SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

May 2019, Vol. 62 052203:1–052203:16
https://doi.org/10.1007/s11432-018-9604-0

FIR system identification with set-valued and precise observations from multiple sensors

Hang ZHANG[1,2], Ting WANG[1] & Yanlong ZHAO[1,2*]

[1]*Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190*, China;*
[2]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100049*, China*

Received 6 June 2018/Accepted 5 September 2018/Published online 2 April 2019

Abstract This paper considers the system identification problem for FIR (finite impulse response) systems with set-valued and precise observations received from multiple sensors. A fusion estimation algorithm based on some suitable identification algorithms for different types of observations is proposed. In particular, least square method is chosen for FIR systems with precise observations, while empirical measure method and EM algorithm are chosen for FIR systems with set-valued observations in the cases of periodic and general system inputs, respectively. Then, the quasi-convex combination estimator (QCCE) fusing the two different estimators by a linear combination with appropriate weights is constructed. Furthermore, the convergence properties are theoretically analyzed in terms of strong consistency and asymptotic efficiency. The fused estimator QCCE is proved to achieve the Cramér-Rao (CR) lower bound asymptotically under periodic inputs. Extensive numerical simulations validate the superiority of the fusion estimation algorithm under both periodic and general inputs.

Keywords system identification, FIR system, set-valued, precise, fusion estimation, quasi-convex combination estimator

Citation Zhang H, Wang T, Zhao Y L. FIR system identification with set-valued and precise observations from multiple sensors. Sci China Inf Sci, 2019, 62(5): 052203, https://doi.org/10.1007/s11432-018-9604-0

1 Introduction

System identification aims to build the mathematical models of dynamical systems based on observed system input and output data [1] and mainly consists of model structure identification and parameter estimation. Data, models and criterions are three basic elements of system identification. Specifically, the model structure is determined beforehand through some priori knowledge and the unknown parameters are estimated based on input-output data in terms of a certain criterion. The finite impulse response (FIR) model structure is favorable and widely used owing to its simplicity in dealing with the parameter estimation problem. With different measuring instruments and methods, we could collect different types of observations such as precise and set-valued (or quantized), which induce different identification algorithms.

The traditional system identification has been studied extensively on the basis of precise observations in the past several decades. Especially in the linear models (e.g., FIR models), a large number of mature identification algorithms and theories have been constructed and developed gradually [2]. Classical methods, such as least square (LS) method [3], maximum likelihood (ML) estimation method [4], state

* Corresponding author (email: ylzhao@amss.ac.cn)

estimation of Kalman filter [5], and Bayesian estimation [6], are applied widely in various fields. Furthermore, LS method and its extended algorithms attract our attention due to their excellent convergent properties including unbiasedness, strong consistency, asymptotic normality and efficiency [7].

System identification with set-valued observations was first investigated for sensor systems [8]. Compared with conventional system identification, set-valued observations offer very limited information on system outputs and consequently introduce new challenges. Recently, a number of set-valued identification algorithms have been developed. They can be classified roughly from system models, e.g., for linear systems, there are FIR models [9–14]; for nonlinear systems, there are Wiener [15] and Hammerstein [16] models. Besides, they can also be categorized in terms of the quantization schemes, e.g., fixed-level quantizer and uniform quantizer. In particular, for FIR models with finite quantization levels, there are empirical measure approach [9, 17] which has good properties of unbiasedness and strong consistency, EM algorithm [10, 11] which has a exponent convergent speed under certain conditions, kernel-based method [12], and quadratic programming-based method [13].

With the development of modernization and informatization, different types of input-output data can be obtained simultaneously for a certain system or some systems with same structures. Here we consider two types of system output data: precise and set-valued. In the system identification community, researchers focus on the identification problem based on a single type of input-output data. In order to get a better estimate effect, those approaches only using the precise data require higher accuracy measuring instruments which cost much, and those only using the set-valued data need more practical samples. Different from the former ones, the current problem is how to improve the estimation based on the available mixed data.

The difficulties of this problem lie in: (1) It is not easy to find an identification algorithm that is suitable to both precise and set-valued data; e.g., LS method can be used for precise data but not be well applied in set-valued data. (2) In general, the identification problem can be transformed into an optimization problem such as the ML method. But it is difficult to deal with the joined likelihood function of the probability density function for precise data and the cumulative distribution function for set-valued data. (3) Though some methods are appropriate for each type of data, how to fuse their respect identification results is troublesome. Ref. [14] studied the blind identification of multi-channel FIR systems using precise and quantized observations. But they could only deal with multiple channel data of same type by giving a special algorithm. Fortunately, information fusion in multi-sensor systems may provide a tool to solve this problem.

Information fusion is a multidisciplinary research area borrowing ideas from many diverse fields such as signal processing, information theory, statistical estimation and inference, and artificial intelligence, which has no universally accepted definition so far [18–20]. In [20], the authors took a review and discussion of many information fusion definitions and give a principled definition: Information fusion is the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human or automated decision making. As one of the most fundamental information fusion technologies, fusion estimation has been an ongoing research issue in multi-sensor information fusion. And information fusion algorithms play a pivotal role in the fusion estimation framework. In general, there are two fundamental methods to deal with the information from multiple sensors [21]. One is the centralized fusion, where all the observed data from different sensors are communicated to a fusion center for processing. The other is the distributed fusion, which consists of local estimation and estimation fusion. In the local estimation stage, individual sensors give local estimates based on their own observed data according to all kinds of estimation criteria such as ML estimation criterion [4], minimum means square error [22], and weight least square [23]. In the estimation fusion stage, these local estimates are gathered to the fusion center to produce an optimal or suboptimal estimate in terms of a certain fusion strategy such as optimal linear estimation fusion [24], Kalman filter method [25], and covariance intersection (CI) fusion method [26].

Considering that the data types of observations received from multiple sensors are different, namely precise and set-valued, it is not suitable to process the mixed data using the centralized fusion method. Thus, distributed fusion method makes it clear that firstly we find the corresponding optimal identification algorithm for each sensor and then fuse the local estimates according to a certain fusion criterion. Finally and most importantly, how to choose the fusion criterion? An easy and effective approach is to take a linear combination by giving appropriate weights to each estimate. Thus, the question is transformed into how to choose the weights for each local estimate. In general, we evaluate the parameter estimates in terms of their convergence properties such as unbiasedness, strong consistency, and efficiency. The Cramér-Rao (CR) lower bound means a lower bound on the variance of any (asymptotically) unbiased estimators [27]. An intuitive view is that we should give larger weights to the estimates with smaller variances as they can provide more accurate information. Furthermore, the fusion algorithm can be constructed based on some identification algorithms for different types of observation data.

In this paper, the FIR systems with set-valued and precise observations received from multiple sensors are considered. For the FIR systems with set-valued observations, we use the empirical measure method under periodic inputs and EM algorithm under general inputs to estimate the model parameters. While for the FIR systems with precise observations, LS method is used to estimate the model parameters regardless of which type of inputs. Then, the quasi-convex combination estimator (QCCE) [9] based on fusing the two kinds of estimators by a linear combination with appropriate weights is constructed. Furthermore, we mainly analyze the convergence properties of the QCCE in terms of strong consistency and asymptotic efficiency under periodic inputs. Because there is no analytical solution under general inputs, we have to adopt an iterative algorithm — EM algorithm, which leads to unavailable analysis of the convergence properties of fused estimator QCCE from a theoretical point of view. However, extensive Monte-Carlo simulation will be conducted to show the performance of fusion estimation algorithm.

The contributions of this paper can be summarized as follows:

- On the basic of the traditional identification algorithms for different types of multiple sensors, a fusion estimation algorithm for FIR systems with set-valued and precise observations is constructed.
- The fused estimator QCCE is proved to be optimal by giving suitable weights under periodic inputs, which indicates that the optimal QCCE has strong consistency and asymptotic efficiency, namely that it can achieve asymptotically the CR lower bound.
- The proposed fusion estimation algorithm can also be applied to the case of general inputs and other identification algorithms for different types of multiple sensors.

The rest of the paper is organized as follows. Section 2 introduces the FIR systems with set-valued and precise observations from multiple sensors and the problem formulation. Section 3 constructs a fusion estimation algorithm for the model parameters and gives a detailed illustration under both periodic and general inputs. Section 4 derives the convergence properties of the fused estimator under periodic inputs according to strong consistency and asymptotic efficiency. Section 5 demonstrates the superiority of the fusion estimation algorithm through extensive numerical simulations. Section 6 concludes the paper briefly and discusses related future topics.

