Asian Journal of Control, Vol. 14, No. 4, pp. 1070–1083, July 2012
Published online 7 June 2011 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/asjc.378

A KIEFER-WOLFOWITZ ALGORITHM BASED ITERATIVE LEARNING CONTROL FOR HAMMERSTEIN-WIENER SYSTEMS

Dong Shen and Han-Fu Chen

ABSTRACT

The iterative learning control (ILC) is considered for the Hammerstein-Wiener (HW) system, which is a cascading system consisting of a static nonlinearity followed by a linear stochastic system and then a static nonlinearity. Except the structure, the system is unknown, but the system output is observed with additive noise. Both the linear and nonlinear parts of the system may be time-varying. The optimal control sequence under the tracking performance is first characterized, which, is however, unavailable since the system is unknown. By using the observations on system output the ILC is generated by a Kiefer-Wolfowitz (KW) algorithm with randomized differences, which aims at minimizing the tracking error. It is proved that ILC converges to the optimal one with probability one and the resulting tracking error tends to its minimal value.

Key Words: Hammerstein-Wiener system, iterative learning control, stochastic approximation.

I. INTRODUCTION

In practical applications, many control tasks are completed within a finite time, and are repeatedly performed, e.g., the robot arms, chemical reactions, etc. For such kind of control problems it is of great importance to effectively use the input and output information obtained in past cycles to adjust the current control input, in order to improve the tracking performance with increasing of the number of cycles. This is the motivation of iterative learning control (ILC), introduced by Arimoto in [1]. Because of its simplicity and effectiveness, ILC has drawn great attention from researchers and obtained many applications in practice [2–4]. In particular, ILC has been applied in various robots [5, 6], motors [7, 8], and industrial/chemical process [9].

The techniques usually used to analyze problems from ILC include the contraction mapping approach [10], the 2-D system method [11–14], and the composite energy function approach [15, 16], etc. However, these techniques are not suitable for analyzing discrete-time stochastic systems.

For discrete-time stochastic systems, most existing results on ILC are for linear systems [17, 18]. The recursive algorithm for learning gain matrix is given in [17] based on the derivative of the covariance matrix of the minimal (in the least-square sense) tracking error with respect to learning gain matrix for linear stochastic systems. This leads to an ILC algorithm. The convergence of the state tracking error to zero in the mean-square sense is established in [17], where the partial information concerning the system matrices is required for the algorithm. The linear stochastic system is also considered in [18], which, requiring no information about system matrices, gives a new type of ILC algorithm based on the Kiefer-Wolfowitz (KW)-algorithm and proves its convergence with probability one to the optimal control under the corresponding

Manuscript received March 25, 2010; revised December 20, 2010; accepted January 8, 2011.

The authors are with the Key Laboratory of Systems and Control, Institute of Systems Science, AMSS, Chinese Academy of Sciences, Beijing 100190, China.

Han-Fu Chen is the corresponding author (e-mail: hfchen@iss.ac.cn).

The work is supported by NSFC (Grants No. 60821091, 60874001).

performance index. The ILC for the affine nonlinear stochastic system is discussed in [19, 20]. To be precise, the recursive algorithm for the learning gain matrix is given in [19] for the case where the noise for the affine nonlinear system comes from the observation only, and the input sequence is proved to converge in the mean-square sense to the desired one. While in [20] the KW-algorithm based ILC is constructed for the case where the system is corrupted by noise in both state and observation. The input sequence is proved to converge with probability one to the optimal one. However, the outputs of the systems considered in [19, 20] depend on the input, in essence, in a linear way, while in practice a nonlinear relationship between input and output is quite common.

A preliminary study of this problem is given in [21], which considers ILC for the affine nonlinear systems with observation noise and with a static nonlinearity such as dead-zone, preload, and saturation at the input side. The same ILC for all these nonlinearities is constructed in [21] with the help of the Robbins-Monro (RM) algorithm from stochastic approximation (SA). The designed control sequence converges to the optimal one with probability one.

The nonlinearity in the paper is allowed to appear not only at the input side but also at the output side. Moreover, the nonlinearity is of general form not restricted to a specific class as mentioned above. Besides, the observation noise is taken into account. Based on the KW-algorithm with randomized differences, ILC is designed for such kind of stochastic Hammerstein-Wiener systems, and the control sequence is proved to converge to the optimal sequence with probability one.

The rest of the paper is arranged as follows. The problem formulation and the basic assumptions to be imposed are given in Section II. The intermediate optimal signal is defined in Section III, where the sufficient conditions are also given for control sequence to be optimal. The ILC algorithm is defined in Section IV, and its convergence analysis is provided there. A simulation example is demonstrated in Section V. Finally, some concluding remarks are given in Section VI.

II. PROBLEM FORMULATION

Consider the stochastic HW system as shown in Fig. 1, where SLS means the stochastic linear system, and $t\in[0,T]$ denotes the time instance in an operation cycle while the subscripts $k=0,1,2,\ldots$ denote the different cycles.

Fig. 1. HW System.

The system is written as follows.

$$\begin{aligned} v_k(t) &= f_t(u_k(t)),\\ x_k(t+1) &= A(t)x_k(t)+B(t)v_k(t)+\varepsilon_k(t+1),\\ z_k(t) &= C(t)x_k(t)+\zeta_k(t),\\ y_k(t) &= g_t(z_k(t))+\xi_k(t), \end{aligned} \tag{1}$$

where $u_k(t)\in\mathbb{R}^p$, $y_k(t)\in\mathbb{R}^q$ are the input and output, respectively, $x_k(t)\in\mathbb{R}^n$ is the state of the linear subsystem, $v_k(t)$ and $z_k(t)$ are the intermediate signals, $\varepsilon_k(t)$ and $\zeta_k(t)$ are the system noises, $\xi_k(t)$ is the observation noise, $f_t(\cdot):\mathbb{R}^p\to\mathbb{R}^p$ and $g_t(\cdot):\mathbb{R}^q\to\mathbb{R}^q$ are the nonlinearities at time t and placed at the input and output sides, respectively, and $A(t)$, $B(t)$, $C(t)$ denote the matrix coefficients of compatible dimensions for the linear subsystem.

Let $\{\mathscr{F}_k\}$ be the nondecreasing $\sigma-$algebras defined by $\mathscr{F}_k\triangleq\sigma(\varepsilon_i(t),\zeta_i(t),\xi_i(t),0\le i\le k,t\in[0,T])$. Define the set of admissible controls as follows

$$U=\Big\{u_{k+1}(t)\in\mathscr{F}_k,\ \sup_k\|u_k(t)\|<\infty \text{ a.s.}\quad \forall t\in[0,T-1],\ \forall k=0,1,2,\ldots\Big\}.$$

The control purpose is to find $\{u_k(t),k=0,1,2,\ldots\}\in U$ in order that the following tracking error is minimized:

$$V_t(\{u_k(t)\})=\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\|y_d(t)-y_k(t)\|^2,\quad \forall t\in[0,T], \tag{2}$$

where $y_d(t)$, $t\in[0,T]$ is the target signal to track.

The following assumptions are imposed on the system. In what follows a sequence of independent and identically distributed random vectors is called simply as an iid sequence.

H1. For any $t\in[0,T]$, the iid sequence $\{\varepsilon_k(t),k=1,2,\ldots\}$ is independent of the iid sequence $\{\zeta_k(t),k=1,2,\ldots\}$, $E\varepsilon_k(t)=0$, $E\zeta_k(t)=0$, and for any integer $m>0$

$$\mathbb{E}\|\varepsilon_k(t)\|^m<\infty,\quad \mathbb{E}\|\zeta_k(t)\|^m<\infty.$$

H2. For any $t\in[0,T]$ the observation noise $\{\xi_k(t)\}$ is iid with $\mathbb{E}\xi_k(t)=0$, $\sup_k \mathbb{E}\xi_k^2(t)<\infty$ and

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k^2(t)=R_t<\infty \quad \text{a.s. } \forall t\in[0,T], \tag{3}$$

where R_t is unknown.

H3. The sequence of initial values $\{x_k(0), k=1,2,\ldots\}$ is iid with $\mathbb{E}x_k(0)=x_d(0)$ and $\mathbb{E}\|x_k(0)\|^m<\infty$ for any integer m, where $x_d(0)$ denotes the initial value of target trajectory. Further, the sequences $\{x_k(0), k=1,2,\ldots\}$, $\{\varepsilon_k(t), k=1,2,\ldots\}$, $\{\zeta_k(t), k=1,2,\ldots\}$, and $\{\xi_k(t), k=1,2,\ldots\}$ are mutually independent.

H4. For any $t\in[0,T]$, $f_t(\cdot)$ and $g_t(\cdot)$ are continuously differentiable: *i.e.*, $F_t(\cdot)\triangleq\nabla f_t(\cdot)$, $G_t(\cdot)\triangleq\nabla g_t(\cdot)$ are continuous.

H5. For any $t\in[0,T]$, as $\|x\|\to\infty$, $\|g_t(x)\|$ is allowed to grow up as fast as a polynomial of any degree, *i.e.*, for any fixed t, there are constants $\alpha>0$, $\beta>0$ such that

$$\|g_t(x)\|\le\alpha(1+\|x\|^{\beta}), \quad x\in R.$$

H6. For any $t\in[0,T]$, the matrix $C(t+1)B(t)$ is of full-column-rank.

