SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

January 2020, Vol. 63 112201:1–112201:12
https://doi.org/10.1007/s11432-018-9845-6

Prediction-based event-triggered identification of quantized input FIR systems with quantized output observations

Jin GUO[1,2*] & Jing-Dong DIAO[1,2]

[1] *School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing* 100083*, China;*
[2] *Key Laboratory of Knowledge Automation for Industrial Processes, Ministry of Education, Beijing* 100083*, China*

Received 24 October 2018/Revised 17 January 2019/Accepted 4 March 2019/Published online 16 December 2019

Abstract This paper addresses the identification of finite impulse response (FIR) systems with both quantized and event-triggered observations. An event-triggered communication scheme for the binary-valued output quantization is introduced to save communication resources. Combining the empirical-measure-based identification technique and the weighted least-squares optimization, an algorithm is proposed to estimate the unknown parameter by full use of the received data and the not-triggered condition. Under quantized inputs, it is shown that the estimate can strongly converge to the real values and the estimator is asymptotically efficient in terms of the Cramér-Rao lower bound. Further, the limit of the average communication rate is derived and the tradeoff between this limit and the estimation performance is discussed. Moreover, the case of multi-threshold quantized observations is considered. Numerical examples are included to illustrate the obtained main results.

Keywords identification, FIR systems, prediction-based event-triggered communication, quantized observations, convergence performance

Citation Guo J, Diao J-D. Prediction-based event-triggered identification of quantized input FIR systems with quantized output observations. Sci China Inf Sci, 2020, 63(1): 112201, https://doi.org/10.1007/s11432-018-9845-6

1 Introduction

Statistical methods are used in the field of system identification to build mathematical models of dynamical systems based on the measured data. The steps include data acquisition, model structure selection, model parameter estimation, and model validation. After the development during more than half a century, significant achievements have been obtained. Their practical applications can be found in almost all areas of science and technology. Many algorithms, such as least-squares algorithms, the maximum likelihood estimation and the minimum-mean-square error estimation, have been introduced and improved gradually [1, 2].

Mathematically, for a given system input sequence, a parameter estimation algorithm can be seen as a mapping from the observation sequence $\{z_k : k \in \mathcal{K}\}$ to the set of possible values of the unknown parameter (denoted by $\mathcal{O}$), which can be described by

$$\mathcal{K} \times \mathcal{Z} \to \mathcal{O},$$

* Corresponding author (email: guojin@ustb.edu.cn)

where $\mathcal{Z}$ represents the value domain of z_k. In practice, the complexities of $\mathcal{K}$ and $\mathcal{Z}$ usually determine resources such as the storage space, transmission bandwidth, and energy consumption. Within the allowable range of estimation accuracy, it is considerably interesting and important to select low complexities of $\mathcal{K}$ and $\mathcal{Z}$, to reduce the communication burden, the computation load, and the memory cost [3–7].

The event-triggered communication scheme is an effective and frequently used way for selecting the desired measured data, which is sampled only when a designed "event" occurs. Because such scheme can simplify the measurement index set $\mathcal{K}$ while guaranteeing a good system performance, it has been widely used in the study of control theory [8–14]. In [8], it was shown that Lebesgue sampling (i.e., event-triggered sampling) could achieve better performance than Riemann sampling (i.e., time-triggered sampling) in case of some simple systems. Ref. [9] focused on the event-triggered gradient-based algorithm for a distributed optimization problem in which the multi-agent system was subjected to the state consensus constraint over directed networks. Ref. [10] investigated the set-valued Kalman filters with multiple sensor measurements for linear time-invariant systems and their application to perform event-based state estimation. Ref. [11] presented basic concepts and recent research directions with respect to the stability of sampled-data systems with aperiodic sampling. Additionally, Ref. [12] used the adaptive dynamic programming to study the event-driven optimal control for uncertain nonlinear systems with external disturbance. Ref. [13] introduced an event-triggered scheme to address the synchronization problem of complex networks with the random switching topologies, and Ref. [14] investigated the encirclement control by employing the newly developed bearing rigidity theory and the event-triggered mechanism.

The quantization is a process from a large set (such as the set of all real numbers) to a smaller set (often a finite set of discrete values); hence, it can substantially simplify the observation value domain $\mathcal{Z}$. This can be directly grasped from the case of binary-valued quantized observation in which $\mathcal{Z} = \{0, 1\}$ [15–17]. With the rapid advancement in micro-sensors, communication technologies and many frontier fields, the system identification with quantized observations has received considerable research attention [16–23]. Ref. [16] studied the system identification with only binary-valued sensors, moreover, the optimal identification errors, time complexity, optimal input design, and impact of disturbances and unmodeled dynamics on identification accuracy and complexity were examined in both stochastic and deterministic information frameworks. Ref. [17] discussed the asymptotically efficient non-truncated identification on finite impulse response (FIR) systems with binary-valued outputs. Under quantized inputs and quantized output observations, the identification of FIR systems was investigated in [18], where estimation algorithms and their convergence performance were established. Identification of Wiener systems with quantized observations was studied in [19, 20]. Ref. [21] addressed the problem of set membership system identification with quantized measurements. Ref. [22] considered linear system identification with batched binary-valued observations, and constructed an iterative parameter estimation algorithm to achieve the maximum likelihood estimate. Ref. [23] used supervised learning algorithms, such as support vector machines, to deal with the identification of systems based on the binary output measurements.

The integrated usage of an event-triggered scheme and quantization is an intuitive methodology for simplifying both $\mathcal{K}$ and $\mathcal{Z}$. However, these complexity reducing methods bring difficulties to algorithm designing. For example, the event-triggered communication scheme breaks the completeness of measurements, and the quantization denotes a many-to-few mapping, which is an inherently nonlinear and irreversible process. Furthermore, the design of the "event" inevitably depends on the quantized value of the measurement, which causes the coupling between the event-triggered scheme and the quantization process. Therefore, the strong correlation, the high quantization nonlinearity, and their coupling create substantial study difficulties and lead to few results in literature on systems identification with both quantized and event-triggered observations.

