Decentralized Iterative Learning Controllers for Nonlinear Large-scale Systems to Track Trajectories with Different Magnitudes

RUAN Xiao-E[1] CHEN Feng-Min[1] WAN Bai-Wu[2]

Abstract In hierarchical steady-state optimization programming for large-scale industrial processes, a feasible technique is to use information of the real system so as to modify the model-based optimum. In this circumstance, a sequence of step function-type control decisions with distinct magnitudes is computed, by which the real system is stimulated consecutively. In this paper, a set of iterative learning controllers is embedded into the procedure of hierarchical steady-state optimization in decentralized mode for a class of large-scale nonlinear industrial processes. The controller for each subsystem is used to generate a sequence of upgraded control signals so as to take responsibilities of the sequential step control decisions with distinct scales. The aim of the learning control design is to consecutively refine the transient performance of the system. By means of the Hausdorff-Young inequality of convolution integral, the convergence of the updating rule is analyzed in the sense of Lebesgue-p norm. Invention of the nonlinearity and the interaction on convergence are discussed. Validity and effectiveness of the proposed control scheme are manifested by some simulations.

Key words Nonlinear large-scale systems, steady-state optimization, iterative learning control, Lebesgue-p norm, convergence

In recent years, iterative learning control (ILC) has been one of the active research directions in intelligent control development for trajectory tracking[1−2]. For nonlinear systems, much attention has been paid, in which the convergence has been analyzed in the sense of Lambda norm[3−5]. Bien′s group has developed a technique to analyze the convergence of proposed iterative learning control in the sense of sup-norm[6−7]. Furthermore, Xie et al.[2, 8] have developed techniques to designate iterative learning control strategies based on geometric analysis or geometric vector plots for fastening convergent speed in the sense of so-called Lambda-Xi norm. The main theme in the above-mentioned investigations is to design iterative learning control schemes for tracking a unique desired trajectory. It is because of the repeatability of the conventional tracking task that the supposed realizability of the prescribed desired trajectory and the satisfaction of Lipschitz condition for the system have played important roles in the convergence analysis. For non-repetitive trajectories tracking, Xu[9] has developed a direct learning scheme to track trajectories with proportional magnitudes, in which the scheme is to generate control profiles based on the predetermined desired control commands. This scheme requires that the predetermined desired control commands are unique desired trajectory tracking-based available beforehand. Also, Saab[10] has studied the problem of tracking slow-varying trajectories, each of which differs from the previous one within a small fluctuation level. In the study, some bounded tracking error is guaranteed in the sense of Lambda-norm under the assumption that the non-repetitive desired trajectories are to be realizable and some disturbances are in presence. Furthermore, Xu[11] has developed a kind of learning rule for general non-repetitive trajectories tracking tasks, which is principally formulated based on some prior measurable system state and prior system knowledge with specific pattern. The mechanism may be regarded as, in some sense, iteratively learning the inversion of the system, and thus the tracking effectiveness sensitively relies on the precision of the system model.

Recently, for the purpose of improving the transient performance of linear large-scale industrial processes, a type of decentralized iterative learning control scheme has been embedded into the procedure of steady-state hierarchical optimization for tracking desired trajectories with distinct magnitudes at different iterations[12]. Due to the drawback as commented in [6−7] that the Lambda-norm, which has been utilized to measure the tracking error in some studies[3−5], may not be a satisfactory measurement and the same situation may emerge when Lambda-Xi norm is adopted as investigated in [8], as well as remarked in [6−7], that it is possible for the algorithm to diverge if the tracking task is defined over the restrained time interval in the sense of sup norm, and Ruan et al.[12] has analyzed the convergence in the sense of Lebesgue-p norm. In the study, the intervention of the system′s inherent nature, such as the distinct scales of the desired trajectories and the interactions among subsystems, on the convergence were discussed. As addressed in [13], some nonlinearities may arise in actual control applications. In this situation, we must deal with a nonlinear model for the control system. In this paper, a decentralized iterative learning control strategy is embedded into nonlinear large-scale systems to track desired trajectories with different magnitudes, and the Lebesgue-p norm is adopted to analyze the convergence of the proposed iterative learning control algorithm.