2 Problem formulation

Consider the FIR system

$$y_k^{(l)} = \big(\phi_k^{(l)}\big)^{\mathrm{T}}\theta + d_k^{(l)}, \quad l = 1, \ldots, L,\ k = 1, \ldots, N_l, \tag{1}$$

where k is the time index, l is the FIR system index with a same structure, i.e., the same unknown but time-invariance model parameter θ and N_l is the data length of the system with index l; $\phi_k^{(l)}, y_k^{(l)}$ and $d_k^{(l)}$ are the system input, output and noise from the system with index l at time k, respectively. For the system whose index is no more than L_0 ($L_0 < L$), the corresponding output can be exactly measured. But for each of the rest systems, the output data is unknown and only measured by a quantized sensor. What can be obtained is the set-valued information denoted by $s_k^{(l)}$ that represents the comparison between the output $y_k^{(l)}$ and one or more thresholds.

$$s_k^{(l)} = Q\big(y_k^{(l)}\big) = \sum_{i=1}^{M} q_i I_{\{c_i^{(l)} \leqslant y_k^{(l)} < c_{i+1}^{(l)}\}}, \quad l = L_0 + 1, \ldots, L,\ k = 1, \ldots, N_l, \tag{2}$$

where $c_i^{(l)}$ $(i=1,\dots,M+1)$ are the fixed thresholds of the quantized sensor for FIR system with index l and $-\infty=c_1^{(l)}<c_2^{(l)}<\cdots<c_M^{(l)}<c_{M+1}^{(l)}=+\infty$. $I_{\{\cdot\}}$ is the indicator function, which is 1 if the condition holds and otherwise 0, and q_i represents the set-value with respect to the set $[c_i^{(l)},c_{i+1}^{(l)})$. For the convenience and uniformness in the formulation, the overall systems can be expressed as

$$\begin{cases} y_k^{(l)}=\big(\phi_k^{(l)}\big)^{\mathrm{T}}\theta+d_k^{(l)}, \\ s_k^{(l)}=\begin{cases} y_k^{(l)}, & l=1,\dots,L_0,\ k=1,\dots,N_l, \\ Q\big(y_k^{(l)}\big), & l=L_0+1,\dots,L,\ k=1,\dots,N_l. \end{cases} \end{cases} \tag{3}$$

In fact, $s_k^{(l)}$ $(l=1,\dots,L_0)$ can be seen as the measurements by a "precise" sensor, which is equivalent to $y_k^{(l)}$. In order to introduce the system identification problem, some assumptions are needed.

Assumption 1. The system input follows the condition that

$$\sum_{k=1}^{N_l}\phi_k^{(l)}\big(\phi_k^{(l)}\big)^{\mathrm{T}}>0,\quad l=1,\dots,L. \tag{4}$$

Assumption 2. For $l=1,2,\dots,L$, $\{d_k^{(l)},k\geqslant 1\}$ is a sequence of independent and identically distributed (i.i.d) variables and $d_1^{(l)}$ is a normally distributed random variable with mean 0 and known covariance σ_l^2. The cumulative distribution and probability density function of $d_1^{(l)}$ are denoted by $F_{\sigma_l}(\cdot)$ and $f_{\sigma_l}(\cdot)$, respectively.

Remark 1. Assumption 1 is the mathematical description of persistent excitation condition, which is a typical assumption in the research of system identification [2]. Under Assumption 2, the noises $d_k^{(l)}$ $(k\geqslant 1, l=L_0+1,\dots,L)$ for set-valued observations can be generalized into the symmetrically distributed random variable whose cumulative distribution function is second order continuous derivable with mean 0 and known covariance.

Without loss of generality, we assume that $L_0=1$ and $L=2$, i.e., only one FIR system with precise observations and one FIR system with set-valued observations are considered in this paper. The goal of the paper is to estimate the unknown parameter θ using the input $\phi_k^{(l)}$ and precise and set-valued observations $s_k^{(l)}$ based on all the FIR systems with different types of sensors.

3 Fusion estimation algorithm

In this section, we first propose a fusion estimation algorithm for parameter θ and then discuss the specific algorithm process under periodic inputs and general inputs, respectively.

Before giving the fusion estimation algorithm, we first introduce the concepts of QCCE and the optimal QCCE [9].

Definition 1. Assuming that $\theta_{N_1}^1,\dots,\theta_{N_m}^m$ are m asymptotically unbiased estimators of a single parameter θ, define $\beta=[\beta_1,\dots,\beta_m]^{\mathrm{T}}$ such that $\beta_1+\cdots+\beta_m=1$. One can construct an estimator $\hat\theta_Q$ of θ by

$$\hat\theta_Q=\sum_{i=1}^m\beta_i\theta_{N_i}^i=\beta^{\mathrm{T}}\Theta,$$

where $\Theta=[\theta_{N_1}^1,\dots,\theta_{N_m}^m]^{\mathrm{T}}$ and β_i need not be nonnegative. $\hat\theta_Q$ is called a QCCE. The estimator that minimizes the asymptotical variance of the estimation error $\hat\theta_Q-\theta$ is called the optimal QCCE.

Remark 2. What is different in [9] is that the m asymptotically unbiased estimators are denoted by $\theta_N^1,\dots,\theta_N^m$, which means the sample size is same for each estimator. But the definition of QCCE can be extended to the case in different sample sizes.

Because $\theta_{N_i}^i$ is asymptotically unbiased, $k=1,\dots,m$, $\hat\theta_Q$ is an asymptotically unbiased estimate of θ. Define the covariance matrix of the estimation error $\Theta-\mathbf{1}\theta$ as

$$V(\Theta)=E(\Theta-\mathbf{1}\theta)(\Theta-\mathbf{1}\theta)^{\mathrm{T}}.$$

Thus, the optimal QCCE can be figured out from

$$\beta = \arg \min_{\beta^{\rm T}\mathbf{1}=1} E(\hat{\theta}_Q - \theta)^2 = \arg \min_{\beta^{\rm T}\mathbf{1}=1} \beta^{\rm T} V(\Theta)\beta. \tag{5}$$

Similarly, the concepts of QCCE and optimal QCCE are directly extended to the case of multi-dimensional parameter θ.

Definition 2. Assuming that $\theta^1_{N_1}, \dots, \theta^m_{N_m}$ are m asymptotically unbiased estimators of multi dimensional parameter θ, define m matrixes $\Pi_1, \dots, \Pi_m$ such that $\Pi_1 + \cdots + \Pi_m = {\rm I}$. One can construct an estimator $\hat{\theta}_Q$ of θ by

$$\hat{\theta}_Q = \sum_{i=1}^{m} \Pi_i \theta^i_{N_i},$$

where Π_i $(1 \leqslant i \leqslant m)$ need not be positive definite. $\hat{\theta}_Q$ is called a QCCE.

As for the concept of optimal QCCE in the multi-dimensional case, there are two criterias can be chosen in terms of the covariance matrix of $\hat{\theta}_Q$, which is defined as

$$\Sigma(\hat{\theta}_Q) = E(\hat{\theta}_Q - \theta)(\hat{\theta}_Q - \theta)^{\rm T}.$$

One is to compare the covariance matrix $\Sigma(\hat{\theta}_Q)$ with CR lower bound. If there exist a group of matrixes $\Pi = \{\Pi_1, \dots, \Pi_m\}$ so that $\Sigma(\hat{\theta}_Q)$ can achieve the CR lower bound, $\hat{\theta}_Q$ is called the optimal QCCE.

The other is to mimic the approach of single parameter to compare the sum of each component's variance of θ, namely, the sum of diagonal elements of the covariance matrix $\Sigma(\hat{\theta}_Q)$, in case that CR lower bound cannot be achieved. Thus, the optimal QCCE can be obtained from

$$\Pi = \arg \min_{\Pi_1+\cdots+\Pi_m={\rm I}} E(\hat{\theta}_Q - \theta)^{\rm T}(\hat{\theta}_Q - \theta).$$

Remark 3. It can be seen easily that the optimal QCCE under CR low bound must be the one whose sum of the diagonal elements of the covariance matrix is the smallest.

The fusion estimation algorithm is presented as follows:

Step 1 (Precise identification). For the FIR system with index 1, based on the inputs $\phi^{(1)}_k$ and precise observations $s^{(1)}_k$ $(k = 1, 2, \dots, N_1)$, the estimator $\hat{\theta}_{N_1}$ of parameter θ is obtained by an appropriate precise identification algorithm (e.g., LS method).

Step 2 (Set-valued identification). For the FIR system with index 2, based on the inputs $\phi^{(2)}_k$ and set-valued observations $s^{(2)}_k$ $(k = 1, 2, \dots, N_2)$, the estimator $\hat{\theta}_{N_2}$ of parameter θ is obtained by an appropriate set-valued identification algorithm.

Step 3 (QCCE infusion). According to the definition of the QCCE, two weight matrixes Π_1 and Π_2 $(\Pi_1 + \Pi_2 = {\rm I})$ are designed to fuse the estimators $\hat{\theta}_{N_1}$ and $\hat{\theta}_{N_2}$ by a linear combination. The fused estimator is denoted by $\hat{\theta}_Q$:

$$\hat{\theta}_Q = \Pi_1 \hat{\theta}_{N_1} + \Pi_2 \hat{\theta}_{N_2}. \tag{6}$$

$\hat{\theta}_Q$ is the optimal QCCE if the estimation error $\hat{\theta}_Q - \theta$ can achieve the CR lower bound or its sum of the diagonal elements of the covariance matrix is the smallest.

