Iterative Learning Control with Unknown Control Direction: A Novel Data-Based Approach

Dong Shen, *Member, IEEE*, and Zhongsheng Hou

***Abstract*— Iterative learning control (ILC) is considered for both deterministic and stochastic systems with unknown control direction. To deal with the unknown control direction, a novel switching mechanism, based only on available system tracking error data, is first proposed. Then two ILC algorithms combined with the novel switching mechanism are designed for both deterministic and stochastic systems. It is proved that the ILC algorithms would switch to the right control direction and stick to it after a finite number of cycles. Moreover, the input sequence converges to the desired one under the deterministic case. The input sequence converges to the optimal one with probability 1 under stochastic case and the resulting tracking error tends to its minimal value.**

***Index Terms*— Data-based control, discrete-time systems, iterative learning control, unknown control direction.**

I. Introduction

ITERATIVE learning control (ILC) is normally designed for those systems that could complete some tasks over a fixed time interval and perform them repeatedly, such as the robot arms, chemical processes, hard-disk drivers, and servo systems. In such systems, the input/output information of past and current cycles, as well as the tracking objective, are used to formulate the input of the next cycle, so that the tracking performance could be improved as the number of cycles increases to infinity. Thus ILC is a typical data-based control method since only data of input, output, and tracking objective, rather than system information, are required in the design. This is a major reason why ILC could handle systems with high nonlinearities and unknown dynamics. Due to the simplicity and efficiency of its algorithm, ILC has drawn much attention from researchers and engineers in the past decades [1]–[3]. A number of analysis techniques including the contraction mapping approach [4], [5], the 2-D system approach [6], [7], and the composite energy function approach [8], [9] have been proposed.

Manuscript received January 4, 2011; accepted November 6, 2011. Date of publication November 24, 2011; date of current version December 13, 2011. This work was supported in part by the Natural Science Foundation of China under Grant 60921061, Grant 90920305, Grant 60834001, and Grant 61120106009.

D. Shen is with the State Key Laboratory of Management and Control for Complex Systems, Beijing Engineering Research Center of Intelligent Systems and Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: dong.shen@ia.ac.cn).

Z. Hou is with the Advanced Control Systems Laboratory, School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China (e-mail: zhshhou@bjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2011.2175947

However, it is required to take, at least part of, the input/output coupling matrices as prior information when designing the ILC algorithms [5], [10]. Roughly speaking, the input/output coupling matrices reflect the control direction, and thus are important in most ILC research. Here the so-called control direction is the multiplier of the control term in the dynamic system equation, representing the motion directions of the system under any given control. Prior knowledge of these multipliers would make control design much easier. For example, Saab considers ILC for stochastic linear systems in [11] and [12], where recursive learning algorithms are derived by minimizing the trace of the input error covariance matrix and mean square convergence of the input sequence is shown. However, system matrices, especially the input/output coupling matrices, are required in the algorithms. ILC is formulated for a class of single-input-single-output (SISO) nonlinear systems with measurement noise in [13] on basis of the Robbins–Monro algorithm [14], which is a kind of stochastic approximation algorithm. This research requires no information about the input/output coupling matrices but their symbols, which are equivalent to the control directions. The objective of this paper is to develop an ILC design procedure that requires no prior knowledge of control directions.

It is difficult and challenging to design ILC with unknown control direction. There are very few related works yet. The technique of Nussbaum-type gain [15] is the most common method in the literature, where the Nussbaum-type gain plays a role as control direction probe. In [16] and [17], ILC for SISO learning control systems have been designed on the basis of this technique. A class of affine nonlinear systems is considered in [16], where the system nonlinearities are assumed bounded. It is shown that the tracking error would converge to a desired bounded domain. Xu *et al*. [17] studies a class of parameterized nonlinear systems with known nonlinear functions but unknown parameters. It is proved that the tracking error converges to zero in the sense of L_T^2. However, the systems in [16] and [17] are both continuous-time deterministic systems to facilitate the application of Nussbaum-type gain. The techniques are unsuitable for discrete-time systems or stochastic systems. Chen *et al*. [18], [19] give a novel ILC algorithm for stochastic discrete-time systems without any information about input/output coupling matrices based on the Kiefer–Wolfowitz algorithm with random difference [20], and show that the algorithm would converge to the optimal control with probability 1. However, it may be inappropriate for practical application due to the random difference technique and drawback technique in [18] and [19]. Unlike these existing methods, we propose a novel switching mechanism to deal with the unknown control direction, which is inspired by Nussbaum-type gain but more suitable for discrete-time systems. Moreover, the switching mechanism depends only on the available data. Specifically, the available data used in this paper is only the tracking error.

By combining with the new switching mechanism, we then design ILC algorithms for both deterministic and stochastic discrete-time linear systems. For expression convenience and making the idea clear, here we consider SISO system only. The multiple-input-multiple-output case could be extended to easily by incorporating the idea of unmixing set [21], as is done in [22]. The main contributions of this paper are listed as follows.

1) A novel data-based switching mechanism for discrete-time systems is first given. Then ILC algorithms are constructed based on this mechanism for discrete-time linear systems with unknown control directions.
2) For deterministic system, the ILC algorithms can regulate the control direction to be the actual one after a finite number of cycles. The input and the output of the system converge to the desired input and the tracking objective, respectively.
3) For stochastic system, the ILC algorithms are also shown to be effective. The input sequence converges to the optimal input almost surely, and the average quadric tracking error is minimized.

The rest of this paper is arranged as follows. The problem formulation is proposed in Section II. The novel switching mechanism for discrete-time systems with unknown control direction is given in Section III. Then, the ILC algorithms combining the switching mechanism for deterministic system and stochastic system with convergence results are given in Section IV. An illustrative example is proposed in Section V. Some concluding remarks are given in Section VI. All detailed proofs of the theorems with some preliminary results are placed in Appendix I–IV for reading convenience.

II. Problem Formulation

A. Deterministic Case

Consider the following SISO system:

$$x_k(t+1) = A(t)x_k(t) + B(t)u_k(t) \tag{1}$$
$$y_k(t) = C(t)x_k(t) \tag{2}$$

where the subscript k denotes different cycles, $k = 1, 2, \ldots,$ and t denotes an arbitrary time in an operation cycle, $t \in [0, T]$. $x_k(t) \in \mathbb{R}^n, u_k(t) \in \mathbb{R}, y_k(t) \in \mathbb{R}$ are system state vector, input and output, respectively. $A(t), B(t)$, and $C(t)$ are unknown time-varying system matrices with appropriate dimensions.

The control purpose is to find input $\{u_k(t), k = 0, 1, 2, \ldots\}$ such that $y_k(t+1) - y_d(t+1) \xrightarrow[k\to\infty]{} 0$, where $y_d(t), t \in [0, T]$ is the tracking objective.

Assume that the tracking objective is realizable, i.e., for an appropriate initial condition $x_d(0)$, there exists a unique control input $u_d(t)$ such that

$$x_d(t+1) = A(t)x_d(t) + B(t)u_d(t) \tag{3}$$
$$y_d(t) = C(t)x_d(t). \tag{4}$$

The following assumptions are to be used.

A1: The initial states are assumed exactly re-initialized, i.e., $x_k(0) = x_d(0)$.

A2: The control direction is unknown, i.e., $\mathrm{sgn}(C(t+1)B(t))$ is unknown. Besides, it is assumed $C(t+1)B(t) \neq 0$.

A1 is a common condition of ILC, which is called the identical initial condition (i.i.c.). For the case where the i.i.c. is not satisfied, refer to [23]–[25].

B. Stochastic Case

Consider the following stochastic system:

$$x_k(t+1) = A(t)x_k(t) + B(t)u_k(t) + w_k(t+1) \tag{5}$$
$$y_k(t) = C(t)x_k(t) + v_k(t) \tag{6}$$

where $w_k(t+1)$ and $v_k(t)$ are system noise and measurement noise, respectively. The other notations have the same meaning as the deterministic case.

The tracking objective is also denoted by $y_d(t)$ to save notations. It is also assumed that $y_d(t)$ is realizable, which means (3) and (4) hold.

Let $\{\mathcal{F}_k\}$ be a family of nondecreasing $\sigma-$algebras defined by $\mathcal{F}_k \triangleq \sigma(y_i(t), x_i(t), w_i(t), 0 \le i \le k, t \in [0, T])$. Define the set of admissible controls as follows:

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

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

$$J_t = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |y_d(t+1) - y_k(t+1)|^2. \tag{7}$$

The following assumptions are to be used.

A3: $\{w_k(t+1), v_l(s), t \in [0, T], s \in [0, T], k = 1, 2, \ldots, l = 1, 2, \ldots\}$ are mutually independent with zero mean and finite second moments

$$\mathbb{E}w_k(t+1) = 0, \ \mathbb{E}v_k(t) = 0, \ t \in [0, T], \quad k = 1, 2, \ldots,$$
$$\mathbb{E}w_k(t+1)w_k^T(t+1) = R_{t+1}^w, \mathbb{E}v_k(t)v_k^T(t) = R_t^v$$

where $\mathbb{E}(\cdot)$ denotes mathematical expectation.

A4: The initial states $\{x_k(0), k = 1, 2, \ldots\}$ are random variables, independent with $\{w_k(t+1), t \in [0, T], k = 1, 2, \ldots\}$ and $\{v_k(t), t \in [0, T], k = 1, 2, \ldots\}$, which satisfy $\mathbb{E}x_k(0) = x_d(0), \mathbb{E}\|x_k(0)\|^2 < \infty$. To facilitate the expression, denote $w_k(0) = x_d(0) - x_k(0)$.

A5: The control direction is unknown, i.e., $\mathrm{sgn}(C(t+1)B(t))$ is unknown. Besides, it is assumed that $C(t+1)B(t) \neq 0$.

III. Switching Mechanism

This section provides the switching mechanism to guarantee that the ILC algorithms proposed in the following section could automatically regulate the control directions and ensure the convergence.

First, let us call a class of sequence as a fast expanding sequence (FES) if it satisfies the following definition.

Definition 1: A sequence $\{\tau_i\}$ is called an FES if the components satisfy $0 = \tau_0 < \tau_1 < \tau_2 < \cdots$ and $\tau_i - \tau_{i-1} \xrightarrow[i\to\infty]{} \infty$.