Remark 1. For convenience of writing let us set $\varepsilon_k(0)=x_k(0)-x_d(0)$. By H3 it is clear that $\{\varepsilon_k(0)\}$ is iid.

The solution to the linear subsystem is

$$x_k(t+1)=\Phi(t,0)x_d(0)+\sum_{i=0}^{t}\Phi(t,i+1)B(i)v_k(i) + \sum_{i=0}^{t+1}\Phi(t,i)\varepsilon_k(i), \tag{4}$$

where $\Phi(j,i)$ satisfies

$$\Phi(j,i)=A(j)\Phi(j-1,i), \quad j\ge i, \quad \Phi(i-1,i)=I.$$

By setting

$$r_k(t+1)\triangleq C(t+1)\left(\Phi(t,0)x_d(0) + \sum_{i=0}^{t-1}\Phi(t,i+1)B(i)v_k(i)\right), \tag{5}$$

$$w_k(t+1)\triangleq C(t+1)\left(\sum_{i=0}^{t+1}\Phi(t,i)\varepsilon_k(i)\right)+\zeta_k(t+1), \tag{6}$$

we then have

$$\begin{aligned} z_k(t+1)&=C(t+1)x_k(t+1)+\zeta_k(t+1)\\ &=C(t+1)B(t)f_t(u_k(t))+r_k(t+1)\\ &\quad+w_k(t+1). \end{aligned} \tag{7}$$

By H1, H3, and Remark 1 it is seen that $\{w_k(t), k=1,2,\ldots\}$ is an iid sequence of random vectors with zero mean and finite moments of any order.

By setting

$$v_k(t)=f_t(u_k(t)) \quad \text{and} \quad x_k^0(0)=x_d(0),$$

$$x_k^0(t+1)=A(t)x_k^0(t)+B(t)v_k(t),$$

$$z_k^0(t)=C(t)x_k^0(t),$$

we see that

$$z_k^0(t+1)=r_k(t+1)+C(t+1)B(t)v_k(t), \tag{8}$$

$$z_k(t)=z_k^0(t)+w_k(t). \tag{9}$$

Therefore,

$$y_k(t)=g_t(z_k^0(t)+w_k(t))+\xi_k(t). \tag{10}$$

For any t let $P_t(x)\triangleq\mathbb{E}\|y_d(t)-g_t(x+w_t)\|^2$, where w_t is a random vector having the same distribution as $w_k(t)$. It is clear that $P_t(x)\ge0$. Set

$$\mathscr{M}_t\triangleq\left\{x_{\min}: P_t(x_{\min})=\min_x P_t(x)\right\}. \tag{11}$$

It is worth noting that $\mathscr{M}_t$ may contain more than one point.

III. OPTIMAL CONTROL

We now characterize the optimal control minimizing the performance index (2).

Let $z_d(i)\in\mathscr{M}_i$, $i=0,1,\ldots,t-1$ be the optimal intermediate signals:

$$P_i(z_d(i))=\min_x P_i(x),$$

and denote the corresponding inputs of SLS by $v_d(i-1)$, where the subscript "d" means "desired". Similar to (5) we have the desired

$$r_d(t+1)\triangleq C(t+1)\left(\Phi(t,0)x_d(0) + \sum_{i=0}^{t-1}\Phi(t,i+1)B(i)v_d(i)\right). \tag{12}$$

Then, by (8) the desired $v_d(t)$ must satisfy the following equation

$$z_d(t+1)-r_d(t+1)=C(t+1)B(t)v_d(t), \tag{13}$$

and hence by H6

$$v_d(t)\triangleq[(C(t+1)B(t))^T C(t+1)B(t)]^{-1} \times(C(t+1)B(t))^T(z_d(t+1)-r_d(t+1)) \tag{14}$$

with the optimal initial

$$v_d(0)=[(C(1)B(0))^T C(1)B(0)]^{-1}(C(1)B(0))^T (z_d(1)-C(1)A(0)x_d(0)).$$

However, $z_d(t)$, $t\in[0,T]$ is not directly available, so $v_d(t)$ given by (14) can neither be directly used in the control process.

By (10) we have

$$y_d(t)-y_k(t)=\phi(t,k)+\varphi(t,k)-\xi_k(t), \tag{15}$$

where $\phi(t,k)$ and $\varphi(t,k)$ are respectively defined as follows

$$\phi(t,k)=y_d(t)-g_t(z_d(t)+w_k(t)), \tag{16}$$

$$\varphi(t,k)=g_t(z_d(t)+w_k(t))-g_t(z_k^0(t)+w_k(t)). \tag{17}$$

Theorem 1. Assume H1-H6 hold. Then for system (1) with tracking index (2), any $\{u_d(t),t\in[0,T]\}$ satisfying $f_t(u_d(t))=v_d(t)$ is the optimal control sequence. Further, any input sequence $\{u_k(t)\}\in U$ with $\delta v_k(t)\triangleq v_d(t)-v_k(t)\to 0, \forall t\in[0,T-1]$ a.s. is the optimal control sequence making the performance index (2) asymptotically minimal.

Proof. Let us first explain the idea of the proof.

Consider the case $t=0$. By (8), (13) it is seen that $\|z_d(1)-z_k^0(1)\|\underset{k\to\infty}{\longrightarrow}0$, whenever $\|v_d(0)-v_k(0)\|\underset{k\to\infty}{\longrightarrow}0$. This implies that

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\|\varphi(1,k)\|^2=0$$

by (17). Hence, from (15) we have

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\|y_d(1)-y_k(1)\|^2$$

$$=\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\|\phi(1,k)\|^2$$

$$+\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\|\xi_k(1)\|^2$$

$$=\mathbb{E}\|y_d(1)-g_1(z_d(1)+w_1)\|^2+R_1$$

$$=\min_x\mathbb{E}\|y_d(1)-g_1(x+w_1)\|^2+R_1.$$

This means that $\{u_k(0)\}$ is optimal, and the optimality for $t>0$ can be shown in a similar way. The detailed proof of the theorem is given in Appendix A. □

Remark 2. Theorem 1 characterizes the optimal control via the values $v_k(t)$ of $f_t(u_k(t))$. However, the optimal control is not directly computable, because neither $z_d(t+1)$ minimizing the tracking error $P_{t+1}(x)$ is available, nor $v_d(t)$ and $u_d(t)$ are computable even with known $z_d(t+1)$, since $C(t+1)B(t)$ and $f_t(\cdot)$ in the system are unknown. In the next section we will use the available information $e_k(t)\triangleq y_d(t)-y_k(t)$ to directly calculate the control approaching to the optimal one.

In order to define the control minimizing the mean of the squared tracking error $\|e_k(t)\|^2$, similar to (8), let us denote

$$\overline{z}_k^0(t+1)=r_d(t+1)+C(t+1)B(t)f_t(u_k(t)), \tag{18}$$

and

$$L_t(u)\triangleq\mathbb{E}(\|y_d(t+1)-g_{t+1}(\overline{z}_k^0(t+1)+w_k(t+1))\|^2|u_k(t)=u)$$

$$=\mathbb{E}\|y_d(t+1)-g_{t+1}(r_d(t+1)+C(t+1)B(t)f_t(u)+w_{t+1})\|^2. \tag{19}$$

Remark 3. Notice that $P_{t+1}(\cdot)$ is defined on the values of $z_k^0(t+1)$ and $L_t(\cdot)$ is defined on the values of $u_k(t)$. This implies that, in general,

$$\min_x P_{t+1}(x)\le\min_u L_t(u),$$

where by Theorem 1 the equality takes place if $f_t(u)=v_d(t)$ is solvable with respect to u.

To demonstrate the functions $P_{t+1}(x)$ and $L_t(u)$ for a system, let us consider the following example.

Let the system be one-dimensional with $A(t)=B(t)=C(t)=1$ and $g_t(x)=x^3$. Assume the noises $\varepsilon_k(t)$, $\zeta_k(t)$ are iid. For simplicity of presentation, assume $w_k(t+1)(=\sum_{i=0}^{t+1}\varepsilon_k(i)+\zeta_k(t+1))\sim\mathcal{N}(0,1^2)$. In other words, $w_{t+1}\sim\mathcal{N}(0,1^2)$, where w_{t+1} is given in the definition of $P_{t+1}(x)$. Then, we have $\mathbb{E}w_{t+1}^2=1$, $\mathbb{E}w_{t+1}^4=3$, and $\mathbb{E}w_{t+1}^6=15$.

For a fixed t assuming $y_d(t+1)=3$ we then have

$$P_{t+1}(x)=\mathbb{E}\|y_d(t+1)-g_{t+1}(x+w_{t+1})\|^2$$

$$=\mathbb{E}(3-(x+w_{t+1})^3)^2$$

$$= 9 + \mathbb{E}(x + w_{t+1})^6 - 6\mathbb{E}(x + w_{t+1})^3$$
$$= x^6 + 15x^4 - 6x^3 + 45x^2 - 18x + 24.$$

The function $P_{t+1}(\cdot)$ achieves its minimum at $x = 0.2026$:

$$\min_x P_{t+1}(x) = P_{t+1}(0.2026) = 22.1757.$$

On the other hand, for a fixed t, $r_d(t+1)$ is defined by $y_d(i)$, $i \in [0,t]$ and is independent of $y_d(t+1)$. So, without loss of generality we may assume $r_d(t+1) = 1$. Let the nonlinearity at the input side be an invertible function $f_t(u) = ue^{|u|}$.

