

• Supplementary File •

## A compensation method for the packet loss deviation in system identification with event-triggered binary-valued observations

Jing-Dong DIAO<sup>1</sup>, Jin GUO<sup>1,2\*</sup> & Chang-Yin SUN<sup>3,4\*</sup>

<sup>1</sup>*School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, PR China;*

<sup>2</sup>*Key Laboratory of Knowledge Automation for Industrial Processes, Ministry of Education, Beijing 100083, PR China;*

<sup>3</sup>*School of Automation, Southeast University, Nanjing 210096, PR China;*

<sup>4</sup>*Key Laboratory of Measurement and Control of Complex System of Engineering, Ministry of Education, Nanjing 210096, PR China*

### Appendix A Proofs of Theorem 1 and Theorem 2

Let  $\mathcal{F}_k$  be the  $\sigma$ -algebra generated by  $d_1, \dots, d_k$ , i.e.,  $\mathcal{F}_k = \sigma(d_i : 1 \leq i \leq k)$ .

**Proof of Theorem 1:** From the system setting, one can see that the trigger indicator  $\gamma_k^e$  only have two possible values, i.e.,  $\gamma_k^e = 0$  or 1. When  $\gamma_k^e = 0$ ,  $s_k = \widehat{s}_k$ . Thus, there exists such an identical relation:

$$\gamma_k^e s_k + (1 - \gamma_k^e) \widehat{s}_k = s_k,$$

wherein one can obtain

$$\gamma_k^e (s_k - \widehat{s}_k) = s_k - \widehat{s}_k. \quad (\text{A1})$$

By (5), it is known that  $k\xi_k - (k-1)\xi_{k-1} = \eta_k$ , and then

$$[k\xi_k - (k-1)\xi_{k-1}] + [(k-1)\xi_{k-1} - (k-2)\xi_{k-2}] + \dots + [2\xi_2 - \xi_1] + [\xi_1 - 0] = k\xi_k = \sum_{i=1}^k \eta_k.$$

Together with (4), one can have

$$\begin{aligned} \xi_k &= \frac{1}{k} \sum_{i=1}^k \{ \gamma_k s_k - \gamma_k \widehat{s}_k + (1-p)\widehat{s}_k \} \\ &= \frac{1}{k} \sum_{i=1}^k \{ \gamma_k^d \gamma_k^e (s_k - \widehat{s}_k) + (1-p)\widehat{s}_k \}. \end{aligned}$$

By virtue of (A1), it follows that

$$\begin{aligned} \xi_k &= \frac{1}{k} \sum_{i=1}^k \{ \gamma_k^d (s_k - \widehat{s}_k) + (1-p)\widehat{s}_k \} \\ &= \frac{1}{k} \sum_{i=1}^k \{ \gamma_k^d s_k - \gamma_k^d \widehat{s}_k + (1-p)\widehat{s}_k \} \\ &= \frac{1}{k} \sum_{i=1}^k \gamma_k^d s_k + \frac{1}{k} \sum_{i=1}^k (1-p - \gamma_k^d) \widehat{s}_k. \end{aligned}$$

Notice that

$$\begin{aligned} E\{(1-p - \gamma_k^d)\widehat{s}_k | \mathcal{F}_{k-1}\} &= \widehat{s}_k E\{1-p - \gamma_k^d | \mathcal{F}_{k-1}\} \\ &= \widehat{s}_k E\{1-p - \gamma_k^d\} \\ &= 0. \end{aligned}$$

---

\* Corresponding author (email: guojin@ustb.edu.cn (Jin GUO), cysun@seu.edu.cn (Chang-Yin SUN))

Therefore,  $\{(1-p-\gamma_k^d)\widehat{s}_k, \mathcal{F}_k, k \geq 1\}$  is a martingale difference sequence.

Note that  $|1-p-\gamma_k^d| \leq 1$ ,  $|\widehat{s}_k| \leq 1$  and then

$$\sum_{k=1}^{\infty} \frac{E((1-p-\gamma_k^d)\widehat{s}_k)^2}{k^2} \leq 4 \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty.$$

By Corollary 2 on Page 397 of [1], we have

$$\frac{1}{k} \sum_{i=1}^k (1-p-\gamma_k^d)\widehat{s}_k \rightarrow 0, \quad w.p.1 \text{ as } k \rightarrow \infty. \quad (\text{A2})$$

Since  $\gamma_k^d$  and  $\widehat{s}_k$  are both i.i.d. and uncorrelated to each other, by the Law of Large Number, one can get

$$\frac{1}{k} \sum_{i=1}^k \gamma_k^d s_k \rightarrow (1-p)F(C-\theta), \quad w.p.1 \text{ as } k \rightarrow \infty.$$