Remark 1: We give some examples satisfying Definition 1: 1) $\tau_i = i^2$; 2) $\tau_i = i\tau_{i-1}$, $i \geq 2$; and 3) $\tau_i = \tau_{i-1}^2$ with $\tau_1 > 1$.

For an SISO system, the control gain would be either a positive constant or a negative one. Thus the set of control directions can actually be denoted as $\{+1, -1\}$. In order to find the right control direction, let us define the following two-valued function as the direction switching function (DSF).

Definition 2: The DSF is defined as

$$S(x) = \begin{cases} -1, & x \in [\tau_{2j}, \tau_{2j+1}) \\ +1, & x \in [\tau_{2j+1}, \tau_{2j+2}) \end{cases}, \quad j \in \mathbb{N} \triangleq \{0, 1, 2, \ldots\}.$$

It is worth noting that the essential requirement of the DSF is that the output chooses $+1$ and -1 alternatively in turn, while the initial value could be either one. In the rest of this paper, the DSF $S(\cdot)$ may be abbreviated as S without causing confusion.

Remark 2: We will introduce this DSF into ILC algorithms to deal with the unknown control direction as shown by (9) and (13) in the next section. Notice that the variable of the DSF should be defined on the basis of the system performance so that the algorithms could regulate the control direction adaptively. This motivates us to form the following parameter as the variable, which is called the switching parameter (SP) here.

Denote $p_k, k \geq 0$ as the SPs with $p_0 = 0$. Then denote $p_{k_0,k} \triangleq p_k - p_{k_0}$, $p_{k_0,\infty} \triangleq \lim_{k\to\infty} p_k - p_{k_0}$.

We first give the conditions on the SPs to make the convergence proofs of our ILC algorithms more apparent. These conditions could make our switching mechanism play the same role in discrete-time systems as what the Nussbaum-type gain does in the continuous-time case. Notice that the outputs of the DSF are $\{+1, -1\}$, one of which corresponds to the right control direction while the other one represents the wrong control direction.

Condition 1: The following conditions are desired for the SPs p_k.

1) Suppose $S(p_k)$ takes the wrong value and no longer switches to the other one since some integer k_0, then $p_{k_0,k} \xrightarrow[k\to\infty]{} \infty$.
2) Suppose $S(p_k)$ takes the right value and no longer switches to the other one since some integer k_0, then $\lim_{k\to\infty} p_{k_0,k} < \infty$ and $\{p_{k_0,\infty}\}$ are uniformly bounded w.r.t. (short for with respect to) k_0.
3) $\{p_k\}$ is nondecreasing.

Remark 3: Note that, if the SP p_k is designed to be a cumulative function of the tracking performance, then Condition 1 1) implies that the accumulation of tracking performance will diverge to infinity provided that $S(p_k)$ always takes the wrong control direction, which means the tracking performance is not improved. On the other hand, Condition 1 2) implies that the accumulation of tracking performance is finite provided that $S(p_k)$ always takes the right control direction, which means the tracking performance is improved along cycle axis.

Definition 3: Define the switching time (ST) as

$$T_j \triangleq \min\{k : p_k \geq \tau_j\}.$$

If $\{k : p_k \geq \tau_j\} = \emptyset$, then $T_j = \infty$.

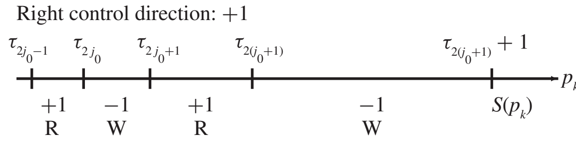

Fig. 1. Illustration of switching mechanism. "R" and "W" denote "the right control direction" and "the wrong control direction," respectively.

By Definition 3 and supposing $S = +1$ is the right control direction, then Condition 1 1) implies that $T_{2j+1} < \infty$ as long as $T_{2j} < \infty$, while Condition 1 2) implies that there exists a j_0 such that $T_{2j_0-1} < \infty$, $T_{2j_0} = \infty$.

We explain the switching mechanism with the help of Fig. 1 now.

Denote the interval $[\tau_j, \tau_{j+1})$ by I_j, $j \geq 0$. As is shown in Fig. 1, suppose the right control direction is $+1$, then $S(p_k)$ picks the right direction when p_k falls into the interval I_{2j-1}, $j \in \mathbb{N}$, and the wrong direction when p_k falls into I_{2j}, $j \in \mathbb{N}$, by Definition 2. Since p_k is nondecreasing by Condition 1 3), p_k would never fall into any interval again once p_k leaves it.

Suppose p_k falls into some interval I_{2j_0} (see Fig. 1) and never gets out it, which means $S = -1$ will always hold since, then, by Condition 1 1), we find that $p_{T_{2j_0},k} \xrightarrow[k\to\infty]{} \infty$. This further yields that p_k will absolutely get out I_{2j_0} and get in I_{2j_0+1}, which contradicts to the assumption that p_k keeps staying in the interval I_{2j_0}. In other words, p_k cannot locate in any interval of I_{2j_0}, $j \in \mathbb{N}$ for ever. That is, the DSF $S(\cdot)$ cannot stick to the wrong control direction.

Suppose p_k falls into interval I_{2j_0+1} and never gets out it, then the DSF equals to $+1$ and by Condition 1 2), $p_{T_{2j_0+1},k}$ would converge to $p_{T_{2j_0+1},\infty}$, which may or may not be larger than the length of I_{2j_0+1}. If so, then p_k will get out I_{2j_0+1} and get in $I_{2(j_0+1)}$. If not, then p_k will stay in I_{2j_0+1} since then, which means that $S(\cdot)$ finds the right control direction and stops switching. However, the former cannot occur infinite times. This is because $p_{T_{2j-1},\infty}$ is uniformly bounded w.r.t. j by Condition 1 2), while $\tau_{2j} - \tau_{2j-1} \to \infty$. Consequently, there must be some integer $j' \in \mathbb{N}$ large enough such that p_k would never get out of the interval $I_{2j'-1}$ once getting into it. That is, the DSF $S(\cdot)$ will switch to the right control direction and then stick to it for ever.

Condition 1 could be interpreted from another perspective, as follows.

The following theorem shows how to guarantee the effectiveness of the switching mechanism.

Theorem 1: If the SP p_k satisfies Condition 1, then p_k converges. Moreover, $S(p_k)$ will switch to the right control direction after finite switches and then stops.

Proof: This theorem holds obviously from the above analysis. ■

IV. ILC Designing and Convergence Analysis

A. Deterministic Case

This subsection gives the ILC algorithms for deterministic linear system with unknown control direction based on the data-based switching mechanism proposed in Section III, while the stochastic case is dealt in the next subsection.

Let $a_k = (1/k)$, $k \geq 1$, and $a_0 = 1$. Design SP $p_k(t+1)$ as

$$p_k(t+1) = \sum_{i=1}^{k} a_i e_i^2(t+1) \tag{8}$$

where $e_k(t) \triangleq y_d(t) - y_k(t)$. As has been noted, $p_0 = 0$. The recursive ILC algorithms are designed for $k = 0, 1, 2, \ldots$

$$u_{k+1}(t) = u_k(t) + a_k S(p_k(t+1)) e_k(t+1) \tag{9}$$

$$p_{k+1}(t+1) = p_k(t+1) + a_{k+1} e_{k+1}^2(t+1) \tag{10}$$

where the initial inputs $\{u_0(t)\}$ are arbitrarily given.

Theorem 2: For the deterministic system (1) and (2), assume A1 and A2 hold and apply recursive ILC algorithms (9) and (10), then $\forall t \in [0, T]$:

1) DSF $S(p_k(t+1))$ finds the right control direction and keeps it after finite switches;
2) the control sequence $\{u_k(t)\}$ converges to the desired control $u_d(t)$ with a rate $|u_d(t) - u_k(t)| = o(k^{-\delta})$, where $\delta > 0$ is a suitable constant;
3) the system output $y_k(t+1)$ converges to $y_d(t+1)$, i.e., $y_d(t+1) - y_k(t+1) \xrightarrow[k\to\infty]{} 0$.

Proof: See Appendix II. ■

B. Stochastic Case

We first give a theorem on the optimal control sequence making the tracking index (7) minimized.

By (3), (4), and A5, we have

$$u_d(t) = \left[(C(t+1)B(t))^T C(t+1)B(t) \right]^{-1} \times (C(t+1)B(t))^T (y_d(t+1) - C(t+1)A(t)x_d(t)).$$

Theorem 3: For the stochastic system (5) (6) and the index (7), assume A3–A5 hold and $y_d(t+1)$ is realizable, then for any arbitrary time t, the index (7) will be minimized if the control sequence $\{u_k(t)\}$ is admissible and satisfies $u_k(i) \xrightarrow[k\to\infty]{} u_d(i), i = 0, 1, \ldots, t$. In this case, $\{u_i(t, k)\}$ is called the optimal control sequence.

Proof: See Appendix III. ■

For stochastic system (5) and (6), because of the system noise and measurement noise, the SP p_k defined by (8) does not satisfy Condition 1 anymore even if the control input is $u_d(t)$. Thus we need to design a new SP for stochastic case

$$q_k(t+1) = \frac{1}{k} \sum_{l=1}^{k} e_l^2(t+1) \tag{11}$$

$$p_k(t+1) = \sup_{s \leq k} \{q_s(t)\} \tag{12}$$

with $p_0 = 0$.

It is obvious that $p_k(t)$ is nondecreasing. Then the recursive ILC algorithms are designed as follows for $k = 0, 1, 2, \ldots$:

$$u_{k+1}(t) = u_k(t) + a_k S(p_k(t+1)) e_k(t+1) \tag{13}$$

$$q_{k+1}(t+1) = q_k(t+1) + \frac{1}{k+1} \times \left(e_{k+1}^2(t+1) - q_k(t+1) \right) \tag{14}$$

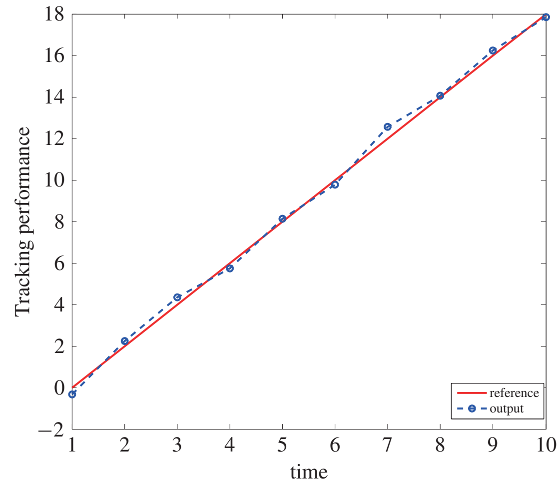

Fig. 2. $y_{300}(t)$ versus $y_d(t)$.