Then, we have

$$L_t(u) = \mathbb{E}(3 - (1 + ue^{|u|} + w_k(t+1))^3)^2$$
$$= u^6 e^{6|u|} + 6u^5 e^{5|u|} + 30u^4 e^{4|u|} + 74u^3 e^{3|u|}$$
$$+ 132u^2 e^{2|u|} + 120ue^{|u|} + 61,$$

which achieves its minimum at $u = -0.4890$:

$$\min_u L_t(u) = L_t(-0.4890) = 22.1757.$$

Thus, we have seen that $P_{t+1}(x)$ and $L_t(u)$ have the same minimal value indeed.

IV. ILC AND ITS CONVERGENCE

In the last section the optimal control has been characterized, but it cannot be actually used, because the matrices $A(t)$, $B(t)$, $C(t)$, and the nonlinear functions $f_t(\cdot)$, $g_t(\cdot)$ are unknown.

For any fixed t we now define the control sequence minimizing $L_t(u)$ by using the available information $\|e_k(t+1)\|^2$ with $e_k(t) \triangleq y_d(t) - y_k(t)$.

For this we apply the KW-algorithm with randomized differences proposed in [22] to seek the minimum of $L_t(u)$. In doing so, the problem is first transformed to an RM-algorithm aiming at searching the roots of the unknown gradient of $L_t(u)$.

To form the randomized differences let us introduce the auxiliary sequence of m-dimensional vectors $\{\Delta_k(t), k = 1, 2, \ldots\}$, $t = 0, \ldots, T$, where $\Delta_k(t) = (\Delta_k^1(t), \ldots, \Delta_k^m(t))^T$ satisfies the following conditions:

1. All components $\Delta_k^i(t)$, $k = 1, 2, \ldots, t = 0, 1, \ldots, T$, $i = 1, \ldots, m$ of $\Delta_k(t)$ are mutually independent with

$$|\Delta_k^i(t)| < d_1, \quad \left|\frac{1}{\Delta_k^i(t)}\right| < d_2,$$

$$\mathbb{E}\left\{\frac{1}{\Delta_k^i(t)}\right\} = \mathbb{E}\Delta_k^i(t) = 0$$

$$\forall k = 1, 2, \ldots, t = 0, 1, \ldots, T, i = 1, \ldots, m, \quad (20)$$

where d_1 and d_2 are some positive constants.

2. Sequences $\{\Delta_k(t)\}$ and $\{\varepsilon_k(t), \zeta_k(t), x_k(0), \xi_k(t), t \in [0,T], k = 0, 1, 2, \ldots\}$ are mutually independent.

Define

$$\overline{\Delta}_k(t) = \left(\frac{1}{\Delta_k^1(t)}, \ldots, \frac{1}{\Delta_k^m(t)}\right)^T,$$
$$t = 0, \ldots, T, k = 0, 1, 2, \ldots. \quad (21)$$

As required by the KW-algorithm with randomized differences [22] we need some sequences of positive real numbers $\{a_k\}$, $\{c_k\}$ and $\{M_k\}$ such that

$$a_k \xrightarrow[k\to\infty]{} 0, \quad \sum_{k=0}^{\infty} a_k = \infty, \quad c_k \xrightarrow[k\to\infty]{} 0,$$

$$\sum_{k=0}^{\infty}\left(\frac{a_k}{c_k}\right)^2 < \infty, \quad M_{k+1} > M_k, \quad M_k \xrightarrow[k\to\infty]{} \infty,$$

where $\{a_k\}$ is the step-size in the algorithm, $\{c_k\}$ is used to form the randomized differences, while $\{M_k\}$ serves as truncation bounds in the algorithm to prevent it from too fast divergence to infinity.

For any $t \in [0,T]$, the initial value $u_0(t)$ is arbitrarily given.

In order to form the randomized differences we define controls in odd cycles of operation differently from those in even cycles of operation. For the odd cycle of operation the control is defined as follows

$$u_{2k+1}(t) = u_{2k}(t) + c_k \Delta_k(t), \quad (22)$$

while for even cycle $u_{2k}(t)$

$$\overline{u}_{2(k+1)}(t) = u_{2k}(t) - a_k \frac{\overline{\Delta}_k(t)}{c_k}(\|e_{2k+1}(t+1)\|^2$$
$$- \|e_{2k}(t+1)\|^2), \quad (23)$$

$$u_{2(k+1)}(t) = \overline{u}_{2(k+1)} \cdot I_{[\|\overline{u}_{2(k+1)}\| \le M_{\sigma_k(t)}]}, \quad (24)$$

$$\sigma_k(t) = \sum_{l=1}^{k-1} I_{[\|\overline{u}_{2(l+1)}\| > M_{\sigma_l(t)}]}, \quad \sigma_0(t) = 0, \quad (25)$$

where $I_{[\text{inequality}]}$ is an indicator function meaning that it equals 1 if the inequality indicated in the bracket is fulfilled, and 0 if the inequality does not hold.

Thus, ILC applied to system (1) is given by (22)–(25), where the random differences are applied to estimate the gradients, *i.e.*, the control directions. The idea of applying random differences in the KW algorithm is widely used, e.g., [22–25]. The method used in the paper originates from [22] and also from [18, 20].

The convergence analysis for ILC will be carried out inductively along time t. For any fixed t the algorithm is first transformed to the RM algorithm, and then the convergence is established with the help of the convergence theorem (Theorem A) of SA given in Appendix C.

Let us transform (22)–(25) to the RM algorithm with expanding truncations.

It is important to note that for any t, the function $L_t(\cdot)$ is independent of controls actually applied to the system. In other words, $L_t(\cdot)$ is an objectively existing function connected with the system (1)(2). We need the following assumption imposed on $L_t(\cdot)$.

H7. There exists a constant c_0 such that $L_t(0) < \inf_{\|u\|=c_0} L_t(u)$.

This condition becomes trivial when $L_t(u) \underset{\|u\|\to\infty}{\longrightarrow} \infty$. It plays an essential role in establishing the boundedness of $\{u_k(t)\}$.

Let

$$N_k(t) \triangleq \|e_k(t+1)\|^2 - L_t(u_k(t)), \tag{26}$$

and denote the derivative of $L_t(\cdot)$ with respect to u by $h_t(\cdot)$:

$$h_t(u_k(t)) = 2F_t^T(u_k(t))[C(t+1)B(t)]^T p_{t+1}(\overline{z}_k^0(t+1)). \tag{27}$$

Let $J_t \triangleq \{u: \ h_t(u) = 0\}$. Then $L_t(\cdot)$ attains its minimum at some points in J_t.

The algorithm (22)–(25) can be rewritten as

$$\overline{u}_{2(k+1)}(t) = u_{2k}(t) - a_k h_t(u_{2k}(t)) + a_k(\theta_{k+1}(t) + \vartheta_{k+1}(t)), \tag{28}$$

$$u_{2(k+1)}(t) = \overline{u}_{2(k+1)} \cdot I_{[\|\overline{u}_{2(k+1)}\| \le M_{\sigma_k(t)}]}, \tag{29}$$

$$\sigma_k(t) = \sum_{l=1}^{k-1} I_{[\|\overline{u}_{2(l+1)}\| > M_{\sigma_l(t)}]}, \quad \sigma_0(t) = 0, \tag{30}$$

where

$$\theta_{k+1}(t) = \frac{L_t(u_{2k}(t) + c_k\Delta_k(t)) - L_t(u_{2k}(t))}{c_k} \overline{\Delta}_k(t) - h_t(u_{2k}(t)), \tag{31}$$

$$\vartheta_{k+1}(t) = \frac{\overline{\Delta}_k(t)}{c_k}(N_{2k+1}(t) - N_{2k}(t)). \tag{32}$$

Theorem 2. Assume H1-H7 hold for system (1) with index (2). Then i) the control sequence $\{u_k(t)\}$ given by (22)–(25) tends to J_t: $d(u_k(t), J_t) \underset{k\to\infty}{\longrightarrow} 0$; and ii) $\{u_k(t)\}$ is the optimal control whenever $L_t(\cdot)$ is a constant at J_t (in particular, if J_t is a singleton).

Proof. See Appendix B. □

Remark 4. Since there are so many unknowns in the system and a random search is applied, it is not realistic to expect a fast convergence rate as can be seen from the numerical example given in the next section. As with most of the theoretical results in control, the results given in this paper have only pointed out the best possibility for practice, but for any application of the proposed algorithm to a practical control process one may have to modify the algorithm according to the real situation. To facilitate practical application, we usually select $a_k = \frac{1}{k}$ and $c_k = \frac{1}{k^\nu}$ with $\nu \in (0, \frac{1}{2})$. Roughly speaking, when ν approaches 0 or $\frac{1}{2}$, the convergence rate decreases. For more information on tuning the convergence rate we refer to [26, Chapter 4].