The remaining part of the paper is organized as follows. Section 1 shows an iterative learning control formulation and the specification of desired trajectories. In Section 2, the convergence of the algorithm is derived in the sense of the Lebesgue-p norm. Section 3 shows the simulation results. Finally, Section 4 concludes the paper.

1 Iterative learning control strategy

We consider a class of stable large-scale systems, each of which consists of a set of nonlinear time-invariant subsystems with the form as

$$\begin{cases} \dot{\boldsymbol{x}}^{(i)}(t) = f^{(i)}(\boldsymbol{x}^{(i)}(t)) + B^{(i)}\boldsymbol{r}^{(i)}(t) + \\ \qquad\qquad \sum\limits_{j=1, j\neq i}^{N} D^{(ij)}\boldsymbol{x}^{(j)}(t) \\ \boldsymbol{y}^{(i)}(t) = C^{(i)}\boldsymbol{x}^{(i)}(t), \quad \boldsymbol{x}^{(i)}(0) = \boldsymbol{0} \end{cases} \tag{1}$$

Received October 10, 2006; in revised form December 10, 2007
Supported by National Natural Science Foundation of China (F030101-60574021) and National "985" Project of China Executed in Xi′an Jiaotong University
1. Department of Mathematics, Faculty of Science, Xi′an Jiaotong University, Xi′an 710049, P. R. China 2. Institute of Systems Engineering, Xi′an Jiaotong University, Xi′an 710049, P. R. China
DOI: 10.3724/SP.J.1004.2008.00426

where $f^{(i)}(\boldsymbol{x}^{(i)}(t))$ is assumed to be a nonlinear second-order differentiable functional vector with a unique original equilibrium, that is, $f^{(i)}(\mathbf{0}) = \mathbf{0}$. The superscript (i) stands for the distinction of subsystem i, $\boldsymbol{x}^{(i)}(t) \in \mathbf{R}^{n^{(i)}}$, $\boldsymbol{r}^{(i)}(t) \in \mathbf{R}^{m^{(i)}}$, and $\boldsymbol{y}^{(i)}(t) \in \mathbf{R}^{m^{(i)}}$ denote $n^{(i)}$-dimensional state, $m^{(i)}$-dimensional control input, and $m^{(i)}$-dimensional output vectors, respectively. Specifically, $\sum_{j=1,j\neq i}^{N} D^{(ij)}\boldsymbol{x}^{(j)}(t)$ represents the state interactions from the other subsystems. $A^{(i)}$, $B^{(i)}$, $C^{(i)}$, and $D^{(ij)}$ are matrices with appropriate dimensions. The iterative learning control (ILC) structure is shown in Fig. 1.