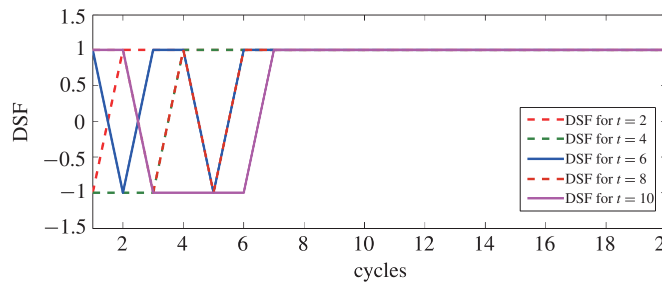

Fig. 3. DSF for $t = 2, 4, 6, 8$, and 10.

$$p_{k+1}(t+1) = \max\{p_k(t+1), q_{k+1}(t+1)\} \tag{15}$$

where $a_0 = 1$ and $a_k = (1/k)$, $k \geq 1$. The initial inputs $\{u_0(t)\}$ could be arbitrarily given.

We have the following convergence theorem.

Theorem 4: For the stochastic system (5) (6) and the index (7), assume A3–A5 hold and apply recursive ILC algorithms (13)–(15), then $\forall t \in [0, T]$:

1) DSF $S(p_k(t+1))$ finds the right control direction and keeps it after finite switches;
2) the control sequence $\{u_k(t)\}$ is optimal with a rate $\|u_d(t) - u_k(t)\| = o(k^{-\delta})$ a.s., where $\delta > 0$ is a suitable constant;
3) (7) is minimized.

Proof: See Appendix IV. ■

V. Illustrative Example

The SISO time-varying system is described as follows:

$$x_k(t+1) = (-1.1 + \sin(t))x_k(t) + (1.3 + 0.5\sin(4t))u_k(t) + w_k(t)$$

$$y_k(t) = (2.7 + 0.5\cos(2t))x_k(t) + v_k(t)$$

where time t is valued in $t \in \{1, 2, \ldots, 10\}$. Assume $w_k(t), v_k(t) \in \mathcal{N}(0, 0.1^2)$. It is easy to see that the right output of the DSFs should be $+1$ by noticing that $(1.3 + 0.5\sin(4t)) \times (2.7 + 0.5\cos(2t)) > 0$, $\forall t$.

The reference signal is $y_d(t) = 2(t-1)$, $t \in \{1, 2, \ldots, 10\}$. The initial control actions are $u_0(t) \triangleq 0$, $\forall t \in \{1, 2, \ldots, 10\}$.

The FES is defined as $\tau_0 = 0$ and $\tau_i = i!$, $i = 1, 2, \ldots$.

The algorithm has run 300 cycles and the tracking results at $k = 300$ are presented in Fig. 2, where the solid line is the reference objective and the cycle-dashed line denotes the system output. Fig. 3 shows the output of the DSF at times $t = 2, 4, 6, 8$, and 10 for the first 20 cycles, where we can see that the DSFs converge to the correct control direction quickly.

If $w_k(t) = v_k(t) = 0$, which means the system is deterministic and the simulation behaves better since no noise is brought in. Due to limitations of space, this case is omitted.

VI. Conclusion

The ILC problems for time-varying discrete-time SISO systems with unknown control direction under both deterministic and stochastic cases were considered in this paper. In order to deal with the unknown control direction, we first proposed a novel switching mechanism to select the control direction adaptively using a two-valued DSF, whose SP was designed based on available data. Then we introduced the DSF into the ILC algorithms to guarantee that the ILC algorithms could find the right control direction after a finite number of switches. The convergence of the control sequences generated by proposed ILC algorithms was proved under suitable conditions. For further research, one can consider the affine nonlinear system or more complex nonlinear system with unknown control direction.

Appendix I
Some Preliminary Theorems

The following Theorem 1.1 about weighted martingale difference estimation refers to [26].

Theorem 1.1: Let $\{X(t), \mathcal{F}_t\}$ be a martingale difference sequence and $\{M(t), \mathcal{F}_t\}$ an adapted sequence, $\|M(t)\| < \infty$, $\forall t \geq 0$. If $\sup_{t\geq 0} \mathbb{E}[\|X(t)\|^2 | \mathcal{F}_{t-1}] < \infty$, a.s., then as $n \to \infty$

$$\sum_{t=0}^{n} M(t)X(t+1) = O\left(\left(\sum_{t=0}^{n} \|M(t)\|^2\right)^{\frac{1}{2}+\eta}\right) \; a.s.\ \forall \eta > 0.$$

The next theorem is a combination of three results on stochastic approximation, namely [27, Ths. 2.2.1 and 3.3.1] and [28, Proposition 1.6].

Consider the RM algorithm

$$x_{k+1} = x_k + a_k(f(x_k) + \epsilon_k) \tag{16}$$

where $a_k = (1/k)$. Let x^0 be root of $f(x)$. The following assumptions will be used.

H1: There exists a uniformly continuous smooth function $V(\cdot) : \mathbb{R} \to \mathbb{R}_+$ such that:

1) $\lim_{\|x\|\to\infty} V(x) = \infty$;
2) there exists $\delta' > 0$ such that

$$\langle \nabla V(x), f(x)\rangle \leq -\delta'$$

for all $\|x\| \geq r$, where ∇ denotes gradient.

H2: For the sample path ω under consideration, the observation noise ϵ_k in (16) can be decomposed into two parts: $\epsilon_k = \epsilon_k^{'} + \epsilon_k^{''}$ such that

$$\sum_{k=1}^{\infty} a_k^{1-\delta} \epsilon_k^{'} < \infty, \qquad \epsilon_k^{''} = O\left(a_k^{\delta}\right) \tag{17}$$

for some $\delta \in (0, 1]$.

H3: $f(\cdot)$ satisfies globally the Lipschitz condition, and is differentiable at x^0 such that, as $x \to x^0$

$$f(x) = F \cdot (x - x^0) + \delta(x),\ \delta\left(x^0\right) = 0,\ \delta(x) = o(\|x - x^0\|). \tag{18}$$

The matrix F is stable, in addition, $F + \delta I$ is also stable, where δ is given by (17).

Theorem 1.2: Assume that H1, H2, and H3 hold. Then $\{x_k\}$ generated by the RM algorithm (16) is bounded. Moreover, $x_k \xrightarrow[k\to\infty]{} x^0$ with the following convergence rate:

$$\left\| x_k - x^0 \right\| = o\left(k^{-\delta}\right) \tag{19}$$

where δ is the one given in (17).

Remark 4: For linear function $f(x) = \beta(x - b)$, $\beta < 0$, it is easy to find such a $V(x)$ that satisfies requirements of H1. Here we give an example. Without loss of generality, assume $b = 0$ and define

$$V(x) = \begin{cases} 2x - 1, & x \geq 1 \\ x^2, & -1 < x < 1 \\ -2x + 1, & x \leq -1. \end{cases}$$

Obviously, this $V(x)$ is a uniformly continuous smooth function. Item 1) of H1 is clearly established. Take $r = 1$, then 2) of H1 holds.

Appendix II
Proof of Theorem 2

This theorem is proved by mathematical induction w.r.t. time t. A roadmap is first given. The case $t = 0$ is considered as the base case, where the three items are proved one by one. Specifically, the first item of Theorem 2 is shown by verifying Condition 1, then the other two hold straightforwardly by simple mathematical derivation. Then these items are affirmed again as the inductive step for time t provided that all conclusions hold for time $0, 1, \ldots, t-1$. In the following, we call $a_n = O(b_n)$ if $b_n \geq 0$ and there exist $M > 0$ such that $|a_n| \leq M b_n$, $\forall n$, and $a_n = o(b_n)$ if $b_n \geq 0$ and $(a_n/b_n) \to 0$, $n \to \infty$.

For simplicity of expression, denote $\delta x_k(t) \triangleq x_d(t) - x_k(t)$, $\delta u_k(t) \triangleq u_d(t) - u_k(t)$.

Basis: Showing that the items hold for the case $t = 0$.

By (1)–(4)

$$e_k(1) = C(1)\delta x_k(1) = C(1)A(0)\delta x_k(0) + C(1)B(0)\delta u_k(0).$$

Denote $m_0 = C(1)B(0) \in \mathbb{R}$, then by A1

$$e_k(1) = m_0 \delta u_k(0). \tag{20}$$

For convenience of statements, without loss of generality, it is assumed $m_0 > 0$ throughout the proof, consequently, the right control direction is $S(\cdot) = +1$. The SP and ST are defined as $p_k(1) = \sum_{j=1}^{k} a_j e_j^2(1)$, $T_j(1) = \min\{k : p_k(1) \geq \tau_j\}$.

To show conclusion 1), it suffices to verify that Condition 1 is satisfied by $p_k(1)$. Notice that the nondecreasing requirement holds obviously.

Let us first verify Condition 1 1), i.e., prove $p_{T_{2j_0}, T_{2j_0}+n}(1) \xrightarrow[n\to\infty]{} \infty$ provided that $T_{2j_0} < \infty$ and the DSF stops switching since the T_{2j_0}th cycle. Notice that under the supposition of Condition 1 1), $S(\cdot) \equiv -1, \forall k \geq T_{2j_0}$.

