International Journal of Control
Vol. 83, No. 3, March 2010, 538–551

Kalman filter-based identification for systems with randomly missing measurements in a network environment

Yang Shi[a]* and Huazhen Fang[b]

[a]*Department of Mechanical Engineering, University of Victoria, Victoria, BC, V8W 3P6 Canada;* [b]*Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK, S7N 5A9, Canada*

(*Received 18 February 2009; final version received 20 August 2009*)

We consider the problem of parameter estimation and output estimation for systems in a transmission control protocol (TCP) based network environment. As a result of networked-induced time delays and packet loss, the input and output data are inevitably subject to randomly missing data. Based on the available incomplete data, we first model the input and output missing data as two separate Bernoulli processes characterised by probabilities of missing data, then a missing output estimator is designed, and finally we develop a recursive algorithm for parameter estimation by modifying the Kalman filter-based algorithm. Under the stochastic framework, convergence properties of both the parameter estimation and output estimation are established. Simulation results illustrate the effectiveness of the proposed algorithms.

Keywords: networked control systems; system identification; Kalman filter; randomly missing data

1. Introduction

To meet the increasing demands of 'teleautomation', modularity, integrated diagnostics, quick maintenance and decentralisation of control, networked control systems (NCSs) have received remarkable attention worldwide in the past decade. An NCS is a feedback control system which has its control loop physically connected via real-time communication networks. Communication networks are used to exchange information (reference input, control input, plant output, etc.) among control system components (sensors, controller, actuators, etc.) (Zhang, Branieky, and Philips 2001). Because network-induced time delays and data packet losses will worsen the control performance and even destroy the stability, many studies have been devoted to controller design for NCSs, e.g. Chow and Tipsuwan (2001), Tipsuwan and Chow (2003), Hespanha, Naghshtabrizi, and Xu (2007) and the references therein. However, it is known that dynamic models are necessary in many control design methodologies, so prior to the development of controllers, system identification, aimed at building models from measured data, must be carried out. As the transmission condition of a network is largely featured by randomness, this article thus discusses how to identify model parameters of a plant subject to randomly missing measurements in a network environment, as illustrated in Figure 1.

In an NCS, the plant plus the actuator and sensor are installed at a remote location. At the near end, an input transmitter sends the input signal to excite the plant, and an output receiver collects the plant output. Both are done through a network. A parameter estimation module identifies the parameters of the plant in an online manner. The network is assumed to operate under *TCP-like protocols* which can guarantee an acknowledgement of received packets; it has been widely used in research on state estimation and control over networks (Schenato, Sinopoli, Franceschetti, Poolla, and Sastry 2007). For such a networked system, both input and output are subject to randomly missing data due to the nature of communication networks. Obviously, identification over networks is more challenging, and classical parameter estimation methods, such as least squares (LS) (Hsia 1977) and stochastic gradient (SG), can no longer be applied *directly*.

Generally, existing research on identification for systems with incomplete input–output data can be divided into two categories.

1.1 *Identification with regularly missing output data*

Systems with regular missing data can also be viewed as multirate systems which have uniform but various input/output sampling rates (Chen and Francis 1995). Such systems may have regular-output-missing feature. In Ding and Chen (2004a), Ding and Chen proposed an auxiliary model-based method and developed a modified recursive least squares (RLS) algorithm to simultaneously estimate the system parameters and

*Corresponding author. Email: yshi@uvic.ca

ISSN 0020–7179 print/ISSN 1366–5820 online

DOI: 10.1080/00207170903273987
http://www.informaworld.com

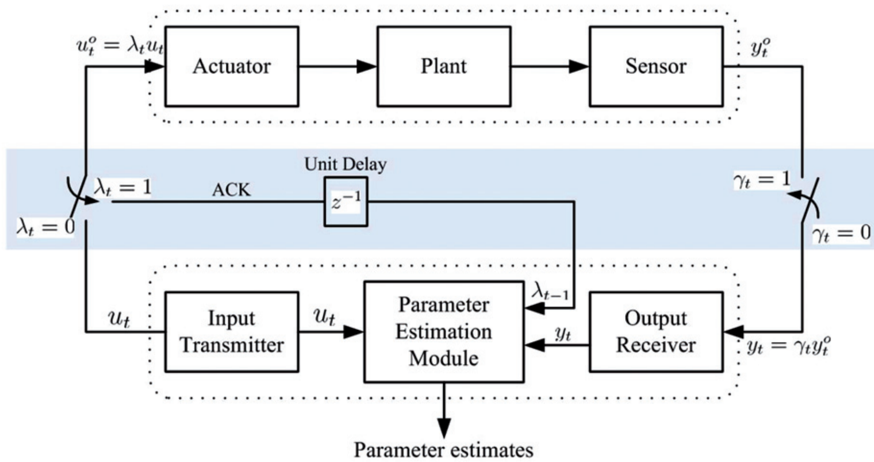

Figure 1. Scenario of identification over a lossy network under TCP-like protocols.

the unavailable outputs of a dual-rate system. Other related research on identification of multirate system models can be referred to Wang, Chen, and Huang (2004), Ding and Chen (2004b), Ding, Liu, and Yang (2008), Li, Shah, and Chen (2001), to name a few.

1.2 *Identification with randomly missing output data*

In many industrial applications non-uniform sampling is widely adopted, i.e. the output is unequally measured at time-varying sampling rates. In Larsson and Söderström (2002), Larsson and Söderström studied the LS-based estimation of continuous-time autoregressive (AR) models from discrete-time data distributed non-uniformly. Parameter estimation of discrete-time AR models with missing observations was studied by using an SG-type method (Mirsaidi, Fleury, and Oksman 1997). Yet in both works no or only qualitative convergence analysis was given. An LS-based scheme was presented in Wallin, Isaksson, and Ljung (2000) for autoregressive exogenous (ARX) model identification from incomplete input/output data; however, this scheme was not recursive for online use, and it is sensitive to data missing pattern. Albertos et al. and Wallin et al. studied the estimation of missing output measurements (Albertos, Sanchis, and Sala 1999; Wallin, Isaksson, and Noréus 2001), which paved the way for further estimating parameters; a systematic description of joint output and parameter estimation with irregular output missing was provided in Sanchis and Albertos (2002). Isaksson used the Kalman filtering and fixed-interval smoothing techniques to reconstruct the missing data first and then identified ARX models by LS, maximum-likelihood (ML) and expectation-maximisation (EM) algorithms (Isaksson 1993); the identification process, however, is one-step but not recursive. A frequency domain solution was proposed in Pintelon and Schoukens (2000) and it treated all missing measurements as parameters, potentially leading to a large amount of parameters to be identified. An l_p norm parameter estimation algorithm was proposed in Chen and Chen (2000) to address parameter estimation from input/output data that has missing values as well as noisy disturbances; further, it was nicely shown that: if the probability of the missing measurements is less than 1/2, then the parameter estimates will converge to the true parameter; however, this algorithm is conservative in two-fold: (1) there was a restrictive assumption that both input and output missing have the same probability, which is not always the case in practical NCSs; (2) the probability of missing data is less than 1/2.

