, \ldots, \omega_{k-2}; n_1, \ldots, n_{k-1}; \gamma_1, \ldots, \gamma_{k-1}),$$

which means $\text{Proj.}[\overline{n}_k | \tilde{Z}_1^{k-1}] = 0$ and thus $\hat{z}_k^- = \lambda\tilde{C}\hat{x}_{k-1}$. Now different from

$$J_{k-1} = \lambda\{P_{k-1}\tilde{C}' + \mathbb{E}[e_{k-1}\tilde{v}_k']\},$$

when $R_v \gg R_n$, we have

$$J_{k-1} = \lambda P_{k-1}\tilde{C}' + \mathbb{E}[e_{k-1}\overline{n}_k'],$$

when $R_v \ll R_n$. Since

$$e_{k-1} \sim (x_0; \omega_0, \ldots, \omega_{k-2}; n_1, \ldots, n_{k-1}; \gamma_1, \ldots, \gamma_{k-1}),$$

one has $\mathbb{E}[e_{k-1}\overline{n}_k'] = 0$ and $J_{k-1} = \lambda P_{k-1}\tilde{C}'$. With some calculations as in the $R_v \gg R_n$ case, we derive the complete recursive estimation in Algorithm 2, which we denote as $(\hat{x}_k, P_k) = \Upsilon(\hat{x}_{k-1}, P_{k-1}, S_{k-1}, \lambda, \overline{\Omega}, \overline{R}_n)$.

Remark 4.6. When $\alpha = 1$, for unknown $\{\gamma_k\}$ and $R_v \ll R_n$, Algorithm 2 reduces to the one developed by Nahi (1969).

Algorithm 2 Estimation algorithm for $R_v \ll R_n$

1: Initial condition: $\overline{\Omega} \triangleq \lambda\alpha QC'$, $\overline{R}_n \triangleq \lambda\alpha^2 CQC' + R_n$, $\tilde{C} \triangleq C[\alpha A + (1-\alpha)I]$, $x_0 \sim \mathcal{N}(0, \Pi_0)$, $S_0 = \Pi_0$, $\hat{x}_0 = 0$ and $P_0 = 0$.
When $k \geq 1$,

2: LMMSE estimation of $z_k^E = \gamma_k z_k$ given Z_1^{k-1}
$$\hat{z}_k^- = \lambda\tilde{C}\hat{x}_{k-1},$$

3: Covariance of x_{k-1} and $\tilde{z}_k$ given Z_1^{k-1}
$$J_{k-1} = \lambda P_{k-1}\tilde{C}',$$

4: Variance of $\tilde{z}_k$ and $\tilde{z}_k$ given Z_1^{k-1}
$$L_k = \lambda\tilde{C}P_{k-1}\tilde{C}' + \lambda(1-\lambda)\tilde{C}(S_{k-1} - P_{k-1})\tilde{C}' + \overline{R}_n,$$

5: Smooth step
$$\hat{x}_{k-1}^+ = \hat{x}_{k-1} + J_{k-1}L_k^{-1}(\gamma_k z_k - \hat{z}_k^-),$$
$$P_{k-1}^+ = P_{k-1} - J_{k-1}L_k^{-1}J_{k-1}',$$

6: Covariance of ω_{k-1} and $\tilde{z}_k$ given Z_1^{k-1}
$$M_k = \overline{\Omega},$$

7: Estimate step
$$\hat{x}_k = A\hat{x}_{k-1}^+ + M_k L_k^{-1}(\gamma_k z_k - \hat{z}_k^-),$$
$$P_k = AP_{k-1}A' + Q - (AJ_{k-1} + M_k)L_k^{-1}(AJ_{k-1} + M_k)',$$

8: Recursion for $S_k \triangleq \mathbb{E}[x_k x_k']$
$$S_k = AS_{k-1}A' + Q.$$

Remark 4.7. Both in Algorithms 1 and 2, one may find that $\{P_k\}$ rely on $\{S_k\}$ which comes from the unknown $\{\gamma_k\}$. The same issue arises in Costa (1994), Nahi (1969) and Sun et al. (2008). Suppose the system is unstable, $\{S_k\}$ diverges. In other words, L_k^{-1} tends to zero and the estimation becomes open-loop prediction as $k \to \infty$. To overcome this difficulty, one may consider adding a reliable control channel to limit the state, e.g.,

$$x_{k+1} = Ax_k - BG_k\hat{x}_k + \omega_k.$$

Similar idea was considered by Gupta, Hassibi, and Murray (2007) and Tatikonda and Mitter (2004).