Subtracting both sides of (9) from the term $u_d(0)$ leads to

$$\delta u_{k+1}(0) = (1 - a_k S(p_k(1)) m_0) \delta u_k(0). \tag{21}$$

Then $\forall i = 1, 2, \ldots$

$$\begin{aligned}(\delta u_{T_{2j_0}+i}(0))^2 &= (1 + a_{T_{2j_0}+i-1} m_0)^2 (\delta u_{T_{2j_0}+i-1}(0))^2 \\ &\geq (\delta u_{T_{2j_0}+i-1}(0))^2 \\ &\geq (\delta u_{T_{2j_0}}(0))^2. \end{aligned} \tag{22}$$

Hence by (20) and (22)

$$\begin{aligned} p_{T_{2j_0}, T_{2j_0}+n}(1) &= \sum_{k=1}^{n} a_{T_{2j_0}+k} e_{T_{2j_0}+k}^2(1) \\ &\geq \sum_{k=1}^{n} a_{T_{2j_0}+k} m_0^2 (\delta u_{T_{2j_0}}(0))^2 \xrightarrow[n\to\infty]{} \infty. \end{aligned}$$

Now let us verify Condition 1 2). For any fixed $j_0 \in \mathbb{N}$, denote $T_0 = T_{2j_0+1}$ for simplicity of notation and assume $T_0 < \infty$. Notice that, under assumption of Condition 1 2), $S(\cdot) \equiv +1$, $\forall k \geq T_0$. Therefore $\forall n > 0$

$$\begin{aligned} p_{T_0, T_0+n} &= \sum_{k=1}^{n} a_{T_0+k} e_{T_0+k}^2(1) \\ &= \sum_{k=1}^{n} a_{T_0+k} m_0^2 \left(\prod_{j=1}^{k-1} (1 - a_{T_0+j} m_0) \cdot \delta u_{T_0}(0) \right)^2. \end{aligned}$$

Noticing $a_k = (1/k)$ and by [29, p. 224, Th. 10], we have

$$\prod_{j=1}^{k-1} (1 - a_{T_0+j} m_0) = O\left(\frac{1}{(T_0 + k)^{m_0}} \right) \tag{23}$$

for sufficiently large k, which means that $\lim_{n\to\infty} p_{T_0, T_0+n} < \infty$. To verify Condition 1, there is only the uniform boundedness of p_{T_0, T_0+n} left to be proved.

Denote $M(T_0) = m_0^2 \sum_{k=1}^{\infty} a_{T_0+k} \left(\prod_{l=0}^{k-1} (1 - a_{T_0+l} m_0)^2 \right)$. Paying attention to (23) we have

$$M(T_0) \leq \frac{\alpha_1}{T_0^{2m_0}} \tag{24}$$

where α_1 is a suitable positive constant independent of T_0. Therefore

$$p_{T_0, \infty} \leq \frac{\alpha_1}{T_0^{2m_0}} \delta u_{T_0}^2(0). \tag{25}$$

From (21) we see that

$$\begin{aligned} \delta u_{T_0}^2(0) &= (1 - a_{T_0-1} S(p_{T_0-1}(1)) m_0)^2 \delta u_{T_0-1}^2(0) \\ &= \left(\prod_{k=1}^{T_0-1} (1 - a_k S(p_k(1)) m_0) \right)^2 \delta u_1^2(0) \\ &\leq \left(\prod_{k=1}^{T_0-1} (1 + a_k m_0) \right)^2 \delta u_1^2(0) \\ &\leq \kappa_1 T_0^{2m_0} \end{aligned} \tag{26}$$

where κ_1 is a suitable positive constant independent of T_0. Then by (25) and (26)

$$p_{T_0, \infty} \leq \frac{\alpha_1}{T_0^{2m_0}} \kappa_1 T_0^{2m_0} = \alpha_1 \kappa_1. \tag{27}$$

The uniform boundedness of $p_{T_0, \infty}$ w.r.t. T_0 holds.

Therefore, by Theorem 1 the first item of Theorem 2 holds for $t = 0$. That is, $S(p_k(1))$ will find the right control direction and keep it after finite switches.

Now it remains to prove the other two items whenever the algorithm stops switching. There is a k_0 such that $S(p_k(1)) = +1$, $\forall k \geq k_0$ by the first item. By (21) it follows that

$$\delta u_{k+1}(0) = (1 - a_k m_0) \delta u_k(0) \quad \forall k \geq k_0.$$

Then the second item of Theorem 2 holds obviously for $t = 0$ by noticing (23). Moreover, the last item also holds for $t = 0$ by (20).

Inductive Step: Showing that the items hold for time t provided that all conclusions hold for time $0, 1, \ldots, t - 1$.

Denote $\Phi_{ij} = A(i) \cdots A(j)$, $i \geq j$, $\Phi_{j,j+1} = I$. By (1)–(4) and A1, we have

$$\delta x_k(t + 1) = \sum_{i=1}^{t+1} \Phi_{ti} B(i - 1) \delta u_k(i - 1). \tag{28}$$

Denote $\rho_k(t) = C(t + 1) \sum_{i=1}^{t} \Phi_{ti} B(i - 1) \delta u_k(i - 1)$, $m_t = C(t + 1) B(t)$. By the induction hypothesis, $\rho_k(t) = o(k^{-\delta})$, where $\delta > 0$ is a constant. Therefore

$$e_k(t + 1) = m_t \delta u_k(t) + \rho_k(t) \tag{29}$$

$$\begin{aligned} \delta u_{k+n}(t) &= \prod_{j=0}^{n-1} (1 - a_{k+j} S_{k+j}(\cdot) m_t) \cdot \delta u_k(t) \\ &- \sum_{j=0}^{n-1} \prod_{i=j}^{n-2} (1 - a_{k+i} S_{k+i}(\cdot) m_t) \cdot a_{k+j} S_{k+j}(\cdot) \rho_{k+j}(t) \end{aligned} \tag{30}$$

where $\prod_{i=a}^{b}(\cdot) \triangleq 1$ if $a > b$ and $S_m(\cdot)$ denotes $S(p_m(t + 1))$. The notation $S_m(\cdot)$ shall be used somewhere in the following without misunderstanding for concise expression.

Without loss of generality, assume $m_t > 0$ throughout the proof. Consequently, the right control direction is $S \equiv +1$. Now we prove that $p_k(t + 1)$ satisfies Condition 1, of which the third item is obviously valid.

Let us take Condition 1 1) into consideration first. Suppose $T_{2j_0} < \infty$ and the DSF stops switching since the T_{2j_0}th cycle, then $S(p_k(t + 1)) = -1$, $\forall k \geq T_{2j_0}$ according to the supposition of Condition 1 1). Thereby

$$\begin{aligned} p_{T_{2j_0}, T_{2j_0}+n}(t + 1) &= \sum_{k=1}^{n} a_{T_{2j_0}+k} e_{T_{2j_0}+k}^2(t + 1) \\ &= \sum_{k=1}^{n} a_{T_{2j_0}+k} \left(m_t \delta u_{T_{2j_0}+k}(t) + \rho_{T_{2j_0}+k}(t) \right)^2 \\ &\geq Q_1 + Q_2 \end{aligned}$$

where

$$Q_1 = \sum_{k=1}^{n} a_{T_{2j_0}+k}(m_t \delta u_{T_{2j_0}+k}(t))^2$$
$$Q_2 = 2m_t \sum_{k=1}^{n} a_{T_{2j_0}+k}\delta u_{T_{2j_0}+k}(t)\rho_{T_{2j_0}+k}(t).$$

By noticing that $(\rho_{T_{2j_0}+k}(t)/\delta u_{T_{2j_0}+k}(t)) \xrightarrow[k\to\infty]{} 0$, it is easy to find that Q_1 plays a leading role in $Q_1 + Q_2$ when n is large enough, i.e., $Q_2 = o(Q_1)$. Thus we will only take Q_1 into consideration. By (30) and assuming Condition 1 1)

$$\begin{aligned} Q_1 &= \sum_{k=1}^{n} a_{T_{2j_0}+k} m_t^2 \left(\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j} m_t) \cdot \delta u_{T_{2j_0}}(t) \right. \\ &\quad \left. + \sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 + a_{T_{2j_0}+i} m_t)\cdot a_{T_{2j_0}+j}\rho_{T_{2j_0}+j}(t)\right)^2 \\ &\geq Q_3 + Q_4 \end{aligned} \tag{31}$$

where

$$Q_3 = \sum_{k=1}^{n} a_{T_{2j_0}+k} m_t^2 \left(\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j} m_t)\delta u_{T_{2j_0}}(t)\right)^2$$
$$\begin{aligned} Q_4 &= 2\sum_{k=1}^{n} a_{T_{2j_0}+k} m_t^2 \left(\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j} m_t)\cdot \delta u_{T_{2j_0}}(t)\right) \\ &\quad \times \left(\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 + a_{T_{2j_0}+i} m_t)\cdot a_{T_{2j_0}+j}\rho_{T_{2j_0}+j}(t)\right). \end{aligned}$$

By [29, p. 224, Th. 10], for sufficiently large k, we find

$$\begin{aligned} &\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j} m_t)\delta u_{T_{2j_0}}(t) \\ &= O((T_{2j_0} + k)^{m_t}) \xrightarrow[k\to\infty]{} \infty \end{aligned} \tag{32}$$

$$\begin{aligned} &2\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 + a_{T_{2j_0}+i} m_t) a_{T_{2j_0}+j}\rho_{T_{2j_0}+j}(t) \\ &= \sum_{j=0}^{k-1} O\left(\left(\frac{T_{2j_0}+k-1}{T_{2j_0}+j}\right)^{m_t}\right)\frac{1}{T_{2j_0}+j}\, o\left(\frac{1}{(T_{2j_0}+j)^{\delta}}\right) \\ &= O\left(\frac{1}{(T_{2j_0}+k)^{\delta}}\right) \xrightarrow[k\to\infty]{} 0. \end{aligned} \tag{33}$$

This fact implies that $Q_4 = o(Q_3)$.

On the other hand, it is clear that

$$Q_3 \geq \sum_{k=1}^{n} a_{T_{2j_0}+k} m_t^2 (\delta u_{T_{2j_0}}(t))^2 \xrightarrow[n\to\infty]{} \infty. \tag{34}$$

Consequently, $Q_1 + Q_2 \xrightarrow[n\to\infty]{} \infty$, which means that $p_{T_{2j_0}, T_{2j_0}+n} \xrightarrow[n\to\infty]{} \infty$. In other words, Condition 1 1) holds.

