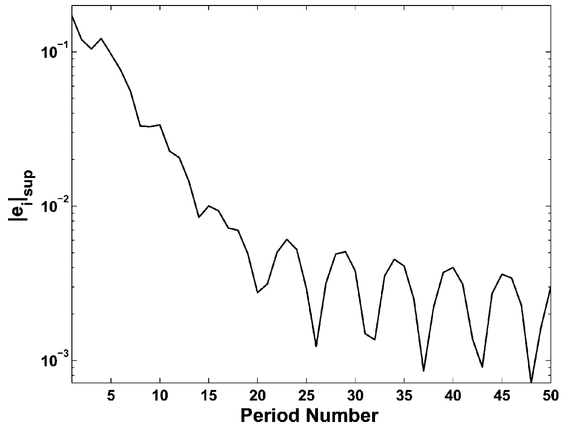

Fig. 3. Error convergence using periodic adaptation (constant b).

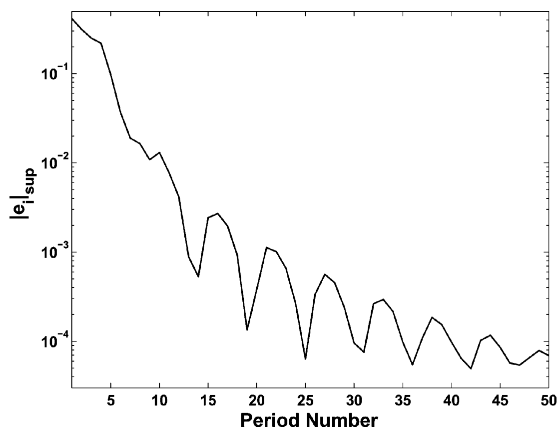

Fig. 4. Error convergence using hybrid adaptation.

First, we try a typical adaptive control using differential updating law, like the one shown in (5). Fig. 1 shows the maximum tracking error over each period. By virtue of the rapid time-varying nature, the tracking error does not converge.

Then we apply the new adaptive method. Fig. 2 shows the maximum tracking error over each period. We can clearly see the effectiveness, as the tracking error has been reduced to less than 4% after 50 periods.

Next, let $b = 3$ be an unknown constant. Still using the same periodic adaptation law, the result is shown in Fig. 3. The result is more or less the same as the preceding case.

Finally, assume that we know *a priori* that b is an unknown constant, the hybrid adaptation law is adopted and the result is shown in Fig. 4. The performance improvement is immediately obvious.

V. Conclusion

To recap, in this note we proposed a new adaptive control approach characterized by periodic parameter adaptation, which complements the existing adaptive control characterized by instantaneous adaptation. By virtue of the periodic adaptation, the new approach is applicable to periodic parameters which can be rapidly time-varying. The only prior knowledge needed in the periodic adaptation is the periodicity. A hybrid differential-periodic adaptation scheme is also proposed when more of the parameter knowledge is available. The validity of the new approach is confirmed through theoretical analysis and numerical simulations.

References

[1] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. Upper Saddle River, NJ: Prentice-Hall, 1989, vol. 3.
[2] S. S. Sastry and M. Boston, *Adaptive Control: Stability, Convergence and Robustness*. Upper Saddle River, NJ: Prentice-Hall, 1989.
[3] K. P. V. Kokotovic, *Foundations of Adaptive Control*. New York: Springer-Verlag, 1991.
[4] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Upper Saddle River, NJ: Prentice-Hall, 1996.
[5] G. Feng and R. Lozano, *Adaptive Control Systems*. Oxford, U.K.: Newnes, 1999.
[6] R. H. Middleton and G. G. C. Goodwin, "Adaptive control of time-varying linear systems," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 150–155, Feb. 1988.
[7] A. M. Annaswamy and K. S. Narendra, "Adaptive control of simple time-varying systems," presented at the IEEE Conf. Decision Control, Tampa, FL, Dec. 1989.
[8] F. M. M. Giri, L. Dugard, and J. M. Dion, "Pole placement direct adaptive control for time-varying ill-modeled plants," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 723–726, June 1990.
[9] K. S. Tsakalis and P. A. Ioannou, *Linear Time-Varying Systems – Control and Adaptation* . Upper Saddle River, NJ: Prentice-Hall, 1993.
[10] R. Marino and P. Tomei, "Adaptive control of linear time-varying systems," *Automatica*, vol. 39, pp. 651–659, 2003.
[11] T. Yoshizawa, *Stability Theory by Liapunov's Second Method*. Tokyo, Japan: Math. Soc. Jpn., 1966.