Parameter estimation subject to missing data is also an important research topic in time series analysis. ML and EM algorithms appear to be popular solutions, by maximising the likelihood functions of linear or nonlinear models, see Jones (1980), Titterington (1984), Charalambous and Logothetis (2000), Ibrahim, Lipsitz, and Horton (2001) and the references therein.

However, to the best of the authors' knowledge, system identification in a lossy network environment has not been fully investigated, which is the focus of this article. The celebrated Kalman filter has been elegantly applied to estimate system parameters (Guo 1990; Cao and Schwartz 2003); however, when the networked system is subject to randomly missing measurements, it cannot be directly applied here. The main contributions of this work are three-fold:

(1) Under the TCP-like protocols, to model the input and output missing as two separate Bernoulli processes, each of which consists of a sequence of independent random variables

taking either one (non-missing) or zero (missing).

(2) To design an output estimator to reconstruct the missing output data, and further to develop a modified Kalman filter-based recursive identification algorithm for networked systems, by taking advantage of TCP protocols.

(3) To well establish the convergence properties of parameter estimation and output estimation under the stochastic framework, mainly by using the martingale tool.

The rest of this article is organised as follows. The problem formulation is described in Section 2. In Section 3, we propose the missing output estimator, and further develop the modified Kalman filter-based recursive algorithm for estimating system parameters. In Section 4, convergence properties of the proposed algorithms are analysed. In Section 5, three illustrative examples are given to demonstrate the effectiveness of the proposed algorithms. Finally, some concluding remarks are given in Section 6. In order to preserve the flow of presentation, proofs of some technical results are postponed to appendices to facilitate the reading.

Notations: The notations used throughout the article are fairly standard. 'E' denotes the expectation. The superscript 'T' stands for matrix transposition; $\lambda_{\max/\min}(X)$ represents the maximum/minimum eigenvalue of X; $|X| = \det(X)$ is the determinant of a square matrix X; $\|X\|^2 = \mathrm{tr}(XX^{\mathrm{T}})$ stands for the trace of XX^{T}. If $\exists\ \delta_0 \in \mathrm{R}^+$ and $k_0 \in \mathrm{Z}^+$, $|f(k)| \le \delta_0 g(k)$ for $k \ge k_0$, then $f(k) = O(g(k))$; if $f(k)/g(k) \to 0$ for $k \to \infty$, then $f(k) = o(g(k))$.

2. Problem formulation

The identification problem in a TCP-based network environment is shown in Figure 1. Let us consider the following output-error model with intermittent input and output information:

$$x_t^o = \frac{B_z}{A_z} u_t^o, \tag{1}$$

$$y_t^o = x_t^o + v_t, \tag{2}$$

$$u_t^o = \lambda_t u_t, \tag{3}$$

$$y_t = \gamma_t y_t^o, \tag{4}$$

where u_t^o is the input to the actuator, u_t is the desired input from the input transmitter, y_t^o is the sensor output and y_t is output transmitted to the output receiver. A_z and B_z are polynomials in the unit delay operator z^{-1}:

$$A_z = 1 + a_1 z^{-1} + a_2^{-2} + \cdots + a_{n_a} z^{-n_a},$$
$$B_z = b_0 + b_1 z^{-1} + b_2^{-2} + \cdots + b_{n_b} z^{-n_b}.$$

The polynomial orders n_a and n_b are assumed to be known. Equations (1) and (2) can be written equivalently as the following linear regression model:

$$x_t^o = \varphi_t^{o\mathrm{T}} \theta, \quad y_t^o = x_t^o + v_t, \tag{5}$$

where

$$\varphi_t^o = \left[-x_{t-1}^o \ -x_{t-2}^o \ \cdots \ -x_{t-n_a}^o \ u_t^o \ u_{t-1}^o \ \cdots \ u_{t-n_b}^o \right]^{\mathrm{T}},$$
$$\theta = \left[a_1 \ a_2 \ \cdots \ a_{n_a} \ b_0 \ b_1 \ \cdots \ b_{n_b} \right]^{\mathrm{T}}.$$

Regression vector φ_t^o contains lagged input and output values of the plant, and the parameter vector θ contains model parameters to be estimated.

As aforementioned, the network is lossy and thus results in randomly input–output missing, so Bernoulli random variables λ_t and γ_t are introduced to characterise the data missing pattern. The probability distributions of λ_t and γ_t are defined as

$$\mathrm{P}(\lambda_t) = \begin{cases} \lambda, \ \text{if } \lambda_t = 1, \\ 1 - \lambda, \ \text{else if } \lambda_t = 0, \end{cases} \quad \text{and}$$
$$\mathrm{P}(\gamma_t) = \begin{cases} \gamma, \ \text{if } \gamma_t = 1, \\ 1 - \gamma, \ \text{else if } \gamma_t = 0. \end{cases}$$

Here, the difference between γ_t and λ_t is noteworthy. Under TCP-like protocols, receipt of a packet u_t is acknowledged by a packet message. Yet the acknowledgement will have a unit-time delay. Therefore, at time instant t, γ_t and λ_{t-1} (instead of λ_t) are known. We can define the following information set:

$$\mathcal{T}_t \triangleq \{\Gamma_t Y_t, U_t, \Lambda_{t-1}\}, \tag{6}$$

where $\Gamma_t Y_t = (\gamma_t y_t, \gamma_{t-1} y_{t-1}, \ldots, \gamma_1 y_1)$, $U_t = (u_t, u_{t-1}, \ldots, u_1)$ and $\Lambda_t = (\lambda_t, \lambda_{t-1}, \ldots, \lambda_1)$. In fact, $\mathcal{T}_t$ provides all the data information that is available for system identification.

To this end, the objectives are posed in the following:

(1) How to estimate the parameter vector θ based on the incomplete input-output data information $\mathcal{T}_t$?

(2) How to evaluate the accuracy of the parameter estimation and the missing output estimation?

In what follows, a recursive algorithm for identification in a TCP-based network environment will be proposed. The algorithm is developed based on the Kalman filter. Further, convergence properties of the proposed algorithm will be analysed.

Remark 2.1: As shown in (3) and (4), the missing input or output data is replaced by zero. It should be noted that other compensation schemes, i.e., replacement by previous or latest available value, can be used as well. However, the use of (3) and (4) could conveniently facilitate the analysis.