Next let us check Condition 1 2). For any fixed $j_0 \in \mathbb{N}$, denote $T_0 = T_{2j_0+1}$ to keep the notation simple. Assume $T_0 < \infty$ and the DSF stops switching since the T_0th cycle. Thus $S(p_k(t+1)) = +1$, $\forall k \geq T_{2j_0-1}$ and then

$$\begin{aligned} p_{T_0, T_0+n}(t+1) &\leq \sum_{k=1}^{\infty} a_{T_0+k} e^2_{T_0+k}(t+1) \\ &= \sum_{k=1}^{\infty} a_{T_0+k}\left(m_t \delta u_{T_0+k}(t) + \rho_{T_0+k}(t)\right)^2 \\ &\leq 2\sum_{k=1}^{\infty} a_{T_0+k}\left(m_t \delta u_{T_0+k}(t)\right)^2 + 2\sum_{k=1}^{\infty} a_{T_0+k}\left(\rho_{T_0+k}(t)\right)^2 \end{aligned}$$

where the last item on the right side is obviously finite by the induction hypothesis, while the first item is estimated as follows:

$$\begin{aligned} &\sum_{k=1}^{\infty} a_{T_0+k}\left(m_t \delta u_{T_0+k}(t)\right)^2 \\ &= \sum_{k=1}^{\infty} a_{T_0+k} m_t^2 \left(\prod_{j=0}^{k-1}(1 - a_{T_0+j} m_t)\cdot \delta u_{T_0}(t) \right. \\ &\qquad \left. - \sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 - a_{T_0+i} m_t)\cdot a_{T_0+j}\rho_{T_0+j}(t)\right)^2 \\ &\leq 2m_t^2(\Delta_1 + \Delta_2) \end{aligned}$$

where

$$\begin{aligned} \Delta_1 &\triangleq \sum_{k=1}^{\infty} a_{T_0+k}\left(\prod_{j=0}^{k-1}(1 - a_{T_0+j} m_t)\cdot \delta u_{T_0}(t)\right)^2 \\ &= \delta u^2_{T_0}(t) \times \sum_{k=1}^{\infty} O\left(\frac{1}{(T_0+k)^{1+2m_t}}\right) < \infty \\ \Delta_2 &\triangleq \sum_{k=1}^{\infty} a_{T_0+k}\left(\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 - a_{T_0+i} m_t) a_{T_0+j}\rho_{T_0+j}(t)\right)^2 \\ &= \sum_{k=1}^{\infty} a_{T_0+k}\left(\frac{1}{(T_0+k)^{m}_t}\sum_{j=0}^{k-1} o\left((T_0+j)^{m_t-1-\delta}\right)\right)^2 \\ &= \sum_{k=1}^{\infty} o\left(\frac{1}{(T_0+k)^{1+2\delta}}\right) < \infty. \end{aligned}$$

Therefore

$$p_{T_0, T_0+n}(t+1) \leq \sum_{k=1}^{\infty} a_{T_0+k} e^2_{T_0+k}(t+1) < \infty.$$

There is only the uniform boundedness of $p_{T_0,\infty}$ w.r.t. T_0 left for us to show, where it suffices to show that the term Δ_1 is uniformly bounded w.r.t. T_0.

Denote $L(T_0) = \sum_{k=1}^{\infty} a_{T_0+k}(\prod_{j=0}^{k-1}(1 - a_{T_0+j} m_t))^2$. By a similar derivation as that of (24) it is clear that there exists a constant α_2 independent of T_0 such that

$$L(T_0) \leq \frac{\alpha_2}{T_0^{2m_t}}. \tag{35}$$

By (30) we know

$$\begin{aligned}\delta u_{T_0}^2(t) = &\Bigg(\prod_{j=1}^{T_0-1}(1-a_jS_j(\cdot)m_t)\cdot\delta u_1(t)\\&-\sum_{j=1}^{T_0-1}\prod_{i=j}^{n-2}(1-a_iS_i(\cdot)m_t)\cdot a_jS_j(\cdot)\rho_j(t)\Bigg)^2\\&\le 2\left(\prod_{j=1}^{T_0-1}(1-a_jS_j(\cdot)m_t)\cdot\delta u_1(t)\right)^2\\&+2\left(\sum_{j=1}^{T_0-1}\prod_{i=j}^{n-2}(1-a_iS_i(\cdot)m_t)\cdot a_jS_j(\cdot)\rho_j(t)\right)^2. \quad (36)\end{aligned}$$

Similar to (26) there exists a constant κ_2 independent of T_0 such that the first term on the right side of (36) satisfies

$$2\left(\prod_{j=1}^{T_0-1}(1-a_jS_j(\cdot)m_t)\cdot\delta u_1(t)\right)^2 \le \kappa_2 T_0^{2m_t}. \quad (37)$$

Now consider the second term on the right side of (36). By the induction hypothesis, there is a constant L such that $\rho_k(t)\le Lk^{-\delta}$. Consequently

$$\begin{aligned}&2\left(\sum_{j=1}^{T_0-1}\prod_{i=j}^{T_0-2}(1-a_iS_i(\cdot)m_t)\cdot a_jS_j(\cdot)\rho_j(t)\right)^2\\\le&2\left(\sum_{j=1}^{T_0-1}\left|\prod_{i=j}^{T_0-2}(1-a_iS_i(\cdot)m_t)\cdot a_jS_j(\cdot)\rho_j(t)\right|\right)^2\\\le&2\left(\sum_{j=1}^{T_0-1}\prod_{i=j}^{T_0-2}(1+a_im_t)\cdot\frac{1}{j}\cdot Lj^{-\delta}\right)^2\\=&O\left(\frac{1}{T_0^{2\delta}}\right). \quad (38)\end{aligned}$$

Combining (35)–(38), it leads to the uniform boundedness of Δ_1.

Then by Theorem 1, the first item of Theorem 2 holds for time t. Thus we can suppose $S(p_k(t+1))$ stops switching after k_0th iteration, then $S(p_{k_0+n})\equiv +1, \forall n>0$ and

$$\begin{aligned}\delta u_{k_0+n}(t) = &\prod_{j=0}^{n-1}(1-a_{k_0+j}m_t)\cdot\delta u_{k_0}(t)\\&-\sum_{j=0}^{n-1}\prod_{i=j}^{n-2}(1-a_{k_0+i}m_t)\cdot a_{k_0+j}\rho_{k_0+j}(t).\end{aligned}$$

Since

$$\prod_{j=0}^{n-1}(1-a_{k_0+j}m_t)\cdot\delta u_{k_0}(t) = o\left((k_0+n)^{-m_t}\right)\xrightarrow[n\to\infty]{} 0$$

$$\begin{aligned}&\sum_{j=0}^{n-1}\prod_{i=j}^{n-2}(1-a_{k_0+i}m_t)\cdot a_{k_0+j}\rho_{k_0+j}(t)\\=&\sum_{j=0}^{n-1}O\left(\left(\frac{k_0+j}{k_0+n-1}\right)^{m_t}\right)a_{k_0+j}\cdot o\left(\frac{1}{(k_0+j)^\delta}\right)\\=&O\left(\frac{1}{(k_0+n-1)^{m_t}}\right)\sum_{j=0}^{n-1}o((k_0+j)^{m_t-1-\delta})\\=&o\left((k_0+n)^{-\delta}\right)\xrightarrow[n\to\infty]{} 0\end{aligned}$$

and thus

$$\delta u_{k_0+n}(t) = o(n^{\max\{-m_t,-\delta\}})\xrightarrow[n\to\infty]{} 0.$$

That is, the second item of Theorem 2 holds for time t. Moreover, the last item of the theorem also holds. By the mathematical induction principle the proof is completed. ∎

Appendix III
Proof of Theorem 3

By (5) and (3)

$$\delta x_k(t+1) = A(t)\delta x_k(t) + B(t)\delta u_k(t) - w_k(t+1).$$

By backwardly iterating this equation, we have

$$\delta x_k(t+1) = \sum_{i=1}^{t+1}\Phi_{ti}B(i-1)\delta u_k(i-1) + \sum_{i=0}^{t+1}\Phi_{ti}w_k(i)$$

and thereby

$$\begin{aligned}y_d(t+1)-y_k(t+1) &= C(t+1)\delta x_k(t+1) - v_k(t+1)\\&= \phi_k(t+1)+\varphi_k(t+1)-v_k(t+1)\end{aligned}$$

where

$$\begin{aligned}\phi_k(t+1) =&C(t+1)\sum_{i=1}^{t+1}\Phi_{ti}B(i-1)\delta u_k(i-1)\\\varphi_k(t+1) =&C(t+1)\sum_{i=0}^{t+1}\Phi_{ti}w_k(i).\end{aligned}$$

By A3 and A4 and noticing $u_k(i)\in\mathcal{F}_{k-1}, i=0,1,\ldots,t$, it is clear that $\phi_k(t+1), \varphi_k(t+1)$, and $v_k(t+1)$ are mutually independent.

By Theorem 1.1

$$\begin{aligned}&\sum_{k=1}^{n}\phi_k(t+1)\left(\varphi_k(t+1)-v_k(t+1)\right)\\=&O\left(\left(\sum_{k=1}^{n}\|\phi_k(t+1)\|^2\right)^{\frac{1}{2}+\eta}\right), \quad \text{a.s.} \quad \forall\eta>0\\&\sum_{k=1}^{n}\varphi_k(t+1)v_k(t+1)\\=&O\left(\left(\sum_{k=1}^{n}\|v_k(t+1)\|^2\right)^{\frac{1}{2}+\eta}\right), \quad \text{a.s.} \quad \forall\eta>0.\end{aligned}$$

Consequently

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

$$= \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\phi_k(t+1)\|^2 + \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\varphi_k(t+1)\|^2$$
$$+ \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\upsilon_k(t+1)\|^2$$
$$= \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\phi_k(t+1)\|^2$$
$$+\mathrm{tr}\left[C(t+1)\sum_{i=0}^{t+1}\Phi_{ti}R_i^w\Phi_{ti}^T C^T(t+1) + R_{t+1}^\upsilon\right]$$

where the last term is independent of control.

Therefore, the minimum of (7) is achieved if and only if the first term on the right side of the above equation is zero. This means that $\{u_k(i), i = 0, 1, \ldots, t, k = 1, 2, \ldots\}$ is optimal if $\delta u_k(i) \xrightarrow[k\to\infty]{} 0, i = 0, 1, \ldots, t$. The proof is completed. ■

Appendix IV
Proof of Theorem 4

This theorem is proved by mathematical induction w.r.t. t. The roadmap is similar to the one in proof of Theorem 2.

Basis: Showing that the items hold for the case $t = 0$.