Output Tracking for Nonlinear Stochastic Systems by Iterative Learning Control

Han-Fu Chen and Hai-Tao Fang

***Abstract*—An iterative learning control (ILC) algorithm, which in essence is a stochastic approximation algorithm, is proposed for output tracking for nonlinear stochastic systems with unknown dynamics and unknown noise statistics. The nonlinear function of the system dynamics is allowed to grow up as fast as a polynomial of any degree, but the system is linear with respect to control. It is proved that the ILC generated by the algorithm a.s. converges to the optimal one at each time $t \in [0, 1, \ldots, N]$ and the output tracking error is asymptotically minimized in the mean square sense as the number of iterates tends to infinity, although the convergence rate is rather slow. The only information used in the algorithm is the noisy observation of the system output and the reference signal $y_d(t)$. When the system state equation is free of noise and the system output is realizable, then the exact state tracking is asymptotically achieved and the tracking error is purely due to the observation noise.**

***Index Terms*—Almost sure (a.s.) convergence, iterative learning control, nonlinear stochastic system, output tracking, stochastic approximation.**

I. Introduction

For control systems, where the same task is performed repeatedly, the system inputs and outputs of previous cycles may be used to improve control performance. This is the motivation to use iterative learning control (ILC), which was first introduced in [2] for robot

Manuscript received June 5, 2002; revised December 6, 2002. Recommended by Associate Editor L. Gerenscer. This work was supported by the National Science Foundation of China and by the Ministry of Science and Technology of China.

The authors are with the Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China (e-mail: hfchen@mail.iss.ac.cn; htfang@mail.iss.ac.cn).

Digital Object Identifier 10.1109/TAC.2004.825613

control. The idea of ILC has naturally attracted a growing attention from researchers in the area of systems and control [8], [9], [11], [14] among others, and people try to apply ILC not only to robots but also to other systems such as, speed control of servomotors [13], extruders [17], etc.

Various important issues such as robustness, stability, reinitialization errors, and convergence among others have been addressed in the aforementioned papers. As concerns performance indexes, not only tracking error but also quadratic criteria are discussed [1], [16]. However, control systems considered in previous papers are deterministic in essence, assuming the boundedness of possibly existing uncertainties, disturbances, and measurement errors. In addition, rather restrictive conditions are imposed to guarantee a satisfactory control performance, for example, the global Lipschitz condition is imposed on the nonlinear dynamics in [8], and the desired trajectory for state is assumed to be given in [14].

There are only a few papers that consider ILC for stochastic systems taking the random nature of systems into account. In [6], the pole assignment problems is solved by learning for single input–single output linear stochastic systems. In [18] and [19], the ILC algorithms leading the tracking error tending to zero in the mean square sense are proposed for linear stochastic systems.

In [5] for the system considered in [18] and [19], a stochastic approximation (SA) based ILC algorithm is proposed, and is proved to be convergent to the optimal control under conditions much weaker than those used in [18] and [19].

In this note, an SA-based ILC algorithm is proposed for nonlinear stochastic systems, where the nonlinear functions are allowed to grow up as fast as polynomials of any degree. The algorithm converges to the optimal control, although the rate is rather slow, and the long run average of the output tracking errors is minimized as the number of iterates increases. For the case where the state equation is free of noise, the exact state tracking is asymptotically achieved and the asymptotic output tracking error is caused purely by the observation noise.

In Section II, the problem is stated, and the optimal control minimizing the output tracking error in the mean square sense for each time $t \in [0, 1, \ldots, N]$ is found. The ILC algorithm is defined in Section III, while its convergence to the optimal one is proved in Section IV, where the exact state tracking is also shown when the state equation is free of noise. A brief conclusion is given in Section V.

II. Optimal Control

Consider the discrete-time nonlinear stochastic system described by the following difference equations:

$$x(t+1,k) = f(t, x(t,k)) + B(t, x(t,k))u(t,k) + w(t+1,k) \tag{1}$$

$$y(t+1,k) = C(t+1)x(t+1,k) + v(t+1,k) \tag{2}$$

where t and k denote time and iteration index, respectively, and $x(t,k) \in \mathbb{R}^p$, $u(t,k) \in \mathbb{R}^r$, $w(t,k) \in \mathbb{R}^p$, and $y(t,k) \in \mathbb{R}^q$, $\forall\, t \in [0, \ldots, N]$ for some positive integer N. Here, $x(t,k)$, $u(t,k)$, and $y(t,k)$ are the system state, input (control), and output, respectively, while $w(t,k)$ and $v(t,k)$ denote the system noise and observation noise, respectively.

Let $\{y_d(t)\}$, $t \in [0, \ldots, N]$, be the desired output trajectory. We want the difference between $y(t,k)$ and $y_d(t)$ in the long run average sense to be minimized as $k \to \infty$, $\forall\, t \in [0, \ldots, N]$.

Let $\{\mathcal{F}_k\}$ be a family of increasing σ-algebras such that $x(t,k)$, $y(t,k)$, $w(t,k)$, and $v(t,k)$ all are $\mathcal{F}_k$-measurable, $\forall\, t \in [0, 1, \ldots, N]$, and $\mathcal{F}_k$ is independent of $\{w(t,l),\ v(t,l),\ l = k+i,\ i = 1, 2, \ldots,\ \forall\, t \in [0, 1, \ldots, N]\}$, where for convenience of writing $w(0,l)$ is defined as $w(0,l) \triangleq x(0,l) - \mathbf{E}x(0,l)$.

The control $u(t,k)$ should iteratively be defined by observations up to but not including the present iterate.