5. Examples

Consider the following illustrative example when $R_v \gg R_n$: $A = [0.98\ 0.01;\ 0\ 0.97]$; $C = [1\ 0;\ 0\ 1]$; $Q = [0.04\ 0;\ 0\ 0.04]$; $R_v = [2.4\ 0;\ 0\ 2.4]$. To verify the filtering algorithms derived in this paper we choose $\alpha = 0.8$, $\lambda = 0.75$ and $\Pi_0 = \text{diag}(0.3, 0.3)$. The state and its estimate for every component is shown in Fig. 3 and the trace of estimation error covariance is shown in Fig. 4.

To study the influence of α on estimation performance, let $P_\infty = \lim_{k\to\infty} P_k$, $S_\infty = \lim_{k\to\infty} S_k$. Then we are interested in the performance improvement rate of linear temporal coding strategy compared with usual scheme $z_k^E = \gamma_k y_k$, defined as follows

$$\beta \triangleq \frac{\text{Tr}(P_\infty(\alpha = 1)) - \text{Tr}(P_\infty(\alpha))}{\text{Tr}(P_\infty(\alpha = 1))}.$$

Fig. 5 shows the relationship between $\text{Tr}(P_\infty)$ and α for $\lambda = 0.3$, where $\alpha^\star = 0.543$ and corresponding $\beta^\star$ is 9.1%. For general packet arrival rate, $\alpha^\star$ versus λ is shown in Fig. 6 and $\beta^\star$ versus λ is given by Fig. 7. Fig. 6 implies when network is more unreliable, one should increase the weight of y_{k-1}. On the other hand, one expects that compared with y_{k-1}, the current measurement y_k always contributes more for decreasing the estimation error. In other words, in $z_k = \alpha y_k + (1-\alpha)y_{k-1}$, we have $\alpha^\star \geq 0.5$.

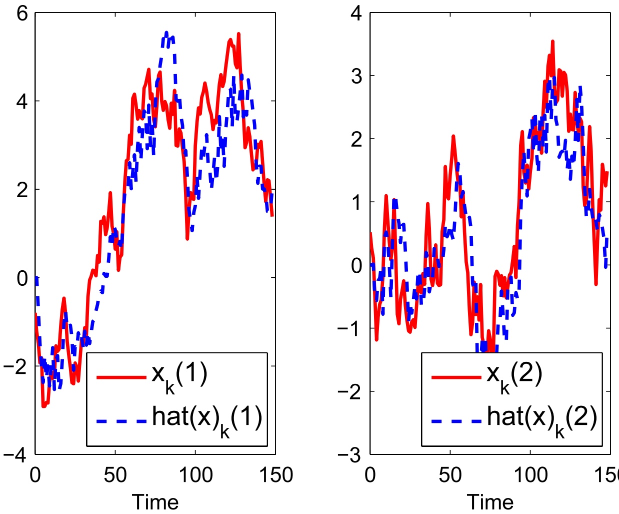

Fig. 3. State and estimate for each component.

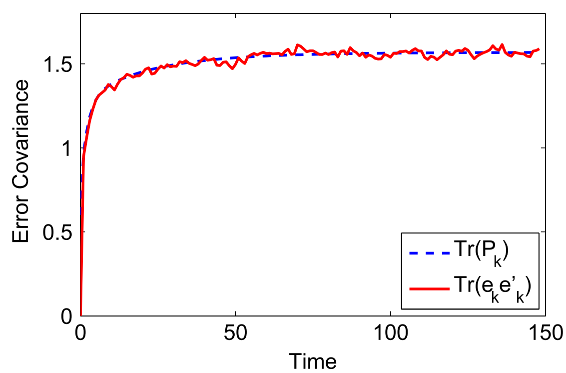

Fig. 4. Trace of the estimation error covariance.