Notice that

$$\begin{aligned} e_k(1) &= y_d(1) - y_k(1) = C(1)\delta x_k(1) - \upsilon_k(1) \\ &\triangleq m_0\delta u_k(0) + \xi_k(1) \end{aligned} \tag{39}$$

where $\xi_k(1) \triangleq C(1)A(0)w_k(0) + C(1)w_k(1) - \upsilon_k(1)$, $m_0 = C(1)B(0)$. By A3 and A4, it is clear that $\xi_k(1)$ is a zero-mean random variable with finite second moment. Denote $R_\xi = C(1)A(0)R_0^w A^T(0)C^T(1) + C(1)R_1^w C^T(1) + R_1^\upsilon$. Without loss of generality, assume $m_0 > 0$; then the right and wrong control directions are $S = +1$ and $S = -1$, respectively. By (13) and (39), we have

$$\begin{aligned} \delta u_{k+1}(0) &= \delta u_k(0) - a_k S(p_k(1))e_k(1) \\ &= (1 - a_k S(p_k(1))m_0)\delta u_k(0) - a_k S(p_k(1))\xi_k(1) \end{aligned} \tag{40}$$

$$\begin{aligned} \delta u_{k+n}(0) &= \prod_{j=0}^{n-1}(1 - a_{k+j}S(p_{k+j}(1))m_0)\delta u_k(0) \\ &- \sum_{j=0}^{n-1}\prod_{i=j}^{n-2}(1 - a_{k+i}S(p_{k+i}(1))m_0)\cdot a_{k+j}S(p_{k+j}(1))\xi_{k+j}(1). \end{aligned} \tag{41}$$

We first verify Condition 1 1). Suppose $T_{2j_0} < \infty$ and the DSF stops switching from the T_{2j_0}th cycle. Then the target is to show $q_{T_{2j_0}+n}(1) \xrightarrow[n\to\infty]{} \infty$. By (14) and (15)

$$\begin{aligned} q_{T_{2j_0}+n}(1) &= \frac{1}{T_{2j_0}+n}\sum_{k=1}^{T_{2j_0}+n} e_k^2(1) \\ &= \frac{T_{2j_0}}{T_{2j_0}+n}q_{T_{2j_0}}(1) + \frac{1}{T_{2j_0}+n}\sum_{k=1}^{n} e_{T_{2j_0}+k}^2(1). \end{aligned} \tag{42}$$

It is clear that the first term on the right side of the last equation tends to zero

$$\frac{T_{2j_0}}{T_{2j_0}+n}q_{T_{2j_0}}(1) \xrightarrow[n\to\infty]{} 0.$$

Noticing $S(p_{T_{2j_0}}(1)) \equiv +1$ under assumption of Condition 1 1) and by (41), the numerator of the second term can be estimated as follows:

$$\begin{aligned} \sum_{k=1}^{n} e_{T_{2j_0}+k}^2(1) &= \sum_{k=1}^{n}\left(m_0\delta u_{T_{2j_0}+k}(0) + \xi_{T_{2j_0}+k}(1)\right)^2 \\ &= \sum_{k=1}^{n}\left(m_0\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j}m_0)\delta u_{T_{2j_0}}(0) + \xi_{T_{2j_0}+k}(1)\right. \\ &\quad \left. + m_0\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 + a_{T_{2j_0}+i}m_0)a_{T_{2j_0}+j}\xi_{T_{2j_0}+j}(1)\right)^2 \\ &= \sum_{k=1}^{n} m_0^2\delta u_{T_{2j_0}}^2(0)\left(\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j}m_0)\right)^2(1 + o(1)) \\ &\quad + \sum_{k=1}^{n}\left(m_0\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 + a_{T_{2j_0}+i}m_0)\cdot a_{T_{2j_0}+j}\xi_{T_{2j_0}+j}(1)\right. \\ &\quad \left. + \xi_{T_{2j_0}+k}(1)\right)^2. \end{aligned}$$

By noticing that

$$\begin{aligned} &\frac{1}{T_{2j_0}+n}\sum_{k=1}^{n}\left(\prod_{j=0}^{k-1}(1 + a_{T_{2j_0}+j}m_0)\right)^2 m_0^2\delta u_{T_{2j_0}}^2(0) \\ &= \frac{1}{T_{2j_0}+n}\sum_{k=1}^{n}\left(O\left(\left(\frac{T_{2j_0}+k}{T_{2j_0}}\right)^{m_0}\right)\right)^2 m_0^2\delta u_{T_{2j_0}}^2(0) \\ &= O\left((T_{2j_0}+n)^{2m_0}\right) \xrightarrow[n\to\infty]{} \infty \end{aligned}$$

we find that the second term on the right side of (42) tends to infinity as $n \to \infty$. Therefore Condition 1 1) is fulfilled.

Now we come to Condition 1 2). Denote $T_0 = T_{2j_0+1}$ and suppose $T_0 < \infty$. In order to show the convergence of $p_{T_0+n}(1)$, it suffices to prove $q_{T_0+n}(1)$ is uniformly bounded w.r.t. n under assumption of Condition 1 2). Similar to (42), we have

$$\begin{aligned} q_{T_0+n}(1) &= \frac{T_0}{T_0+n}q_{T_0}(1) + \frac{1}{T_0+n}\sum_{k=1}^{n} e_{T_0+k}^2(1) \\ &\le q_{T_0}(1) + \frac{1}{T_0+n}\sum_{k=1}^{n} e_{T_0+k}^2(1). \end{aligned} \tag{43}$$

By involving that $S(p_{T_0+j}(1)) \equiv -1, \forall j > 0$

$$\begin{aligned} \textstyle\sum_{k=1}^{n} e_{T_0+k}^2(1) &= \textstyle\sum_{k=1}^{n}\Big(m_0\prod_{j=0}^{k-1}(1 - a_{T_0+j}m_0)\delta u_{T_0}(0) \\ &\textstyle -m_0\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1 - a_{T_0+i}m_0)\cdot a_{T_0+j}\xi_{T_0+j}(1) + \xi_{T_0+k}(1)\Big)^2 \\ &\le 3(\Omega_1 + \Omega_2 + \Omega_3) \end{aligned}$$

where

$$\begin{aligned} \Omega_1 &\triangleq \sum_{k=1}^{n}\left(m_0\prod_{j=0}^{k-1}(1 - a_{T_0+j}m_0)\delta u_{T_0}(0)\right)^2 \\ &= \sum_{k=1}^{n} O\left(\frac{1}{(T_0+k)^{2m_0}}\right) = o(n) \end{aligned}$$

$$\Omega_2 \triangleq \sum_{k=1}^{n}\Bigg(m_0\sum_{j=0}^{k-1}\underbrace{\prod_{i=j}^{k-2}(1-a_{T_0+i}m_0)\cdot a_{T_0+j}}_{\triangleq M_j}\xi_{T_0+j}(1)\Bigg)^2$$

$$=\sum_{k=1}^{n}\left(O\left(\left(\sum_{j=0}^{k-1}M_j^2\right)^{\frac{1}{2}+\eta}\right)\right)^2$$

(By Theorem 1.1)

$$=\sum_{k=1}^{n}\left(O\left(\left(\sum_{j=0}^{k-1}a_{T_0+j}^2\right)^{\frac{1}{2}+\eta}\right)\right)^2$$

$$=\sum_{k=1}^{n}O(1)=O(n)$$

$$\Omega_3 \triangleq \sum_{k=1}^{n}\left(\xi_{T_0+k}(1)\right)^2=O(n).$$

(By the law of large numbers, $(1/n)\sum_{k=1}^{n}\left(\xi_{T_0+k}(1)\right)^2 \xrightarrow[n\to\infty]{} R_\xi$).

Therefore, the uniform boundedness of $q_{T_0+n}(1)$ holds w.r.t. n and moreover $p_{T_0+n}(1)$ converges. For the uniform boundedness of $p_{T_0,\infty}$ w.r.t. T_0, it suffices to prove the terms containing δu_{T_0} are uniformly bounded w.r.t. T_0. Similar to (25), there exists a constant α_3 independent of T_0 such that

$$\overline{\lim_{n\to\infty}}\frac{1}{T_0+n}\Omega_1 \le \frac{\alpha_3}{T_0^{2m_0}}\delta u_{T_0}^2(0). \tag{44}$$

By (41)

$$\delta u_{T_0}^2(0)=\Bigg(\prod_{j=1}^{T_0-1}(1-a_kS(p_k(1))m_0)\delta u_1(0)$$
$$-\sum_{j=1}^{T_0-1}\prod_{i=j}^{T_0-2}(1-a_iS(p_i(1))m_0)\cdot a_jS(p_j(1))\xi_j(1)\Bigg)^2$$
$$\le 2\left(\prod_{j=1}^{T_0-1}(1-a_kS(p_k(1))m_0)\delta u_1(0)\right)^2$$
$$+2\left(\sum_{j=1}^{T_0-1}\prod_{i=j}^{T_0-2}(1-a_iS(p_i(1))m_0)\cdot a_jS(p_j(1))\xi_j(1)\right)^2$$
$$\triangleq 2\Omega_4+2\Omega_5. \tag{45}$$

Similar to (26), there is a constant κ_3 such that

$$\Omega_4 \le \kappa_3 T_0^{2m_0}. \tag{46}$$

On the other hand

$$\Omega_5 = O\left(\left(\sum_{j=1}^{T_0-1}\left(\prod_{i=j}^{T_0-2}(1+a_im_0)\cdot a_j\right)^2\right)^{1+2\eta}\right)$$
$$=O\left(\left(\sum_{j=1}^{T_0-1}O\left(\left(\frac{T_0}{j}\right)^{2m_0}\right)\times\frac{1}{j^2}\right)^{1+2\eta}\right)$$
$$=O\left(\frac{1}{T_0^{1+2\eta}}\right). \tag{47}$$

Combining (44)–(47) leads to the uniform boundedness of $\overline{\lim}_{n\to\infty}(1/T_0+n)\Omega_1$ w.r.t. T_0. Thus all items of Condition 1 are satisfied and therefore the first item of Theorem 4 holds for $t=0$ by Theorem 1.

Suppose the DSF $S(\cdot)$ stops switching from the k_0th cycle, i.e., $S(\cdot)\equiv +1, \forall k\ge k_0$. By (13) we have $\forall k \ge k_0$

$$u_{k+1}(0)=u_k(0)+a_km_0(u_d(0)-u_k(0))+a_k\xi_k(1). \tag{48}$$

It is clear that (48) is the stochastic approximation algorithm [27], where $f(x)=m_0(u_d(0)-x)$ is the regression function. It is obvious that the second item of Theorem 4 holds for $t=0$ provided that Theorem 1.2 holds. Thus we have to verify all the assumptions for this theorem.