3. Modified Kalman filter-based recursive estimation algorithm

Consider the model in (5). It is shown by Guo (1990) and Cao and Schwartz (2003) that the Kalman filter can be used to perform parameter estimation. By using the auxiliary model, we have the following Kalman filter-based algorithm:

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t^a(y_t^o - \varphi_t^{a\mathrm{T}}\hat{\theta}_{t-1}), \tag{7}$$

$$K_t^a = \frac{P_{t-1}^a \varphi_t^a}{r_v + \varphi_t^{a\mathrm{T}} P_{t-1}^a \varphi_t^a}, \tag{8}$$

$$P_t^a = P_{t-1}^a - \frac{P_{t-1}^a \varphi_t^a \varphi_t^{a\mathrm{T}} P_{t-1}^a}{r_v + \varphi_t^{a\mathrm{T}} P_{t-1}^a \varphi_t^a}, \tag{9}$$

$$x_t^a = \varphi_t^{a\mathrm{T}}\hat{\theta}_t, \tag{10}$$

$$\varphi_t^a = \left[-x_{t-1}^a \;\; -x_{t-2}^a \;\; \cdots \;\; -x_{t-n_a}^a \;\; u_t^o \;\; u_{t-1}^o \;\; \cdots \;\; u_{t-n_b}^o \right]^{\mathrm{T}}, \tag{11}$$

where $\hat{\theta}_t$ represents the estimated parameter vector at instant t.

However, due to missing input–output data (as shown in (3) and (4)), the information available for identification is constrained within $\mathcal{T}_t$. Hence, the above algorithm will fail if directly used to estimate the parameters of the networked plant. Thus we are motivated to design a new algorithm accounting for the problem of randomly missing data in a network environment.

From the definition of $\mathcal{T}_t$ and Figure 1, let us take a closer look at how to deal with the problem of randomly input and output missing: (1) Thanks to TCP acknowledgement mechanism, the input to the actuator $u_t^o = \lambda_t u_t$ is available, but there exists a unit-time delay. For the input missing, we propose to delay identification for a unit-time, and if it is missed, then the input is regarded as zero. (2) The sensor output y_t^o is subject to random loss when transmitted through network to the output receiver. Therefore, an output estimator is to be designed.

Let us begin with output estimator design. Inspired by the work on output prediction by Albertos et al. (1999), the following estimator is to be used:

$$z_t = y_t + (1-\gamma_t)\hat{y}_t, \tag{12}$$

where

$$\hat{y}_t = \varphi_t^{\mathrm{T}}\hat{\theta}_t$$

and φ_t will be defined later in (17). The above estimator adaptively reconstructs the missing output data. It has an intuitive yet efficient structure, and its convergence properties will be analysed in Section 4.

Now we are in good position to derive the recursive parameter estimation algorithm. Replacing y_t^o in the algorithm (7)–(11) by z_t, and incorporating the random variable γ_t leads to

$$\hat{\theta}_{t+1}|\mathcal{T}_{t+1} = \hat{\theta}_t + K_{t+1}(z_t - \varphi_t^{\mathrm{T}}\hat{\theta}_t), \tag{13}$$

$$K_{t+1}|\mathcal{T}_{t+1} = \frac{P_t\varphi_t}{r_v + \varphi_t^{\mathrm{T}}P_t\varphi_t}, \tag{14}$$

$$P_{t+1}|\mathcal{T}_{t+1} = P_t - \gamma_t\frac{P_t\varphi_t\varphi_t^{\mathrm{T}}P_t}{r_v + \varphi_t^{\mathrm{T}}P_t\varphi_t}, \tag{15}$$

$$x_t|\mathcal{T}_{t+1} = \varphi_t^{\mathrm{T}}\hat{\theta}_{t+1}, \tag{16}$$

$$\begin{aligned}\varphi_t|\mathcal{T}_{t+1} = \big[&-x_{t-1} \;\; -x_{t-2} \;\; \cdots \;\; -x_{t-n_a} \;\; \lambda_t u_t \\ &\lambda_{t-1}u_{t-1} \;\; \cdots \;\; \lambda_{t-n_b}u_{t-n_b}\big]^{\mathrm{T}}.\end{aligned} \tag{17}$$

For ease of presentation, we shall drop $\mathcal{T}_{t+1}$ and use $\hat{\theta}_{t+1}$, K_{t+1}, P_{t+1}, x_t and φ_t alone instead.

Remark 3.1: A closer scrutiny of (13) shows that it can be decomposed into two equations:

$$\begin{cases} \hat{\theta}_{t+1} = \hat{\theta}_t + K_{t+1}(y_t - \varphi_t^{\mathrm{T}}\hat{\theta}_t), & \text{if } \gamma_t = 1; \\ \hat{\theta}_{t+1} = \hat{\theta}_t, & \text{else if } \gamma_t = 0. \end{cases}$$

Similarly, we can decompose (15) in a similar way. It is noted that $\hat{\theta}_t$ and P_t are updated only when the output is available, and will remain unchanged otherwise.

Remark 3.2: The algorithm in (13)–(17) can be applied to cases when b_0 is either zero or non-zero, through incorporating the unit-time delay into the identification procedure. In fact, if $b_0 = 0$, then the unit delay may not be needed. However, the TCP protocol is necessary in either case to obtain the information about missing input data.

4. Convergence analysis

To begin our quest of establishing convergence properties of the estimation algorithms, some necessary preliminaries are presented first. The convergence analysis of the parameter estimation in (13)–(15) and the output estimation in (12) are to be carried out under the stochastic framework, inspired by Chen and Guo (1991), Ding and Chen (2003) and Ding and Chen (2004a).

4.1 *Preliminaries*

Some basic facts about the positive definite matrices will be used in this section and are summarised in the following lemma.

Lemma 4.1 (Horn and Johnson 1991)**:** *Let A, B be $n \times n$ positive definite matrices with the relation $A \leq B$, and C be a $n \times m$ matrix. Then*

$$\lambda_{\min}(A)I \leq A \leq \lambda_{\max}(A)I, \tag{18}$$

$$C^{\mathrm{T}}AC \leq C^{\mathrm{T}}BC, \tag{19}$$

where I is the identity matrix. If the eigenvalues of A and B are arranged in the same order, then

$$\lambda_k(A) \leq \lambda_k(B) \tag{20}$$

for $k = 1, 2, \ldots, n$.

To facilitate the convergence analysis, we equivalently rewrite the proposed algorithm, as shown in the following lemma.

Lemma 4.2: *The algorithm* (13)–(15) *can be written equivalently as follows*:

$$\hat{\theta}_{t+1} = \hat{\theta}_t + r_v^{-1} P_{t+1}\varphi_t(z_t - \varphi_t^{\mathrm{T}}\hat{\theta}_t), \tag{21}$$

$$P_{t+1}^{-1} = P_t^{-1} + \gamma_t r_v^{-1}\varphi_t\varphi_t^{\mathrm{T}}. \tag{22}$$

Proof: The proof is straightforward and thus is omitted here. □

Further, define

$$(P_{t+1}^o)^{-1} = p_0 I + r_v^{-1}\varphi_t^o\varphi_t^{o\mathrm{T}},$$
$$r_t^o = \mathrm{tr}\big((P_t^o)^{-1}\big),$$
$$r_t = \mathrm{tr}\big(P_t^{-1}\big),$$

where p_0 is a positive real value large enough. The relations between r_t^o and $|P_t^{o-1}|$, and between r_t and $|P_t^{-1}|$, are established in the following lemma.