3.1 Algorithms for periodic inputs

In this subsection, we will give the specific algorithm process under periodic inputs. The FIR systems with precise and binary-valued sensors are firstly considered. The model can be rewritten as

$$\begin{cases} y^{(l)}_k = \left(\phi^{(l)}_k\right)^{\rm T}\theta + d^{(l)}_k, & l = 1, 2, \\ s^{(1)}_k = y^{(1)}_k, & k = 1, \dots, N_1, \\ s^{(2)}_k = I_{\{y^{(2)}_k \leqslant c\}}, & k = 1, \dots, N_2. \end{cases} \tag{7}$$

To investigate the fusion estimation problem, we start with the simplest model with a single parameter θ. Assume that the system input is a constant for any FIR system, i.e.,

$$\phi_k^{(1)} \equiv u_1,\ k=1,2,\ldots,N_1;\quad \phi_k^{(2)} \equiv u_2,\ k=1,2,\ldots,N_2.$$

The LS estimator using the LS method based on the precise observations $s_k^{(1)}$ and inputs $\phi_k^{(1)} \equiv u_1$ ($k=1,2,\ldots,N_1$) is

$$\hat{\theta}_{N_1} = \frac{1}{N_1}\sum_{k=1}^{N_1} s_k^{(1)}/u_1. \tag{8}$$

The set-valued estimator using the empirical measure method based on the set-valued observations $s_k^{(2)}$ and inputs $\phi_k^{(2)} \equiv u_2$ ($k=1,2,\ldots,N_2$) is

$$\hat{\theta}_{N_2} = F_{\sigma_2}\left(c - \frac{1}{N_2}\sum_{k=1}^{N_2} s_k^{(2)}\right)/u_2. \tag{9}$$

Remark 4. The ML method and moment estimation method can be also applied for set-valued identification, which can derive a same estimator (9).

Because the parameter θ is scalar, the weights matrixes Π_1 and Π_2 degenerate to two real numbers. Then, we can figure out the optimal QCCE by solving the optimization problem (5).

Now, we consider the model when the parameter is multi-dimensional. In fact, the estimation problem in multi-dimensional periodic inputs can be transformed into the one in one-dimensional periodic inputs.

Assumption 3. For $l=1,2$, the FIR system inputs have an n_l-periodic form, that is, $\forall\, k \geqslant 1, \phi_k^{(l)} = \phi_{k+n_l}^{(l)}$. The number of observations N_l is divisible by n_l, which satisfies the equation $N_l/n_l = \bar{N}_l$.

Define the input matrixes of two FIR systems and two parameters $\gamma = (\gamma_1,\ldots,\gamma_{n_1})^{\rm T}$ and $\beta = (\beta_1,\ldots,\beta_{n_2})^{\rm T}$ as follows:

$$\Phi_1 = \left(\phi_1^{(1)},\phi_2^{(1)},\ldots,\phi_{n_1}^{(1)}\right)^{\rm T},\quad \Phi_2 = \left(\phi_1^{(2)},\phi_2^{(2)},\ldots,\phi_{n_2}^{(2)}\right)^{\rm T},$$

$$\gamma = \Phi_1\cdot\theta = \begin{bmatrix} (\phi_1^{(1)})^{\rm T} \\ \vdots \\ (\phi_{n_1}^{(1)})^{\rm T}\end{bmatrix}\cdot\theta = \begin{bmatrix} (\phi_1^{(1)})^{\rm T}\theta \\ \vdots \\ (\phi_{n_1}^{(1)})^{\rm T}\theta\end{bmatrix} = \begin{bmatrix}\gamma_1 \\ \vdots \\ \gamma_{n_1}\end{bmatrix}, \tag{10}$$

$$\beta = \Phi_2\cdot\theta = \begin{bmatrix} (\phi_1^{(2)})^{\rm T} \\ \vdots \\ (\phi_{n_2}^{(2)})^{\rm T}\end{bmatrix}\cdot\theta = \begin{bmatrix} (\phi_1^{(2)})^{\rm T}\theta \\ \vdots \\ (\phi_{n_2}^{(2)})^{\rm T}\theta\end{bmatrix} = \begin{bmatrix}\beta_1 \\ \vdots \\ \beta_{n_2}\end{bmatrix}. \tag{11}$$

Under Assumption 1, $\Phi_1^{\rm T}\Phi_1$ and $\Phi_2^{\rm T}\Phi_2$ are both reversible. If the estimates of parameters γ and β denoted by $\hat{\gamma}$ and $\hat{\beta}$ can be obtained, we can compute the estimates of θ for FIR systems with precise and set-valued observations, respectively.

$$\hat{\theta}_{N_1} = (\Phi_1^{\rm T}\Phi_1)^{-1}\Phi_1^{\rm T}\hat{\gamma}_{\bar{N}_1}, \tag{12}$$

$$\hat{\theta}_{N_2} = (\Phi_2^{\rm T}\Phi_2)^{-1}\Phi_2^{\rm T}\hat{\beta}_{\bar{N}_2}. \tag{13}$$

Remark 5. Because the estimate of each component of parameter γ or β is only relevant to $\bar{N}_1$ or $\bar{N}_2$ observations, we adopt the more accurate representations $\hat{\gamma}_{\bar{N}_1}$ or $\hat{\beta}_{\bar{N}_2}$.

Thus, we only need to solve the estimation problem of parameters γ and β. Actually, each component of γ or β can be regarded as the single parameter in the one-dimensional model aforementioned.

For the i-th ($1 \leqslant i \leqslant \bar{N}_1$) component γ_i of parameter γ, the LS estimator using the LS method based on the precise observations $s_k^{(1)}$ and inputs $\phi_k^{(1)}$ ($k = 1, 2, \ldots, N_1$) is

$$\hat{\gamma}_{i,\bar{N}_i} = \frac{1}{\bar{N}_1}\sum_{k=1}^{\bar{N}_1} s^{(1)}_{(k-1)n_1+i}. \tag{14}$$

For the i-th ($1 \leqslant i \leqslant \bar{N}_2$) component β_i of parameter β, the set-valued estimator using the empirical measure method based on the binary-valued observations $s_k^{(2)}$ and inputs $\phi_k^{(2)}$ ($k = 1, 2, \ldots, N_2$) is

$$\hat{\beta}_{i,\bar{N}_i} = F_{\sigma_2}\left(c - \frac{1}{\bar{N}_2}\sum_{k=1}^{\bar{N}_2} s^{(2)}_{(k-1)n_2+i}\right). \tag{15}$$

For the convenience of description, we rewrite the estimate equations of γ and β in the form of vector. Firstly define the observations in the form of vector as follows:

$$Y_i^{(1)} = \left(s^{(1)}_{(i-1)n_1+1}, \ldots, s^{(1)}_{in_1}\right)^{\mathrm{T}}, \quad i = 1, \ldots, \bar{N}_1, \tag{16}$$

$$Y_i^{(2)} = \left(s^{(2)}_{(i-1)n_2+1}, \ldots, s^{(2)}_{in_2}\right)^{\mathrm{T}}, \quad i = 1, \ldots, \bar{N}_2. \tag{17}$$

And then, the estimators $\hat{\gamma}_{\bar{N}_1}$ and $\hat{\beta}_{\bar{N}_2}$ can be expressed as

$$\hat{\gamma}_{\bar{N}_i} = \frac{1}{\bar{N}_1}\sum_{k=1}^{\bar{N}_1} Y_i^{(1)}, \tag{18}$$

$$\hat{\beta}_{\bar{N}_i} = F_{\sigma_2}\left(c - \frac{1}{\bar{N}_2}\sum_{k=1}^{\bar{N}_2} Y_i^{(2)}\right). \tag{19}$$

In consequence, the estimators $\hat{\theta}_{N_1}$ and $\hat{\theta}_{N_2}$ are

$$\hat{\theta}_{N_1} = \frac{1}{\bar{N}_1}(\Phi_1^{\mathrm{T}}\Phi_1)^{-1}\Phi_1^{\mathrm{T}}\sum_{k=1}^{\bar{N}_1} Y_i^{(1)}, \tag{20}$$

$$\hat{\theta}_{N_2} = (\Phi_2^{\mathrm{T}}\Phi_2)^{-1}\Phi_2^{\mathrm{T}}F_{\sigma_2}\left(c - \frac{1}{\bar{N}_2}\sum_{k=1}^{\bar{N}_2} Y_i^{(2)}\right). \tag{21}$$

The remaining task is to fuse the two estimators to obtain the optimal QCCE by selecting the corresponding appropriate weight matrixes Π_1 and Π_2 in terms of two criterions as mentioned earlier.