The regression function is $f(x)=m_0(u_d(0)-x), m_0>0$, thus H3 obviously holds. Then by Remark 4 following Theorem 1.2, the existence of $V(x)$ is obvious; that is, H1 is satisfied. Since $\{\xi_k(1)\}$ are mutually independent zero-mean random variables with finite second moments of k, then for any $\delta<(1/2)$, we have $\sum_{k=1}^{\infty}a_k^{1-\delta}\xi_k(1)<\infty$, which implies the validation of H2.

Therefore the second item of Theorem 4 holds for $t=0$ and, moreover, the last item holds by Theorem 3.

Inductive Step: Showing that the items hold for time t provided that conclusions hold for time $0,1,\ldots,t-1$.

Similar to the derivation of Theorem 2, we have

$$\begin{aligned}u_{k+1}(t)&=u_k(t)+a_kS(p_k(t+1))e_k(t+1)\\&=u_k(t)+a_kS_k(\cdot)[C(t+1)\delta x_k(t+1)-v_k(t+1)]\\&=u_k(t)+a_kS_k(\cdot)[m_t\delta u_k(t)+\rho_k(t)-\zeta_k(t+1)]\end{aligned}$$

where $\rho_k(t)\triangleq C(t+1)\sum_{i=1}^{t}\Phi_{ti}B(i-1)\delta u_k(i-1)$, $\zeta_k(t+1)\triangleq C(t+1)\sum_{i=0}^{t+1}\Phi_{ti}w_k(i)+v_k(t+1)$. Here, $m_t\triangleq C(t+1)B(t)$ denotes the control direction. Without loss of generality, suppose $m_t>0$, thus the right and wrong control directions are $S=+1$ and $S=-1$, respectively.

Then we have

$$\delta u_{k+n}(t)=\prod_{j=0}^{n-1}(1-a_{k+j}S_{k+j}(\cdot)m_t)\delta u_k(t)$$
$$+\sum_{j=0}^{n-1}\left(\prod_{i=j}^{n-2}(1-a_{k+i}S_{k+i}(\cdot)m_t)\right)a_{k+j}S_{k+j}(\cdot)\rho_{k+j}(t)$$
$$-\sum_{j=0}^{n-1}\left(\prod_{i=j}^{n-2}(1-a_{k+i}S_{k+i}(\cdot)m_t)\right)a_{k+j}S_{k+j}(\cdot)\zeta_{k+j}(t+1).$$

By the induction hypothesis, $\rho_k(t)=o(k^{-\delta})$, $\delta>0$. $\{\zeta_k(t)\}$ are mutually independent zero-mean random variables with finite second moments, and the covariance matrix is $R_{t+1}^{\zeta}=C(t+1)\sum_{i=0}^{t+1}\Phi_{ti}R_i^{w}\Phi_{ti}^{T}C^{T}(t+1)+R_{t+1}^{v}$.

Consider the validation of Condition 1 1) first. Assume $T_{2j_0}<\infty$ and the DSF stops switching directions from the T_{2j_0}th cycle, then it suffices to show $q_{T_{2j_0}+n}(t+1)\xrightarrow[n\to\infty]{}\infty$.

Notice that

$$q_{T_{2j_0}+n}(t+1) = \frac{1}{T_{2j_0}+n}\sum_{k=1}^{T_{2j_0}+n} e_k^2(t+1)$$
$$= \frac{T_{2j_0}}{T_{2j_0}+n} q_{T_{2j_0}}(t+1) + \frac{1}{T_{2j_0}+n}\sum_{k=1}^{n} e_{T_{2j_0}+k}^2(t+1)$$

where the first term $(T_{2j_0}/T_{2j_0}+n)q_{T_{2j_0}}(t+1) \xrightarrow[n\to\infty]{} 0$, while the numerator of the other term on the right side of last equation can be estimated as follows:

$$\sum_{k=1}^{n} e_{T_{2j_0}+k}^2(t+1)$$
$$= \sum_{k=1}^{n}\left[m_t \delta u_{T_{2j_0}+k}(t) + \rho_{T_{2j_0}+k}(t) - \zeta_{T_{2j_0}+k}(t+1)\right]^2$$
$$= \sum_{k=1}^{n}\left[m_t \delta u_{T_{2j_0}+k}(t) + \rho_{T_{2j_0}+k}(t)\right]^2(1+o(1))$$
$$+ \sum_{k=1}^{n}\left[\zeta_{T_{2j_0}+k}(t+1)\right]^2$$
$$\geq \sum_{k=1}^{n}\left[m_t \delta u_{T_{2j_0}+k}(t) + \rho_{T_{2j_0}+k}(t)\right]^2(1+o(1))$$
$$\geq \sum_{k=1}^{n}(m_t \delta u_{T_{2j_0}+k}(t))^2(1+o(1))$$
$$+2\sum_{k=1}^{n} m_t \delta u_{T_{2j_0}+k}(t)\rho_{T_{2j_0}+k}(t)(1+o(1))$$

where the first term on the right side plays a leading role obviously. For clarity of statements, let us omit the coefficients and take the following in account: For sufficiently large n

$$\sum_{k=1}^{n}\left(\delta u_{T_{2j_0}+k}(t)\right)^2$$
$$= \sum_{k=1}^{n}\left(\prod_{j=0}^{k-1}(1+a_{T_{2j_0}+j}m_t)\delta u_{T_{2j_0}}(t)\right.$$
$$+\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1+a_{T_{2j_0}+i}m_t)a_{T_{2j_0}+j}\rho_{T_{2j_0}+j}(t)$$
$$\left.+\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1+a_{T_{2j_0}+i}m_t)a_{T_{2j_0}+j}\zeta_{T_{2j_0}+j}(t+1)\right)^2$$
$$\geq \sum_{k=1}^{n}\left(\prod_{j=0}^{k-1}(1+a_{T_{2j_0}+j}m_t)\delta u_{T_{2j_0}}(t)\right.$$
$$\left.+\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1+a_{T_{2j_0}+i}m_t)a_{T_{2j_0}+j}\rho_{T_{2j_0}+j}(t)\right)^2(1+o(1))$$
$$\geq (Q_5+Q_6)(1+o(1))$$

where

$$Q_5 = \sum_{k=1}^{n}\left(\prod_{j=0}^{k-1}(1+a_{T_{2j_0}+j}m_t)\delta u_{T_{2j_0}}(t)\right)^2$$
$$Q_6 = 2\sum_{k=1}^{n}\left(\prod_{j=0}^{k-1}(1+a_{T_{2j_0}+j}m_t)\delta u_{T_{2j_0}}(t)\right)$$
$$\times\left(\sum_{j=0}^{k-1}\prod_{i=j}^{k-2}(1+a_{T_{2j_0}+i}m_t)a_{T_{2j_0}+j}\rho_{T_{2j_0}+j}(t)\right).$$

By (32) and (33), it is easy to know Q_5 plays a leading role in Q_5+Q_6. Since

$$\frac{1}{T_{2j_0}+n}Q_5 = \frac{1}{T_{2j_0}+n}\sum_{k=1}^{n}\left(O\left(\left(\frac{T_{2j_0}+k}{T_{2j_0}}\right)^{m_t}\right)\right)^2$$
$$= O\left((T_{2j_0}+n)^{2m_t}\right) \xrightarrow[n\to\infty]{} \infty$$

it is obvious that

$$\frac{1}{T_{2j_0}+n}\sum_{k=1}^{n}\left(\delta u_{T_{2j_0}+k}(t)\right)^2 \xrightarrow[n\to\infty]{} \infty$$

which implies

$$\frac{1}{T_{2j_0}+n}\sum_{k=1}^{n} e_{T_{2j_0}+k}^2(t+1) \xrightarrow[n\to\infty]{} \infty.$$

That is, $q_{T_{2j_0}+n} \xrightarrow[n\to\infty]{} \infty$. Condition 1 1) is verified.

Now it comes to verifying Condition 1 2). Denote $T_0 = T_{2j_0+1}$ for simplicity of notations. Suppose $T_0 < \infty$ and the DSF stops switching from the T_0th cycle whence $S(p_k(t+1)) \equiv 1, \forall k \geq T_0$. The aim is to show the uniform boundedness of $q_{T_0+n}(t+1)$ w.r.t. n by (14)(15), which further implies that $p_{T_0+n}(t+1)$ converges. Similar to (43) we have

$$q_{T_0+n}(t+1) = \frac{1}{T_0+n}\sum_{k=1}^{T_0+n} e_k^2(t+1)$$
$$\leq q_{T_0}(t+1) + \frac{1}{T_0+n}\sum_{k=1}^{n} e_{T_0+k}^2(t+1)$$

where

$$\frac{1}{T_0+n}\sum_{k=1}^{n} e_{T_0+k}^2(t+1)$$
$$= \frac{1}{T_0+n}\sum_{k=1}^{n}\left[m_t\delta u_{T_0+k}(t) + \rho_{T_0+k}(t) - \zeta_{T_0+k}(t+1)\right]^2$$
$$\leq \frac{3}{T_0+n}\sum_{k=1}^{n}\Big[(m_t\delta u_{T_0+k}(t))^2$$
$$+(\rho_{T_0+k}(t))^2 + (\zeta_{T_0+k}(t+1))^2\Big].$$