Lemma 4.3: *The following relations hold*:

$$\ln\big|(P_t^o)^{-1}\big| = O\big(\ln r_t^o\big), \quad \ln \mathrm{E}|P_t^{-1}| = O(\ln \mathrm{E}r_t). \tag{23}$$

Proof: It can be proved by following the similar line in Ding and Chen (2004a). □

The next lemma shows the convergence of three infinite series that will be useful later.

Lemma 4.4: *The following inequalities hold*:

$$\sum_{i=1}^{t} r_v^{-1}\mathrm{E}\big(\gamma_i\varphi_i^{\mathrm{T}}P_{i+1}\varphi_i\big) \leq \ln \mathrm{E}|P_{t+1}^{-1}| + n_0 \ln p_0 \ \textit{almost surely (a.s.)} \tag{24}$$

$$\sum_{i=1}^{\infty} r_v^{-1}\frac{\mathrm{E}\big(\gamma_i\varphi_i^{\mathrm{T}}P_{i+1}\varphi_i\big)}{\big(\ln \mathrm{E}|P_{i+1}^{-1}|\big)^c} < \infty \ \textit{a.s.}, \tag{25}$$

$$\sum_{i=1}^{\infty} r_v^{-1}\frac{\mathrm{E}\big(\gamma_i\varphi_i^{\mathrm{T}}P_{i+1}\varphi_i\big)}{\ln \mathrm{E}|P_{i+1}^{-1}|\big(\ln\ln \mathrm{E}|P_{i+1}^{-1}|\big)^c} < \infty \ \textit{a.s.}, \tag{26}$$

where $n_0 = n_a + n_b$ and $c > 1$.

Proof: The proof can be done along the similar way as Lemma 2 in Ding and Chen (2004b) and is omitted here. □

The following is the well-known martingale convergence theorem that lays the foundation for the upcoming convergence analysis.

Theorem 4.5 (Goodwin and Sin 1984)**:** *Let $\{X_t\}$ be a sequence of non-negative random variables adapted to an increasing σ-algebras $\{\mathcal{F}_t\}$. If*

$$\mathrm{E}(X_{t+1}|\mathcal{F}_t) \leq (1+\epsilon_t)X_t - \alpha_t + \beta_t, \ \textit{a.s.}$$

where $\alpha_t \geq 0$, $\beta_t \geq 0$ and $\mathrm{E}X_0 < \infty$, $\sum_{i=0}^{\infty}|\epsilon_i| < \infty$, $\sum_{i=0}^{\infty}\beta_i < \infty$ a.s., then X_t converges almost surely to a finite random variable and

$$\lim_{N\to\infty}\sum_{i=0}^{N}\alpha_i < \infty, \ \textit{a.s.}$$

4.2 *Convergence of parameter estimation*

We will prove that the parameter estimation algorithm (13)–(15) is convergent by constructing a martingale process satisfying the conditions of Theorem 4.5. This is the essential of the next theorem.

Theorem 4.6: *Assume that the driven noise $\{v_t, \mathcal{F}_t\}$ is a martingale difference sequence adapted to a family of increasing σ-algebras $\{\mathcal{F}_t\}$. For the system considered in* (5), *the following assumptions are made*:

(A1) $\mathrm{E}(v_t|\mathcal{F}_{t-1}) = 0$, *a.s.*,

(A2) $\mathrm{E}\big(v_t^2|\mathcal{F}_{t-1}\big) = r_v < \infty$, *a.s.*,

(A3) $\exists\, \alpha_0, \beta_0, c_0 \in R^+$ *and* $t_0 \in N^+$,

$$\alpha_0 I \leq \frac{1}{t}\sum_{i=1}^{t}\varphi_i^o\varphi_i^{o\mathrm{T}} \leq \beta_0 t^{c_0} I, \ \textit{for}\ t \geq t_0,$$

(A4) $G_z = \dfrac{1}{A_z} - \dfrac{1}{2}$ *is strictly positive real.*

Then the square parameter estimation error, $\|\hat{\theta}_t - \theta\|^2$, *produced by the algorithm* (13)–(15), *satisfies*

$$\text{(C1)} \quad \|\hat{\theta}_t - \theta\|^2 = O\left[\frac{(\ln t)^c}{t}\right] \to 0 \ a.s., \ c > 1,$$

$$\text{(C2)} \quad \|\hat{\theta}_t - \theta\|^2 = O\left[\frac{\ln t(\ln\ln t)^c}{t}\right] \to 0 \ a.s., \ c > 1.$$

Proof: The proof is given in Appendix A. □

Remark 4.7: Assumptions A1 and A2 show that $\{v_t\}$ is an independent noise sequence with zero mean and bounded time-varying variance. Yet it is noteworthy that the exact r_v might be difficult to determine in practice. One simple solution is to assume $r_v = 1$. By doing this, the consequent convergence proof will be very similar, with the only difference in the convergence speed.

Remark 4.8: The assumption A3 refers to the persistent excitation (PE) condition that is a standard assumption; its practical meaning is to have rich enough excitation signals to drive and further identify the system.

Remark 4.9: For systems with complete input-output measurements, under the stochastic framework, the convergence analysis of LS identification algorithms and gradient-based algorithms have been discussed extensively (Guo 1990; Chen and Guo 1991; Ding and Chen 2004a, b; Ding et al. 2008). In Theorem 4.6, following the similar line, the results have been extended to the convergence properties of the modified Kalman filter-based parameter estimation algorithm for systems with randomly missing measurements in a TCP-like network environment. It is worthwhile noting that γ_t is involved in the developed algorithm to characterise the random missing measurements, which also poses challenges on the proof of Theorem 4.6; in this sense, the proof procedure is different from existing results in terms of incorporating the data missing at random into account.

Remark 4.10: Theorem 4.6 reveals that the parameter estimation error of the algorithm (13)–(15), even in the presence of output missing will converge to zero at the speed of $O[(\ln t)^c/t]$.

4.3 *Convergence of output estimation*

It is also important to analyse the convergence properties of the output estimation. To establish this, we give the next theorem.

Theorem 4.11: *Assume that the assumptions* (A1)–(A4) *hold, and*

(A5) *The input is bounded, i.e.* $u_t^2 < \infty$ *for any* t.