Define the set of admissible controls as follows:

$$U = \left\{ u(t,k) \in \mathcal{F}_{k-1}, \sup_k u(t,k) < \infty \text{ a.s.}, \quad t = 0, 1, \ldots, N-1,\ k = 0, 1, 2, \ldots \right\}. \tag{3}$$

The control objective is to find $\{u^0(t,k), k = 0, 1, \ldots\} \in U$ such that

$$\inf_{\{u(t,j), j=0,1,\ldots\} \in U} J(t+1, \{u(t,j), j = 0, 1, \ldots\}) = J(t+1, \{u^0(t,k), k = 0, 1, \ldots\}) \tag{4}$$

where

$$J(t+1, \{u(t,j), j = 0, 1, \ldots\}) \triangleq \limsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \|y(t+1,k) - y_d(t+1)\|^2. \tag{5}$$

The following conditions will be needed.

A1) For any $t \in [0, \ldots, N]$, $f(t,x)$ and $B(t,x)$ are continuous in x, and there are real numbers l, c, and b such that $\|f(t,x)\| + \|B(t,x)\| \le c\|x\|^l + b$ as $\|x\| \to \infty\ \forall\, t \in [0, \ldots, N]$, where $\|A\| = \sqrt{\sum_{ij} |a_{ij}|^2}$ with a_{ij} being the elements of a matrix A.

A2) $q \ge r$ and for any $x \in \mathbb{R}^p$ and $t \in [0, \ldots, N]$

$$P(t,x) \triangleq B^T(t,x)C^T(t+1)C(t+1)B(t,x) \tag{6}$$

is positive definite.

A3) For any $t \in [0, \ldots, N]$, both sequences of random vectors $\{w(t,k),\ k = 1, 2, \ldots\}$ and $\{v(t,k),\ k = 1, 2, \ldots\}$ are i.i.d. and mutually independent with zero mean and $\mathbf{E}\|w(t,k)\|^m < \infty$, $\mathbf{E}\|v(t,k)\|^{2(1+\gamma)} < \infty$ for any integer $m > 0$ and some $\gamma \in (0, 1]$. The covariance matrices $\mathbf{E}w(t,k)w(t,k)^T \triangleq R_t^w$ and $\mathbf{E}v(t,k)v(t,k)^T \triangleq R_t^v$ are unknown $\forall\, k = 1, 2, \ldots$.

A4) $\{x(0,k), k = 1, 2, \ldots\}$ is a sequence of i.i.d. random vectors, independent of both $\{w(t,k), t \in [1, \ldots, N], k = 1, 2, \ldots\}$ and $\{v(t,k), t \in [0, \ldots, N], k = 1, 2, \ldots\}$ with $\mathbf{E}\|x(0,k)\|^m < \infty$ for any integer $m > 0$. The covariance matrix R_0^x of $\{x(0,k)\}$ is unknown.

If A1), A2), and A4) hold, then the vector $u^0(t)$ is well-defined for $t = 0$:

$$u^0(t) = -[\mathbf{E}P(t, x^0(t,k))]^{-1} \cdot \left\{ \mathbf{E}[B^T(t, x^0(t,k))C^T(t+1)f(t, x^0(t,k))] - \mathbf{E}[B^T(t, x^0(t,k))]C^T(t+1)y_d(t+1) \right\} \tag{7}$$

which is optimal in the sense that

$$\mathbf{E}\|C(t+1)[f(t, x^0(t,k)) + B(t, x^0(t,k))u^0(t)] - y_d(t+1)\| = \min_u \mathbf{E}\|C(t+1)[f(t, x^0(t,k)) + B(t, x^0(t,k))u] - y_d(t+1)\|. \tag{8}$$

Set $x^0(0,k) \equiv x(0,k)$ and define

$$x^0(1,k) = f(0, x^0(0,k)) + B(0, x^0(0,k))u^0(0) + w(1,k). \quad (9)$$

Since both $\{x^0(0,k) = x(0,k)\}$ and $\{w(1,k)\}$ are sequences of i.i.d. random vectors and they are mutually independent, $\{x^0(1,k)\}$ is a sequence of i.i.d. random vectors and $x^0(1,k)$ is independent of $\mathcal{F}_{k-1}$.

Inductively, assume that $\{x^0(t,k)\}$ is a sequence of i.i.d. random vectors and that $x^0(t,k) \in \mathcal{F}_k$ but $x^0(t,k)$ is independent of $\mathcal{F}_{k-1}$. Then, under A1), A2), and A4), the deterministic vector $u^0(t)$ given by (7) is well defined and (8) takes place. Denote

$$\begin{aligned} x^0(t+1,k) &= f(t, x^0(t,k)) + B(t, x^0(t,k))u^0(t) \\ &\quad + w(t+1,k) \quad (10)\\ x(t+1,k) &= f(t, x(t,k)) + B(t, x(t,k))u(t,k) \\ &\quad + w(t+1,k) \quad (11)\\ \varepsilon(t,k) &= x(t,k) - x^0(t,k) \quad (12) \end{aligned}$$

and

$$\delta u(t,k) = u(t,k) - u^0(t) \quad (13)$$

$t = 0, 1, \ldots, N$, $k = 1, 2, \ldots$.

It is clear that $\{x^0(t+1,k)\}$ is a sequence of i.i.d. random vectors, and $x^0(t+1,k) \in \mathcal{F}_k$ but $x^0(t+1,k)$ is independent of $\mathcal{F}_{k-1}$.