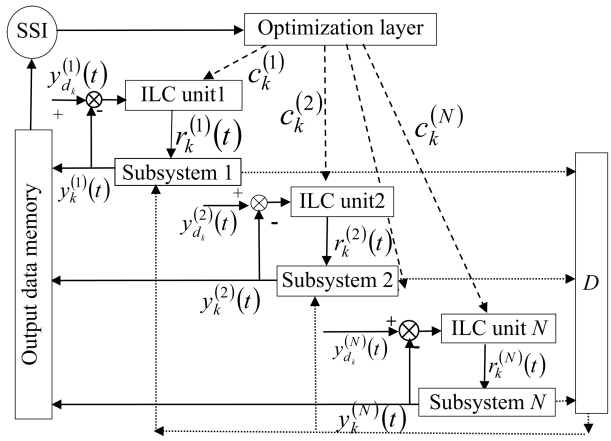


Fig. 1 ILC structure for a large-scale system

The structure indicates that the iterative learning controller for a large-scale system is designed in a decentralized mode. In Fig. 1, the optimization layer consists of a coordinator and a set of local making-decision units in a hierarchical structure as detailed in [12], from which the control decisions change sequence $\boldsymbol{c}_1^{(i)} = [c_1^{(i1)}, c_1^{(i2)}, \cdots, c_1^{(im^{(i)})}]$, $\boldsymbol{c}_2^{(i)} = [c_2^{(i1)}, c_2^{(i2)}, \cdots, c_2^{(im^{(i)})}]$, $\cdots$, $\boldsymbol{c}_k^{(i)} = [c_k^{(i1)}, c_k^{(i2)}, \cdots, c_k^{(im^{(i)})}]$, of subsystem i, for $i = 1, 2, \cdots, N$, and $k = 1, 2, \cdots$, is supposed to be carried out consecutively while the steady state hierarchical optimization using the feedback steady-state information (SSI) is undergoing. Here, the subscript k is used to represent the feedback index, at which the model-based local control decision for subsystem i is improved as $\boldsymbol{c}_k^{(i)}$ for $i = 1, 2, \cdots, N$, after the model is modified by the real feedback information. $D = [D^{(ij)}]$ stands for the interconnected block matrix, which consists of block entries $D^{(ij)}$ as formulated in dynamic equations (1), for $i, j = 1, 2, \cdots, N, i \neq j$ and block entries $D^{(ii)} = 0$ specified, for $i = 1, 2, \cdots, N$.

It is noted that $\boldsymbol{c}_k^{(i)}$ is the set-point change vector of control decision of subsystem i at k-th implementation. This implies that we need not operate the real system to get the real steady-state information if $\boldsymbol{c}_k^{(i)} = \mathbf{0}$. Therefore, it is reasonable to assume that all of the set-point changes vectors, or called control decisions, $\boldsymbol{c}_k^{(i)} \neq 0$, for all $i = 1, 2, \cdots, N$ and $k = 1, 2, \cdots$. In particular, when partial components of the control decision $\boldsymbol{c}_k^{(i)}$ at k-th implementation are in absence, we consider the case as a singular occurrence and it shall be specifically remarked later. Without loss of generality, we further assume that all of the components of the control decision $\boldsymbol{c}_k^{(i)}$ are in presence, mathematically speaking, $0 < |\boldsymbol{c}_k^{(il)}| < \infty$, for $i = 1, 2, \cdots, N$, $l = 1, 2, \cdots, m^{(i)}$, and $k = 1, 2, \cdots$. Under such an assumption, we denote a set of diagonal matrices as follows for constructing the updating law.

$$Q_1^{(i)} = \mathrm{diag}[c_1^{(i1)}, c_1^{(i2)}, \cdots, c_1^{(im^{(i)})}]$$

$$Q_2^{(i)} = \mathrm{diag}[c_2^{(i1)}, c_2^{(i2)}, \cdots, c_2^{(im^{(i)})}]$$

$$\cdots\cdots$$

$$Q_k^{(i)} = \mathrm{diag}[c_k^{(i1)}, c_k^{(i2)}, \cdots, c_k^{(im^{(i)})}], (k = 1, 2, \cdots)$$

We propose an iterative learning control updating law for subsystem i to generate a sequential constrained input vectors $\boldsymbol{r}_1^{(i)}(t), \boldsymbol{r}_2^{(i)}(t), \cdots, \boldsymbol{r}_k^{(i)}(t), \cdots$, consecutively in response to the control decision sequence $\boldsymbol{c}_1^{(i)}$, $\boldsymbol{c}_2^{(i)}$, $\cdots$, $\boldsymbol{c}_k^{(i)}$, $\cdots$ $(k = 1, 2, \cdots)$. The iterative learning control updating rule, represented as an outcome of the ILC unit i in Fig. 1, for $i = 1, 2, \cdots, N$, is formulated as follows.

$$\boldsymbol{r}_k^{(i)}(t) = Q_k^{(i)}(Q_{k-1}^{(i)})^{-1}[\boldsymbol{r}_{k-1}^{(i)}(t) + \Gamma_p^{(i)}\boldsymbol{e}_{k-1}^{(i)}(t) + \Gamma_d^{(i)}\dot{\boldsymbol{e}}_{k-1}^{(i)}(t)]$$