This paper focuses on FIR systems to investigate the identification with quantized and event-triggered observations. An event-triggered communication scheme is introduced for the binary-valued output quantization. Combining the empirical-measure-based identification technique and the weighted least-squares optimization, an identification algorithm is proposed to estimate the unknown parameter by making the most of the received data, the triggered indicator, the quantization threshold, and the statistical property of the system noise. The algorithm is proved to be strongly convergent under quantized inputs. Its

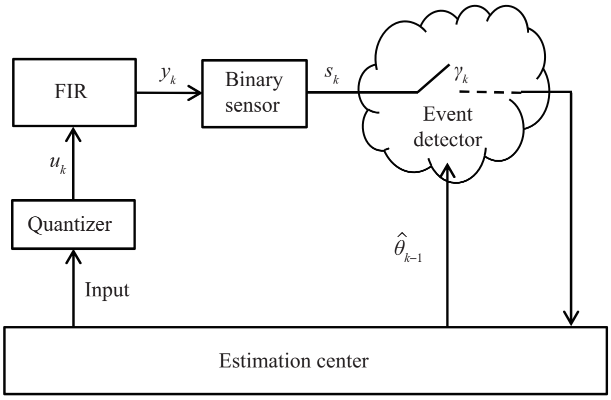

Figure 1 System setup.

mean-square convergence rate and asymptotic efficiency are also established in terms of the Cramér-Rao lower bound. The limit of the average communication rate is derived, and some of its properties are presented. Subsequently, the tradeoff between the communication cost and the estimation performance is formulated as a constrained minimization problem and discussed. The method and results are extended to the case of multi-threshold quantized observations as well.

Compared with the existing study on the system identification with quantized observations, the main contributions of this paper can be given as follows. (i) With respect to the research purpose, majority of the existing studies focuses on the design of the identification algorithm and the convergence analysis. This paper intends to present event-triggered schemes not only for the parameter identification but also for reducing communication resources. (ii) With respect to the algorithm structure, the available information in this study is both event-triggered and quantized, which exhibits stronger correlation than the data in the previous studies, which were only quantized. (iii) With respect to the research content, the previous studies presented the identification algorithm and the convergence performance. This paper not only establishes these, but also proposes a new event-triggered scheme, derives its communication rate, and discusses how to balance the communication rate and the identification precision.

The remainder of the paper is arranged into the following sections. Section 2 describes the identification problem with binary-valued quantized and event-triggered output observations under quantized inputs. Section 3 designs the parameter estimation algorithm and establishes the strong convergence and the asymptotic efficiency of the algorithm. Section 4 discusses the tradeoff between the communication cost and the estimation performance. Section 5 considers the case of multi-threshold quantized observations. Section 6 simulates a numerical example to demonstrate the effectiveness of the algorithm and the obtained main results. Section 7 discusses the findings of the paper as well as the future work.

2 Problem formulation

Consider a single-input-single-output FIR system

$$y_k = a_1 u_k + \cdots + a_n u_{k-n+1} + d_k = \phi_k^{\mathrm{T}} \theta + d_k,\ k = 1, 2, \ldots, \tag{1}$$

where $\phi_k = [u_k, \ldots, u_{k-n+1}]^{\mathrm{T}}$ is the regressor, $\theta = [a_1, \ldots, a_n]^{\mathrm{T}}$ is the parameter vector to be identified, and the superscript T indicates the transpose of a vector or a matrix.

As depicted in Figure 1, the input is finitely quantized with r possible values, i.e., $u_k \in \mathcal{U} = \{\mu_1, \ldots, \mu_r\}$, resulting that the sequence $\{\phi_k^{\mathrm{T}}\}$ can only take values in $l = r^n$ possible (row vector) patterns denoted by

$$\mathcal{P} = \{\pi_1, \ldots, \pi_l\}.$$

For example, $\pi_1 = [\mu_1, \ldots, \mu_1, \mu_1]$, $\pi_2 = [\mu_1, \ldots, \mu_1, \mu_2]$. The output y_k is measured by a binary-valued sensor with threshold $C \in (-\infty, +\infty)$, which can be represented by

$$s_k = I_{\{y_k \leqslant C\}} = \begin{cases} 1, & y_k \leqslant C, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $I_{\{y_k \in A\}}$ is the indicator function for a given set A, i.e., $I_{\{y_k \in A\}} = 1$ if $y_k \in A$; otherwise, $I_{\{y_k \in A\}} = 0$.

Further, an event detector is employed to decide whether the binary-valued output observation s_k is transmitted to the estimator center (EC), and can also access $\widehat{\theta}_{k-1}$, which is either broadcasted by the EC or computed by itself, where $\widehat{\theta}_{k-1}$ denotes the estimate of θ at time $k-1$. Let γ_k denote the transmission indicator; i.e., $\gamma_k = 1$ indicates that the detector is triggered, and the sensor data s_k are transmitted to the EC at time k, whereas $\gamma_k = 0$ means the communication is denied. In this paper, γ_k is given by

$$\gamma_k = \begin{cases} 1, & s_k \neq \widehat{s}_k, \\ 0, & s_k = \widehat{s}_k, \end{cases} \tag{3}$$

where $\widehat{s}_k = I_{\{\phi_k^{\mathrm{T}} \widehat{\theta}_{k-1} \leqslant C\}}$ can be considered as the prediction of s_k based on the estimate $\widehat{\theta}_{k-1}$.

For the EC, the available information can be denoted as $\{\phi_k\}$, $\{\gamma_k\}$ and $\{\gamma_k s_k\}$, based on which an estimation algorithm is constructed for θ and the convergence properties are established. Further more, the limit of the average communication rate is obtained, and the tradeoff between it and the estimation quantity is discussed.

Assumption 1. The system noise $\{d_k\}$ is a sequence of i.i.d. (independent and identically distributed) random variables. The cumulative distribution function $F(\cdot)$ of d_1 is invertible, and the inverse function is twice continuously differentiable. The moment generating function of d_1 exists.

Remark 1. (i) The model order is assumed to be known at this instance. It can be obtained through parameter estimation method with quantized observations, together with the assistances such as cross-validation, information criteria, the F-test. (ii) The cumulative distribution function of d_k is supposed to be known in Assumption 1. We can draw on the parameterization technique in [18] to deal with the case in which $F(\cdot)$ is unknown. (iii) As can be observed from (3), the design idea of the event-triggered condition is based on the difference between the observation and its prediction value. If the prediction is right, the detector is not triggered, and the current observation is not transmitted to the EC. Otherwise, the detector is triggered, and s_k is transmitted.

3 Identification algorithm and convergence performance

For a given pattern $\pi_j \in \mathcal{P}$, let $\mathbb{K}_{k,j}$ denote the index set where $\phi_i = \pi_j^{\mathrm{T}}$, $i = 1, \ldots, k$. Therefore, $\mathbb{K}_{k,j}$ is a subset of $\{1, 2 \ldots, k\}$, given by

$$i \in \mathbb{K}_{k,j} \quad \text{if and only if} \quad \phi_i = \pi_j^{\mathrm{T}}, \quad i = 1, \ldots, k.$$

$\tau_{k,j}$ is used to denote the number of elements in $\mathbb{K}_{k,j}$, which implies

$$\tau_{k,j} = \sum_{i=1}^{k} I_{\{\phi_i^{\mathrm{T}} = \pi_j\}}, \quad j = 1, \ldots, l. \tag{4}$$

Thus, $\{\phi_1^{\mathrm{T}}, \ldots, \phi_k^{\mathrm{T}}\}$ contains $\tau_{k,j}$ copies of the pattern π_j and we have $\sum_{j=1}^{l} \tau_{k,j} = k$.