V. SIMULATION EXAMPLE

Consider the following nonlinear stochastic system

$$v_k(t) = f(u_k(t))$$

$$x_k(t+1) = \begin{bmatrix} 1.04 + 0.8\sin\left(\frac{t}{100}\right) & 0.4 \\ -0.4 & 1.2 - 0.8\cos\left(\frac{t}{100}\right) \end{bmatrix} x_k(t) + \begin{bmatrix} -1 \\ 1.1 \end{bmatrix} v_k(t) + \begin{bmatrix} \varepsilon_k^{(1)}(t+1) \\ \varepsilon_k^{(2)}(t+1) \end{bmatrix}$$

$$z_k(t) = [1 \quad 1] x_k(t) + \zeta_k(t)$$

$$y_k(t) = g(z_k(t)) + \xi_k(t)$$

where $\{\varepsilon_k^{(1)}(t)\}$, $\{\varepsilon_k^{(2)}(t)\}$, $\{\zeta_k(t)\}$ and $\{\xi_k(t)\}$ are mutually independent random variables with normal distribution $\mathcal{N}(0, \sigma^2)$. Let the nonlinearities be time-invariant:

$$f(u) = \begin{cases} u - \frac{1}{2}, & u \ge 1 \\ \frac{1}{2}u^2, & 0 \le u < 1 \\ -\frac{1}{2}u^2, & -1 \le u < 0 \\ u + \frac{1}{2}, & u < -1 \end{cases}$$

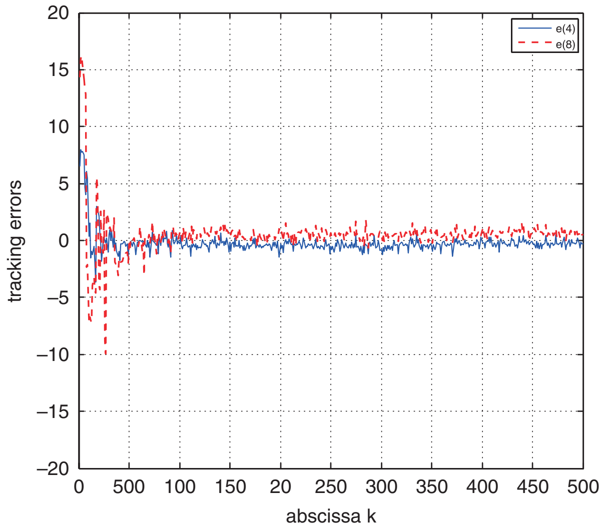

Fig. 2. Tracking errors $e_k(t)$ for $t=4$ and 8, where all ξ_k, ζ_k, $\varepsilon_k \in \mathcal{N}(0, \sigma^2)$, $\sigma=0.1$.

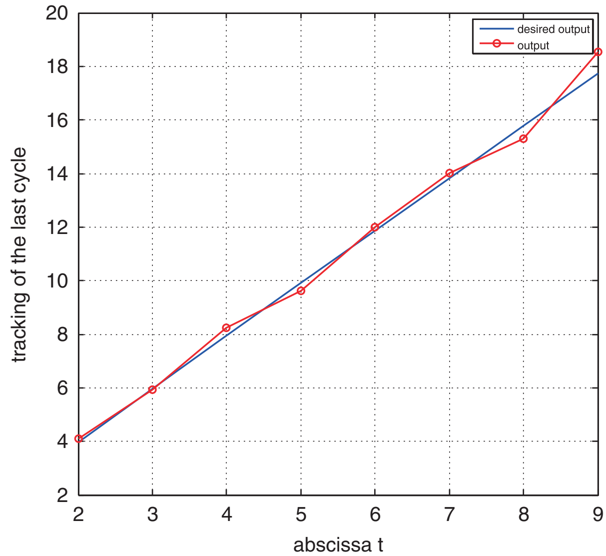

Fig. 3. Comparison of $y_{500}(t)$ with $y_d(t)$ for $t=2,\ldots,9$, where all ξ_k, ζ_k, $\varepsilon_k \in \mathcal{N}(0, \sigma^2)$, $\sigma=0.1$.

and

$$g(x) = 6x + \sin\frac{x}{100}.$$

Let the reference trajectory be $y_d(t) = \frac{1}{3}t(6-0.01t)$, $t=[1,\ldots,10]$, and let the initial control $u_0(t) \equiv 0, \forall t \in [0, 10]$.

In the algorithm (28)–(31) we set

$$a_k = \frac{1}{k+1}, \quad c_k = \frac{1}{(k+1)^{0.3}}, \quad M_k = 3^k.$$

Let the random variables Δ_k^i be uniformly distributed on $[-1, -0.5] \cup [0.5, 1]$ for $i=1,\ldots,m$.

For each $t \in [1, 10]$ the algorithm runs 500 cycles. The simulation results are presented in Figs 2–7. Figs 2 and 3 are for the case where all ξ_k, ζ_k, $\varepsilon_k \in \mathcal{N}(0, \sigma^2)$ with $\sigma=0.1$, while Figs 4 and 5 are for the case $\sigma=0.3$. For a comparison the case where $\zeta_k=0$, $\varepsilon_k=0$, and $\xi_k \in \mathcal{N}(0, \sigma^2)$ with $\sigma=0.3$ is also computed, and the simulation results are presented in Figs 6 and 7.

In these figures, arbitrarily chosen samples are demonstrated. In Figs 2, 4, and 6, the solid lines represent the tracking errors $e_k(4)$, while the dashed lines $e_k(8)$. In Figs 3, 5, and 7 the solid lines are the reference signal $y_d(t)$ and the solid lines with circles are the output trajectory at the 500th cycle.

From the figures we see that (i) the ILC designed in the paper gives satisfactory results; (ii) with increasing σ the tracking accuracy is decreasing quite significantly; and (iii) the internal noises ζ_k and ε_k play a major role in such a worsening.

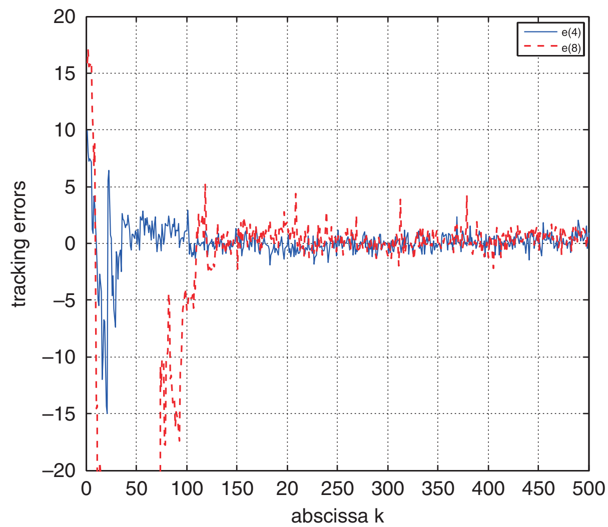

Fig. 4. Tracking errors $e_k(t)$ for $t=4$ and 8, where all ξ_k, ζ_k, $\varepsilon_k \in \mathcal{N}(0, \sigma^2)$, $\sigma=0.3$.

VI. CONCLUDING REMARKS

The ILC is considered for Hammerstein-Wiener systems with random noises in both state equation and observations. The optimal intermediate signal is first expressed, and the conditions that should be satisfied by the optimal control are given. Then, the ILC is constructed based on the KW algorithm with randomized differences. Its convergence and optimality are established under some conditions.

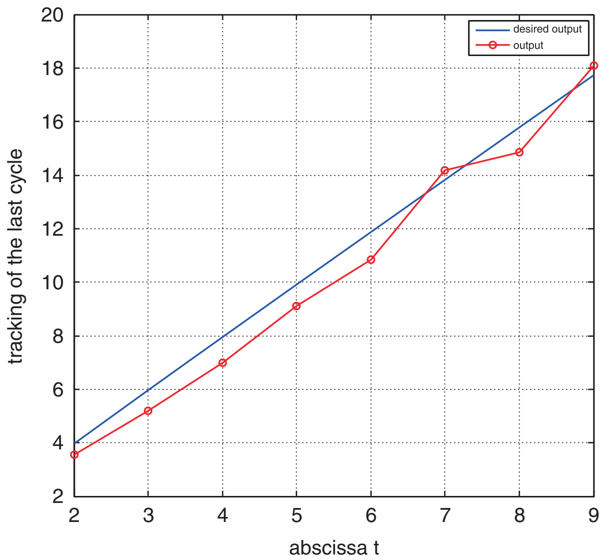

Fig. 5. Comparison of $y_{500}(t)$ with $y_d(t)$ for $t=2,\dots,9$, where all ξ_k, ζ_k, $\varepsilon_k \in \mathcal{N}(0,\sigma^2)$, $\sigma=0.3$.