$$t \in [0, T], k = 1, 2, \cdots \tag{2}$$

where $\boldsymbol{r}_1^{(i)}(t) = \boldsymbol{c}_1^{(i)}$, $\Gamma_p^{(i)}$ and $\Gamma_d^{(i)}$ are $m^{(i)}$-dimension proportional and derivative diagonal learning matrices, respectively. Here, T refers to the settling time of the transient response whereas system (1) is driven by a control input vector with each component being a unit step function. In algorithm (2), $\boldsymbol{e}_{k-1}^{(i)}(t) = \boldsymbol{y}_{d_{k-1}}^{(i)}(t) - \boldsymbol{y}_{k-1}^{(i)}(t)$ denotes the output error vector, which presents the discrepancy between the desired reference trajectory vector $\boldsymbol{y}_{d_{k-1}}^{(i)}(t)$ and the output vector $\boldsymbol{y}_{k-1}^{(i)}(t)$. $\dot{\boldsymbol{e}}_{k-1}^{(i)}(t) = \mathrm{d}\boldsymbol{e}_{k-1}^{(i)}(t)/\mathrm{d}t$ stands for the derivative of output error vector $\boldsymbol{e}_{k-1}^{(i)}(t)$ with respect to variable t. The desired reference trajectory vector $\boldsymbol{y}_{d_{k-1}}^{(i)}(t)$ with respect to the control decision vector $\boldsymbol{c}_{k-1}^{(i)}$ will be specified later. Note that in the ILC algorithm (2) $(k = 1, 2, \cdots)$, an amplified matrix $Q_k^{(i)}(Q_{k-1}^{(i)})^{-1}$ is introduced so as to guarantee that the steady-state value of the input $\boldsymbol{r}_k^{(i)}(t)$ is equal to the control decision $\boldsymbol{c}_k^{(i)}$, that is, $\boldsymbol{r}_k^{(i)}(\infty) = \boldsymbol{c}_k^{(i)}$, whereas $\boldsymbol{e}_k^{(i)}(\infty) = 0$ $(k = 1, 2, \cdots)$.

To specify the desired reference trajectories, we select the trajectory vector $\boldsymbol{y}_{d_0}^{(i)}(t) = [y_{d_0}^{(i1)}(t), y_{d_0}^{(i2)}(t), \cdots, y_{d_0}^{(im^{(i)})}(t)]$ appropriately to meet the desired requirement that all of whose components have satisfactory characteristics, such as with no or less overshooting, quick transient responding, and short settling time. For example, we may choose $y_{d_0}^{(il)}(t) = 1 - \exp(-a^{(il)}t)$ or $y_{d_0}^{(il)}(t) = 1 - b^{(il)}(1+t)^{-3}$ with $y_{d_0}^{(il)}(0) = 0$ and $y_{d_0}^{(il)}(\infty) = 1$ for all $i = 1, 2, \cdots, N$ and $l = 1, 2, \cdots, m^{(i)}$. Generally, we may not accurately estimate the steady-state value of the system if we are not aware of the precise knowledge of the system. In this circumstance, we adjust the scale of the desired trajectory $\boldsymbol{y}_{d_{k-1}}^{(i)}(t), (k = 1, 2, \cdots)$ referring not to the magnitude of the given control decision but to the real steady-state output vector formulated as

$$\boldsymbol{s}_{k-1}^{(i)} = \lim_{t\to\infty} \boldsymbol{y}_{k-1}^{(i)}(t) = [s_{k-1}^{(i1)}, s_{k-1}^{(i2)}, \cdots, s_{k-1}^{(im^{(i)})}]^{\mathrm{T}}$$

conducted by the constrained control input $\boldsymbol{r}_{k-1}^{(i)}(t), (k = 1, 2, \cdots)$. To express the desired trajectory $\boldsymbol{y}_{d_{k-1}}^{(i)}(t)$ at

$(k-1)$-th implementation compactly, we set $S_{k-1}^{(i)} = \mathrm{diag}[s_{k-1}^{(i1)}, s_{k-1}^{(i2)}, \cdots, s_{k-1}^{(im^{(i)})}]$, called steady-state value matrix, for all $i = 1, 2, \cdots, N$ and $l = 1, 2, \cdots, m^{(i)}$. Hence, the desired trajectory vector $\boldsymbol{y}_{d_{k-1}}^{(i)}(t)$ can be specified as

$$\boldsymbol{y}_{d_{k-1}}^{(i)}(t) = S_{k-1}^{(i)} \boldsymbol{y}_{d_0}^{(i)}(t) \tag{3}$$