Assumption 2 ([18])**.** The input sequence u is a deterministic signal, and there exists $\beta_j \geqslant 0$ such that $\lim_{k\to\infty} \tau_{k,j}/k = \beta_j$, $j \in L = \{1, \ldots, l\}$. Without loss of generality, suppose that $\beta_j \neq 0$ for $j \in L_0 = \{1, \ldots, l_0\}$ and that $\beta_j = 0$ for $j = l_0 + 1, \ldots, l$.

Definition 1 ([18]). The pattern π_j is said to be persistent if $\beta_j > 0$. The input u is said to be persistently exciting if the matrix

$$\Psi = \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_{l_0} \end{bmatrix} \in \mathbb{R}^{l_0 \times n} \tag{5}$$

is full column rank.

Define $j_k^0 = \sum_{j=1}^{l} j I_{\{\phi_k^{\rm T}=\pi_j\}}$ and $\Psi_k = [\sqrt{\tau_{k,1}}\pi_1^{\rm T}, \ldots, \sqrt{\tau_{k,l}}\pi_l^{\rm T}]^{\rm T}$. Supposed that $\Psi_k^{\rm T}\Lambda\Psi_k$ is full rank where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_l) > 0$ is a weighting matrix; thus, an identification algorithm is introduced by

$$\eta_k = \gamma_k s_k + (1-\gamma_k)\widehat{s}_k, \tag{6}$$

$$\xi_{k,j} = \begin{cases} \xi_{k-1,j}, & j \neq j_k^0, \\ (1-\frac{1}{\tau_{k,j}})\xi_{k-1,j} + \frac{1}{\tau_{k,j}}\eta_k, & j = j_k^0, \end{cases} \tag{7}$$

$$\widehat{W}_k = \left[\sqrt{\tau_{k,1}}\left(C - F^{-1}(\xi_{k,1})\right), \ldots, \sqrt{\tau_{k,l}}\left(C - F^{-1}(\xi_{k,l})\right)\right]^{\rm T}, \tag{8}$$

$$\widehat{\theta}_k = \left(\frac{1}{k}\Psi_k^{\rm T}\Lambda\Psi_k\right)^{-1}\frac{1}{k}\Psi_k^{\rm T}\Lambda\widehat{W}_k, \tag{9}$$

where the initial value $\xi_{0,j} \in (0,1)$ for $j \in L$, C is the threshold in (2) and the distribution function $F(\cdot)$ is given by Assumption 1.

Represent the covariance matrix of the estimation error by Σ_k, i.e., $\Sigma_k = E(\widehat{\theta}_k - \theta)(\widehat{\theta}_k - \theta)^{\rm T}$, and let $H^* = \mathrm{diag}\left(\beta_1\rho_1^{-2}, \ldots, \beta_{l_0}\rho_{l_0}^{-2}\right)$, where $\rho_j^2 = \varrho(C - \pi_j\theta)$ for $j \in L$, $\varrho(z) = \frac{F(z)(1-F(z))}{f^2(z)}$ for $z \in \mathbb{R}$, and $f(\cdot)$ is the density function of the system noise, that is, $f(z) = \mathrm{d}F(z)/\mathrm{d}z$.

Theorem 1. Consider system (1) with binary-valued observations (2) and triggering mechanism (3) under Assumptions 1 and 2. If the input u is persistently exciting, then

(i) $\widehat{\theta}_k$ from the algorithm (6)–(9) converges strongly to the true value θ, i.e., $\widehat{\theta}_k \to \theta$, w.p.1 (with probability 1) as $k \to \infty$;

(ii) $k\Sigma_k \to (\Psi^{\rm T}H_1\Psi)^{-1}\Psi^{\rm T}H_2\Psi(\Psi^{\rm T}H_1\Psi)^{-1}$ as $k \to \infty$, where $H_1 = \mathrm{diag}(\lambda_1\beta_1, \ldots, \lambda_{l_0}\beta_{l_0})$ and $H_2 = \mathrm{diag}(\beta_1\lambda_1^2\rho_1^2, \ldots, \beta_{l_0}\lambda_{l_0}^2\rho_{l_0}^2)$.

Proof. If $\gamma_k = 1$, then by (6) it follows that $\eta_k = \gamma_k s_k + (1-\gamma_k)\widehat{s}_k = s_k$. If $\gamma_k = 0$, then from (3) it is known that $\widehat{s}_k = s_k$, and then $\eta_k = \gamma_k s_k + (1-\gamma_k)\widehat{s}_k = s_k$. In conclusion, we always have $\eta_k = s_k$, which together with (7) indicates that

$$\xi_{k,j} = \frac{1}{\tau_{k,j}}\sum_{i\in\mathbb{K}_{k,j}} s_i.$$

This implies that (6)–(9) is just the identification algorithm in [18] for the single threshold case, and the proof can be obtained by virtue of Theorems 5 and 7 of [18].

Lemma 1. The Cramér-Rao lower bound for estimating θ based on $\{s_1, \ldots, s_k\}$ is

$$\Sigma_k^{\rm CR} = \left(\sum_{j=1}^{l}\tau_{k,j}\pi_j^{\rm T}\pi_j\rho_j^{-2}\right)^{-1},$$

where $\tau_{k,j}$ is given by (4) for $j \in L$.

Proof. The lemma follows from setting $m = 1$ in Theorem 9 of [18].

Theorem 2. Under the condition of Theorem 1, if the weighting matrix Λ in (9) is selected as

$$\mathrm{diag}(\rho_1^{-2}, \ldots, \rho_{l_0}^{-2}, \lambda_{l_0+1}, \ldots, \lambda_l),$$

then the estimate $\widehat{\theta}_k$ is asymptotically efficient in the sense that

$$k\left(\Sigma_k - \Sigma_k^{\rm CR}\right) \to 0 \ \text{ as } \ k \to \infty.$$

Proof. Note that $H_1 = \mathrm{diag}(\lambda_1\beta_1, \ldots, \lambda_{l_0}\beta_{l_0})$ and $H_2 = \mathrm{diag}(\beta_1\lambda_1^2\rho_1^2, \ldots, \beta_{l_0}\lambda_{l_0}^2\rho_{l_0}^2)$. Because $\lambda_j = \rho_j^{-2}$ for $j \in L_0$, it can be verified that $H_1 = H_2 = H^*$ and $(\Psi^{\mathrm{T}} H_1 \Psi)^{-1}\Psi^{\mathrm{T}} H_2 \Psi(\Psi^{\mathrm{T}} H_1 \Psi)^{-1} = (\Psi^{\mathrm{T}} H^* \Psi)^{-1}$. By virtue of Theorem 1 and (10), we have

$$k\Sigma_k \to \left(\Psi^{\mathrm{T}} H^* \Psi\right)^{-1} \quad \text{as} \quad k \to \infty. \tag{10}$$