By A1), A3), and A4), inductively, it is seen that

$$\mathbf{E}\|x^0(t,k)\|^m < \infty \qquad \forall\, m > 0 \quad \forall\, t \in [0, 1, \ldots, N] \quad \forall\, k = 1, 2, \ldots. \quad (14)$$

Therefore, by induction, from (7) and (8) the vectors $u^0(0)$, $u^0(1)$, $\ldots$, $u^0(N-1)$ are well defined, but they are not available because $f(\cdot,\cdot)$, $B(\cdot,\cdot)$, and $C(\cdot)$ are unknown.

Theorem 1: Assume A1)–A4) hold. Then, $\{u^0(t)\}$ defined by (7) is optimal for the performance index (4). Further, any $\{u(t,k),\ k = 0, 1, 2, \ldots\} \in U$ with $\delta u(t,k) \underset{k\to\infty}{\to} 0$ a.s., $\forall\, t = 0, 1, \ldots, N-1$ is also optimal, where $\delta u(t,k)$ is given by (13).

Instead of the detailed proof, we only outline the key points.

First, it is shown that

$$\mathbf{E}\left(\|\varepsilon(t,k)\|^m | \mathcal{F}_{k-1}\right) \underset{k\to\infty}{\to} 0 \text{ a.s.} \qquad \forall\, m > 0 \quad (15)$$

if $\delta u(s,k) \underset{k\to\infty}{\to} 0$ a.s., $\forall\, s = 0, \ldots, t-1$, $t \in [1, 2, \ldots, N]$. This is done by induction by noticing that for the initial step $\varepsilon(1,k) = B(0, x^0(0,k))\delta u(0,k)$

$$\begin{aligned} &\mathbf{E}(\|\varepsilon(1,k)\|^m \mid \mathcal{F}_{k-1}) \\ &\le \|\delta u(0,k)\|^m \mathbf{E}(\|B(0, x^0(0,k))\|^m) \\ &\to 0 \text{ a.s.} \qquad \forall\, m > 0 \end{aligned}$$

and by A1) $\varepsilon(1,k)$ is continuous with respect to $x^0(0,k)$ and $\delta u(0,k)$.

Second, expressing the tracking error at time $t = 0$

$$\begin{aligned} y(1,k) - y_d(1) &= \phi(1,k) + C(1) \\ &\quad \times B(0, x(0,k))\delta u(0,k) + \varphi(1,k) \quad (16) \end{aligned}$$

where $\varphi(1,k)$ and $\phi(1,k)$ are the values of the following functions evaluated at $t = 0$:

$$\varphi(t+1,k) \triangleq C(t+1)w(t+1,k) + v(t+1,k) \quad (17)$$

$$\begin{aligned} \phi(t+1,k) \triangleq\, &C(t+1)[f(t, x^0(t,k)) \\ &+ B(t, x^0(t,k))u^0(t)] - y_d(t+1) \quad (18) \end{aligned}$$

we find that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|y(1,k) - y_d(1)\|^2 \\ &= tr(C(1)R_1^w C^T(1) + R_1^v) + \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\phi(1,k)\|^2 \\ &+ \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|C(1)B(0, x(0,k))\delta u(0,k)\|^2 \quad (19) \end{aligned}$$

the minimum of which is achieved when $\delta u(0,k) \underset{k\to\infty}{\to} 0$ a.s., i.e., $u(0,k) \underset{k\to\infty}{\to} u^0(0)$ a.s.

Finally, the proof is completed by induction.

By (15), it follows that

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|y(t+1,k) - y_d(t+1)\|^2 \\ &= \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\delta y_u^0(t+1,k)\|^2 \quad (20) \end{aligned}$$

where

$$\begin{aligned} \delta y_u^0(t+1,k) &= \phi(t+1,k) + C(t+1) \\ &\quad \times B(t, x^0(t,k))\delta u(t,k) + \varphi(t+1,k). \quad (21) \end{aligned}$$

From (20) and (21), it is concluded that $u(t,k) \in U$ with $\delta u(t,k) \underset{k\to\infty}{\to} 0$ is optimal. ■

Remark 1: Since $u^0(t)$ is deterministic, the minimum (4) has no change if U is restricted to

$$U^0 = \{u(t,k) \equiv u(t) \text{ being deterministic and independent of } k\}.$$

If $u(t,k) \equiv u(t)$ in (1), then we write

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|y(t+1,k) - y_d(t+1)\| \triangleq J(t+1, u(t))$$

and the problem is reduced to optimizing $J(t+1, u(t))$ with respect to $u(t)$, if all $u(s,k)$, $s = 0, \ldots, t-1$, are optimal, i.e., $u(s,k) = u^0(s) + \delta u(s,k)$ with $\delta u(s,k) \underset{k\to\infty}{\to} 0$ a.s., $s = 0, \ldots, t-1$.

From (20) and (21) it follows that

$$\begin{aligned} J(t+1, u(t)) &= tr(C(t+1)R_{t+1}^w C^T(t+1) + R_{t+1}^v) \\ &\quad + \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \|\phi(t+1,k)\|^2. \quad (22) \end{aligned}$$

III. ILC

We now define the ILC algorithm to generate the control sequence $\{u(t,k),\ k=1,2,\ldots\}$ for each $t \in [0,\ldots,N]$ such that the performance (4) is minimized. The ILC algorithm to be defined in essence is a Kiefer–Wolfowitz (KW) algorithm with expanding truncations and with randomized differences [7], [20]. Applying random directions in the KW algorithm was first introduced in [15] and later was reintroduced and called as SPSA in [20], which has caused a series of subsequent works, e.g., [10] among others. As concerns the method of convergence analysis, in contrast to [20] where the ODE method is used, here we apply the one that is presented in [4] and called the trajectory-subsequence (TS) method.