We reform the nonlinear function $f^{(i)}(\boldsymbol{x}^i(t))$ in system (1) by utilizing Maclaurin series expansion as

$$f^{(i)}(\boldsymbol{x}^{(i)}(t)) = f^{(i)}(\boldsymbol{x}^{(i)}(t)) - f^{(i)}(\mathbf{0}) =$$

$$\frac{\partial f^{(i)}}{\partial \boldsymbol{x}^{(i)}}|_{\boldsymbol{x}^{(i)}=\mathbf{0}} \boldsymbol{x}^{(i)}(t) + h^{(i)}(\boldsymbol{x}^{(i)}(t)) =$$

$$F^{(i)} \boldsymbol{x}^{(i)}(t) + h^{(i)}(\boldsymbol{x}^{(i)}(t)) \tag{4}$$

where $F^{(i)} = \frac{\partial f^{(i)}}{\partial \boldsymbol{x}^{(i)}}|_{\boldsymbol{x}^{(i)}=\mathbf{0}}$ refers to the Jacobian matrix, namely, the first-order derivative matrix, of the functional vector with respect to the state vector $\boldsymbol{x}^{(i)}(t)$ at the original equilibrium, whereas $h^{(i)}(\boldsymbol{x}^{(i)}(t))$, which stands for the nonlinear term of the functional vector $f^{(i)}(\boldsymbol{x}^i(t))$, is a functional vector with respect to the state vector $\boldsymbol{x}^{(i)}(t)$, in which the power of the components of the state vector $\boldsymbol{x}^{(i)}(t)$ is higher than one. Besides the notations in (4) and (6), we additionally denote

$$F = \begin{bmatrix} F^{(1)} & D^{(12)} & \cdots & D^{(1N)} \\ D^{(21)} & F^{(2)} & \cdots & D^{(2N)} \\ \vdots & \vdots & \ddots & \vdots \\ D^{(N1)} & D^{(N2)} & \cdots & F^{(N)} \end{bmatrix}$$

$$h(\boldsymbol{x}(t)) = [(h^{(1)}(\boldsymbol{x}^{(1)}(t)))^{\mathrm{T}}, \cdots, (h^{(N)}(\boldsymbol{x}^{(N)}(t)))^{\mathrm{T}}]^{\mathrm{T}}$$

$$B = \mathrm{blockdiag}[B^{(1)}, B^{(2)}, \cdots, B^{(N)}]$$

$$C = \mathrm{blockdiag}[C^{(1)}, C^{(2)}, \cdots, C^{(N)}]$$

$$\begin{aligned} \boldsymbol{x}(t) &= [(\boldsymbol{x}^{(1)}(t))^{\mathrm{T}}, (\boldsymbol{x}^{(2)}(t))^{\mathrm{T}}, \cdots, (\boldsymbol{x}^{(N)}(t))^{\mathrm{T}}]^{\mathrm{T}} \\ \boldsymbol{y}(t) &= [(\boldsymbol{y}^{(1)}(t))^{\mathrm{T}}, (\boldsymbol{y}^{(2)}(t))^{\mathrm{T}}, \cdots, (\boldsymbol{y}^{(N)}(t))^{\mathrm{T}}]^{\mathrm{T}} \\ \boldsymbol{r}(t) &= [(\boldsymbol{r}^{(1)}(t))^{\mathrm{T}}, (\boldsymbol{r}^{(2)}(t))^{\mathrm{T}}, \cdots, (\boldsymbol{r}^{(N)}(t))^{\mathrm{T}}]^{\mathrm{T}} \end{aligned} \tag{5}$$

Additionally, we denote

$$\boldsymbol{r}_k(t) = [(\boldsymbol{r}_k^{(1)}(t))^{\mathrm{T}}, (\boldsymbol{r}_k^{(2)}(t))^{\mathrm{T}}, \cdots, (\boldsymbol{r}_k^{(N)}(t))^{\mathrm{T}}]^{\mathrm{T}}$$

$$\boldsymbol{e}_k(t) = [(\boldsymbol{e}_k^{(1)}(t))^{\mathrm{T}}, (\boldsymbol{e}_k^{(2)}(t))^{\mathrm{T}}, \cdots, (\boldsymbol{e}_k^{(N)}(t))^{\mathrm{T}}]^{\mathrm{T}}$$