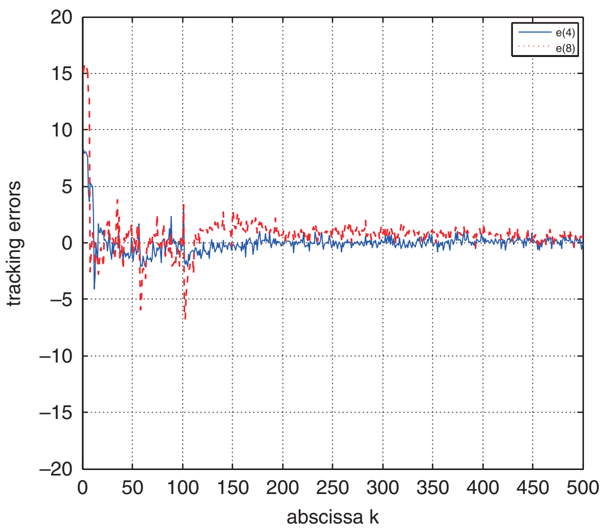

Fig. 6. Tracking errors $e_k(t)$ for $t=4$ and 8, where $\zeta_k=0$, $\varepsilon_k=0$, and $\xi_k \in \mathcal{N}(0,\sigma^2)$, $\sigma=0.3$. $\sigma=0.3$.

For further research it is of interest to consider how to guarantee that the roots of $h_t(\cdot)$ correspond to the global minimum of $L_t(\cdot)$ and how to guarantee H7 in terms of conditions on $A(t)$, $B(t)$, $C(t)$, $f_t(\cdot)$, and $g_t(\cdot)$.

For applications of the proposed algorithm, its rather slow convergence rate is the concern. For improving control performance one may have to reduce the randomness of the algorithm, in particular, to reduce the noise level of the internal noises.

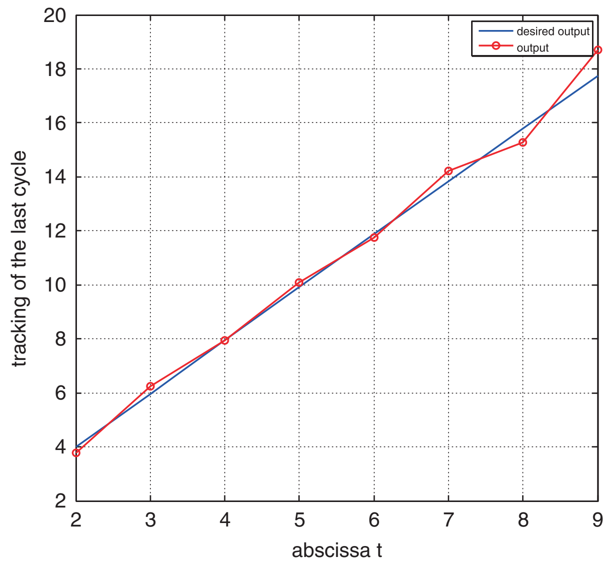

Fig. 7. Comparison of $y_{500}(t)$ with $y_d(t)$ for $t=2,\dots,9$, where $\zeta_k=0$, $\varepsilon_k=0$, and $\xi_k \in \mathcal{N}(0,\sigma^2)$, $\sigma=0.3$.

APPENDIX A

Proof of Theorem 1

For proving the theorem we first show that if $\delta v_k(s) \underset{k\to\infty}{\longrightarrow} 0$ and $\sup_k \|v_k(s)\| < \infty$, $s=0,1,\dots,t-1$, then

$$\mathbb{E}(\|\varphi(t,k)\|^m | \mathscr{F}_{k-1}) \underset{k\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \quad \forall m>0. \tag{A1}$$

By noticing (5), (8), (12), and (13), from $\delta v_k(s) \underset{k\to\infty}{\longrightarrow} 0$, $s=0,1,\dots,t-1$, it follows that

$$z_d(t) - z_k^0(t) \underset{k\to\infty}{\longrightarrow} 0.$$

By (17) we have

$$\begin{aligned}\mathbb{E}(\|\varphi(t,k)\|^m | \mathscr{F}_{k-1}) &= \mathbb{E}(\|g_t(z_d(t)+w_k(t)) - g_t(z_k^0(t) \\ &\quad + w_k(t))\|^m | \mathscr{F}_{k-1}),\end{aligned} \tag{A2}$$

where $\mathbb{E}\|g_t(x+w_k(t))\|^m < \infty, \forall x \in \mathbb{R}^q$ by H1 and H5. Noticing that $\mathbb{E}\|g_t(x+w_k(t))\|^m$ is continuous in x, $z_k^0(t) \in \mathscr{F}_{k-1}$ by $\delta v_k(t) \in \mathscr{F}_{k-1}$, and $w_k(t) \in \mathscr{F}_k$ is independent of $z_k^0(t)$ we find

$$\begin{aligned}&\mathbb{E}(\|g_t(z_k^0(t)+w_k(t))\|^m \,|\, \mathscr{F}_{k-1}) \\ &\quad = \mathbb{E}(\|g_t(x+w_k(t))\|^m \,|\, \mathscr{F}_{k-1})|_{x=z_k^0(t)}.\end{aligned} \tag{A3}$$

Therefore, by the dominated convergence theorem applied to the conditional mathematical expectation, the right-hand side of (A2) tends to zero as $k \to \infty$.

We now inductively prove the optimality of $u_d(t)$.

By the definition of $\mathscr{F}_k$, we see that $\mathscr{F}_k$ is independent of $\{\xi_l(t), \varepsilon_l(t), \zeta_l(t), l = k+i, i = 1, 2, \ldots, \forall t \in [0, T]\}$ and $\{\xi_k(t), \mathscr{F}_k\}$ is a martingale difference sequence (MDS).

To prove the optimality it suffices to show that the minimum of (2) is achieved whenever $\delta v_k(k) \underset{k\to\infty}{\longrightarrow} 0$.

For $t = 0$ we have

$$\begin{aligned} y_d(1) - y_k(1) &= y_d(1) - g_1(z_k^0(1) + w_k(1)) - \xi_k(1) \\ &= \phi(1,k) + \varphi(1,k) - \xi_k(1), \end{aligned}$$

where $\phi(1,k)$ and $\varphi(1,k)$ are defined by (16) and (17), respectively. Therefore, it follows that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|y_d(1) - y_k(1)\|^2 \\ &= \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\phi(1,k)\|^2 \\ &\quad + \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\xi_k(1)\|^2 \\ &\quad + \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\varphi(1,k)\|^2 \\ &\quad + 2\Bigg(\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \phi^T(1,k)\varphi(1,k) \\ &\quad - \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \phi^T(1,k)\xi_k(1) \\ &\quad - \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \varphi^T(1,k)\xi_k(1)\Bigg). \end{aligned} \tag{A4}$$

Since $\phi(1,k)$ is iid, by the law of large numbers we have

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\phi(1,k)\|^2 \\ &= \mathbb{E}\|y_d(1) - g_1(z_d(1) + w_1)\|^2 \\ &= \min_x \mathbb{E}\|y_d(1) - g_1(x + w_1)\|^2. \end{aligned}$$

For the second term on the right-hand side of (A4) it is clear that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\xi_k(1)\|^2 = R_1.$$

When $\delta v_k(0) \underset{k\to\infty}{\longrightarrow} 0$, by (A1) we have

$$\mathbb{E}(\|\varphi(1,k)\|^m | \mathscr{F}_{k-1}) \underset{k\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \quad \forall m > 0, \tag{A5}$$

and by the convergence theorem for MDS [27]

$$\sum_{k=1}^{\infty} \frac{1}{k} \{\|\varphi(1,k)\|^2 - \mathbb{E}(\|\varphi(1,k)\|^2 | \mathscr{F}_{k-1})\} < \infty.$$

From here by the Kronecker lemma it follows that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\varphi(1,k)\|^2 = 0 \quad \text{a.s.} \tag{A6}$$

Let us consider the three cross terms on the right-hand side of (A4).

It is clear that $\phi^T(1,k)\xi_k(1)$ is an iid sequence and $\phi(1,k)$ is independent of $\xi_k(1)$ and $\mathbb{E}\xi_k(1) = 0$. This implies by the law of large numbers

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \phi^T(1,k)\xi_k(1) = 0 \quad \text{a.s.} \tag{A7}$$

Noticing that $\xi_k(1)$, $\mathscr{F}_{k-1}$, and $w_k(1)$ are mutually independent, we find that

$$\begin{aligned} &\mathbb{E}(\varphi^T(1,k)\xi_k(1) | \mathscr{F}_{k-1}) \\ &\quad = \mathbb{E}(\mathbb{E}(\varphi^T(1,k)\xi_k(1) | \sigma\{\mathscr{F}_{k-1}, w_k(1)\}) | \mathscr{F}_{k-1}) \\ &\quad = \mathbb{E}(\varphi^T(1,k)\mathbb{E}(\xi_k(1) | \sigma\{\mathscr{F}_{k-1}, w_k(1)\}) | \mathscr{F}_{k-1}) \\ &\quad = 0. \end{aligned}$$

Since $\{\varphi^T(1,k)\xi_k(1), \mathscr{F}_k\}$ is an MDS, by the convergence theorem for MDS we have

$$\sum_{k=1}^{\infty} \frac{1}{k} \varphi^T(1,k)\xi_k(1) < \infty,$$

and hence

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \varphi^T(1,k)\xi_k(1) = 0 \quad \text{a.s.} \tag{A8}$$

Though $\phi(1,k)$ and $\varphi(1,k)$ are not independent, similar to the proof for (A6) we can still establish the following two formulas

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} y_d^T(1)\varphi(1,k) = 0 \quad \text{a.s.}$$

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g_1^T(z_d(1) + w_k(1))\varphi(1,k) = 0 \quad \text{a.s.}$$

Therefore, we have

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \phi^T(1,k)\varphi(1,k) = 0 \quad \text{a.s.} \tag{A9}$$

From (A4)–(A9) it is seen that the index (2) is minimized at time $t=1$:

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\|y_d(1)-y_k(1)\|^2 = \min_x \mathbb{E}\|y_d(1)-g_1(x+w_1)\|^2+R_1.$$

Thus, the optimality of the input at $t=0$ has been proved.