Now, we consider the FIR systems with precise and multi-valued observations. The LS estimator $\hat{\theta}_{N_1}$ for precise identification has no changes. The only difference from the previous problem is how to figure out the set-valued estimator $\hat{\theta}_{N_2}$ based on the multi-valued observations instead of binary-valued observations for set-valued identification. Actually, the multi-valued sensor with many thresholds is equivalent to many binary-valued sensors with one threshold. Then, the scalar observation $s_k^{(2)}$ can be represented by a vector $S_k^{(2)} = [s_k^{(2)}(1), \ldots, s_k^{(2)}(M-1)]^{\mathrm{T}}$ of $(M-1)$ dimensions, where $s_k^{(2)}(i) = I_{\{-\infty < y_k^{(2)} \leqslant c_{i+1}\}}$, $i = 1, \ldots, M-1$. Thus, we obtain $M-1$ set-valued estimators of parameter θ and then fuse them to form a new set-valued estimator $\hat{\theta}_{N_2}$ that is what we want in the light of the idea of QCCE. Furthermore, it can be proved that its optimality can be achieved by giving some special weights [9].

Remark 6. Owing to space constraints, this paper does not give the detailed derivation process for set-valued identification with multi-valued observations, which can be found in [9]. The subsequent processing about fusion estimation is the same as that in the case of binary-valued observations.

3.2 Algorithms for general inputs

In this subsection, we will discuss the specific algorithm process under general inputs. LS method is still the reasonable choice for the FIR system identification with precise observations. As for the FIR system identification with set-valued observations, we adopt the ML method. However, as there is no explicit solution of ML estimate that exists under periodic inputs, we have to choose certain iterative algorithms for solving the ML problem. Here we adopt the EM algorithm that has been proved to achieve a convergent rate of exponent in [11].

Similar to the case under periodic inputs, the FIR systems (7) with precise and binary-valued observations are firstly considered. Define the input matrixes and observation vectors of two FIR systems as follows:

$$\Psi_1 = \left(\phi_1^{(1)}, \ldots, \phi_{N_1}^{(1)}\right)^{\mathrm{T}}, \quad S_1 = \left(s_1^{(1)}, \ldots, s_{N_1}^{(1)}\right)^{\mathrm{T}},$$
$$\Psi_2 = \left(\phi_1^{(2)}, \ldots, \phi_{N_2}^{(2)}\right)^{\mathrm{T}}, \quad S_2 = \left(s_1^{(2)}, \ldots, s_{N_2}^{(2)}\right)^{\mathrm{T}},$$

where $\Psi_1^{\mathrm{T}}\Psi_1$ and $\Psi_2^{\mathrm{T}}\Psi_2$ are both reversible according to the Assumption 1.

The LS estimator using the LS method based on the inputs $\phi_k^{(1)}$ and precise observations $s_k^{(1)}$ $(k = 1, 2, \ldots, N_1)$ is

$$\hat{\theta}_{N_1} = (\Psi_1^{\mathrm{T}}\Psi_1)^{-1}\Psi_1^{\mathrm{T}} S_1. \tag{22}$$

The ML estimator using the EM algorithm based on the inputs $\phi_k^{(2)}$ and binary-valued observations $s_k^{(2)}$ $(k = 1, 2, \ldots, N_2)$ is

$$\begin{aligned}\hat{\theta}_{N_2}(t+1) = {} & \hat{\theta}_{N_2}(t) - \left(\sum_{k=1}^{N_2} \phi_k^{(2)} \left(\phi_k^{(2)}\right)^{\mathrm{T}}\right)^{-1} \left(\sum_{k=1}^{N_2} \sigma_2^2 \phi_k^{(2)} \cdot f_{\sigma_2}\left(c - \left(\phi_k^{(2)}\right)^{\mathrm{T}} \hat{\theta}_{N_2}(t)\right)\right. \\ & \left.\left[\frac{I_{\{s_k^{(2)}=1\}}}{F_{\sigma_2}\left(c - (\phi_k^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right)} - \frac{I_{\{s_k^{(2)}=0\}}}{1 - F_{\sigma_2}\left(c - (\phi_k^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right)}\right]\right),\end{aligned} \tag{23}$$

and the matrix form is

$$\hat{\theta}_{N_2}(t+1) = \hat{\theta}_{N_2}(t) - (\Psi_2^{\mathrm{T}}\Psi_2)^{-1}\Psi_2^{\mathrm{T}}\hat{S}_2(t), \tag{24}$$

where t represents the iteration time and

$$\hat{S}_2(t) = \sigma_2^2 \begin{pmatrix} f_{\sigma_2}\left(c - \left(\phi_1^{(2)}\right)^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right) \left[\dfrac{I_{\{s_1^{(2)}=1\}}}{F_{\sigma_2}\left(c - (\phi_1^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right)} - \dfrac{I_{\{s_1^{(2)}=0\}}}{1 - F_{\sigma_2}\left(c - (\phi_1^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right)}\right] \\ \vdots \\ f_{\sigma_2}\left(c - \left(\phi_{N_2}^{(2)}\right)^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right) \left[\dfrac{I_{\{s_{N_2}^{(2)}=1\}}}{F_{\sigma_2}\left(c - (\phi_{N_2}^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right)} - \dfrac{I_{\{s_{N_2}^{(2)}=0\}}}{1 - F_{\sigma_2}\left(c - (\phi_{N_2}^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t)\right)}\right] \end{pmatrix}. \tag{25}$$

Furthermore, the EM algorithm can be easily extended to the case of multi-valued observations [28]. Consider the FIR systems (3), the iteration estimation process is as follows:

$$\hat{\theta}_{N_2}(t+1) = \hat{\theta}_{N_2}(t) - \left(\sum_{k=1}^{N_2} \phi_k^{(2)} \left(\phi_k^{(2)}\right)^{\mathrm{T}}\right)^{-1} \left(\sum_{k=1}^{N_2} \sigma_2^2 \phi_k^{(2)} \cdot \left[\sum_{j=1}^{M} I_{\{s_k^{(2)}=q_j\}} \cdot \frac{f_{\sigma_2}(k, j+1) - f_{\sigma_2}(k, j)}{F_{\sigma_2}(k, j+1) - F_{\sigma_2}(k, j)}\right]\right), \tag{26}$$

where $f_{\sigma_2}(k, j) = f_{\sigma_2}(c_j - (\phi_k^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t))$, $F_{\sigma_2}(k, j) = F_{\sigma_2}(c_j - (\phi_k^{(2)})^{\mathrm{T}}\hat{\theta}_{N_2}(t))$ and then the matrix form can be written as

$$\hat{\theta}_{N_2}(t+1) = \hat{\theta}_{N_2}(t) - (\Psi_2^{\mathrm{T}}\Psi_2)^{-1}\Psi_2^{\mathrm{T}}\tilde{S}_2(t), \tag{27}$$

$$\tilde{S}_2(t) = \sigma_2^2 \begin{pmatrix} \sum_{j=1}^{M} I_{\{s_1^{(2)}=q_j\}} \cdot \frac{f_{\sigma_2}(1,j+1) - f_{\sigma_2}(1,j)}{F_{\sigma_2}(1,j+1) - F_{\sigma_2}(1,j)} \\ \vdots \\ \sum_{j=1}^{M} I_{\{s_{N_2}^{(2)}=q_j\}} \cdot \frac{f_{\sigma_2}(N_2,j+1) - f_{\sigma_2}(N_2,j)}{F_{\sigma_2}(N_2,j+1) - F_{\sigma_2}(N_2,j)} \end{pmatrix}. \tag{28}$$

Similarly, the key point of the proposed fusion estimation algorithm lies in choosing two suitable weight matrixes for the QCCE so as to achieve the optimality as far as possible. However, the iteration algorithm makes it not easy to solve this problem, which we will talk about in Section 4.

Remark 7. In the set-valued identification algorithms mentioned above, it is tacitly approved that all the thresholds of the quantization sensor are known constants. In some situations, the thresholds are not available.

In fact, for the binary-valued sensor, even if the threshold c is unknown, we can still detect it by a simple transformation process. We replace the input $\check{\phi}_k^{(2)} = ((\phi_k^{(2)})^{\mathrm{T}}, -1)^{\mathrm{T}}$ and parameter $\check{\theta} = (\theta^{\mathrm{T}}, c)^{\mathrm{T}}$. Thus, the new threshold is fixed as 0 and c can be estimated together with parameter θ by the corresponding set-valued algorithm.

As for the multi-valued sensor, if all the thresholds $c_2, \ldots, c_M$ are unknown, we can adopt the same method for the binary-valued sensor under periodic inputs because the algorithm is essentially to split the multi-valued observations into binary-valued observations and fuse all the binary-valued estimators. But under general inputs, we have to take steps to estimate the thresholds at the iteration process, such as firstly estimating model parameter θ based on the EM algorithm (27) and then estimating the thresholds using the gradient descent method [28].

4 Convergence property

In this section, we will evaluate the fusion estimation algorithm in terms of strong consistency and asymptotic efficiency given the weight matrixes.