Then there exists a positive integer t_0 *such that for any* $t \geq t_0$ *the output estimation error* $z_t - y_t$ *satisfies*

$$\text{(C3)} \quad \sum_{i=t_0}^{t} (z_i - y_i^o)^2 = O\left[(\ln t)^{c+1}\right], \quad a.s., \quad c > 1,$$

$$\text{(C4)} \quad \frac{1}{t}\sum_{i=t_0}^{t} (z_i - y_i^o)^2 = O\left[\frac{(\ln t)^{c+1}}{t}\right] \to 0, \quad a.s., \quad c > 1.$$

Proof: The proof is given in Appendix B. □

Remark 4.12: The proposed output estimator possesses a simple structure, yet it is effective: the estimation error is proven to converge to zero in average sense at the speed of $O[(\ln t)^{c+1}/t]$.

Remark 4.13: The assumption (A5) is very mild. The input excitation usually is not possible to be infinity in practice. In addition, because of the network-induced missing data, the boundedness of the actual input to the plant is further guaranteed.

5. Numerical simulation

In this section, the proposed algorithm is examined through simulation studies.

The proposed algorithm (13)–(15) is applied to the input–output data collected from a second-order SISO plant placed in a network environment:

$$\begin{aligned} y_t^o &= \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u_t^o + v_t, \\ u_t^o &= \lambda_t u_t, \\ y_t &= \gamma_t y_t^o. \end{aligned}$$

The desired input $\{u_t\}$ is a uniformly distributed sequence with zero mean and unit variance. As aforementioned, due to the network-induced randomly missing data, the actual input to the plant is $\{u_t^o\}$, an intermittent version of $\{u_t\}$. In a similar way, $\{y_t^o\}$ is the output response of the plant, yet it is $\{y_t\}$ that is finally received and used for identification. $\{v_t\}$ is a Gaussian white noise sequence with zero mean and variance r_v. The parameter vector $\theta = [a_1 \ a_2 \ b_0 \ b_1]^{\mathrm{T}}$ is to be estimated. Here, θ is supposed to be

$$\theta = [0.523 \ 0.349 \ 0.440 \ 0.762]^{\mathrm{T}}.$$

In the following simulation studies, we carry out experiments for three different cases regarding the data completeness and the data missing pattern.

Example 1: $\lambda = 0.8$ and $\gamma = 0.7$. In this case, about 20% of the input data and 30% of the output data are missing.

The estimated parameters and corresponding estimation errors of four unknown parameters are shown in Figure 2 and Table 1, respectively. It is observed that

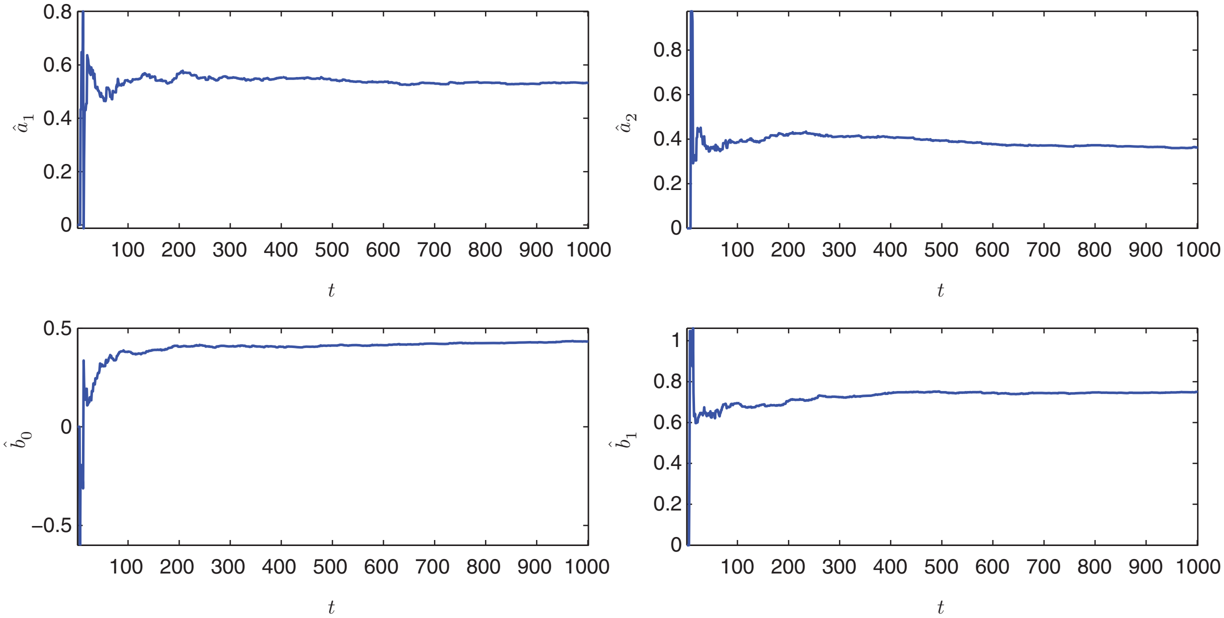

Figure 2. Example 1: estimates of the unknown parameters a_1, a_2, b_0 and b_1 ($\lambda = 0.8$ and $\gamma = 0.7$).

Table 1. Example 1: intermediate parameter estimates and estimation errors ($\lambda = 0.8$ and $\gamma = 0.7$).

t	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_0$	$\hat{b}_1$	$\delta_{par}(\%)$
10	0.648322	0.974209	−0.262872	0.946401	89.3932
63	0.507489	0.355992	0.352118	0.643831	13.7077
125	0.551276	0.403267	0.367646	0.675491	11.8644
250	0.553247	0.422744	0.411101	0.715296	8.95045
500	0.543309	0.393619	0.412163	0.748159	5.3677
1000	0.532908	0.363081	0.433015	0.749644	2.06329
True values	0.523	0.349	0.440	0.762	

all the parameter estimates gradually converge to their true values as t increases.

To further quantify the estimation accuracy, define the relative parameter estimation error as

$$\delta_{par}\ (\%) = \frac{\|\hat{\theta}_t - \theta\|}{\|\theta\|} \times 100\%.$$

It is shown in Figure 3 that δ_{par} has a clear tendency to approach zero. To examine the output estimation performance, a comparison between the estimated outputs and true outputs during the time range $501 \le t \le 550$ is illustrated in Figure 4: dashed lines illustrate time instants when data missing occurs, and corresponding small asterisks represent the estimated outputs at these time instants. In addition, define the average output estimation error

$$\delta_{out} = \frac{1}{t}\sum_{i=1}^{t}\left(z_i - y_i^o\right)^2.$$

The curve of the average output estimation error is shown in Figure 5. From both Figures 4 and 5, we can observe that the output estimation also exhibits good performance.