By Assumption 2, it is known that $\beta_j = 0$ for $j = l_0 + 1, \ldots, l$. According to Lemma 1, it can be seen that

$$\begin{aligned} k\Sigma_k^{\mathrm{CR}} &= \left(\sum_{j=1}^{l} \frac{\tau_{k,j}}{k}\pi_j^{\mathrm{T}}\pi_j\rho_j^{-2}\right)^{-1} \\ &\to \left(\sum_{j=1}^{l} \beta_j\pi_j^{\mathrm{T}}\pi_j\rho_j^{-2}\right)^{-1} = \left(\sum_{j=1}^{l_0} \beta_j\pi_j^{\mathrm{T}}\pi_j\rho_j^{-2}\right)^{-1} \\ &= \left(\Psi^{\mathrm{T}} H^* \Psi\right)^{-1}, \quad \text{as} \quad k \to \infty, \end{aligned}$$

which together with (10) completes the proof.

4 Tradeoff between the estimation performance and the communication cost

Because $\lim_{k\to\infty} k\Sigma_k = \lim_{k\to\infty} k\Sigma_k^{\mathrm{CR}} = (\Psi^{\mathrm{T}} H^* \Psi)^{-1}$ by Theorem 2, we can use $\left(\Psi^{\mathrm{T}} H^* \Psi\right)^{-1}$ as the estimation performance index. By the definition of H^*, it can be seen that

$$H^* = \mathrm{diag}\left(\beta_1\varrho^{-1}(C - \pi_1\theta), \ldots, \beta_{l_0}\varrho^{-1}(C - \pi_{l_0}\theta)\right),$$

which indicates that H^* is a function of C and $H^* = H^*(C)$.

Define the average communication rate of the event trigger (3):

$$\overline{\gamma}_k = \frac{1}{k}\sum_{i=1}^{k}\gamma_i, \quad k = 1, 2, \ldots. \tag{11}$$

To a certain extent, the limit of $\overline{\gamma}_k$ can reveal the ability of (3) in saving the communication resource, and be employed as the communication cost index. These will be provided further. For convenience, let

$$\widetilde{F}(z) = I_{\{z<0\}}F(z) + I_{z\geqslant 0}(1 - F(z)), \quad z \in \mathbb{R}, \tag{12}$$

where the function $F(\cdot)$ comes from Assumption 1.

Lemma 2. Consider an MDS (martingale difference sequence) $\{\chi_k, \mathcal{F}_k, k \geqslant 1\}$. If $E(\sum_{i=1}^k \chi_i)^2 < \infty$ and $\sum_{k=1}^{\infty} \frac{E\chi_k^2}{k^2} < \infty$, then

$$\frac{1}{k}\sum_{i=1}^{k}\chi_i \to 0, \ \text{w.p.1} \quad \text{as} \quad k \to \infty.$$

Proof. The argument of Corollary 2 on Page 397 of [24] carries over verbatim.

Theorem 3. Under the condition of Theorem 1, the average communication rate from the event trigger (3) is convergent, i.e.,

$$\overline{\gamma}_k \to \overline{\gamma}(C) = \sum_{j=1}^{l_0}\beta_j\widetilde{F}\left(C - \pi_j\theta\right), \quad \text{w.p.1} \quad \text{as} \quad k \to \infty. \tag{13}$$

Proof. Let $\mathcal{F}_{k-1}$ be the σ-algebra generated by $d_1, \ldots, d_{k-1}$, i.e., $\mathcal{F}_{k-1} = \sigma(d_i : 1 \leqslant i \leqslant k-1)$. Owing to (3), it can be seen that

$$\widehat{\gamma}_k = \Pr(\gamma_k = 1|\mathcal{F}_{k-1})$$

$$= \Pr(s_k \neq \widehat{s}_k | \mathscr{F}_{k-1})$$
$$= \Pr(\widehat{s}_k \neq 1 | \mathscr{F}_{k-1}) \Pr(s_k = 1) + \Pr(\widehat{s}_k \neq 0 | \mathscr{F}_{k-1}) \Pr(s_k = 0)$$
$$= I_{\{\phi_k^{\mathrm{T}} \widehat{\theta}_{k-1} > C\}} F(C - \phi_k^{\mathrm{T}} \theta) + I_{\{\phi_k^{\mathrm{T}} \widehat{\theta}_{k-1} \leqslant C\}} (1 - F(C - \phi_k^{\mathrm{T}} \theta)).$$

This implies that $E\{\gamma_k - \widehat{\gamma}_k | \mathcal{F}_{k-1}\} = 0$. Therefore, $\{\gamma_k - \widehat{\gamma}_k, \mathcal{F}_k, k \geqslant 1\}$ is an MDS.

Note that $\gamma_k \leqslant 1$, $\widehat{\gamma}_k \leqslant 1$ and then $\sum_{k=1}^{\infty} \frac{E(\gamma_k - \widehat{\gamma}_k)^2}{k^2} \leqslant 4 \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$. By Lemma 2, we have

$$\frac{\sum_{i=1}^{k} (\gamma_i - \widehat{\gamma}_i)}{k} \to 0, \quad \text{w.p.1} \quad \text{as} \quad k \to \infty. \tag{14}$$

By virtue of Theorem 1 and with $\widehat{\gamma}_k^0 = I_{\{\phi_k^{\mathrm{T}} \theta > C\}} F(C - \phi_k^{\mathrm{T}} \theta) + I_{\{\phi_k^{\mathrm{T}} \theta \leqslant C\}} (1 - F(C - \phi_k^{\mathrm{T}} \theta)) = \widetilde{F}(C - \phi_k^{\mathrm{T}} \theta)$ by (12), one can have $\widehat{\gamma}_k - \widehat{\gamma}_k^0 \to 0$ as $k \to \infty$, which together with Eq. (14) gives that

$$\frac{\sum_{i=1}^{k} (\gamma_i - \widehat{\gamma}_i^0)}{k} \to 0, \quad \text{w.p.1} \quad \text{as} \quad k \to \infty. \tag{15}$$

In view of (4), it can be seen that

$$\begin{aligned}
\sum_{i=1}^{k} \widehat{\gamma}_i^0 &= \sum_{i=1}^{k} \sum_{j=1}^{l} I_{\{\phi_i^{\mathrm{T}} = \pi_j\}} \widetilde{F}(C - \pi_j \theta) \\
&= \sum_{j=1}^{l} \left(\sum_{i=1}^{k} I_{\{\phi_i^{\mathrm{T}} = \pi_j\}} \right) \widetilde{F}(C - \pi_j \theta) \\
&= \sum_{j=1}^{l} \tau_{k,j} \widetilde{F}(C - \pi_j \theta).
\end{aligned}$$

By Assumption 2, it follows that $\frac{\sum_{i=1}^{k} \widehat{\gamma}_i^0}{k} = \sum_{j=1}^{l} \frac{\tau_{k,j}}{k} \widetilde{F}(C - \pi_j \theta) \to \sum_{j=1}^{l} \beta_j \widetilde{F}(C - \pi_j \theta)$ as $k \to \infty$. From (11) and (15), the proof can be obtained.