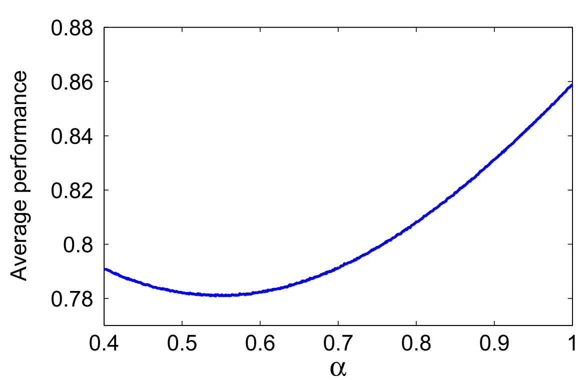

Fig. 5. $\mathrm{Tr}(P_\infty)$ versus α for $R_v \gg R_n$, $\lambda = 0.3$.

Fig. 7 shows that it is better to choose linear temporal coding when the network is more unreliable which meets with our intuition. However when the packet arrival rate approaches to zero, the improvement rate contributed by this scheme drops which is also understandable. For example, consider the extreme case when $\lambda = 0$, then it is clear that the proposed scheme does not provide any benefit.

When $R_v \ll R_n$, for scalar system, we can solve the following optimization problem

$$\min_{\alpha\in[0,1]} P_\infty$$

$$\text{s.t. } P_\infty = \Upsilon(P_\infty, S_\infty, \lambda, \overline{\Omega}, \overline{R}_n). \tag{38}$$

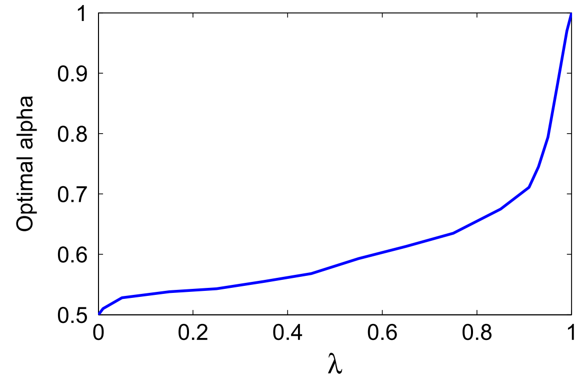

Fig. 6. $\alpha^\star$ versus λ for $R_v \gg R_n$.

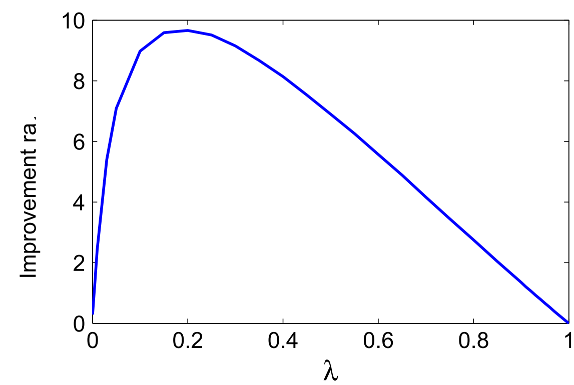

Fig. 7. $\beta^\star$ versus λ for $R_v \gg R_n$.

Let $\frac{\partial P_\infty}{\partial \alpha} = 0$, with some direct calculation, we get

$$\begin{aligned}\alpha^\star &= \frac{Q[\lambda P_\infty + (1-\lambda)S_\infty] + \frac{R_n}{\lambda C^2}[A^2P_\infty - AP_\infty + Q]}{Q\{(1-\lambda)A(P_\infty - S_\infty) + [\lambda P_\infty + (1-\lambda)S_\infty]\}} \\ &= 1 + \frac{\frac{R_n}{\lambda C^2}[A^2P_\infty - AP_\infty + Q] + (1-\lambda)AQ(S_\infty - P_\infty)}{Q\{[\lambda I + (1-\lambda)A]P_\infty + (1-\lambda)(I-A)S_\infty\}}.\end{aligned} \tag{39}$$