Example 2: $\lambda = 0.4$ **and** $\gamma = 0.2$. In the second case, about 60% of the input data and 80% of the output data are missing, respectively. It is clear that in this example, the missing data scenario is much worse than that in Example 1.

Estimates of four parameters are shown in Figure 6 and Table 2, respectively. Even though the available output measurements are more scarce than those in Example 1, it is still observed that all the parameter estimates gradually converge to their true values as t increases.

The relative estimation error, δ_{par}, shown in Figure 7, is still approaching to zero. Performance of the output estimator in terms of both the difference from the true outputs and average output estimation error is illustrated in Figures 8 and 9 (dashed red curve), respectively.

By comparing Figure 3 and Figure 7, and Figure 5 and Figure 9, we note that: (1) the estimation

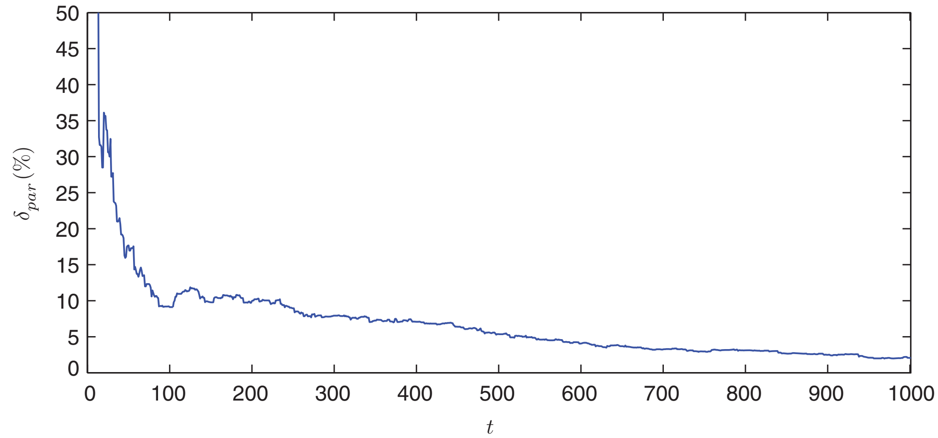

Figure 3. Example 1: relative parameter estimation error versus time ($\lambda = 0.8$ and $\gamma = 0.7$).

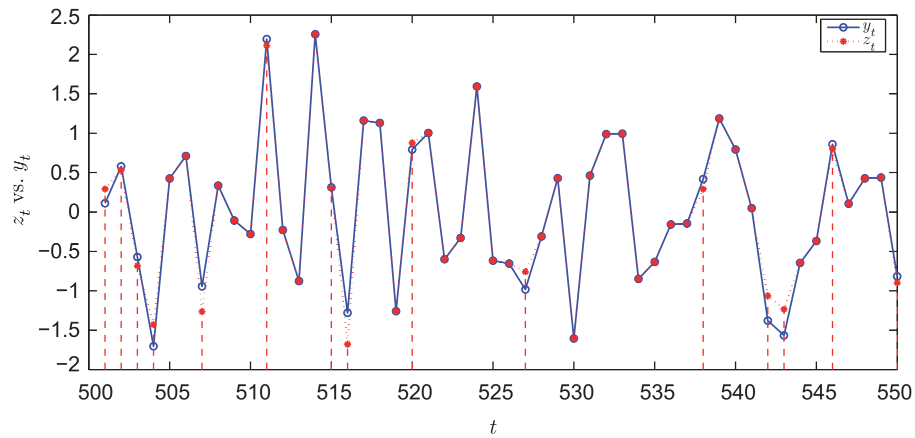

Figure 4. Example 1: comparison between estimated and true outputs ($\lambda = 0.8$ and $\gamma = 0.7$).

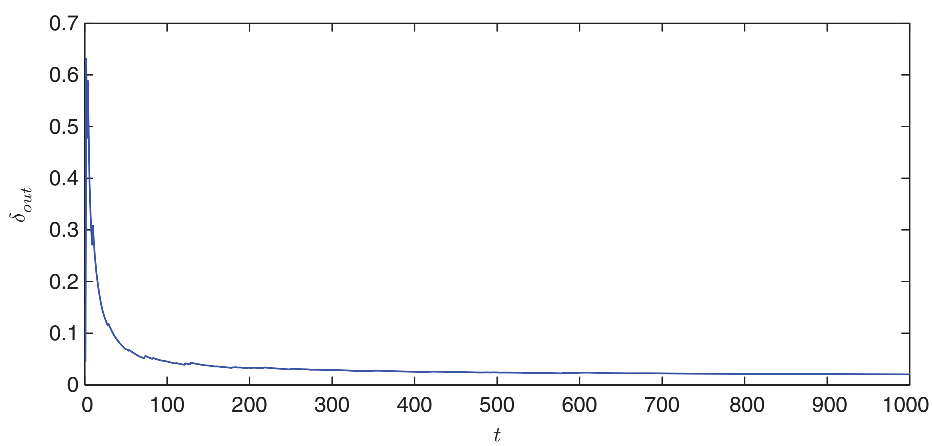

Figure 5. Example 1: average output estimation error versus time ($\lambda = 0.8$ and $\gamma = 0.7$).

performance in Example 1 is better than that in Example 2, because there were less data missed in Example 1; (2) to achieve the same level of estimation accuracy, more data would be needed in Example 2, when more measurements are missing in this case; (3) the estimation performance depends on the data completeness that can be characterised by both λ and γ.

Example 3: Different data missing patterns. It is also paramount to explore the influence of missing data patterns, e.g. $\{\lambda_t\}$ and $\{\gamma_t\}$, on the estimation performance.