Proposition 1. Under the condition of Theorem 1, $\overline{\gamma}(C)$ is monotonically increasing on $(-\infty, \min_{1 \leqslant j \leqslant l_0} \pi_j \theta]$ and monotonically decreasing on $[\max_{1 \leqslant j \leqslant l_0} \pi_j \theta, \infty)$ with respect to C.

Proof. Because $F(z)$ is the cumulative distribution function of d_1, $F(z)$ is monotonically increasing on $(-\infty, \infty)$. In view of (12), it is known that

$$\widetilde{F}(z) = \begin{cases} F(z), & z < 0, \\ 1 - F(z), & z \geqslant 0, \end{cases} \tag{16}$$

and hence from (13) one can have

$$\overline{\gamma}(C) = \begin{cases} \sum_{j=1}^{l_0} \beta_j F(C - \pi_j \theta), & C \leqslant \min_{1 \leqslant j \leqslant l_0} \pi_j \theta, \\ \sum_{j=1}^{l_0} \beta_j (1 - F(C - \pi_j \theta)), & C \geqslant \max_{1 \leqslant j \leqslant l_0} \pi_j \theta. \end{cases}$$

Due to $\beta_j > 0$ for $j \in L_0$, the proof follows.

Example 1. Suppose $\theta = [-2, 5]^{\mathrm{T}}$. The noise is normally distribution with zero mean and standard deviation $\sigma = 5$. The quantized input is taken as $\pi_1 = [-1, -1]$, $\pi_2 = [-1, 2]$, $\pi_3 = [2, -1]$, $\pi_4 = [2, 2]$ and $\beta_1 = 0.1$, $\beta_2 = 0.3$, $\beta_3 = 0.3$, $\beta_4 = 0.3$ with $l_0 = 4$, which implies that $\max_{1 \leqslant j \leqslant l_0} \pi_j \theta = 12$, $\min_{1 \leqslant j \leqslant l_0} \pi_j \theta = -9$. By (13), we have

$$\overline{\gamma}(C) = 0.1 \widetilde{F}(C + 3) + 0.3 \widetilde{F}(C - 12) + 0.3 \widetilde{F}(C + 9) + 0.3 \widetilde{F}(C - 6),$$

whose graph is given by Figure 2. One can see that $\overline{\gamma}(C)$ is indeed monotonically decreasing on $[12, \infty)$ and monotonically increasing on $(-\infty, -9]$.

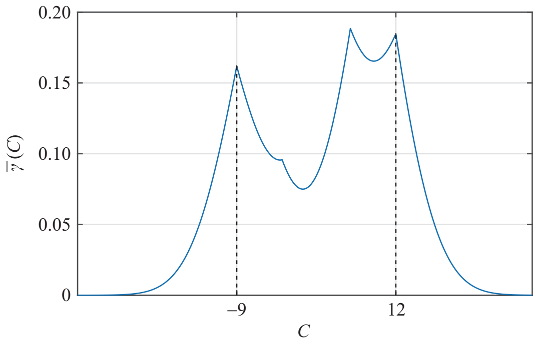

Figure 2 (Color online) The graph of $\overline{\gamma}(C)$ with respect to C.

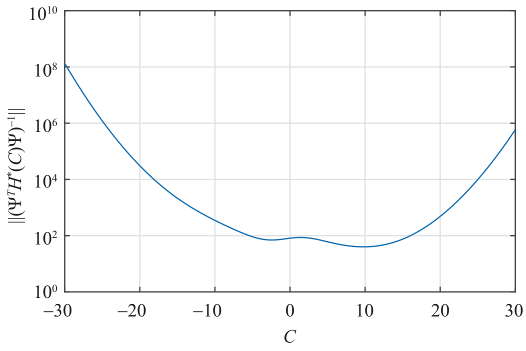

Figure 3 (Color online) The graph of $\|(\Psi^{\mathrm{T}} H^{*}(C)\Psi)^{-1}\|$ with respect to C.

In Example 1, we have seen that $\overline{\gamma}(C) \to 0$ as $C \to \infty$ or $C \to -\infty$. With

$$\Psi = \begin{bmatrix} -1 & -1 \\ -1 & 2 \\ 2 & -1 \\ 2 & 2 \end{bmatrix},$$

Figure 3 shows the graph of $\|(\Psi^{\mathrm{T}} H^{*}(C)\Psi)^{-1}\|$, where $\|\cdot\|$ denotes the Frobenius Norm and one can see that $\|(\Psi^{\mathrm{T}} H^{*}(C)\Psi)^{-1}\| \to \infty$ as $C \to \infty$ or $C \to -\infty$.

Then, how to balance the estimation performance and the communication cost is an interesting issue. This can be considered as a constrained minimization problem

$$\min_C \overline{\gamma}(C) \quad \text{s.t. } (\Psi^{\mathrm{T}} H^{*}(C)\Psi)^{-1} \leqslant \Delta. \tag{17}$$

Thus, for a given estimation accuracy Δ, Eq. (17) aims to solve the minimum communication cost. Generally, it is hard to obtain an explicit solution.

Let $\Omega_\Delta = \{C : (\Psi^{\mathrm{T}} H^{*}(C)\Psi)^{-1} \leqslant \Delta\}$, $\underline{C}_\Delta = \inf_{C\in\Omega_\Delta} C$ and $\overline{C}_\Delta = \sup_{C\in\Omega_\Delta} C$. Then we have the following proposition.

Proposition 2. Under the condition of Theorem 1, the following assertions hold.