Assume $v_d(s)$, $s=0,1,2,\ldots,t-1$ are all the optimal intermediate signals. We now show that $v_d(t)$ is also an optimal intermediate signal.

By the induction assumption, $r_k(t+1)-r_d(t+1) \xrightarrow[k\to\infty]{} 0$. Therefore, when $v_k(t)-v_d(t)\xrightarrow[k\to\infty]{}0$, $z_k^0(t+1)-z_d(t+1)\xrightarrow[k\to\infty]{}0$, and hence

$$\mathbb{E}(\|\varphi(t+1,k)\|^m|\mathscr{F}_{k-1})\xrightarrow[k\to\infty]{}0 \quad \text{a.s.}$$

Completely similar to the case $t=0$, it is shown that the index reaches its minimum and $v_d(t)$ is the optimal intermediate signal.

APPENDIX B

Proof of Theorem 2

The proof consists in verification of conditions S1-S4 of Theorem A given in Appendix C. For this we note that for a fixed t, $-h_t(\cdot)$, $\theta_{k+1}(t)+\vartheta_{k+1}(t)$ correspond to $f(\cdot)$ and ε_{k+1} in (C2) in Appendix C, and x^* in (C1) is zero in (28).

First, S1 holds by selection of $\{a_k\}$, while S3 holds by H4-H6. It is clear that S2 is also satisfied by choosing $v(\cdot)=L_t(\cdot)$. Consequently, to prove $\delta v_k(t)\xrightarrow[k\to\infty]{}0$ $\forall t\in[0,T-1]$ it suffices to show S4.

We now verify S4 by induction along t.

We modify $\mathscr{F}_k$ defined in Section II, since now $\Delta_k(t)$ has to be taken into account. Define

$$\mathscr{F}_{2k}\triangleq\sigma\{\varepsilon_l(t),\zeta_l(t),x_l(0),\xi_l(t),l\le 2k, \Delta_m(t),m\le k,t=0,\ldots,T\}, \tag{B1}$$

$$\mathscr{F}_{2k+1}=\sigma\{\varepsilon_l(t),\zeta_l(t),x_l(0),\xi_l(t),l\le 2k+1, \Delta_m(t),m\le k,t=0,\ldots,T\}. \tag{B2}$$

It is clear that $u_k(t)\in\mathscr{F}_{k-1}$.

Further, by setting $\mathscr{G}_{k+1}\triangleq\mathscr{F}_{2k+1}$, we see that $u_{2k}(t)$ is $\mathscr{G}_k-$measurable.

Let us first consider the case $t=0$. Similar to the proof of Theorem 1 in [22], it is shown that

$$\lim_{T\to 0}\limsup_{k\to\infty}\frac{1}{T}\left\|\sum_{i=2n_k}^{m(2n_k,T_k)}a_i\theta_{i+1}(0)I_{[\|u_{2k}(0)\|\le K]}\right\| =0,\quad \forall T_k\in[0,T] \tag{B3}$$

along the subscripts $\{2n_k\}$ of any convergent subsequence $\{u_{2n_k}(0)\}$, where $m(k,T)\triangleq\max\{m:\sum_{i=k}^m a_i\le T\}$.

We now show that

$$\lim_{T\to 0}\limsup_{k\to\infty}\frac{1}{T}\left\|\sum_{i=2n_k}^{m(2n_k,T_k)}a_i\vartheta_{i+1}(0)I_{[\|u_{2i}(0)\|\le K]}\right\| =0,\quad \forall T_k\in[0,T] \tag{B4}$$

along the subscripts $\{2n_k\}$ of any convergent subsequence $\{u_{2n_k}(0)\}$.

Noticing $z_k^0(1)=\overline{z}_k^0(1)$, we have

$$\begin{aligned}N_k(0)&=\|e_k(1)\|^2-L_0(u_k(0))\\&=\|y_d(1)-g_1(z_k^0(1)+w_k(1))\\&\quad-\xi_k(1)\|^2-L_0(u_k(0))\\&=\eta_k(1)+\|\xi_k(1)\|^2-2[y_d(1)\\&\quad-g_1(z_k^0(1)+w_k(1))]^T\xi_k(1),\end{aligned}$$

where

$$\eta_k(1)=\|y_d(1)-g_1(z_k^0(1)+w_k(1))\|^2-L_0(u_k(0)).$$

Consequently, we have

$$\begin{aligned}\vartheta_{k+1}(0)&=\frac{\overline{\Delta}_k(0)}{c_k}(N_{2k+1}(0)-N_{2k}(0))\\&=\pi_{k1}+\pi_{k2}+\pi_{k3},\end{aligned}$$

where

$$\pi_{k1}=\frac{\eta_{2k+1}(1)-\eta_{2k}(1)}{c_k}\overline{\Delta}_k(0), \tag{B5}$$

$$\pi_{k2}=\frac{\|\xi_{2k+1}(1)\|^2-\|\xi_{2k}(1)\|^2}{c_k}\overline{\Delta}_k(0), \tag{B6}$$

$$\begin{aligned}\pi_{k3}&=2\frac{[y_d(1)-g_1(z_{2k}^0(1)+w_{2k}(1))]^T\xi_{2k}(1)}{c_k}\overline{\Delta}_k(0)\\&\quad-2\frac{[y_d(1)-g_1(z_{2k+1}^0(1)+w_{2k+1}(1))]^T\xi_{2k+1}(1)}{c_k}\overline{\Delta}_k(0).\end{aligned} \tag{B7}$$

Since $\xi_k(1)$ is independent of $\overline{\Delta}_k(0)$ and $\mathbb{E}\overline{\Delta}_k(0)=0$, we have

$$\mathbb{E}\{\pi_{k2}|\mathscr{G}_k\}=0. \tag{B8}$$

Then, by the convergence theorem for MDS it follows that

$$\sum_{k=1}^{\infty}\frac{a_k}{c_k}(\|\xi_{2k+1}(1)\|^2-\|\xi_{2k}(1)\|^2)\overline{\Delta}_k(0) I_{[\|u_{2k}(0)\|\le K]}<\infty, \tag{B9}$$

which means

$$\sum_{k=1}^{\infty}a_k\pi_{k2}I_{[\|u_{2k}(0)\|\le K]}<\infty. \tag{B10}$$

Similarly, by noticing that $\eta_{2k}(1)$ is independent of $\overline{\Delta}_k(0)$ it follows that

$$\sum_{k=1}^{\infty}\frac{a_k}{c_k}\eta_{2k}(1)\overline{\Delta}_k(0)I_{[\|u_{2k}(0)\|\le K]}<\infty. \tag{B11}$$

Since w_{2k+1} is independent of $\mathscr{F}_{2k}$, we have

$$\mathbb{E}\{\eta_{2k+1}(1)\overline{\Delta}_k(0)|\mathscr{F}_{2k}\}=0, \tag{B12}$$

and by the convergence theorem for MDS

$$\sum_{k=1}^{\infty}\frac{a_k}{c_k}\eta_{2k+1}(1)\overline{\Delta}_k(0)I_{[\|u_{2k}(0)\|\le K]}<\infty, \tag{B13}$$

and hence

$$\sum_{k=1}^{\infty}a_k\pi_{k1}I_{[\|u_{2k}(0)\|\le K]}<\infty. \tag{B14}$$

By noticing that $y_d(1)-g_1(z_k^0(1)+w_k(1))$ and $\overline{\Delta}_k(0)$ are independent of $\xi_k(1)$ and $\mathbb{E}\{\xi_k(1)|\mathscr{F}_{k-1}\}=0$, by the convergence theorem for MDS it follows that

$$\sum_{k=1}^{\infty}\frac{a_k}{c_k}[y_d(1)-g_1(z_{2k}^0(1)+w_{2k}(1))]^T \xi_{2k}(1)\overline{\Delta}_k(0)I_{[\|u_{2k}(0)\|\le K]}<\infty, \tag{B15}$$

$$\sum_{k=1}^{\infty}\frac{a_k}{c_k}[y_d(1)-g_1(z_{2k+1}^0(1)+w_{2k+1}(1))]^T \xi_{2k+1}(1)\overline{\Delta}_k(0)I_{[\|u_{2k}(0)\|\le K]}<\infty. \tag{B16}$$

Therefore,

$$\sum_{k=1}^{\infty}a_k\pi_{k3}I_{[\|u_{2k}(0)\|\le K]}<\infty, \tag{B17}$$

and hence (B4) holds.