Theorem 1 (Strong consistency)**.** Consider the FIR systems (7) under periodic inputs. Under Assumptions 1–3, the fused estimate $\hat{\theta}_Q$ given by the fusion estimation algorithm that consist of (6), (20) and (21) is strongly consistent in that

$$\hat{\theta}_Q \to \theta, \quad \text{w.p.1}, \quad N_1 \to \infty, \ N_2 \to \infty.$$

Proof. As is known to all, the LS method has the strong consistency under Assumptions 1 and 2, i.e.,

$$\hat{\theta}_{N_1} \to \theta, \quad \text{w.p.1}, \quad N_1 \to \infty.$$

In reference to Theorem 3.5 in [9], the set-valued estimator $\hat{\theta}_{N_1}$ is strongly consistent under Assumptions 1–3, i.e.,

$$\hat{\theta}_{N_2} \to \theta, \quad \text{w.p.1}, \quad N_2 \to \infty.$$

Thus, on the condition that $\Pi_1 + \Pi_2 = \mathrm{I}$ and the strong consistency of $\hat{\theta}_{N_1}$ and $\hat{\theta}_{N_2}$, it is easy to prove that

$$\hat{\theta}_Q = \Pi_1 \hat{\theta}_{N_1} + \Pi_2 \hat{\theta}_{N_2} \to \Pi_1 \theta + \Pi_2 \theta = \mathrm{I}\theta = \theta, \quad \text{w.p.1}, \quad N_1 \to \infty, \ N_2 \to \infty.$$

The assertion is proved.

Theorem 2 (CR lower bound)**.** Consider the FIR systems (7). Under Assumptions 1 and 2, the following conclusions hold:

(i) For periodic inputs, under Assumption 3, the CR lower bound of parameter θ based on the precise and binary-valued observations and the corresponding inputs is

$$\Sigma_{\mathrm{CR}} = \left(\frac{\bar{N}_1}{\sigma_1^2} \Phi_1^{\mathrm{T}} \Phi_1 + \bar{N}_2 \Phi_2^{\mathrm{T}} \Lambda \Phi_2 \right)^{-1}. \tag{29}$$

In particular, when the two FIR systems have the same inputs, the CR lower bound of parameter θ can be simplified as

$$\Sigma_{\mathrm{CR}} = \frac{1}{\bar{N}_1}\left(\Phi_1^{\mathrm{T}}\left[\frac{1}{\sigma_1^2}I+\Lambda\right]\Phi_1\right)^{-1}, \tag{30}$$

where I represents the $n_2\times n_2$ identity matrix, and Λ is an $n_2\times n_2$ diagonal matrix as follows:

$$\Lambda = \begin{pmatrix} \frac{f_{\sigma_2}^2\left(c-(\phi_1^{(2)})^{\mathrm{T}}\theta\right)}{F_{\sigma_2}\left(c-(\phi_1^{(2)})^{\mathrm{T}}\theta\right)\left[1-F_{\sigma_2}\left(c-(\phi_1^{(2)})^{\mathrm{T}}\theta\right)\right]} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \frac{f_{\sigma_2}^2\left(c-(\phi_{n_2}^{(2)})^{\mathrm{T}}\theta\right)}{F_{\sigma_2}\left(c-(\phi_{n_2}^{(2)})^{\mathrm{T}}\theta\right)\left[1-F_{\sigma_2}\left(c-(\phi_{n_2}^{(2)})^{\mathrm{T}}\theta\right)\right]} \end{pmatrix}. \tag{31}$$

(ii) For general inputs, the CR lower bound of parameter θ based on precise and binary-valued observations and the corresponding inputs is

$$\Sigma_{\mathrm{CR}}^{*} = \left(\frac{1}{\sigma_1^2}\Psi_1^{\mathrm{T}}\Psi_1+\Psi_2^{\mathrm{T}}\Lambda^{*}\Psi_2\right)^{-1}. \tag{32}$$

In particular, when the two FIR systems have the same inputs, the CR lower bound of parameter θ can be simplified as

$$\Sigma_{\mathrm{CR}}^{*} = \left(\Psi_1^{\mathrm{T}}\left[\frac{1}{\sigma_1^2}I^{*}+\Lambda^{*}\right]\Psi_1\right)^{-1}, \tag{33}$$

where I^{*} represents the $N_2\times N_2$ identity matrix, and Λ^{*} is an $N_2\times N_2$ diagonal matrix as follows:

$$\Lambda^{*} = \begin{pmatrix} \frac{f_{\sigma_2}^2\left(c-(\phi_1^{(2)})^{\mathrm{T}}\theta\right)}{F_{\sigma_2}\left(c-(\phi_1^{(2)})^{\mathrm{T}}\theta\right)\left[1-F_{\sigma_2}\left(c-(\phi_1^{(2)})^{\mathrm{T}}\theta\right)\right]} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \frac{f_{\sigma_2}^2\left(c-(\phi_{N_2}^{(2)})^{\mathrm{T}}\theta\right)}{F_{\sigma_2}\left(c-(\phi_{N_2}^{(2)})^{\mathrm{T}}\theta\right)\left[1-F_{\sigma_2}\left(c-(\phi_{N_2}^{(2)})^{\mathrm{T}}\theta\right)\right]} \end{pmatrix}. \tag{34}$$

Proof. (i) Under the condition that periodic inputs $\phi_k^{(1)}$ $(k=1,\ldots,N_1)$ and $\phi_k^{(2)}$ $(k=1,\ldots,N_2)$ and the noises $d_k^{(1)}$ $(k=1,\ldots,N_1)$ and $d_k^{(2)}$ $(k=1,\ldots,N_2)$ are i.i.d, the joint conditional probability distribution function (i.e., likelihood function) of the precise observations $s_k^{(1)}$ $(k=1,\ldots,N_1)$ and binary-valued observations $s_k^{(2)}$ $(k=1,\ldots,N_2)$ is

$$\begin{aligned} L(\theta) &= \prod_{k=1}^{N_1} P\left(s_k^{(1)}|\theta,\phi_k^{(1)}\right)\prod_{k=1}^{N_2} P\left(s_k^{(2)}|\theta,\phi_k^{(2)}\right) \\ &= \prod_{k=1}^{n_1}\prod_{i=1}^{\bar{N}_1} P\left(s_{(i-1)n_1+k}^{(1)}|\theta,\phi_k^{(1)}\right)\prod_{k=1}^{n_2}\prod_{k=1}^{\bar{N}_2} P\left(s_{(i-1)n_2+k}^{(2)}|\theta,\phi_k^{(2)}\right) \\ &= \prod_{k=1}^{n_1}\prod_{i=1}^{\bar{N}_1} f_{\sigma_1}\left(s_{(i-1)n_1+k}^{(1)}-\left(\phi_k^{(1)}\right)^{\mathrm{T}}\theta\right) \\ &\quad\cdot\prod_{k=1}^{n_2}\prod_{k=1}^{\bar{N}_2}\left[I_{\{s_{(i-1)n_2+k}^{(2)}=1\}}F_{\sigma_2}\left(c-\left(\phi_k^{(2)}\right)^{\mathrm{T}}\theta\right)+I_{\{s_{(i-1)n_2+k}^{(2)}=0\}}\left(1-F_{\sigma_2}\left(c-\left(\phi_k^{(2)}\right)^{\mathrm{T}}\theta\right)\right)\right]. \end{aligned}$$

Then the log-likelihood function is

$$l(\theta) = \sum_{k=1}^{n_1}\sum_{i=1}^{\bar{N}_1}\log\left[f_{\sigma_1}\left(s_{(i-1)n_1+k}^{(1)}-\left(\phi_k^{(1)}\right)^{\mathrm{T}}\theta\right)\right]$$

$$
\begin{aligned}
&+\sum_{k=1}^{n_2}\sum_{k=1}^{\bar N_2}\log\left[I_{\{s^{(2)}_{(i-1)n_2+k}=1\}}F_{\sigma_2}\left(c-\left(\phi_k^{(2)}\right)^{\mathrm T}\theta\right)+I_{\{s^{(2)}_{(i-1)n_2+k}=0\}}\left(1-F_{\sigma_2}\left(c-\left(\phi_k^{(2)}\right)^{\mathrm T}\theta\right)\right)\right]\\
&=-N_1\log(\sqrt{2\pi}\sigma_1)+\sum_{k=1}^{n_1}\sum_{i=1}^{\bar N_1}-\frac{1}{2\sigma_1^2}\left(s^{(1)}_{(i-1)n_1+k}-(\phi_k^1)^{\mathrm T}\theta\right)^2\\
&+\sum_{k=1}^{n_2}\sum_{k=1}^{\bar N_2}s^{(2)}_{(i-1)n_2+k}\log F_{\sigma_2}\left(c-\left(\phi_k^{(2)}\right)^{\mathrm T}\theta\right)+\left(1-s^{(2)}_{(i-1)n_2+k}\right)\log\left(1-F_{\sigma_2}\left(c-\left(\phi_k^{(2)}\right)^{\mathrm T}\theta\right)\right).
\end{aligned}
$$