(i) If $\underline{C}_\Delta \geqslant \max_{1\leqslant j\leqslant l_0} \pi_j\theta$, then $\inf_{C\in\Omega_\Delta} \overline{\gamma}(C) = \sum_{j=1}^{l_0} \beta_j \widetilde{F}\left(\overline{C}_\Delta - \pi_j\theta\right)$;

(ii) If $\overline{C}_\Delta \leqslant \min_{1\leqslant j\leqslant l_0} \pi_j\theta$, then $\inf_{C\in\Omega_\Delta} \overline{\gamma}(C) = \sum_{j=1}^{l_0} \beta_j \widetilde{F}\left(\underline{C}_\Delta - \pi_j\theta\right)$;

(iii) In general, we have $\inf_{C\in\Omega_\Delta} \overline{\gamma}(C) \geqslant \sum_{j=1}^{l_0} \beta_j \min\left\{\widetilde{F}\left(\underline{C}_\Delta - \pi_j\theta\right), \widetilde{F}\left(\overline{C}_\Delta - \pi_j\theta\right)\right\}$.

Proof. (i) If $\underline{C}_\Delta \geqslant \max_{1\leqslant j\leqslant l_0} \pi_j\theta$, then, according to Proposition 1, it is known that $\overline{\gamma}(C)$ is monotonically decreasing on $[\underline{C}_\Delta, \overline{C}_\Delta]$, which gives the desired result.

(ii) The proof is similar to the one of (i).

(iii) By (13), it is known that

$$\inf_{C\in\Omega_\Delta} \overline{\gamma}(C) \geqslant \sum_{j=1}^{l_0} \beta_j \inf_{C\in\Omega_\Delta} \widetilde{F}\left(C - \pi_j\theta\right). \tag{18}$$

On account of (16), we have $\inf_{C\in\Omega_\Delta} \widetilde{F}\left(C - \pi_j\theta\right) = \min\{\widetilde{F}\left(\underline{C}_\Delta - \pi_j\theta\right), \widetilde{F}\left(\overline{C}_\Delta - \pi_j\theta\right)\}$, which together with Eq. (18) completes the proof.

5 Multi-threshold quantized observations

This section considers the case of multi-threshold quantized output observations, where y_k is measured by a sensor of m thresholds $-\infty < C_1 < \cdots < C_m < \infty$. With $C_0 = -\infty$ and $C_{m+1} = \infty$, the observation sensor can be represented by

$$s_k = \sum_{q=1}^{m+1} q I_{\{y_k \in (C_{q-1}, C_q]\}}. \tag{19}$$

Hence, $s_k = q$ implies that $y_k \in (C_{q-1}, C_q]$ for $q = 1, \dots, m+1$. An alternative representation of (19) is to view the multi-threshold quantized observation as a vector-valued binary observation in which each vector component represents the output of one threshold, by defining

$$\widetilde{s}_k = [s_k^1, \dots, s_k^m]^{\mathrm{T}}, \quad \text{where} \;\; s_k^q = I_{\{-\infty < y_k \leqslant C_q\}}, \quad q = 1, \dots, m. \tag{20}$$

The triggering condition γ_k is given by

$$\gamma_k = \begin{cases} 1, \; s_k \neq \widehat{s}_k, \\ 0, \; s_k = \widehat{s}_k, \end{cases} \tag{21}$$

$$\widehat{s}_k = \sum_{q=1}^{m+1} q I_{\{\phi_k^{\mathrm{T}} \widehat{\theta}_{k-1} \in (C_{q-1}, C_q]\}}. \tag{22}$$

Let $\gamma_k s_k + (1-\gamma_k)\widehat{s}_k = \eta_k$, which is available for the following algorithm design:

$$\eta_k^q = \begin{cases} 1, \; \eta_k \leqslant q, \\ 0, \; \text{otherwise}, \end{cases} \quad q = 1, \dots, m.$$

Then it can be seen that $[\eta_k^1, \dots, \eta_k^m]^{\mathrm{T}} = \widetilde{s}_k$ by (20)–(22). As a result, we can use $\eta_k^1, \dots, \eta_k^m$ to construct m algorithms as (6)–(9), and obtain m estimators of θ, denoted by $\widehat{\theta}_k^1, \dots, \widehat{\theta}_k^m$. Define $\nu = [\nu_1, \dots, \nu_m]^{\mathrm{T}}$ such that $\nu_1 + \cdots + \nu_m = 1$. One can construct an estimator of θ by

$$\widehat{\theta}_k = \sum_{q=1}^{m} \nu_q \widehat{\theta}_k^q,$$

which is the QCCE (quasi-convex combination estimator) proposed by [25]. Under quantized inputs, the corresponding convergence performance and the asymptotic efficiency can be given as in [18]. We omit them to avoid unnecessary duplication.

The limit of the average communication rate from (21) is established by the following theorem.

Theorem 4. Consider system (1) with quantized observations (19) and triggering mechanism (21) under Assumptions 1 and 2. If the input u is persistently exciting, then $\overline{\gamma}_k$ from (21) converges to

$$\overline{\gamma} = \sum_{j=1}^{l_0} \beta_j \sum_{q=1}^{m+1} \Big(I_{\{\pi_j\theta \notin (C_{q-1}, C_q]\}} \left(F(C_q - \pi_j\theta) - F(C_{q-1} - \pi_j\theta) \right) \Big), \quad \text{w.p.1.}$$

Proof. Define $\widehat{\gamma}_k^0 = \sum_{q=1}^{m+1} I_{\{\phi_k^{\mathrm{T}}\theta \notin (C_{q-1}, C_q]\}} (F(C_q - \phi_k^{\mathrm{T}}\theta) - F(C_{q-1} - \phi_k^{\mathrm{T}}\theta))$. By (4), one can have

$$\begin{aligned} \sum_{i=1}^{k} \widehat{\gamma}_i^0 &= \sum_{i=1}^{k} \sum_{j=1}^{l} I_{\{\phi_k^{\mathrm{T}} = \pi_j\}} \sum_{q=1}^{m+1} \Big(I_{\{\pi_j\theta \notin (C_{q-1}, C_q]\}} \left(F(C_q - \pi_j\theta) - F(C_{q-1} - \pi_j\theta) \right) \Big) \\ &= \sum_{j=1}^{l} \tau_{k,j} \sum_{q=1}^{m+1} \Big(I_{\{\pi_j\theta \notin (C_{q-1}, C_q]\}} \left(F(C_q - \pi_j\theta) - F(C_{q-1} - \pi_j\theta) \right) \Big), \end{aligned}$$

which together with Assumption 2 yields that

$$\frac{1}{k} \sum_{i=1}^{k} \widehat{\gamma}_i^0 = \sum_{j=1}^{l} \frac{\tau_{k,j}}{k} \sum_{q=1}^{m+1} \Big(I_{\{\pi_j\theta \notin (C_{q-1}, C_q]\}} \left(F(C_q - \pi_j\theta) - F(C_{q-1} - \pi_j\theta) \right) \Big)$$

$$\to \sum_{j=1}^{l_0} \beta_j \sum_{q=1}^{m+1} \Big(I_{\{\pi_j\theta \notin (C_{q-1}, C_q]\}} \left(F(C_q - \pi_j\theta) - F(C_{q-1} - \pi_j\theta) \right) \Big), \quad \text{w.p.1} \quad \text{as} \quad k \to \infty. \quad (23)$$