From Algorithm 2, we have

$$P_\infty = AP_\infty A' + Q - (AJ_\infty + M_\infty)L_\infty^{-1}(AJ_\infty + M_\infty)'. \tag{40}$$

For a stable scalar system,

$$A^2P_\infty + Q = P_\infty + \frac{(AJ_\infty + M_\infty)^2}{L_\infty} \geq P_\infty > AP_\infty,$$

which means $A^2P_\infty - AP_\infty + Q > 0$. From $S_\infty = AS_\infty A' + Q$ and (40), we have $S_\infty - P_\infty > 0$. From (39) and recall that $\alpha \in [0, 1]$, we get $\alpha^\star = 1$. In other words, different from the $R_v \gg R_n$ case, when $R_v \ll R_n$, there is no benefit to use linear temporal coding strategy. Consider the following system whose parameters are: $A = 0.97$, $C = 1$, $Q = 0.04$, $R_n = 2.4$, $\Pi_0 = 0.3$. When $\lambda = 0.3$, the relationship between P_∞ and α is shown in Fig. 8.

6. Conclusion

In this paper, we considered the LMMSE estimation problem over an unreliable network. We used the linear temporal coding strategy to improve the estimation performance when the sensor data is communicated to a remote estimator over a packet-dropping network. To cope with colored noise from measurement combination, after comparing with the standard state-augmentation approach and the measurement differencing approach, we choose orthogonal projection principle and innovation sequence approach to get the recursive estimation algorithm.

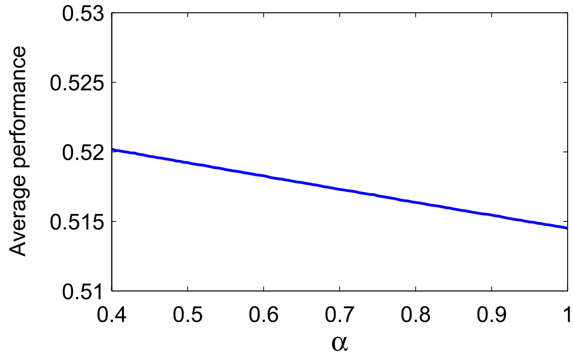

Fig. 8. P_∞ versus α for $R_v \ll R_n$, $\lambda = 0.3$.

For large measurement noise variance case we show the optimal weight parameter relies on packet arrival rate via an illustrative example. By contrast, for scalar system we prove there is no benefit to choose this strategy for the large communication noise case.

There are many interesting directions for continuing this work. When $R_v \gg R_n$, derive a closed-form expression for $\alpha^\star$ with respect to system parameters and network condition; find a general scheme that outperforms the one developed in this paper, e.g., $z_k = \alpha_0^\star y_k + \alpha_1^\star y_{k-1} + \cdots + \alpha_D^\star y_{k-D}$; when $R_v \ll R_n$, extend the proof to higher-order system and experimentally evaluate the theory developed in this paper.