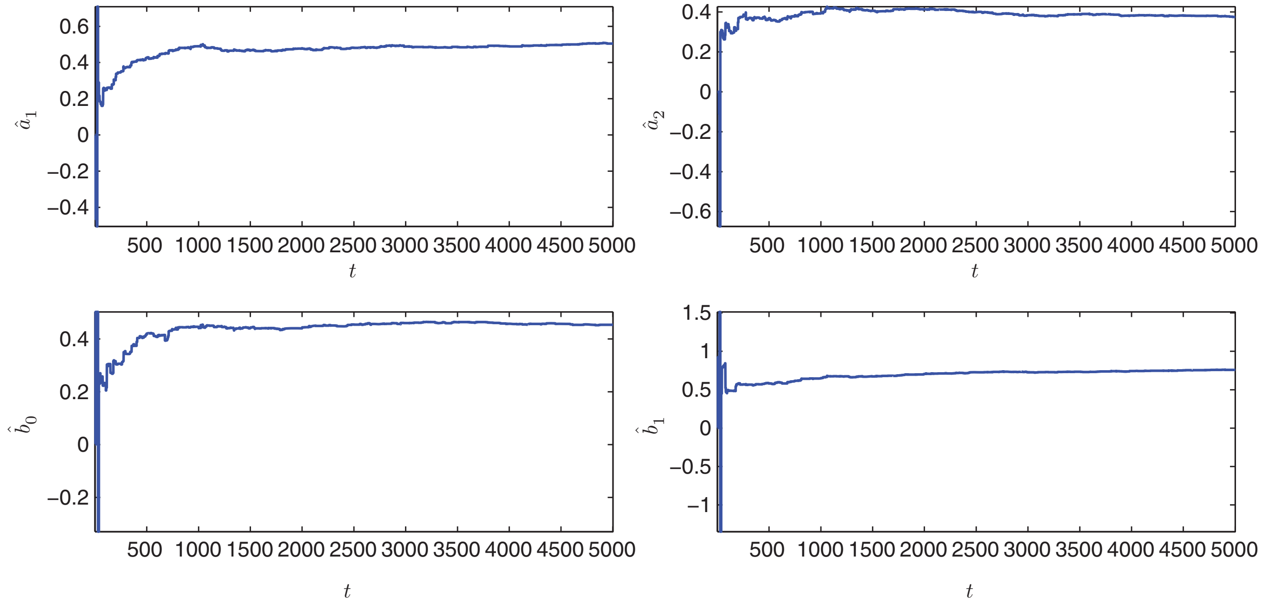

Figure 6. Example 2: estimates of the unknown parameters a_1, a_2, b_0 and b_1 ($\lambda = 0.4$ and $\gamma = 0.2$).

Table 2. Example 2: intermediate parameter estimates and estimation errors ($\lambda = 0.4$ and $\gamma = 0.2$).

t	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_0$	$\hat{b}_1$	$\delta_{par}(\%)$
10	−0.464422	0.000000	0.502433	0.924343	98.1655
313	0.375595	0.371118	0.342766	0.566215	24.4651
625	0.441897	0.361778	0.411453	0.587723	18.0081
1250	0.465504	0.409613	0.44603	0.674206	11.2201
2500	0.48064	0.399537	0.454141	0.7259	7.07334
5000	0.504581	0.37537	0.453457	0.758708	3.23836
True values	0.523	0.349	0.440	0.762	

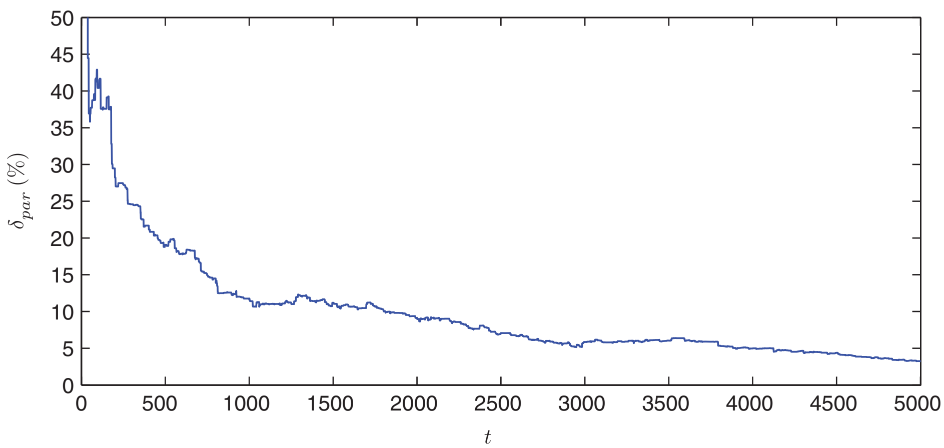

Figure 7. Example 2: relative parameter estimation error versus time ($\lambda = 0.4$ and $\gamma = 0.2$).

In this example, we implement the proposed algorithm twice with same $\lambda = 0.8$ and $\gamma = 0.7$, but the input and output availability sequences $\{\lambda_t\}$ and $\{\gamma_t\}$s are randomly generated and thus different.

The resulting relative parameter estimation errors and average output estimation errors are visually compared in Figures 10 and 11. Obviously, even with identical λ and γ, estimation performance may still vary. In fact, not only the data completeness, but also the missing data patterns are significant factors for identification of systems with randomly missing measurements in a network environment.

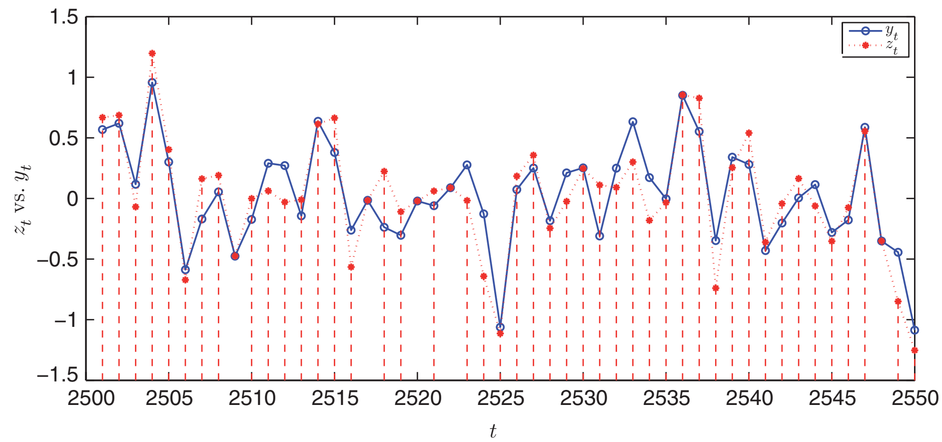

Figure 8. Example 2: comparison between estimated and true outputs (λ = 0.4 and γ = 0.2).

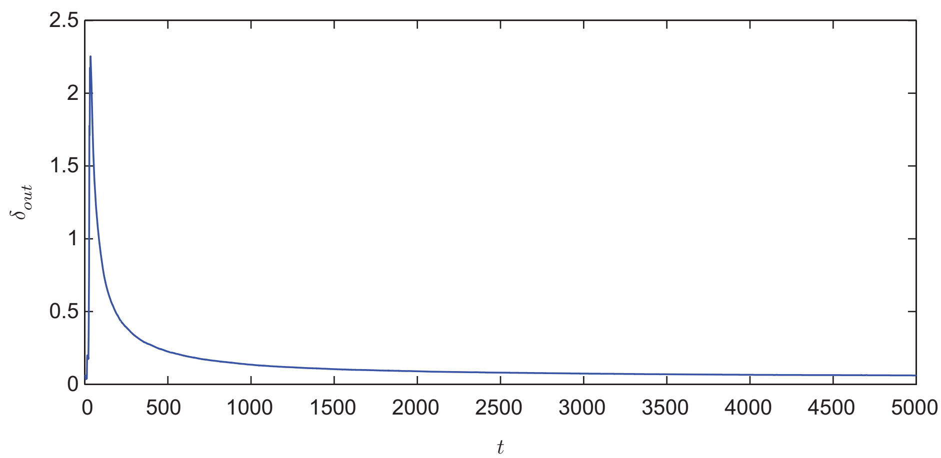

Figure 9. Example 2: average output estimation error versus time (λ = 0.4 and γ = 0.2).