By (1) and Assumption 1, we have

$$\begin{aligned}
\Pr(s_k = q) &= \Pr(C_{q-1} < y_k \leqslant C_q) \\
&= \Pr\left(C_{q-1} < \phi_k^{\mathrm{T}}\theta + d_k \leqslant C_q\right) \\
&= F(C_q - \phi_k^{\mathrm{T}}\theta) - F(C_{q-1} - \phi_k^{\mathrm{T}}\theta).
\end{aligned}$$

Noticing that $\widehat{s}_k \neq q$ if and only if $\phi_k^{\mathrm{T}}\widehat{\theta}_{k-1} \notin (C_{q-1}, C_q]$ by (22), it is known that

$$\begin{aligned}
\widehat{\gamma}_k &= E[\gamma_k | \mathscr{F}_{k-1}] \\
&= \Pr(s_k \neq \widehat{s}_k | \mathscr{F}_{k-1}) \\
&= \sum_{q=0}^{m} \Pr(\widehat{s}_k \neq q | \mathscr{F}_{k-1}) \Pr(s_k = q) \\
&= \sum_{q=1}^{m+1} I_{\{\widehat{s}_k \neq q\}} \left(F(C_q - \phi_k^{\mathrm{T}}\theta) - F(C_{q-1} - \phi_k^{\mathrm{T}}\theta) \right) \\
&= \sum_{q=1}^{m+1} \Big(I_{\{\phi_k^{\mathrm{T}}\widehat{\theta}_{k-1} \notin (C_{q-1}, C_q]\}} \left(F(C_q - \phi_k^{\mathrm{T}}\theta) - F(C_{q-1} - \phi_k^{\mathrm{T}}\theta) \right) \Big)
\end{aligned}$$

and $E\{r_k - \widehat{r}_k | \mathcal{F}_{k-1}\} = 0$. Therefore, $\{r_k - \widehat{r}_k, \mathcal{F}_k, k \geqslant 1\}$ is an MDS. Recalling Lemma 2, we have

$$\frac{\sum_{i=1}^{k} (r_i - \widehat{r}_i)}{k} \to 0, \quad \text{w.p.1} \quad \text{as} \quad k \to \infty. \quad (24)$$

By the convergence of $\widehat{\theta}_k$, it can be obtained that $\widehat{r}_k - \widehat{r}_k^0 \to 0$ w.p.1 as $k \to \infty$, which together with (23) and (24) can complete the proof.

6 Numerical simulation

Consider a gain system $y_k = u_k\theta + d_k$, where the true value $\theta = 3$ and $\{d_k\}$ is a sequence of i.i.d. normal random variables with zero mean and standard deviation $\sigma = 6$. The output is measured by a binary-valued sensor with the threshold $C = 2$, and the input is quantized and takes value from $\mathcal{U} = \{\pi_1, \pi_2, \pi_3\} = \{-1.5, 2, 1\}$. Because $\theta \in \mathbb{R}$, we have $\mathcal{P} = \mathcal{U}$. At k, assume that

$$\tau_{k,1} = k - \tau_{k,2} - \tau_{k,3}, \quad \tau_{k,2} = \lceil 0.6(k - \tau_{k,3}) \rceil, \quad \tau_{k,3} = \min\{110, \lfloor \lceil \log k \rceil \rfloor\}.$$

Thus, it is known that $\beta_1 = 0.4$, $\beta_2 = 0.6$, $\pi_1 = -1.5$ and $\pi_2 = 2$ are persistent, and then $\Psi = [-1.5, 2]^{\mathrm{T}}$.

The event-triggered communication scheme (3) and the algorithm (6)–(9) are used to estimate θ with $\Lambda = \mathrm{diag}(0.0114, 0.015, 0.001)$ and $\xi_{0,j} = 1/2$ for $j \in L$. The convergence is shown by Figure 4. The average of 20 trajectories of $k(\widehat{\theta}_k - \theta)(\widehat{\theta}_k - \theta)^{\mathrm{T}}$ is employed to approximate $k\Sigma_k$, and the difference between it and $k\Sigma_k^{\mathrm{CR}}$ gradually decreases in Figure 5, illustrating the asymptotic efficiency of the algorithm.

The transmission during the time interval $[200, 250]$ is depicted in Figure 6, where only 12 measurements are transmitted to the EC. It can be calculated that $\overline{\gamma} = \sum_{j=1}^{l_0} \beta_j \widetilde{F}(C - \pi_j\theta) = 0.2072$. In Figure 7, we can see that $\overline{\gamma}_k \to 0.2072$ as $k \to \infty$, which is in accord with Theorem 3 by (11).

Figure 8 shows the graphs of $\overline{\gamma}(C)$ and $(\Psi^{\mathrm{T}} H^*(C)\Psi)^{-1}$ with respect to C, where the left and right vertical axes denote $\overline{\gamma}(C)$ and $(\Psi^{\mathrm{T}} H^*(C)\Psi)^{-1}$, respectively. Let the estimation accuracy requirement $\Delta = 30$. Then one can see that $\Omega_\Delta = [-3.1, 11.1]$, and the minimum value of $\overline{\gamma}(C)$ on Ω_Δ is 0.1205.

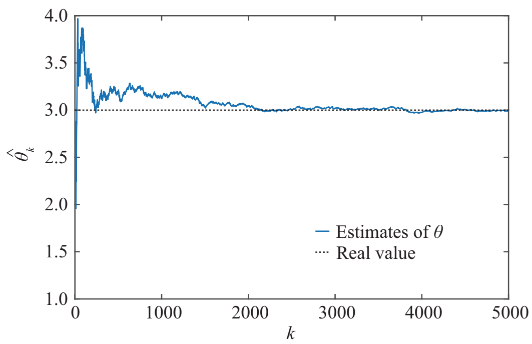

Figure 4 (Color online) Convergence of $\widehat{\theta}_k$.

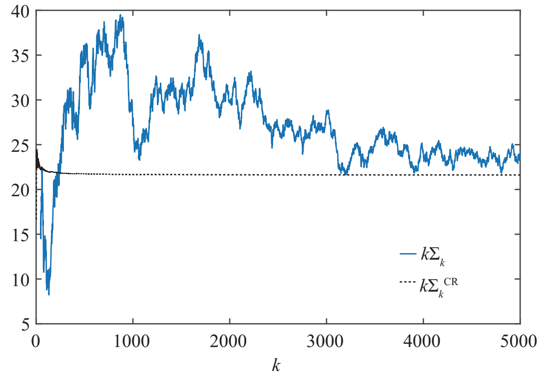

Figure 5 (Color online) Asymptotic efficiency of $\widehat{\theta}_k$.