To generate $u(t,k)$ we will use the tool vector sequences $\{\Delta_k(t),\ k=1,2,\ldots\}$, $t=0,\ldots,N$, where $\Delta_k(t) = (\Delta_k^1(t),\ldots,\Delta_k^r(t))^T$ is an r-dimensional random vector satisfying the following conditions.

1) All components $\Delta_k^i(t)$, $k=1,2,\ldots$, $t=0,1,\ldots,N$, $i=1,\ldots,r$ are mutually independent and identically distributed random variables such that

$$\left|\Delta_k^i(t)\right| < d_1 \quad \left|\frac{1}{\Delta_k^i(t)}\right| < d_2 \quad \mathbf{E}\left\{\frac{1}{\Delta_k^i(t)}\right\} = 0 \tag{23}$$

for any $k=0,1,\ldots,t=0,\ldots,N,i=1,\ldots,r$, where d_1 and d_2 are positive constants.

2) The sequence $\{\Delta_k(t)\}$ is independent of $\{w(t,k)$ and $v(t,k)\ t \in [0,\ldots,N],\ k=0,1,2,\ldots\}$.

Define the r-dimensional vector

$$\bar{\Delta}_k(t) = \left(\frac{1}{\Delta_k^1(t)},\ldots,\frac{1}{\Delta_k^r(t)}\right)^T, \qquad t=0,\ldots,N,\ k=0,1,\ldots. \tag{24}$$

Let $\{a_k\}$, $\{c_k\}$, and $\{M_k\}$ be sequences of positive real numbers satisfying the following conditions: $a_k \underset{k\to\infty}{\longrightarrow} 0$, $\sum_{k=0}^{\infty} a_k = \infty$, $c_k \underset{k\to\infty}{\longrightarrow} 0$, $\sum_{k=0}^{\infty}(a_k/c_k)^{1+\gamma} < \infty$, and $M_{k+1} > M_k$, $\forall\ k = 0,1,2,\ldots$, $M_k \underset{k\to\infty}{\longrightarrow} \infty$, where γ is the one given in A3).

For any $t \in [0,\ldots,N]$ and $k=1,2,\ldots$, we denote the output tracking error by

$$\delta y(t,k) = y(t,k) - y_d(t)$$

where $y(t,k)$ is the observation given by (2).

For any $t \in [0,\ldots,N]$ the initial value $u(t,0)$ is arbitrarily given. The control at an odd number of iterates is defined to equal its value of the last iterate disturbed by a small random vector $c_k\Delta_k(t)$, i.e.,

$$u(t,2k+1) \triangleq u(t,2k) + c_k\Delta_k(t). \tag{25}$$

The control $u(t,2k)$ at an even number of iterates is recursively defined by

$$\begin{aligned}\bar{u}(t,2(k+1)) &= u(t,2k) - a_k\frac{\bar{\Delta}_k(t)}{c_k}\\ &\quad\cdot\left(\|\delta y(t+1,2k+1)\|^2\right.\\ &\quad\left.-\ \|\delta y(t+1,2k)\|^2\right)\end{aligned} \tag{26}$$

$$\begin{aligned}u(t,2(k+1)) &= \bar{u}(t,2(k+1))\\ &\quad\times I_{\{\|\bar{u}(t,2(k+1))\|\le M_{\sigma_k(t)}\}}\end{aligned} \tag{27}$$

and

$$\sigma_k(t) = \sum_{l=1}^{k-1} I_{\{\|\bar{u}(t,2(l+1))\|>M_{\sigma_l(t)}\}},\ \sigma_0(t)=0. \tag{28}$$

The algorithm (26)–(28), together with (25), defines the ILC $u(t,k)$, $k=0,1,2,\ldots$, $t\in[0,\ldots,N]$ to be applied to (1) as system input.

Although the randomized differences are used here like in [7], results obtained in [7] cannot be applied to the present case. This is because the main effort in the convergence analysis in [7] is devoted to dealing with the approximation error caused by replacing the gradient of the objective function with its finite difference, while here the main concern is the observation noise

$$z(t,k) \triangleq \|\delta y(t+1,k)\|^2 - F(t,u(t,k)) \tag{29}$$

where

$$\begin{aligned}F(t,u(t,k)) &= (u^0(t)-u(t,k))^T\\ &\quad\times\mathbf{E}P(t,x^0(t,x))(u^0(t)-u(t,k)).\end{aligned}$$

IV. Convergence of ILC Algorithm

We need only to show that $u(t,k) \underset{k\to\infty}{\longrightarrow} u^0(t)$ a.s. $\forall\ t \in [0,\ldots,N]$, where $u^0(t)$ is given by (7).

Theorem 2: Assume A1)–A4) hold and $\{y_d(t), t=1,\ldots,N\}$ is the desired output. Then, $u(t,k)$ defined by (25)–(28) converges to the optimal control $u^0(t)$ a.s. as $k\to\infty$ for any $t\in[0,\ldots,N]$.