References

Anderson, B., & Moore, J. (1979). *Optimal filtering*. Englewood Cliffs, NJ: Prentice Hall.
Bryson, A., & Henrikson, L. (1968). Estimation using sampled data containing sequentially correlated noise. *Journal of Spacecraft and Rockets*, *5*(6), 662–665.
Costa, O. (1994). Linear minimum mean square error estimation for discrete-time Markovian jump linear systems. *IEEE Transactions on Automatic Control*, *39*(8), 1685–1689.
Costa, O., Fragoso, M., & Marques, R. (2005). *Discrete time Markov jump linear systems*. Springer Verlag.
Culler, D., Estrin, D., & Srivastava, M. (2004). Guest editors' introduction: overview of sensor networks. *Computer*, *37*(8), 41–49.
Dey, S., Leong, A., & Evans, J. (2009). Kalman filtering with faded measurements. *Automatica*, *45*(10), 2223–2233.
Fletcher, A., Rangan, S., & Goyal, V. (2004). Estimation from lossy sensor data: jump linear modeling and kalman filtering. In *Proceedings of the 3rd international symposium on information processing in sensor networks* (pp. 251–258). ACM.
Gupta, V., Hassibi, B., & Murray, R. (2007). Optimal lqg control across packet-dropping links. *Systems & Control Letters*, *56*(6), 439–446.
Henrikson, L. (1968). Sequentially correlated measurement noise with applications to inertial navigation. *Ph.D. Thesis*. Harvard University.
Hespanha, J., Naghshtabrizi, P., & Xu, Y. (2007). A survey of recent results in networked control systems. *Proceedings of the IEEE*, *95*(1), 138–162.
Jiang, P., Zhou, J., & Zhu, Y. (2010). Globally optimal kalman filtering with finite-time correlated noises. In *Proceedings of the IEEE confonference on decision and control* (pp. 5007–5012). Atlanta, GA, USA.
Liang, Y., Chen, T., & Pan, Q. (2010). Optimal linear state estimator with multiple packet dropouts. *IEEE Transactions on Automatic Control*, *55*(6), 1428–1433.
Mahalik, N. (2007). *Sensor networks and configuration: fundamentals, standards, platforms and applications*. Springer Verlag.
Mesquita, A., Hespanha, J., & Nair, G. (2009). Redundant data transmission in control/estimation over wireless networks. In *Proceedings of the American control conference* (pp. 3378–3383). St. Louis, MO, USA.
Mostofi, Y., & Murray, R. (2009). To drop or not to drop: design principles for kalman filtering over wireless fading channels. *IEEE Transactions on Automatic Control*, *54*(2), 376–381.
Nahi, N. (1969). Optimal recursive estimation with uncertain observation. *IEEE Transactions on Information Theory*, *15*(4), 457–462.
Robinson, C., & Kumar, P. (2007). Sending the most recent observation is not optimal in networked control: linear temporal coding and towards the design of a control specific transport protocol. In *Proceedings of the IEEE confonference on decision and control* (pp. 334–339). New Orleans, LA, USA.
Sahebsara, M., Chen, T., & Shah, S. (2007). Optimal filtering with random sensor delay, multiple packet dropout and uncertain observations. *International Journal of Control*, *80*(2), 292–301.
Sun, S., Xie, L., Xiao, W., & Soh, Y. (2008). Optimal linear estimation for systems with multiple packet dropouts. *Automatica*, *44*(5), 1333–1342.
Tatikonda, S., & Mitter, S. (2004). Control over noisy channels. *IEEE Transactions on Automatic Control*, *49*(7), 1196–1201.
Zhang, W., Yu, L., & Feng, G. (2011). Optimal linear estimation for networked systems with communication constraints. *Automatica*, *47*(9), 1992–2000.

Lidong He received his B.S. degree in Mechanical Engineering from Zhejiang Ocean University, Zhoushan, China, in 2005 and Master's degree in Control Theory and Control Engineering from Northeastern University, Shenyang, China, in 2008. He is currently a Ph.D. candidate at Shanghai Jiao Tong University. His research interests include networked control systems and wireless sensor networks.

Dongfang Han received his Ph.D. degree in Mathematics from Shantou University, Shantou, China, in 2008. He worked as a Research Associate at the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, from January 2010 to December 2011.

He is currently a lecturer at the School of Mathematics and Statistics, South-Central University for Nationalities, Wuhan, China. His research interests include nonlinear systems, time-delayed systems and networked control systems.

Xiaofan Wang received the Ph.D. degree from Southeast University, China in 1996. He has been a Professor in the Department of Automation, Shanghai Jiao Tong University (SJTU) since 2002 and a Distinguished Professor of SJTU since 2008.

He received the 2002 National Science Foundation for Distinguished Young Scholars of P.R. China and the 2005 Guillemin-Cauer Best Transactions Paper Award from the IEEE Circuits and Systems Society. His current research interests include analysis and control of complex dynamical networks. He has (co)authored 3 books and more than 70 papers. He is a senior member of IEEE.

Ling Shi received his B.S. degree in Electrical and Electronic Engineering from the Hong Kong University of Science and Technology in 2002 and Ph.D. degree in Control and Dynamical Systems from California Institute of Technology in 2008. He is currently an Assistant Professor at the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology. His research interests include networked control systems, wireless sensor networks and distributed control.