The first partial derivative of $l(\theta)$ is

$$
\begin{aligned}
\frac{\partial l(\theta)}{\partial\theta}=&-\frac{1}{\sigma_1^2}\sum_{k=1}^{n_1}\sum_{i=1}^{\bar N_1}\left(s^{(1)}_{(i-1)n_1+k}-\left(\phi_k^{(1)}\right)^{\mathrm T}\theta\right)\phi_k^{(1)}\\
&+\sum_{k=1}^{n_2}\sum_{k=1}^{\bar N_2}\left[s^{(2)}_{(i-1)n_2+k}\frac{-f_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}{F_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}+\left(1-s^{(2)}_{(i-1)n_2+k}\right)\frac{f_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}{1-F_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}\right]\phi_k^{(2)}.
\end{aligned}
$$

We continue the partial process to obtain the second partial derivative of $l(\theta)$:

$$
\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial\theta^2}=&-\frac{1}{\sigma_1^2}\sum_{k=1}^{n_1}\sum_{i=1}^{\bar N_1}\phi_k^{(1)}\left(\phi_k^{(1)}\right)^{\mathrm T}\\
&+\sum_{k=1}^{n_2}\sum_{k=1}^{\bar N_2}\left[s^{(2)}_{(i-1)n_2+k}\left(\frac{-\partial f_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)/\partial\theta}{F_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}-\frac{\partial f^2_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)/\partial\theta}{F^2_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}\right)\right]\phi_k^{(2)}\left(\phi_k^{(2)}\right)^{\mathrm T}\\
&-\sum_{k=1}^{n_2}\sum_{k=1}^{\bar N_2}\left[\left(1-s^{(2)}_{(i-1)n_2+k}\right)\left(\frac{\partial f_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)/\partial\theta}{1-F_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}+\frac{\partial f^2_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)/\partial\theta}{\big(1-F^2_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)\big)^2}\right)\right]\phi_k^{(2)}\left(\phi_k^{(2)}\right)^{\mathrm T}.
\end{aligned}
$$

Notice that $Es^{(2)}_{(i-1)n_2+k}=F_{\sigma_2}(c-(\phi_k^{(2)})^{\mathrm T}\theta)$, $1\leqslant i\leqslant\bar N_2$, $1\leqslant k\leqslant n_2$. As a result, the CR lower bound can be calculated as

$$
\begin{aligned}
\Sigma_{\mathrm{CR}}&=-\left(E\frac{\partial^2 l(\theta)}{\partial\theta^2}\right)^{-1}\\
&=\left(\frac{\bar N_1}{\sigma_1^2}\sum_{k=1}^{n_1}\phi_k^{(1)}\left(\phi_k^{(1)}\right)^{\mathrm T}+\bar N_2\sum_{k=1}^{n_2}\frac{f^2_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)}{F_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)\big[1-F_{\sigma_2}\big(c-(\phi_k^{(2)})^{\mathrm T}\theta\big)\big]}\phi_k^{(2)}\left(\phi_k^{(2)}\right)^{\mathrm T}\right)^{-1},
\end{aligned}
$$

and then, the matrix form can be expressed as

$$
\Sigma_{\mathrm{CR}}=\left(\frac{\bar N_1}{\sigma_1^2}\Phi_1^{\mathrm T}\Phi_1+\bar N_2\Phi_2^{\mathrm T}\Lambda\Phi_2\right)^{-1}.
$$

In particular, the inputs of two FIR systems are same, i.e., $\Phi_1=\Phi_2$, $\bar N_1=\bar N_2$, so it is not difficult to obtain (30).

(ii) For the case of general inputs, the proof process is similar to that in (i).

Theorem 3 (Asymptotic efficiency)**.** Consider the FIR systems (7) under periodic inputs. Under Assumptions 1–3, define two matrixes Σ_1 and Σ_2 as follows:

$$
\Sigma_1=\frac{\sigma_1^2}{\bar N_1}(\Phi_1^{\mathrm T}\Phi_1)^{-1},\quad \Sigma_2=\frac{1}{\bar N_2}(\Phi_2^{\mathrm T}\Lambda\Phi_2)^{-1}, \tag{35}
$$

where Λ is defined in Theorem 2. By selecting two weight matrixes

$$
\Pi_1=(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\Sigma_1^{-1},\quad \Pi_2=(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\Sigma_2^{-1}, \tag{36}
$$

the fused estimate $\hat\theta_Q$ is the optimal QCCE in that it is asymptotically efficient, namely,

$$
\Sigma_{\mathrm{CR}}^{-1}\big[E(\hat\theta_Q-\theta)(\hat\theta_Q-\theta)^{\mathrm T}-\Sigma_{\mathrm{CR}}\big]\to 0,\quad N_2\to\infty. \tag{37}
$$

Proof. In fact, Σ_1 and Σ_2 are the CR lower bound of parameter θ based on the only precise observations and the only binary-valued observations, respectively. And it can be easily proved that LS estimator $\hat{\theta}_{N_1}$ can achieve the CR lower bound Σ_1 and the set-valued estimator $\hat{\theta}_{N_2}$ is asymptotically efficient, that is to say,

$$E(\hat{\theta}_{N_1}-\theta)(\hat{\theta}_{N_1}-\theta)^{\rm T}=\Sigma_1,\quad E(\hat{\theta}_{N_2}-\theta)(\hat{\theta}_{N_2}-\theta)^{\rm T}\to\Sigma_2,\quad N_2\to\infty.$$

Then, we have

$$\begin{aligned}E(\hat{\theta}_Q-\theta)(\hat{\theta}_Q-\theta)^{\rm T}&=E\big(\Pi_1\hat{\theta}_{N_1}+\Pi_1\hat{\theta}_{N_2}-(\Pi_1+\Pi_2)\theta)(\Pi_1\hat{\theta}_{N_1}+\Pi_1\hat{\theta}_{N_2}-(\Pi_1+\Pi_2)\theta\big)^{\rm T}\\&=\Pi_1E(\hat{\theta}_{N_1}-\theta)(\hat{\theta}_{N_1}-\theta)^{\rm T}\Pi_1^{\rm T}+\Pi_2E(\hat{\theta}_{N_2}-\theta)(\hat{\theta}_{N_2}-\theta)^{\rm T}\Pi_2^{\rm T}\\&\quad+\Pi_1E(\hat{\theta}_{N_1}-\theta)(\hat{\theta}_{N_2}-\theta)^{\rm T}\Pi_2^{\rm T}+\Pi_2E(\hat{\theta}_{N_2}-\theta)(\hat{\theta}_{N_1}-\theta)^{\rm T}\Pi_1^{\rm T}.\end{aligned}$$

Because the noises $d_k^{(1)}$ $(1\leqslant k\leqslant N_1)$ and $d_k^{(2)}$ $(1\leqslant k\leqslant N_2)$ are i.i.d and the LS estimator $\hat{\theta}_{N_1}$ is unbiased, we can obtain that

$$\begin{aligned}\Pi_1E(\hat{\theta}_{N_1}-\theta)(\hat{\theta}_{N_2}-\theta)^{\rm T}\Pi_2^{\rm T}=\Pi_1E(\hat{\theta}_{N_1}-\theta)E(\hat{\theta}_{N_2}-\theta)^{\rm T}\Pi_2^{\rm T}=0,\\\Pi_2E(\hat{\theta}_{N_2}-\theta)(\hat{\theta}_{N_1}-\theta)^{\rm T}\Pi_1^{\rm T}=\Pi_2E(\hat{\theta}_{N_2}-\theta)E(\hat{\theta}_{N_1}-\theta)^{\rm T}\Pi_1^{\rm T}=0.\end{aligned}$$

Thus,

$$\begin{aligned}E(\hat{\theta}_Q-\theta)(\hat{\theta}_Q-\theta)^{\rm T}&=\Pi_1E(\hat{\theta}_{N_1}-\theta)(\hat{\theta}_{N_1}-\theta)^{\rm T}\Pi_1^{\rm T}+\Pi_2E(\hat{\theta}_{N_2}-\theta)(\hat{\theta}_{N_2}-\theta)^{\rm T}\Pi_2^{\rm T}\\&\to(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\Sigma_1^{-1}\Sigma_1\Sigma_1^{-1}(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\\&\quad+(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\Sigma_2^{-1}\Sigma_2\Sigma_2^{-1}(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\ (N_2\to\infty)\\&=(\Sigma_1^{-1}+\Sigma_2^{-1})^{-1}\triangleq\Sigma_{\rm CR}.\end{aligned}$$

Until now, the proof is completed.