Proof: The proof is completed by three steps.

1) We first transform the algorithm (25)–(28) into a Robbins–Monro algorithm with expanding truncations

$$\begin{aligned}\bar{u}(t,2(k+1)) &= u(t,2k) - a_k\left\{2\mathbf{E}[P(t,x^0(t,k))]\right.\\ &\quad\cdot(u(t,2k)-u^0(t))\\ &\quad\left.+\sum_{i=1}^{3}\xi_i(t,k+1)\right\}\end{aligned} \tag{30}$$

$$u(t,2(k+1)) = \bar{u}(t,2(k+1))I_{\{\|\bar{u}(t,2(k+1))\|\le M_{\sigma_k(t)}\}} \tag{31}$$

$$\sigma_k(t) = \sum_{l=1}^{k-1} I_{\{\|\bar{u}(t,2(l+1))\|>M_{\sigma_l(t)}\}} \tag{32}$$

where $\{\xi_i(t,k), i=1,2,3\}$ are treated as noise terms

$$\begin{aligned}\xi_1(t,k+1) &= 2(\bar{\Delta}_k(t)\Delta_k^T(t)-I)\mathbf{E}[P(t,x^0(t,k)]\\ &\quad\cdot[u(t,2k)-u^0(t)]\end{aligned} \tag{33}$$

$$\xi_2(t,k+1) = c_k\Delta_k^T(t)\mathbf{E}[P(t,x^0(t,k)]\Delta_k(t)\bar{\Delta}_k(t) \tag{34}$$

$$\xi_3(t,k+1) = \frac{\bar{\Delta}_k(t)}{c_k}[z(t,2k+1)-z(t,2k)]. \tag{35}$$

Since $c_k \underset{k\to\infty}{\longrightarrow} 0$ and Δ_k and $\bar{\Delta}_k$ are bounded, it is clear that

$$\xi_2(t,k+1) \to 0 \text{ a.s..}$$

As $\mathcal{F}_k$, introduced in Section II, we take

$$\mathcal{F}_{2k} = \sigma\{w(t,l), v(t,l), l \le 2k \quad \Delta_m(t), m \le k, t = 0,\dots,N\} \tag{36}$$
$$\mathcal{F}_{2k+1} = \sigma\{w(t,l), v(t,l), l \le 2k+1,\ \Delta_m(t), m \le k \quad t = 0,\dots,N\}. \tag{37}$$

It is clear that $u(t,k)$ is $\mathcal{F}_{k-1}$-measurable. Let us denote $\mathcal{F}_{2k+1}$ by $\mathcal{G}_{k+1}$. Then $u(t,2k)$ is $\mathcal{G}_k$-measurable, and

$$\mathbf{E}(\xi_1(t,k+1)|\mathcal{G}_k) = 0. \tag{38}$$

By the martingale convergence theorem (MCT) it follows that for any fixed $D > 0$

$$\sum_{k=0}^{\infty} a_k \xi_1(t,k+1) I_{\{\|u(t,2k)\|<D\}} < \infty. \tag{39}$$

To prove $u(t,k) \underset{k\to\infty}{\to} u^0(t)$ a.s., by [3, Th. 1] or [4, Th. 2.3.1], we need only to verify that for any $D > 0$

$$\lim_{T\to 0} \limsup_{n\to\infty} \frac{1}{T} \left\| \sum_{k=n}^{m(n,T)} a_k(\xi_3(t,k+1)) I_{\{\|u(t,2k\}\|<D\}} \right\| = 0 \text{ a.s.} \tag{40}$$

where $m(n,T) \triangleq \max\{m : \sum_{k=n}^{m} a_k < T\}$.

2) The proof of the theorem is completed by verifying (40) inductively. Let us first prove (40) for $t = 0$.
Setting

$$Q(1,k) = P(0,x(0,k)) - \mathbf{E}P(0,x(0,k)) \tag{41}$$
$$\zeta_k(1) = \|\phi(1,2k+1) + \varphi(1,2k+1)\|^2 - \|\phi(1,2k) + \varphi(1,2k)\|^2 \tag{42}$$
$$h_{2k}(1) = B^T(0,x(0,2k))C^T(1)[\phi(1,2k) + \varphi(1,2k)] + \frac{1}{2}Q(1,2k)\delta u(0,2k) \tag{43}$$
$$h_{2k+1}(1) = B^T(0,x(0,2k+1))C^T(1) \cdot [\phi(1,2k+1) + \varphi(1,2k+1)] + \frac{1}{2}Q(1,2k+1)\delta u(0,2k) \tag{44}$$

we have

$$\begin{aligned} z(0,2k+1) - z(0,2k) &= \zeta_k(1) + 2\delta u^T(0,2k)(h_{2k+1}(1) - h_{2k}(1)) \\ &\quad + c_k^2 \Delta_k^T(0) Q(1,2k+1)\Delta_k(0) + c_k \Delta_k^T(0) \\ &\quad \times (2h_{2k+1}(1) + Q(1,2k+1)\delta u(0,2k)). \end{aligned} \tag{45}$$