Together with (A2), it can be seen that

$$\xi_k \rightarrow (1-p)F(C-\theta), \quad w.p.1 \text{ as } k \rightarrow \infty,$$

which completes the proof together with (6).  $\square$

**Proof of Theorem 2:** i) In view of (2), we have

$$\begin{aligned} \widehat{\gamma}_k^e &:= \Pr(\gamma_k^e = 1 | \mathcal{F}_{k-1}) \\ &= \Pr(s_k \neq \widehat{s}_k | \mathcal{F}_{k-1}) \\ &= \Pr(\widehat{s}_k \neq 1 | \mathcal{F}_{k-1}) \Pr(s_k = 1) + \Pr(\widehat{s}_k \neq 0 | \mathcal{F}_{k-1}) \Pr(s_k = 0) \\ &= I_{\{\widehat{\theta}_{k-1} > C\}} F(C-\theta) + I_{\{\widehat{\theta}_{k-1} \leq C\}} (1 - F(C-\theta)). \end{aligned}$$

This implies that  $E\{\gamma_k^e - \widehat{\gamma}_k^e | \mathcal{F}_{k-1}\} = 0$ , and hence  $\{\gamma_k^e - \widehat{\gamma}_k^e, \mathcal{F}_k, k \geq 1\}$  is a martingale difference sequence.

By  $\gamma_k^e \leq 1$ ,  $\widehat{\gamma}_k^e \leq 1$ ,  $\sum_{k=1}^{\infty} \frac{E(\gamma_k^e - \widehat{\gamma}_k^e)^2}{k^2} \leq 4 \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$  and Corollary 2 on Page 397 of [1] again, we have

$$\frac{\sum_{i=1}^k (\gamma_i^e - \widehat{\gamma}_i^e)}{k} \rightarrow 0, \quad w.p.1 \text{ as } k \rightarrow \infty. \quad (\text{A3})$$

By virtue of Theorem 1 and  $\bar{\gamma}^e := I_{\{\theta > C\}} F(C-\theta) + I_{\{\theta \leq C\}} (1 - F(C-\theta)) = \widetilde{F}(C-\theta)$ , one can have  $\widehat{\gamma}_k^e - \bar{\gamma}^e \rightarrow 0$  as  $k \rightarrow \infty$ , which together with (A3) gives that

$$\frac{\sum_{i=1}^k (\gamma_i^e - \bar{\gamma}^e)}{k} \rightarrow 0, \quad w.p.1 \text{ as } k \rightarrow \infty. \quad (\text{A4})$$

The proof can be obtained by the above.

ii) By Assumption 2,  $\gamma_k^e$  and  $\gamma_k^d$  are uncorrelated, then one can get

$$\begin{aligned} \widehat{\gamma}_k &:= \Pr(\gamma_k = 1 | \mathcal{F}_{k-1}) \\ &= \Pr(\gamma_k^e = 1, \gamma_k^d = 1 | \mathcal{F}_{k-1}) \\ &= \Pr(\gamma_k^d = 1) \Pr(\gamma_k^e = 1 | \mathcal{F}_{k-1}) \\ &= (1-p)\widehat{\gamma}_k^e, \end{aligned}$$

and  $E\{\gamma_k^e \gamma_k^d - \widehat{\gamma}_k | \mathcal{F}_{k-1}\} = 0$ . Consequently,  $\{\gamma_k^e \gamma_k^d - \widehat{\gamma}_k, \mathcal{F}_k, k \geq 1\}$  is also a martingale difference sequence. Repeating the proving process of i), it follows that

$$\frac{\sum_{i=1}^k (\gamma_i^e \gamma_i^d - \bar{\gamma})}{k} \rightarrow 0, \quad w.p.1 \text{ as } k \rightarrow \infty,$$

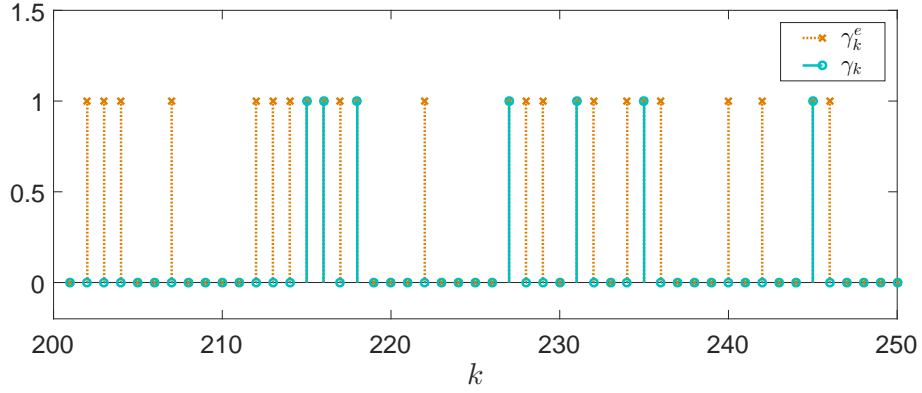
which leads to  $\frac{1}{k} \sum_{i=1}^k \gamma_i^e \gamma_i^d \rightarrow \bar{\gamma} = (1-p)\widetilde{F}(C-\theta)$ , w.p.1 as  $k \rightarrow \infty$ .  $\square$