For the case of general inputs, we also choose the two weight matrixes Π_1^* and Π_2^* for LS estimator and set-valued estimator in terms of their own CR lower bound, i.e.,

$$\Pi_1^*=(\Sigma_{1*}^{-1}+\Sigma_{2*}^{-1})^{-1}\Sigma_{1*}^{-1},\quad \Pi_2^*=(\Sigma_{1*}^{-1}+\Sigma_{2*}^{-1})^{-1}\Sigma_{2*}^{-1},$$

where Σ_{1*} and Σ_{2*} are defined as

$$\Sigma_{1*}=\left(\frac{1}{\sigma_1^2}\Psi_1^{\rm T}\Psi_1\right)^{-1},\quad \Sigma_{2*}=(\Psi_2^{\rm T}\Lambda^*\Psi_2)^{-1}.$$

Because the fusion estimation algorithm is based on the EM iterative algorithm for set-valued identification, we cannot analyze its convergence properties from a theoretical point of view and decide if the fused QCCE is the optimal QCCE. However, extensive numerical simulations will be conducted to show its performance.

5 Simulation

In this section, we will conduct extensive simulations in different situations that include input sources (same or different) and input types (periodic or general) to evaluate the convergent properties of the proposed fusion estimation algorithm.

Consider the FIR systems (7). The goal is to estimate the parameter θ based on all the precise observations and set-valued observations and the corresponding inputs. For periodic inputs, the fusion estimation algorithm that is consisted of (6), (20) and (21) is adopted. As for general inputs, the fusion estimation algorithm that is consisted of (6), (22) and (24) is adopted.

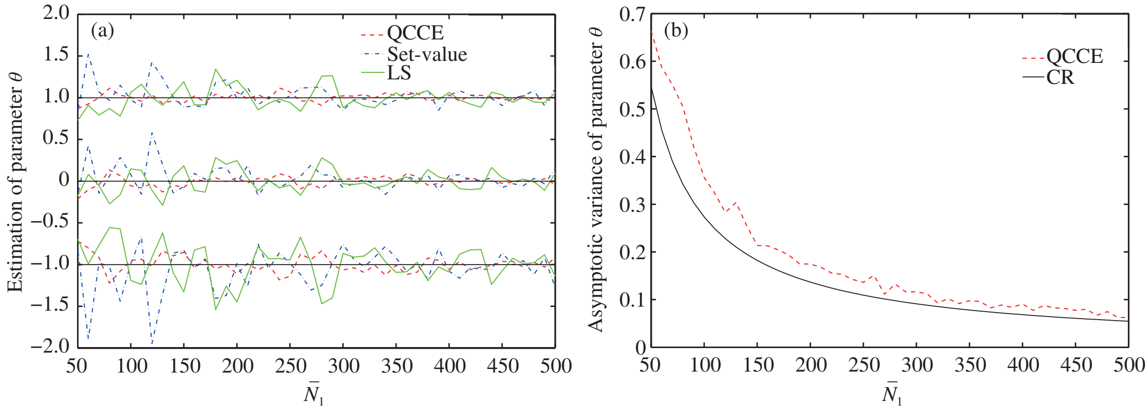

Figure 1 (Color online) Two FIR systems have same periodic inputs. (a) The estimation results of parameter θ including $\hat{\theta}_{N_1}$ (LS), $\hat{\theta}_{N_2}$ (set-value) and $\hat{\theta}_Q$ (QCCE); (b) the asymptotic efficiency of estimation error of the QCCE by the evaluation criterion $E(\hat{\theta}_Q-\theta)^{\mathrm{T}}(\hat{\theta}_Q-\theta)$ with 500 replicated simulations.

5.1 Periodic inputs

(i) Same inputs. Let $\theta=(-1,0,1)^{\mathrm{T}}$ and $c=1$. The system noises $d_k^{(1)}$ and $d_k^{(2)}$, $k\geqslant 1$ satisfy Assumption 2 with $\sigma_1=3$ and $\sigma_2=2$, respectively. The two input matrixes are

$$\Phi_1=\Phi_2=(\phi_1,\phi_2,\phi_3)^{\mathrm{T}}=\begin{pmatrix}1&0&2\\1&1&0\\2&3&1\end{pmatrix}.$$

By varying sample size $\bar{N}_1(=\bar{N}_2)$ from 50 to 500, we compare the performance of the QCCE $\hat{\theta}_Q$ with the set-valued estimate $\hat{\theta}_{N_2}$ and the LS estimate $\hat{\theta}_{N_1}$ (see Figure 1).

(ii) Different inputs. Let $\theta=(-1,0,1)^{\mathrm{T}}$ and $c=1$. The system noises $d_k^{(1)}$ and $d_k^{(2)}$, $k\geqslant 1$ satisfy Assumption 2 with $\sigma_1=1$ and $\sigma_2=2$, respectively. The two input matrixes are

$$\Phi_1=(\phi_1^1,\phi_2^1,\phi_3^1)^{\mathrm{T}}=\begin{pmatrix}1&0&2\\1&1&0\\2&3&1\end{pmatrix},\quad \Phi_2=(\phi_1^2,\phi_2^2,\phi_3^2)^{\mathrm{T}}=\begin{pmatrix}1&2&1\\0&1&1\\2&1&0\end{pmatrix}.$$

We fix $\bar{N}_1=30$ and vary $\bar{N}_2$ from 10 to 300, and then compare the performance of the QCCE $\hat{\theta}_Q$ with the set-valued estimate $\hat{\theta}_{N_2}$ (see Figure 2).

In general, when two systems have same periodic inputs, which indicates that the total samples are same, the identification effect based on the precise observations is better than that based on the set-valued observations. But the system noise variance or SNR (signal to noise ratio) is also not neglected. Figures 1(a) and 2(a) show the fused QCCE can converge to the true value faster than only utilizing LS method for precise identification or empirical measurement method for set-valued identification regardless of the noise and input source. And Figure 2(a) also shows the identification effect will have a significant improvement based on some set-valued observations by adding few precise observations. It can be inferred from Figures 1(b) and 2(b) that the fused QCCE is optimal and asymptotic in that its estimation error can achieve the CR lower bound.

5.2 General inputs

(i) Same inputs. Let $\theta=(1,3)^{\mathrm{T}}$ and $c=4$. The system noises $d_k^{(1)}$ and $d_k^{(2)}$, $k\geqslant 1$ satisfy Assumption 2 with $\sigma_1=3$ and $\sigma_2=2$, respectively. The input matrix is randomly generated from a uniform distribution in the interval [0,2]. Similarly, by varying sample size $\bar{N}_1(=\bar{N}_2)$ from 20 to 300, we compare the performance of the QCCE $\hat{\theta}_Q$ with the set-valued estimate $\hat{\theta}_{N_2}$ and the LS estimate $\hat{\theta}_{N_1}$ (see Figure 3).

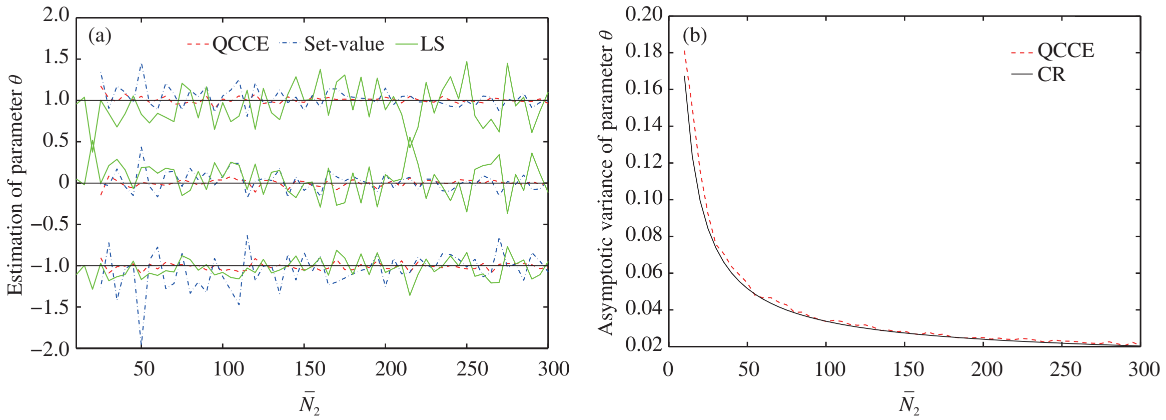

Figure 2 (Color online) Two FIR systems have different periodic inputs and the sample size of precise observations is fixed. (a) The estimation results of parameter θ including $\hat{\theta}_{N_2}$ (set-value) and $\hat{\theta}_Q$ (QCCE); (b) the asymptotic efficiency of estimation error of the QCCE by the evaluation criterion $E(\hat{\theta}_Q-\theta)^{\mathrm{T}}(\hat{\theta}_Q-\theta)$ with 500 replicated simulations.