By Theorem A in Appendix C, we have

$$d(u_{2k}(0),J_0)\underset{k\to\infty}{\longrightarrow}0, \tag{B18}$$

which incorporating with (28) and by taking notice of $c_k\Delta_k(0)\underset{k\to\infty}{\longrightarrow}0$ implies

$$d(u_k(0),J_0)\underset{k\to\infty}{\longrightarrow}0. \tag{B19}$$

This means that $\delta v_k(0)\underset{k\to\infty}{\longrightarrow}0$. Thus, the designed control sequence is optimal for $t=0$.

Assume that the conclusion of the theorem is valid for $s=0,\dots,t-1$. We now proceed to show that the theorem is also true for t, *i.e.*, to prove $\delta v_k(t)\underset{k\to\infty}{\longrightarrow}0$. For this it suffices to show that $\theta_k(t)$ and $\vartheta_k(t)$ given by (31), (32) satisfy S4.

Completely similar to the case $t=0$ it can be shown that $\theta_k(t)$ satisfies S4. Let us verify S4 for $\vartheta_k(t)$.

By the inductive assumption $\delta v_k(s)\underset{k\to\infty}{\longrightarrow}0$, $s=0,\dots,t-1$ we have

$$r_k(t+1)-r_d(t+1)\underset{k\to\infty}{\longrightarrow}0, \tag{B20}$$

or equivalently,

$$z_k^0(t+1)-\overline{z}_k^0(t+1)\underset{k\to\infty}{\longrightarrow}0. \tag{B21}$$

Notice that

$$\begin{aligned}N_k(t)&=\|e_k(t+1)\|^2-L_t(u_k(t))\\&=\|y_d(t+1)-g_{t+1}(z_k^0(t+1)\\&\quad+w_k(t+1))-\xi_k(t+1)\|^2-L_t(u_k(t))\\&=\eta_k(t+1)+\|\xi_k(t+1)\|^2+\chi_k(t+1)\\&\quad+\tau_k(t+1),\end{aligned}$$

where

$$\begin{aligned}\eta_k(t+1)&=\|y_d(t+1)-g_{t+1}(z_k^0(t+1)\\&\qquad+w_k(t+1))\|^2-P_{t+1}(z_k^0(t+1)),\\ \chi_k(t+1)&=\mathbb{E}\|y_d(t+1)-g_{t+1}(z_k^0(t+1)\\&\qquad+w_k(t+1))\|^2-L_t(u_k(t)),\\ \tau_k(t+1)&=2[y_d(t+1)-g_{t+1}(z_k^0(t+1)\\&\qquad+w_k(t+1))]^T\xi_k(t+1).\end{aligned}$$

We then have

$$\vartheta_{k+1}(t) = \frac{\overline{\Delta}_k(t)}{c_k}(N_{2k+1}(t) - N_{2k}(t)) = \delta_{k1} + \delta_{k2} + \delta_{k3},$$

where

$$\delta_{k1} = \frac{\eta_{2k+1}(t+1) - \eta_{2k}(t+1) + \|\xi_{2k+1}(t+1)\|^2 - \|\xi_{2k}(t+1)\|^2}{c_k}\overline{\Delta}_k(t),$$

$$\delta_{k2} = \frac{\chi_{2k+1}(t+1) - \chi_{2k}(t+1)}{c_k}\overline{\Delta}_k(t),$$

$$\delta_{k3} = \frac{\tau_{2k+1}(t+1) - \tau_{2k}(t+1)}{c_k}\overline{\Delta}_k(t).$$

Proceeding similarly as for the case $t=0$, we find that $\sum_{k=1}^{\infty} \frac{a_k}{c_k}(\eta_{2k+1}(t+1) - \eta_{2k}(t+1))\overline{\Delta}_k(t) I_{[\|u_{2k}(t)\| \le K]} < \infty.$ Since $\xi_k(t+1)$ is independent of $\Delta_k(t)$, we have

$$\sum_{k=1}^{\infty} \frac{a_k}{c_k}(\|\xi_{2k+1}(t+1)\|^2 - \|\xi_{2k}(t+1)\|^2) \overline{\Delta}_k(t) I_{[\|u_{2k}(t)\| \le K]} < \infty,$$

and hence

$$\sum_{k=1}^{\infty} a_k \delta_{k1} I_{[\|u_{2k}(t)\| \le K]} < \infty.$$

Similar to (B15), (B16) we see

$$\sum_{k=1}^{\infty} a_k \delta_{k3} I_{[\|u_{2k}(t)\| \le K]} < \infty.$$

So, for proving S4, it suffices to show

$$\lim_{T\to 0} \limsup_{k\to\infty} \frac{1}{T} \left\| \sum_{i=2n_k}^{m(2n_k, T_k)} a_i \delta_{k2} I_{[\|u_{2i}(t)\| \le K]} \right\| = 0, \quad \forall T_k \in [0, T]. \tag{B22}$$

For this we need only to show that (B22) holds for each component of δ_{k2}. Notice that the ith component of δ_{k2} is expressed as

$$\delta_{k2}^i = \frac{P_{t+1}(z_{2k+1}^0(t+1)) - P_{t+1}(z_{2k}^0(t+1))}{c_k \Delta_k^i(t)} - \frac{L_t(u_{2k+1}(t)) - L_t(u_{2k}(t))}{c_k \Delta_k^i(t)}. \tag{B23}$$

Similar to the proof given in [22] we find that

$$\frac{L_t(u_{2k+1}(t)) - L_t(u_{2k}(t))}{c_k \Delta_k^i(t)} = h_t^i(u_{2k}(t)) + \beta_k(t), \tag{B24}$$

where $\beta_k(t)$ satisfies

$$\lim_{T\to 0} \limsup_{k\to\infty} \frac{1}{T} \left\| \sum_{i=2n_k}^{m(2n_k, T_k)} a_i \beta_i(t) I_{[\|u_{2i}(t)\| \le K]} \right\| = 0, \quad \forall T_k \in [0, T]. \tag{B25}$$

The first term at the right-hand side of (B23) is

$$\begin{aligned} &P_{t+1}(z_{2k+1}^0(t+1)) - P_{t+1}(z_{2k}^0(t+1)) \\ &= P_{t+1}(z_{2k+1}^0(t+1)) - P_{t+1}(r_{2k+1}(t+1) \\ &\quad + C(t+1)B(t)f_t(u_{2k}(t))) + P_{t+1}(r_{2k+1}(t+1) \\ &\quad + C(t+1)B(t)f_t(u_{2k}(t))) - P_{t+1}(z_{2k}^0(t+1)). \end{aligned}$$

Noticing that $d_k(t) \triangleq P_{t+1}(r_{2k+1}(t+1) + C(t+1)B(t) f_t(u_{2k}(t))) - P_{t+1}(z_{2k}^0(t+1))$ is independent of $\Delta_k^i(t)$, by the convergence theorem for MDS we have

$$\sum_{k=1}^{\infty} \frac{a_k}{c_k \Delta_k^i(t)} d_k(t) I_{[\|u_{2k}(t)\| \le K]} < \infty. \tag{B26}$$

Similar to (B24) we obtain

$$\frac{P_{t+1}(z_{2k+1}^0(t+1)) - P_{t+1}(r_{2k+1}(t+1) + C(t+1)B(t)f_t(u_{2k}(t)))}{c_k \Delta_k^i(t)} = h_t^{*i}(u_{2k}(t)) + \beta_k'(t), \tag{B27}$$

where $\beta_k'(t)$ satisfies

$$\lim_{T\to 0} \limsup_{k\to\infty} \frac{1}{T} \left\| \sum_{i=2n_k}^{m(2n_k, T_k)} a_i \beta_i'(t) I_{[\|u_{2i}(t)\| \le K]} \right\| = 0, \quad \forall T_k \in [0, T]. \tag{B28}$$

Paying attention to (25) (B24) (B27) and to the proof of [22], we find that $h_t^i(u_{2k}(t))$ continuously depends on $r_d(t+1)$ completely in the same way as how $h_t^{*i}(u_{2k}(t))$ depends on $r_k(t+1)$. Then, by the inductive assumption (B20) we arrive at

$$[h_{t, z_{2k}^0(t)}^i(u_{2k}(t)) - h_t^i(u_{2k}(t))] I_{[\|u_{2k}(t)\| \le K]} \xrightarrow[k\to\infty]{} 0. \tag{B29}$$

Finally, (B22) follows from (B23)–(B29).

By Theorem A in Appendix C we conclude that $d(u_{2k}(t), J_t) \xrightarrow[k\to\infty]{} 0$. Since $c_k \Delta_k(t) \xrightarrow[k\to\infty]{} 0$ by (20) we have $d(u_k(t), J_t) \xrightarrow[k\to\infty]{} 0$, and hence $\delta v_k(t) \xrightarrow[k\to\infty]{} 0$.

This means that the control sequence converges to the optimal one at time t. This completes the proof.

APPENDIX C

General convergence theorem for stochastic approximation

Analysis carried out in this paper is essentially based on the following convergence theorem for stochastic approximation with expanding truncations [26].