First note that

$$\frac{1}{T_0+n}\sum_{k=1}^{n}(m_t\delta u_{T_0+k}(t))^2$$

$$
\begin{aligned}
&= \frac{1}{T_0+n}\sum_{k=1}^{n} m_t^2 \Bigg[\prod_{j=0}^{k-1}(1-a_{T_0+j}m_t)\delta u_{T_0}(t) \\
&\quad + \sum_{j=0}^{k-1}\left(\prod_{i=j}^{k-2}(1-a_{T_0+i}m_t)\right) a_{T_0+j}\rho_{T_0+j}(t) \\
&\quad - \sum_{j=0}^{k-1}\left(\prod_{i=j}^{k-2}(1-a_{T_0+i}m_t)\right) a_{T_0+j}\zeta_{T_0+j}(t+1)\Bigg]^2 \\
&\le \frac{3m_t^2}{T_0+n}(\Delta_3+\Delta_4+\Delta_5) \qquad (49)
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta_3 &= \sum_{k=1}^{n}\left(\prod_{j=0}^{k-1}(1-a_{T_0+j}m_t)\delta u_{T_0}(t)\right)^2 \\
&= \sum_{k=1}^{n} O\left(\frac{1}{(T_0+k)^{2m_t}}\right) = o(n) \qquad (50)
\end{aligned}
$$

$$
\begin{aligned}
\Delta_4 &= \sum_{k=1}^{n}\left(\sum_{j=0}^{k-1}\left(\prod_{i=j}^{k-2}(1-a_{T_0+i}m_t)\right)a_{T_0+j}\rho_{T_0+j}(t)\right)^2 \\
&= \sum_{k=1}^{n}\left(\sum_{j=0}^{k-1} O\left(\left(\frac{T_0+j}{T_0+k-1}\right)^{m_t}\right)\frac{1}{T_0+j}\,o\left(\frac{1}{(T_0+j)^{\delta}}\right)\right)^2 \\
&= \sum_{k=1}^{n} O\left(\frac{1}{(T_0+k-1)^{2\delta}}\right) = o(n) \qquad (51)
\end{aligned}
$$

$$
\begin{aligned}
\Delta_5 &= \sum_{k=1}^{n}\left(\sum_{j=0}^{k-1}\left(\prod_{i=j}^{k-2}(1-a_{T_0+i}m_t)\right)a_{T_0+j}\zeta_{T_0+j}(t+1)\right)^2 \\
&= \sum_{k=1}^{n}\left(O\left(\left(\sum_{j=0}^{k-1}\left(\prod_{i=j}^{k-2}(1-a_{T_0+i}m_t)a_{T_0+j}\right)^2\right)^{\frac{1}{2}+\eta}\right)\right)^2 \\
&= \sum_{k=1}^{n}\left(O\left(\left(\sum_{j=0}^{k-1}\left(a_{T_0+j}\right)^2\right)^{\frac{1}{2}+\eta}\right)\right)^2 \\
&= \sum_{k=1}^{n} O(1) = O(n). \qquad (52)
\end{aligned}
$$

From (49)–(52)

$$\frac{1}{T_0+n}\sum_{k=1}^{n}(m_t\delta u_{T_0+k}(t))^2 = O(1).$$

Moreover

$$
\begin{aligned}
&\frac{1}{T_0+n}\sum_{k=1}^{n}(\rho_{T_0+k}(t))^2 = \frac{1}{T_0+n}\sum_{k=1}^{n} o\left(\frac{1}{(T_0+k)^{\delta}}\right) = o(1) \\
&\frac{1}{T_0+n}\sum_{k=1}^{n}(\zeta_{T_0+k}(t+1))^2 = R_{t+1}^{\zeta} < \infty.
\end{aligned}
$$

Therefore $q_{T_0+n}(t+1)$ is uniformly bounded w.r.t. n, which leads to the convergence of $p_{T_0+n}(t+1)$ as $n \to \infty$. On the other hand, completely similar to (44)–(47), and (38), it is easy to deduce that Δ_3 is uniformly bounded w.r.t. T_0 and, consequently, $p_{T_0,\infty}$ is uniformly bounded w.r.t. T_0. Condition 1 1) holds. The nondecreasing of $p_k(t+1)$ is clearly satisfied. By Theorem 1, the first item of Theorem 4 holds.

Suppose that $S(p_k(t+1))$ stops switching from the k_0th cycle, i.e., $S \equiv +1, \forall k \ge k_0$. Then $\forall k \ge k_0$

$$u_{k+1}(t) = u_k(t) + a_k[m_t\delta u_k(t) + \rho_k(t) - \zeta_k(t+1)]$$

where $f(x) = m_t(u_d(t) - x), m_t > 0$ is the regression function. Both $\{\rho_k(t)\}$ and $\{\zeta_k(t+1)\}$ are noises. The former satisfies $\rho_k(t) = o(k^{-\delta})$ by the induction hypothesis, and the latter are mutually independent zero-mean random variables with finite second moments of k. Similar to the derivation of the case $t = 0$, it is easy to see that all assumptions for Theorem 1.2 are fulfilled. Therefore the second item of Theorem 4 holds for time t. Moreover, by Theorem 3 the last item is true. By the mathematical induction principle, the proof is completed. ■

References

[1] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot. Syst.*, vol. 1, no. 2, pp. 123–140, 1984.

[2] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Syst. Mag.*, vol. 26, no. 3, pp. 96–114, Jun. 2006.

[3] H.-S. Ahn, Y. Q. Chen, and K. L. Moore, "Iterative learning control: Brief survey and categorization," *IEEE Trans. Syst., Man, Cybern., Part C: Appl. Rev.*, vol. 37, no. 6, pp. 1099–1121, Nov. 2007.

[4] S. N. Huang, 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, Jul. 2002.

[5] K. L. Moore, "Iterative learning control for deterministic systems," in *Advances in Industrial Control*, vol. 22. London, U.K.: Springer-Verlag, 1993.

[6] Y. Fang 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, May 2003.

[7] W. Paszke, R. Merry, and R. de Molengraft, "Iterative learning control by 2-D system theory applied to a motion system," in *Proc. Amer. Control Conf.*, Jul. 2007, pp. 5484–5489.

[8] J.-X. Xu 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, Nov. 2002.

[9] J.-X. Xu and J. Xu, "On iterative learning from different tracking tasks in the presence of time-varying uncertainties," *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.*, vol. 34, no. 1, pp. 589–597, Feb. 2004.

[10] H.-S. Ahn, K. L. Moore, and Y. Q. Chen, *Iterative Learning Control: Robustness and Monotonic Convergence for Interval Systems*. London, U.K.: Springer-Verlag, 2007.

[11] S. S. Saab, "A discrete-time stochastic learning control algorithm," *IEEE Trans. Autom. Control*, vol. 46, no. 6, pp. 877–887, Jun. 2001.

[12] S. S. Saab, "On a discrete-time stochastic learning control algorithm," *IEEE Trans. Autom. Control*, vol. 46, no. 8, pp. 1333–1336, Aug. 2001.

[13] D. Shen and H. F. Chen, "Iterative learning control for a class of nonlinear systems," *J. Syst. Sci. Math. Sci.*, vol. 28, no. 9, pp. 1053–1064, 2008.

[14] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat.*, vol. 22, no. 3, pp. 400–407, Sep. 1957.

[15] R. D. Nussbaum, "Some remarks on a conjecture in parameter adaptive control," *Syst. Control Lett.*, vol. 3, no. 5, pp. 243–246, Nov. 1983.

[16] H. Chen and P. Jiang, "Adaptive iterative learning control for nonlinear systems with unknown control gain," *ASME J. Dyn. Syst., Meas. Control*, vol. 126, no. 4, pp. 916–920, Dec. 2004.

[17] J.-X. Xu and R. Yan, "Iterative learning control design without a priori knowledge of the control direction," *Automatica*, vol. 40, no. 10, pp. 1803–1809, Oct. 2004.

[18] H. F. Chen, "Almost sure convergence of iterative learning control for stochastic systems," *Sci. China (Ser. F)*, vol. 46, no. 1, pp. 69–79, 2003.
[19] H. F. Chen 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, Apr. 2004.
[20] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, vol. 23, no. 3, pp. 462–466, 1952.
[21] A. Ilchmann, "High-gain adaptive stabilization of multivariable linear systems-revisited," *Syst. Control Lett.*, vol. 18, no. 5, pp. 355–364, May 1992.
[22] P. Jiang, H. Chen, and C. A. Bamforth, "A universal iterative learning stabilizer for a class of MIMO systems," *Automatica*, vol. 42, no. 6, pp. 973–981, Jun. 2006.
[23] M. X. Sun and D. W. Wang, "Initial shift issues on discrete-time iterative learning control with system relative degree," *IEEE Trans. Autom. Control*, vol. 48, no. 1, pp. 144–148, Jan. 2003.
[24] J.-X. Xu and R. Yan, "On initial conditions in iterative learning control," *IEEE Trans. Autom. Control*, vol. 50, no. 9, pp. 1349–1354, Sep. 2005.
[25] K.-H. Park, "An average operator-based PD-type iterative learning control for variable initial state error," *IEEE Trans. Autom. Control*, vol. 50, no. 6, pp. 865–869, Jun. 2005.
[26] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*. Boston, MA: Birkhäuser, 1991.
[27] H. F. Chen, *Stochastic Approximation and Its Applications*. Dordrecht, The Netherlands: Kluwer, 2002.
[28] M. Benaim, "A dynamical system approach to stochastic approximations," *SIAM J. Control Optim.*, vol. 34, no. 2, pp. 437–472, 1996.
[29] K. Knopp, *Theory and Application of Infinite Series*. London, U.K.: Blackie, 1928.

Dong Shen (M'10) received the B.S. degree in mathematics from Shandong University, Jinan, China, in 2005. He received the Ph.D. degree in mathematics from the Academy of Mathematics and System Science, Chinese Academy of Sciences (CAS), Beijing, China, in 2010.

He is currently a Post-Doctoral Fellow with the Institute of Automation, CAS. His current research interests include iterative learning controls, stochastic control and optimization.

Zhongsheng Hou received the Bachelors and Masters degrees from the Jilin University of Technology, Changchun, China, in 1983 and 1988, respectively, and the Ph.D. degree from Northeastern University, Shenyang, China, in 1994.

He was a Post-Doctoral Fellow with the Harbin Institute of Technology, Harbin, China, from 1995 to 1997, and a Visiting Scholar with Yale University, New Haven, CT, from 2002 to 2003. In 1997, he joined Beijing Jiaotong University, Beijing, China, and is currently a Full Professor and the Founding Director of the Advanced Control Systems Laboratory and the Head of the Department of Automatic Control in the School of Electronic and Information Engineering. He is the author of two monographs: *Nonparametric Model and Its Adaptive Control Theory* (Science Press of China, 1999, in Chinese) and the coming *Model Free Adaptive Control Theory and Applications* (2012). He has over 80 peer-reviewed journal papers published and 100 papers in prestigious conference proceedings. His current research interests include the fields of data-driven control, model free adaptive control, learning control, and intelligent transportation systems.

Dr. Hou has served as a Committee Member for over 40 international conferences and Chinese conferences and as an Associate Editor and a Guest Editor for a few international journals.