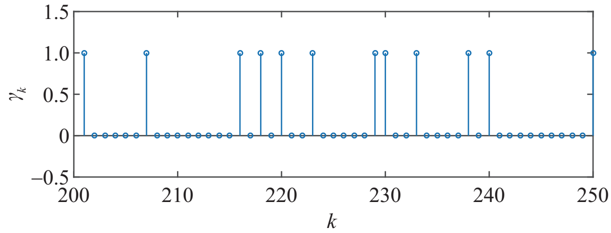

Figure 6 (Color online) Transmission during the time interval $[200, 250]$.

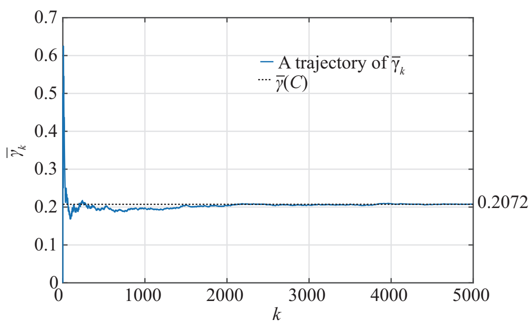

Figure 7 (Color online) Convergence of the average communication rate.

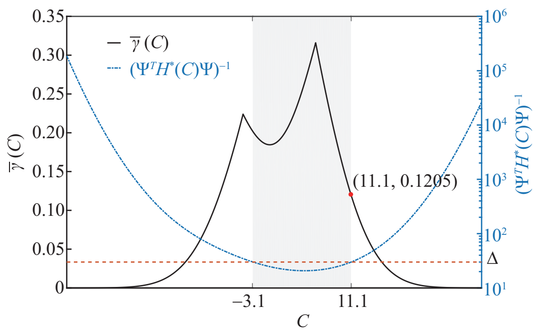

Figure 8 (Color online) Tradeoff between the communication cost $\overline{\gamma}(C)$ and the estimation performance $(\Psi^{\rm T} H^*(C)\Psi)^{-1}$.

7 Concluding remarks

The conservation of communication resources is indispensable in the current era of information networks, and it is interesting to consider the lowest measurement complexity to obtain the desired system performance. This paper introduced event-triggered communication schemes for the identification of FIR systems with quantized inputs and quantized output observations to reduce the communication burden. Beginning with the binary-valued quantization, an identification algorithm was proposed to estimate the unknown parameter, and the convergence performance of the algorithm was established. The tradeoff between the estimation quality and the communication cost was also discussed. Finally, the case of multi-threshold quantized observations was considered.

Future studies can further develop more effective event-triggered communication schemes and identification algorithms for obtaining an extensive variety of system models as well as noise characterizations.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61773054).

References

1 Söderström T, Stoica P. System Identification. Upper Saddle River: Prentice Hall, 1989
2 Wang L Y, Zhao W X. System identification: new paradigms, challenges, and opportunities. Acta Autom Sin, 2013, 39: 933–942
3 Akyildiz I F, Su W, Sankarasubramaniam Y, et al. Wireless sensor networks: a survey. Comput Netw, 2002, 38: 393–422
4 Hespanha J P, Naghshtabrizi P, Xu Y. A survey of recent results in networked control systems. Proc IEEE, 2007, 95: 138–162
5 Guo J, Mu B, Wang L Y, et al. Decision-based system identification and adaptive resource allocation. IEEE Trans Automat Contr, 2017, 62: 2166–2179
6 Ma C Q, Li T, Zhang J F. Consensus control for leader-following multi-agent systems with measurement noises. J Syst Sci Complex, 2010, 23: 35–49
7 Ma C Q, Zhang J F. On formability of linear continuous multi-agent systems. J Syst Sci Complex, 2012, 25: 13–29
8 Aström K J, Bernhardsson B M. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In: Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, 2002
9 Wang A, Liao X, Dong T. Event-triggered gradient-based distributed optimisation for multi-agent systems with state consensus constraint. IET Control Theory Appl, 2018, 12: 1515–1519
10 Shi D, Chen T, Shi L. On set-valued Kalman filtering and its application to event-based state estimation. IEEE Trans Automat Contr, 2015, 60: 1275–1290
11 Hetel L, Fiter C, Omran H, et al. Recent developments on the stability of systems with aperiodic sampling: an overview. Automatica, 2017, 76: 309–335
12 Wang A, Liao X, Dong T. Event-driven optimal control for uncertain nonlinear systems with external disturbance via adaptive dynamic programming. Neurocomputing, 2018, 281: 188–195
13 Wang A, Dong T, Liao X. Event-triggered synchronization strategy for complex dynamical networks with the Markovian switching topologies. Neural Netw, 2016, 74: 52–57
14 Yu Y G, Zeng Z W, Li Z K, et al. Event-triggered encirclement control of multi-agent systems with bearing rigidity. Sci China Inf Sci, 2017, 60: 110203
15 Zheng C, Li L, Wang L Y, et al. How much information is needed in quantized nonlinear control? Sci China Inf Sci, 2018, 61: 092205
16 Wang L Y, Zhang J F, Yin G G. System identification using binary sensors. IEEE Trans Automat Contr, 2003, 48: 1892–1907
17 Wang T, Tan J W, Zhao Y L. Asymptotically efficient non-truncated identification for FIR systems with binary-valued outputs. Sci China Inf Sci, 2018, 61: 129208
18 Guo J, Wang L Y, Yin G, et al. Asymptotically efficient identification of FIR systems with quantized observations and general quantized inputs. Automatica, 2015, 57: 113–122
19 Zhao Y, Wang L Y, Yin G G, et al. Identification of Wiener systems with binary-valued output observations. Automatica, 2007, 43: 1752–1765
20 Guo J, Wang L Y, Yin G, et al. Identification of Wiener systems with quantized inputs and binary-valued output observations. Automatica, 2017, 78: 280–286
21 Casini M, Garulli A, Vicino A. Input design in worst-case system identification with quantized measurements. Automatica, 2012, 48: 2997–3007
22 Zhao Y, Bi W, Wang T. Iterative parameter estimate with batched binary-valued observations. Sci China Inf Sci, 2016, 59: 052201
23 Goudjil A, Pouliquen M, Pigeon E. Identification of systems using binary sensors via support vector machines. In: Proceedings of IEEE 54th Annual Conference on Decision and Control, Osaka, 2015
24 Chow Y S, Teicher H. Probability Theory: Independence, Interchangeability, Martingales. 2nd ed. New York: Springer-Verlag, 1997
25 Wang L Y, Yin G G. Asymptotically efficient parameter estimation using quantized output observations. Automatica, 2007, 43: 1178–1191