## Appendix B Simulation Results

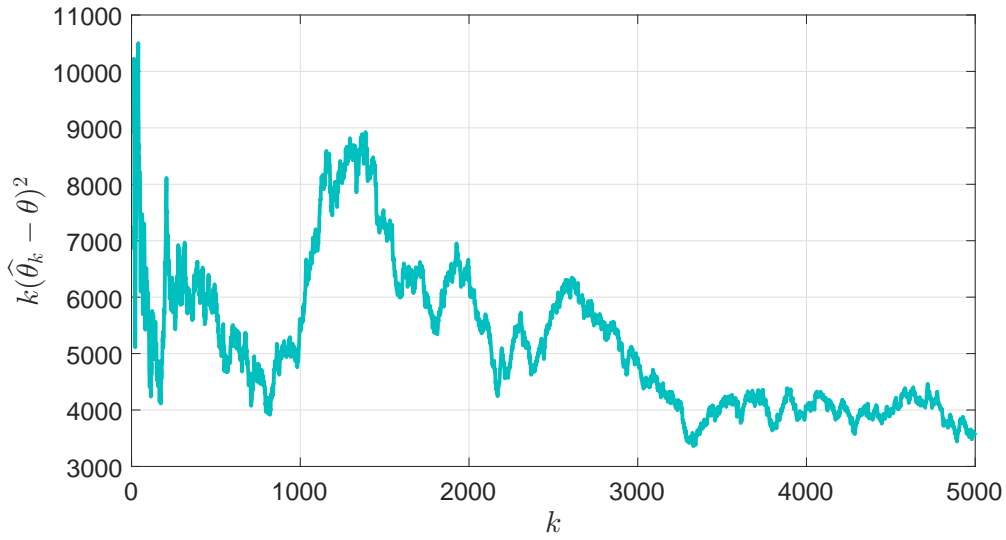
Fig. B1 shows the triggered transmission and data packet receiving moment on the time interval [200, 250], we can see that the receiving number is less than transmission times due to packet loss. Fig. B2 displays the average of 20 trajectories of  $k(\widehat{\theta}_k - \theta)^2$ , which implies that the algorithm maybe have the convergence rate  $E(\widehat{\theta}_k - \theta)^2 = O(1/k)$ . Moreover, if we do not consider the packet loss, the estimate will have bias as illustrated in Fig. B3.

## References

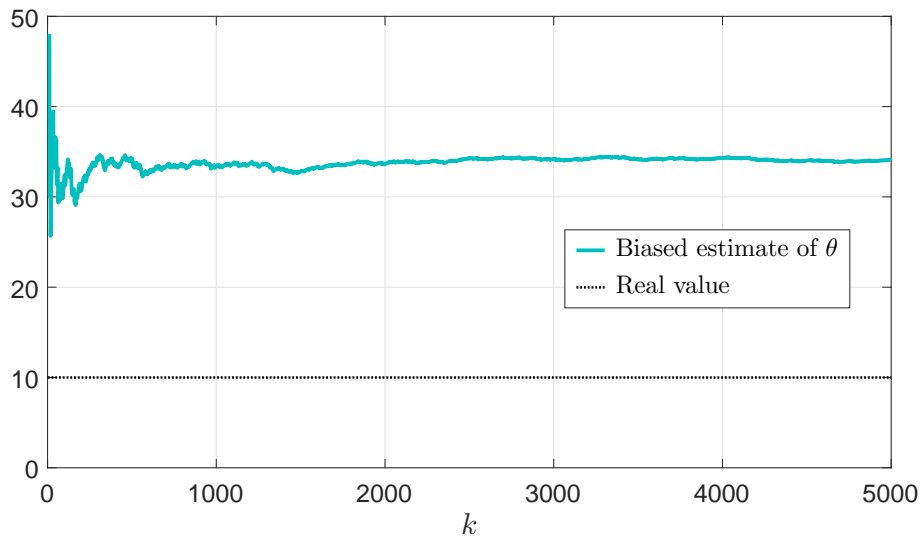
- 1 Chow Y S, Teicher H. Probability Theory: Independence, Interchangeability, Martingales. Springer-Verlag, New York, 2nd Edition, 1997



**Figure B1** Triggered transmission and receiving moment on the time interval  $[200, 250]$ , the orange dotted line represents  $\gamma_k^e$ , the green solid line represents  $\gamma_k$ .



**Figure B2** The average of 20 trajectories of  $k(\hat{\theta}_k - \theta)^2$ .



**Figure B3** Biased estimation without considering packet loss.