It can be shown that $\{\xi_3(0,k+1), \mathcal{G}_{k+1}\}$ is a martingale difference sequence. By A1), the divergence rate of $\|f(0,x)\| + \|B(0,x)\|$ as $\|x\| \to \infty$ is not faster than a polynomial, and hence by A3) and A4) we have

$$\sup_k \mathbf{E}\left(\left\|\xi_3(0,k+1) I_{\{\|u(0,2k\}\|<D\}}\right\|^{1+\gamma} |\mathcal{G}_k \right) < \infty \qquad \forall\, D > 0. \tag{46}$$

Since $\sum_{k=0}^{\infty} \left(a_k^{1-\delta}/c_k\right)^{1+\gamma} < \infty$ implies $\sum_{k=0}^{\infty} (a_k/c_k)^{1+\gamma} < \infty$, by MCT we have

$$\sum_{k=0}^{\infty} a_k(\xi_3(0,k+1)) I_{\{\|u(0,2k\}\|<D\}} < \infty, \text{ a.s.}$$

$\forall\, i = 1,\dots,r$, which verifies (40) for $t = 0$.

3) Inductively, we now assume $u(s,k) - u^0(s) \underset{k\to\infty}{\to} 0$ a.s., $s = 0,1,\dots,t-1$. We proceed to verify (40) for t.

Let

$$\begin{aligned} \delta y_u^0(t+1,k) &\triangleq C(t+1)\left[f\left(t,x^0(t,k)\right) + B(t,x^0(t,k))u(t,k)\right] \\ &\quad + \varphi(t+1,k) - y_d(t+1) \end{aligned} \tag{47}$$
$$\tilde{z}(t+1,2k+1) = \|\delta y_u^0(t+1,2k+1)\|^2 - F(t,u(t,2k) + c_k\Delta_k(t))$$

and

$$\tilde{z}(t+1,2k) = \|\delta y_u^0(t+1,2k)\|^2 - F(t,u(t,2k)).$$

It can be shown that for any $D > 0$ and $i = 1,\dots,r$

$$\sum_{k=0}^{\infty} \frac{a_k\left(\tilde{z}(t+1,2k+1) - \tilde{z}(t+1,2k)\right)}{c_k \Delta_k^i(t)} \times I_{\{\|u(t,2k\}\|<D\}} < \infty, \text{ a.s.} \tag{48}$$

Furthermore, we have

$$\begin{aligned} z(t+1,2k+1) &- z(t+1,2k) \\ &= \tilde{z}(t+1,2k+1) - \tilde{z}(t+1,2k) + \eta_1(t,k) \\ &\quad + 2c_k\eta_2(t,k) + c_k^2\eta_3(t,k) \end{aligned} \tag{49}$$

where

$$\begin{aligned} \eta_1(t,k) &= \|\beta_1(t,2k+1) + \beta_2(t,2k+1)u(t,2k)\|^2 \\ &\quad - \|\beta_1(t,2k) + \beta_2(t,2k)u(t,2k)\|^2 \\ &\quad + 2[\beta_1(t,2k+1) + \beta_2(t,2k+1)u(t,2k)]^T \\ &\quad \times \delta\bar{y}_u^0(t+1,2k+1) \\ &\quad - 2[\beta_1(t,2k) + \beta_2(t,2k)u(t,2k)]^T \\ &\quad \times \delta y_u^0(t+1,2k) \\ \eta_2(t,k) &= 2\Delta_k^T(t)\Big\{\beta_2^T(t,2k+1)[\beta_1(t,2k+1) \\ &\quad + 2\delta\bar{y}_u^0(t+1,2k+1)] \\ &\quad + B^T(t,x^0(t,2k+1))C^T(t+1) \\ &\quad \cdot [\beta_1(t,2k+1) + \beta_2(t,2k+1)u(t,2k)]\Big\} \\ \eta_3(t,k) &= \Delta_k^T(t)\beta_2^T(t,2k+1) \\ &\quad \cdot [\beta_2(t,2k+1) + C(t+1) \\ &\quad \times B(t,x^0(t,2k+1))]\Delta_k(t). \\ \beta_1(t,k) &= C(t+1)[f(t,x(t,k)) - f(t,x^0(t,k))] \\ \beta_2(t,k) &= C(t+1)[B(t,x(t,k)) - B(t,x^0(t,k))] \end{aligned}$$

and

$$\delta \bar{y}_u^0(t+1, 2k+1) = \delta y_u^0(t+1, 2k+1) - c_k C(t+1) B(t, x^0(t, 2k+1)) \Delta_k(t).$$

By (48), for verifying (40) for t, it suffices to show that

$$\sum_{k=1}^{\infty} \frac{a_k}{c_k \Delta_k^i(t)} \eta_1(t,k) I_{\{\|u(t,2k)\|<D\}} < \infty \text{ a.s. } \forall\, D > 0 \qquad (50)$$

which is not difficult to be shown by using MCT.

Thus, we have verified (40) for t and in the meantime have completed the induction.