Let $f(\cdot):\mathbb{R}^p \to \mathbb{R}^p$ and $J \triangleq \{x: f(x)=0\}$. Take a sequence of positive real numbers M_k satisfying

$$M_{k+1}>M_k, \quad M_k \underset{k\to\infty}{\longrightarrow} \infty.$$

Consider the following algorithm

$$x_{k+1} = (x_k + a_k y_{k+1}) I_{[\|x_k + a_k y_{k+1}\| \le M_{\sigma_k}]} + x^* I_{[\|x_k + a_k y_{k+1}\| > M_{\sigma_k}]} \tag{C1}$$

$$y_{k+1} = f(x_k) + \varepsilon_{k+1} \tag{C2}$$

$$\sigma_k = \sum_{i=1}^{k-1} I_{[\|x_i + a_i y_{i+1}\| > M_{\sigma_i}]}, \quad \sigma_0 = 0. \tag{C3}$$

Theorem 1. Assume the following S1-S3 hold.

S1. $a_k>0$, $a_k \underset{k\to\infty}{\longrightarrow} 0$, and $\sum_{k=1}^{\infty} a_k = \infty$.

S2. There exists a continuously differentiable function $v(\cdot):\mathbb{R}^p \to R$ such that

$$\sup_{\delta \le d(x,J) \le \Delta} f^T(x) \frac{\partial v(x)}{\partial x} < 0$$

for any $\Delta>\delta>0$, where $d(x,J)=\inf_y\{\|x-y\|, y\in J\}$, $v(J)$ is nowhere dense. Further, there exists a constant $c_0>0$ such that $\|x^*\|<c_0$ and $v(x^*)<\inf_{\|x\|=c_0} v(x)$.

S3. The function $f(\cdot)$ is measurable and locally bounded.

Then with any initial value x_0, x_k defined by (78)–(80) converges to J at any sample path where S4 holds:

S4. Along the subscripts $\{n_k\}$ of any convergent subsequence x_{n_k}

$$\lim_{T\to 0} \limsup_{k\to\infty} \frac{1}{T} \left\| \sum_{i=n_k}^{m(n_k,T_k)} a_i \varepsilon_{i+1} I_{[\|x_i\| \le K]} \right\| = 0,$$

$$\forall T_k \in [0, T]$$

if K is sufficiently large, where

$$m(k,T) \triangleq \max\left\{ m : \sum_{i=k}^{m} a_i \le T \right\}.$$

REFERENCES

1. Arimoto, S., S. Kawamura, and F. Miyazaki, “Bettering operation of robots by learning,” *J. Robot. Syst.*, Vol. 1, No. 2, pp. 123–140 (1984).
2. Bristow, D. A., M. Tharayil, and A. G. Alleyne, “A survey of iterative learning control: a learning-based method for high-performance tracking control,” *IEEE Control Syst. Mag.*, Vol. 26, No. 3, pp. 96–114 (2006).
3. Ahn, H.-S., Y. Q. Chen, and K. L. Moore, “Iterative learning control: survey and categorization from 1998 to 2004,” *IEEE Trans. Syst. Man Cybern. Part C*, Vol. 37, No. 6, pp. 1099–1121 (2007).
4. Wang, Y., F. Gao, and F. J. Doyle III, “Survey on iterative learning control, repetive control and run-to-run control,” *J. Process Control*, Vol. 19, No. 10, pp. 1589–1600 (2009).
5. Hamamoto, K. and T. Sugie, “Iterative learning control for robot manipulators using the finite dimensional input subspace,” *IEEE Trans. Autom. Control*, Vol. 18, No. 4, pp. 632–635 (2002).
6. Sakagami, N., M. Inoue, and S. Kawamura, “Theoretical and experimental studies on iterative learning control for underwater robots,” *Int. J. Offshore Polar Eng.*, Vol. 13, No. 2, pp. 120–127 (2003).
7. Sahoo, S. K., S. K. Panda, and J.-X. Xu, “Iterative learning control based direct instantaneous torque control of switched reluctance motors.” *Proc. IEEE 35th Annu. Power Electron. Spec. Conf.*, Aachen, Germany, pp. 4832–4837 (2004).
8. Xu, J.-X., S. K. Panda, Y. J. Pan, T. H. Lee, and B. H. Lam, “A modular control scheme for PMSM speed control with pulsating torque minimization,” *IEEE Trans. Ind. Electron.*, Vol. 51, No. 3, pp. 526–536 (2004).
9. Xiong, Z. and J. Zhang, “Batch-to-batch optimal control of nonlinear batch processes based on incrementally updated models,” *IEE Proc. D: Control Theory Appl.*, Vol. 151, No. 2, pp. 158–165 (2004).
10. Huang, S. N., K. K. Tan, and T. H. Lee, “Necessary and sufficient condition for convergence of iterative learning algorithm,” *Automatica*, Vol. 38, No. 7, pp. 1257–1260 (2002).
11. Fang, Y. and T. W. S. Chow, “2-D analysis for iterative learning controller for discrete-time

systems with variable initial conditions," *IEEE Trans. Circuits Syst. I-Fundam. Theory Appl.*, Vol. 50, No. 5, pp. 722–727 (2003).
12. Paszke, W., R. Merry, and R. van de Molengraft, "Iterative learning control by two-dimensional system theory applied to a motion system," *Proc. Am. Control Conf.*, New York, NY, pp. 5484–5489 (2007).
13. Meng, D., Y. Jia, J. Du, and F. Yu, "Necessary and sufficient stability condition of LTV iterative learning control systems using a 2-D approach, *Asian J. Control*, Vol. 13, No. 1, pp. 25–37 (2011).
14. Ashraf, S., E. Muhammad, and A. Al-Habaibeh, "Identification based adaptive iterative learning controller," *Asian J. Control*, Vol. 12, No. 5, pp. 607–615 (2010).
15. Xu, J.-X. and Y. Tan, "A composite energy function-based learning control approach for nonlinear systems with time-varying parametric uncertainties," *IEEE Trans. Autom. Control*, Vol. 47, No. 11, pp. 1940–1945 (2002).
16. Xu, J.-X. and J. Xu, "On iterative learning from different tracking tasks in the presence of time-varying uncertainties," *IEEE Trans. Syst. Man Cybern. Part B*, Vol. 34, No. 1, pp. 589–597 (2004).
17. Saab, S. S., "A discrete-time stochastic learning control algorithm," *IEEE Trans. Autom. Control*, Vol. 46, No. 6, pp. 877–887 (2001).
18. Chen, H.-F., "Almost surely convergence of iterative learning control for stochastic systems," *Sci. China (Ser. F)*, Vol. 46, No. 1, pp. 69–79 (2003).
19. Saab, S. S., "Selection of the learning gain matrix of an iterative learning control algorithm in presence of measurement noise," *IEEE Trans. Autom. Control*, Vol. 50, No. 11, pp. 1761–1774 (2005).
20. Chen, H.-F. and H. T. Fang, "Output tracking for nonlinear stochastic systems by iterative learning control," *IEEE Trans. Autom. Control*, Vol. 49, No. 4, pp. 583–588 (2004).
21. Shen, D. and H.-F. Chen, "ILC for a class of nonlinear systems," *J. Syst. Sci. Math.*, Vol. 28, No. 9, pp. 1053–1064 (2008).
22. Chen, H.-F., T. E. Duncan, and B. Pasik-Duncan, "A Kiefer-Wolfowitz algorithm with randomized differences," *IEEE Trans. Autom. Control*, Vol. 44, No. 3, pp. 442–453 (1999).
23. Koronacki, J., "Random-seeking methods for the stochastic uncertained optinization," *Int. J. Control*, Vol. 21, pp. 517–527 (1975).
24. Spall, J. C., "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Trans. Autom. Control*, Vol. 37, No. 2, pp. 332–341 (1992).
25. Gerencsér, L. and Z. S. Vágó, "The mathematics of noise-free SPSA." *Proc. 40th IEEE Conf. Decis. Control*, Orlando, FL, pp. 4400–4405 (2001).
26. Chen, H. F., *Stochastic Approximation and its Applications*, Kluwer, Dordrecht, The Netherlands (2002).
27. Chen, H. F. and L. Guo, *Identification and Stochastic Adaptive Control*, Birkhauser, Boston, MA (1991).

Dong Shen received his BS degree in Mathematics from Shandong University, China, in 2005. He received his PhD degree in Mathematics from the Academy of Mathematics and System Science, CAS, in 2010. Now he is a postdoctoral in Institute of Automation, CAS. His research interests include iterative learning control, stochastic control, and optimization.

Han-Fu Chen graduated from the Leningrad (St. Petersburg) State University, Leningrad, Russia. He joined the Institute of Mathematics, Chinese Academy of Sciences (CAS), Beijing, in 1961. Since 1979, he has been with the Institute of Systems Science, CAS. He is a Professor of the Key Laboratory of Systems and Control of CAS. He has authored and coauthored more than 180 journal papers and seven books. His research interests are mainly in stochastic systems, including system identification, adaptive control, and stochastic approximation and its applications to systems, control, and signal processing. Prof. Chen is a Fellow of the IFAC and IEEE. He is also a Member of TWAS and CAS.