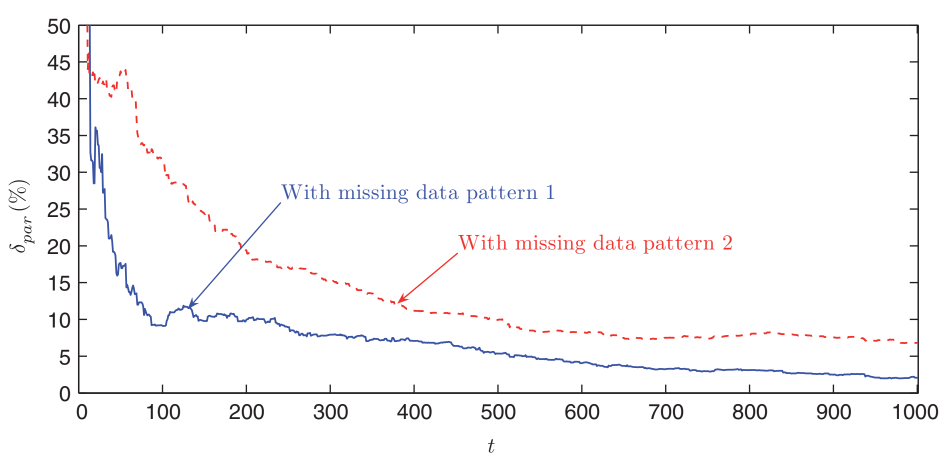

Figure 10. Example 3: relative parameter estimation errors for two cases with different data missing patterns (λ = 0.8 and γ = 0.7).

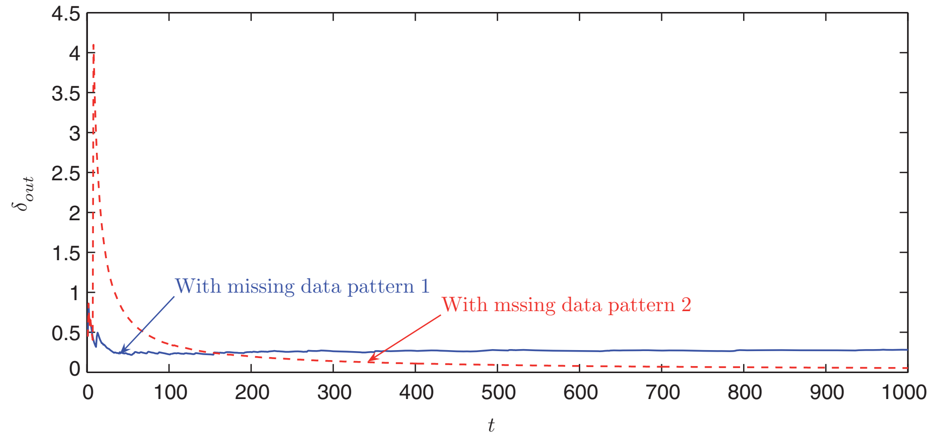

Figure 11. Example 3: average output estimation errors for two cases with different data missing patterns ($\lambda = 0.8$ and $\gamma = 0.7$).

6. Conclusions

In this article, we have studied the problem of parameter estimation of systems placed in a TCP-based network environment. For such a problem, missing input and output data is an important concern as data transmitted over a network may encounter time delays and packet losses. Randomly missing input and output are modelled as two Bernoulli processes. A missing output estimator is designed, and further a modified Kalman filter-based recursive estimation algorithm is developed. Convergence properties for both parameter estimation and output estimation are established. Simulation examples verify the effectiveness of the proposed algorithm and also illustrate that the data completeness and the data missing pattern would affect the estimation performance. It is worthwhile noting that the proposed algorithm can handle two practical cases: (1) the input and output have different probabilities of missing data; (2) the probabilities of missing data may be very large. Thus, the design in this work makes an important step forward in addressing practical issues of network-induced randomly missing data for identification of systems over lossy networks. The proposed method could be potentially extended to multiple input and multiple output (MIMO) systems.

It is worth noting that the idea behind this article could be extended to other widely applied communication protocols, such as UDP, Profibus, factory instrumentation protocol (FIP) (Hristu-Varsakelis and Levine 2005) and so on. In addition, although in this article the orders of the system model are assumed to be fixed and known, the framework of this article could be extended to the case where model orders are also needed to be identified. Moreover, it is desirable to further develop adaptive control schemes for NCSs based on the proposed parameter estimation algorithms. This article considers missing data that follow the Bernoulli distribution. Yet its convergence properties in the non-Bernoulli cases are still unknown. These topics are worth further studying.

References

Albertos, P., Sanchis, R., and Sala, A. (1999), 'Output Prediction Under Scarce Data Operation: Control Applications', *Automatica*, 35, 1671–1681.

Cao, L., and Schwartz, H.M. (2003), 'Exponential Convergence of the Kalman Filter Based Parameter Estimation Algorithm', *International Journal of Adaptive Control and Signal Processing*, 17, 763–783.

Charalambous, C.D., and Logothetis, A. (2000), 'Maximum Likelihood Parameter Estimation from Incomplete Data via the Sensitivity Equations: The Continuous-time Case', *IEEE Transactions on Automatic Control*, 45, 928–934.

Chen, H.F., and Guo, L. (1991), *Identification and Stochastic Adaptive Control*, Boston: Birkhäuser.

Chen, J.M., and Chen, B.S. (2000), 'System Parameter Identification with Input/Output Noisy Data and Missing Measurements', *IEEE Transactions on Signal Processing*, 48, 1548–1558.

Chen, T., and Francis, B. (1995), *Optimal Sampled-Data Control Systems*, London: Springer.

Chow, M.Y., and Tipsuwan, Y. (2001), 'Network Based Control Systems: A Tutorial', in *Proceedings of IECON01: 27th Annual Conference on IEEE Industrial Electronics Society*, pp. 1593–1602.

Ding, F., and Chen, T. (2003), 'Parameter Identification and Intersample Output Estimation of a Class of Dual-Rate Systems', in *Proceedings on 42nd IEEE Conference on Decision and Control*, pp. 5555–5560.

Ding, F., and Chen, T. (2004a), 'Combined Parameter and Output Estimation of Dual-Rate Systems using an Auxiliary Model', *Automatica*, 40, 1739–1748.

Ding, F., and Chen, T. (2004b), 'Identification of Dual-Rate Systems Based on Finite Impulse Response Models', *International Journal of Adaptive Control and Signal Processing*, 18, 589–598.