We say that $\{y_d(t)\}$ is a realizable trajectory, if there are $\{u_d(t)\}$ and an initial value $x_d(0)$ such that

$$x_d(t+1) = f(t, x_d(t)) + B(t, x_d(t)) u_d(t) \qquad (51)$$

$$y_d(t+1) = C(t+1) x_d(t+1). \qquad (52)$$

Theorem 3: If for (1) and (2) $w(i,k) \equiv 0$, $i \in [0, N]$, $\forall\, k = 0, 1, 2, \ldots$, $\{y_d(t)\}$ is a realizable trajectory, and A1)-A4) are satisfied, then the exact state tracking is asymptotically achieved

$$x(t,k) \underset{k\to\infty}{\rightarrow} x_d(t), \text{ a.s. } \forall\, t \in [0, N]$$

and the output tracking error is purely due to the observation noise

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|\delta y(t,k)\|^2 = tr R_{t+1}^v. \qquad (53)$$

Proof: By A3), we derive the following expression:

$$u_d(t) = [P(t, x_d(t))]^{-1} B^T(t, x_d(t)) C^T(t+1) \cdot (y_d(t+1) - C(t+1) f(t, x_d(t)))$$

$t = 0, 1, \ldots, N$. If $w(0,k) \equiv 0$, then $x(0,k) \equiv x_d(0)$. By (7), we have $u^0(0) = u_d(0)$, and hence

$$x(1,k) - x^0(1,k) \underset{k\to\infty}{\rightarrow} 0, \text{ a.s.}$$

and $x^0(1,k) \equiv x_d(1)$.

Then, the proof is completed by induction.

V. Conclusion

For nonlinear stochastic systems we have proposed a stochastic approximation based ILC algorithm and have shown its convergence to the optimal control minimizing the output tracking error in the mean square sense. The conditions used are quite general: The nonlinear dynamics is allowed to grow up as fast as a polynomial of any degree and no noise statistics are required to be known. The only information used in the algorithm is the noisy observation of the system output. To the authors' knowledge this note provides the first results on a.s. convergence of ILC for nonlinear stochastic systems. The limitation of the results presented in the note consists in that: 1) as shown by simulation the convergence speed is rather slow; 2) the system is required to be linear with respect to control; and 3) conditions on the noise are rather restrictive. This belongs to further research.

References

[1] N. Amann, D. H. Owens, and E. Rogers, "Iterative learning control for discrete-time system with exponential rate of convergence," *IEE Proc. Control Theory Applications*, vol. 143, no. 2, pp. 217–224, 1996.

[2] S. Arimoto, S. Karamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot. Syst.*, vol. 1, pp. 123–140, 1984.

[3] H. F. Chen, "Stochastic approximation with state-dependent noise," *Sci. China (Series E)*, vol. 43, pp. 531–541, 2000.

[4] ——, *Stochastic Approximation and Its Applications*. Dordrecht, The Netherlands: Kluwer, 2002.

[5] ——, "Almost surely convergence of iterative learning control for stochastic systems," *Sci. China (Series F)*, vol. 46, no. 1, pp. 1–13, 2003.

[6] H. F. Chen and X. R. Cao, "Pole assignment of stochastic systems with unknown coefficients," *Sci. China (Series E)*, vol. 43, pp. 313–323, 2000.

[7] H. F. Chen, T. E. Duncan, and B. Pasik-Duncan, "A Kiefer–Wolfowitz algorithm with randomized differences," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 442–453, Mar. 1999.

[8] Y. Chen, C. Wen, Z. Gong, and M. Sun, "An iterative learning controller with initial state learning," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 371–376, Mar. 1999.

[9] C. J. Chien, "A discrete iterative learning control for a class of nonlinear time-varying systems," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 748–752, June 1998.

[10] L. Gerencsér and Z. S. Vágó, "The mathematics of noise-free SPSA," in *Proc. 40th IEEE Conf. Decision Control*, Orlando, FL, 2001, pp. 4400–4405.

[11] D. Gorinevsky, "An approach to parametric nonlinear least square optimization and application to task-level learning control," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 912–927, July 1997.

[12] R. Horowitz, W. Messner, and J. B. Moore, "Exponential convergence of a learning controller for robot manipulators," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 890–894, June 1991.

[13] Y. H. Kim and I. J. Ha, "A learning approach to precision speed control of servomotors and its application to VCR," *IEEE Trans. Contr. Syst. Technol.*, vol. 7, pp. 466–477, Mar. 1999.

[14] ——, "Asymptotic state tracking in a class of nonlinear systems via learning-based inversion," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 2011–2027, Dec. 2000.

[15] J. Koronacki, "Random-seeking methods for the stochastic uncertained optinization," *Int. J. Control*, vol. 21, pp. 517–527, 1975.

[16] J. H. Lee, K. S. Lee, and W. C. Kim, "Model-based iterative learning control with quadratic criterion for time-varying linear systems," *Automatica*, vol. 36, pp. 641–658, 2000.

[17] M. Pandit and K. H. Buchheit, "Optimizing iterative learning control of cyclic production process with application to extruders," *IEEE Trans. Control Syst. Technol.*, vol. 7, pp. 384–390, Mar. 1999.

[18] S. S. Saab, "A discrete-time stochastic learning control algorithm," *IEEE Trans. Automat. Contr.*, vol. 46, pp. 877–887, June 2001.

[19] ——, "On a discrete-time stochastic learning control algorithm," *IEEE Trans. Automat. Contr.*, vol. 46, pp. 1333–1336, Sept. 2001.

[20] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 332–341, Feb. 1992.