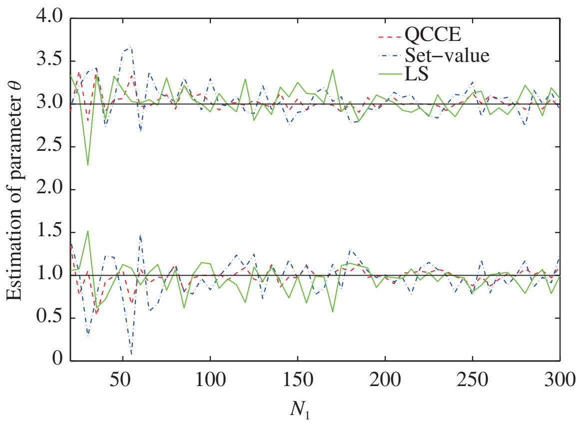

Figure 3 (Color online) The estimation results of parameter θ including $\hat{\theta}_{N_1}$ (LS), $\hat{\theta}_{N_2}$ (set-value) and $\hat{\theta}_Q$ (QCCE) when two FIR systems have same general inputs.

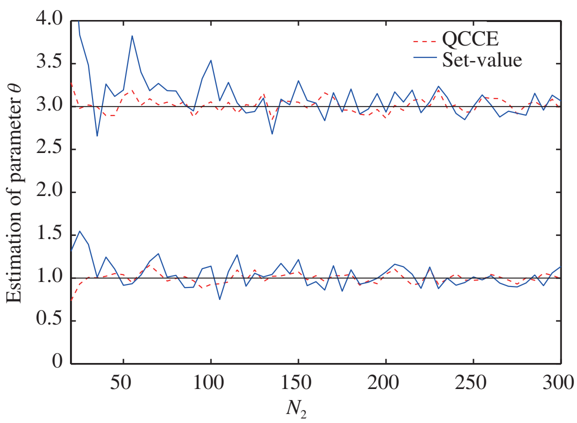

Figure 4 (Color online) The estimation results of parameter θ including $\hat{\theta}_{N_2}$ (set-value) and $\hat{\theta}_Q$ (QCCE) when two FIR systems have different general inputs and the sample size of precise observations is fixed.

(ii) Different inputs. Let $\theta=(1,3)^{\mathrm{T}}$ and $c=2$. The system noises $d_k^{(1)}$ and $d_k^{(2)}$, $k\geqslant 1$ satisfy Assumption 2 with $\sigma_1=2$ and $\sigma_2=2$, respectively. The two input matrixes are both randomly generated from a uniform distribution in the intervals [0,2] and [−2,2], respectively. Similarly, We fix $\bar{N}_1=50$ and vary $\bar{N}_2$ from 20 to 300, and then compare the performance of the QCCE $\hat{\theta}_Q$ with the set-valued estimate $\hat{\theta}_{N_2}$ (see Figure 4).

Figures 3 and 4 display the estimation results when two FIR systems have same and different general inputs, respectively. We can obtain the similar conclusion that the fused QCCE can converge to the true value faster than only utilizing LS method for precise identification or EM method for set-valued identification. Furthermore, the QCCE can get twice the result with half the effort only depending on few precise observations and some set-valued observations.

In short, the QCCE for the fusion estimation algorithm has a faster convergent speed and certain robustness for system noises by making full use of both advantages of identification methods for precise and set-valued observations.

6 Summary

This paper proposed a fusion estimation algorithm for the system identification problem of FIR systems with set-valued and precise observations received from multiple sensors. According to the principle of distributed fusion, individual sensors give their local estimates based on their own observations of different data types by some excellent identification algorithms in the local estimation stage, respectively. Then, we fuse the two estimators by a linear combination with appropriate weights to obtain the QCCE. In addition, it is proved theoretically that the fusion estimation algorithm that makes full use of all the available information has better convergent properties than a single system identification algorithm using only the observations received from single sensor under periodic inputs. Furthermore, extensive numerical simulations show the good performance of the fusion estimation algorithm even under general inputs.

Generally speaking, given the model structure, the effect of parameter estimation mainly depends on the available data and the identification algorithm. Different identification algorithms correspond to different types of data. In the system identification community, researchers' attention is focused on developing the effective identification algorithms with single data type of observations. Information fusion theory provides us a unified framework for obtaining better parameter estimation with all the available data of different types utilized fully, which brings a new perspective for system identification problems.

Furthermore, sample size and data accuracy are two main factors taken into account when analyzing the data. In order to get a good estimation effect, a certain number of samples are needed for precise observations but more samples are necessary for set-valued observations due to the limitation of less accurate information. The proposed fusion estimation algorithm in this paper can absorb the advantages of the two identification algorithms. That is to say, it can achieve a better estimation effect with just a few precise and many set-valued observations. In addition, though the paper only considers the system identification problem for FIR systems, the framework of the proposed estimation fusion algorithm can be easily extended to other systems. How to find out the corresponding algorithms for each sensor in the local estimation stage and how to choose an appropriate fusion strategy are the problems of great significance and worthy of further research.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2016YFB0901902), National Natural Science Foundation of China (Grant Nos. 61803370, 61622309), and National Key Basic Research Program of China (973 Program) (Grant No. 2014CB845301).

References

1 Ljung L. Perspectives on system identification. Annu Rev Control, 2008, 34: 1–12
2 Ljung L. System Identification: Theory for the User. 2nd ed. Englewood Cliffs: Prentice-Hall, 1999
3 Clarke D W. Generalized least squares estimation of the parameters of a dynamic model. In: Proceeding of IFAC Symposium on Identification in Automatic Control Systems, Prague, 1967
4 Astrom K J. Maximum likelihood and prediction error methods. Automatica, 1980, 16: 551–574
5 Kalman R E. A new approach to linear filtering and prediction problems. J Basic Eng, 1960, 82: 35–45
6 Ho Y, Lee R. A Bayesian approach to problems in stochastic estimation and control. IEEE Trans Autom Control, 1964, 9: 333–339
7 Anderson T W, Taylor J B. Some experimental results on the statistical properties of least squares estimates in control problems. Econometrica, 1976, 44: 1289–1302
8 Wang L Y, Zhang J F, Yin G G. System identification using binary sensors. IEEE Trans Autom Control, 2003, 48: 1892–1907
9 Wang L Y, Zhang J F, Yin G G, et al. System Identification with Quantized Observations. Boston: Birkhäuser, 2010
10 Godoy B I, Goodwin G C, Agüero J C, et al. On identification of FIR systems having quantized output data. Automatica, 2011, 47: 1905–1915
11 Zhao Y L, Bi W J, Wang T. Iterative parameter estimate with batched binary-valued observations. Sci China Inf Sci, 2016, 59: 052201
12 Bottegal G, Hjalmarsson H, Pillonetto G. A new kernel-based approach to system identification with quantized output data. Automatica, 2017, 85: 145–152
13 Wang J D, Zhang Q H. Identification of FIR systems based on quantized output measurements: a quadratic programming-based method. IEEE Trans Autom Control, 2015, 60: 1439–1444
14 Yu C P, Zhang C S, Xie L H. Blind system identification using precise and quantized observations. Automatica, 2013, 49: 2822–2830

15 Zhao Y L, Wang L Y, Yin G G, et al. Identification of Wiener systems with binary-valued output observations. Automatica, 2007, 43: 1752–1765

16 Zhao Y L, Zhang J F, Wang L Y, et al. Identification of Hammerstein systems with quantized observations. SIAM J Control Optim, 2010, 48: 4352–4376

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 White F E. Data Fusion Lexicon. Joint Directors of Laboratories, Technical Panel for C^3. 1991. https://apps.dtic.mil/dtic/tr/fulltext/u2/a529661.pdf

19 Lawrence A K. Sensor and Data Fusion Concepts and Applications. 2nd ed. Bellingham: SPIE Optical Engineering Press, 1999

20 Boström H, Andler S F, Brohede M, et al. On the definition of information fusion as a field of research. Neoplasia, 2008, 13: 98–107

21 Han C Z, Zhu H Y, Duan Z S. Multi-Source Information Fusion. Beijing: Tsinghua University Press, 2006

22 Costa O L V. Linear minimum mean square error estimation for discrete-time Markovian jump linear systems. IEEE Trans Autom Control, 1994, 39: 1685–1689

23 Marelli D E, Fu M. Distributed weighted least-squares estimation with fast convergence for large-scale systems. Automatica, 2015, 51: 27–39

24 Li X R. Optimal linear estimation fusion for multisensor dynamic systems. In: Proceedings of the Workshop on Multiple Hypothesis Tracking – A Tribute to Sam Blackman, 2003

25 Sun S L, Deng Z L. Multi-sensor optimal information fusion Kalman filter. Automatica, 2004, 40: 1017–1023

26 Deng Z L, Zhang P, Qi W J, et al. The accuracy comparison of multisensor covariance intersection fuser and three weighting fusers. Inf Fusion, 2013, 14: 177–185

27 Ferguson T S. A Course in Large Sample Theory. New York: Chapman and Hall, 1996

28 Bi W J, Kang G L, Zhao Y L, et al. SVSI: fast and powerful set-valued system identification approach to identifying rare variants in sequencing studies for ordered categorical traits. Ann Human Genets, 2015, 79: 294–309