$$\boldsymbol{y}_{d_k}(t) = [(\boldsymbol{y}_{d_k}^{(1)}(t))^{\mathrm{T}}, (\boldsymbol{y}_{d_k}^{(2)}(t))^{\mathrm{T}}, \cdots, (\boldsymbol{y}_{d_k}^{(N)}(t))^{\mathrm{T}}]^{\mathrm{T}}$$

$$\boldsymbol{c}_k = [(\boldsymbol{c}_k^{(1)})^{\mathrm{T}}, (\boldsymbol{c}_k^{(2)})^{\mathrm{T}}, \cdots, (\boldsymbol{c}_k^{(N)})^{\mathrm{T}}]^{\mathrm{T}}$$

$$S_{k-1} = \mathrm{blockdiag}[S_{k-1}^{(1)}, S_{k-1}^{(2)}, \cdots, S_{k-1}^{(N)}]$$

$$Q_k = \mathrm{blockdiag}[Q_k^{(1)}, Q_k^{(2)}, \cdots, Q_k^{(N)}]$$

$$\Gamma_p = \mathrm{blockdiag}[\Gamma_p^{(1)}, \Gamma_p^{(2)}, \cdots, \Gamma_p^{(N)}]$$

$$\Gamma_d = \mathrm{blockdiag}[\Gamma_d^{(1)}, \Gamma_d^{(2)}, \cdots, \Gamma_d^{(N)}] \tag{6}$$

The updating law (2) can be reformed compactly as

$$\boldsymbol{r}_k(t) = Q_k Q_{k-1}^{-1}[\boldsymbol{r}_{k-1}(t) + \Gamma_p \boldsymbol{e}_{k-1}(t) + \Gamma_d \dot{\boldsymbol{e}}_{k-1}(t)] \tag{7}$$

The nonlinear large-scale system consisting of the nonlinear subsystems (1) is thus expressed as

$$\begin{cases} \dot{\boldsymbol{x}}(t) = F\boldsymbol{x}(t) + h(\boldsymbol{x}(t)) + B\boldsymbol{r}(t) \\ \boldsymbol{y}(t) = C\boldsymbol{x}(t), \quad \boldsymbol{x}(\mathbf{0}) = \mathbf{0} \end{cases} \tag{8}$$

Provided that the control signal $\boldsymbol{r}(t)$ in the above nonlinear system (8) is taken by control profile generated by iterative learning control rule (7), the corresponding large-scale system (8) becomes

$$\begin{cases} \dot{\boldsymbol{x}}_k(t) = F\boldsymbol{x}_k(t) + h(\boldsymbol{x}_k(t)) + B\boldsymbol{r}_k(t) \\ \boldsymbol{y}_k(t) = C\boldsymbol{x}_k(t), \quad \boldsymbol{x}_k(\mathbf{0}) = \mathbf{0} \end{cases} \tag{9}$$

Here, k represents the iteration suffix as mentioned earlier. Besides, $\boldsymbol{x}_k(t) \in \mathbf{R}^{\sum_{i=1}^N n^{(i)}}$, $\boldsymbol{r}_k(t) \in \mathbf{R}^{\sum_{i=1}^N m^{(i)}}$, and $\boldsymbol{y}_k(t) \in \mathbf{R}^{\sum_{i=1}^N m^{(i)}}$ refer to the state, control input, and output vectors of large-scale system (8) at k-th implementation, respectively.

In the sequel, for a continuous function $f(t) : [0, T] \to \mathbf{R}$, the Lebesgue-$p$ norm is defined as

$$\| f(\cdot) \|_p = [\int_0^T |f(t)|^p \mathrm{d}t]^{\frac{1}{p}}, 1 \le p \le \infty$$

As a consequence, for a continuous function vector $\boldsymbol{g}(t) = [g^{(1)}(t), g^{(2)}(t), \cdots, g^{(m)}(t)]^{\mathrm{T}} : [0, T] \to \mathbf{R}^m$, its Lebesgue-$p$ norm is formulated as

$$\||\boldsymbol{g}(\cdot)|\|_p = [\int_0^T (\max_{j=1,2,\cdots,m} |g^{(j)}(t)|)^p \mathrm{d}t]^{\frac{1}{p}}, 1 \le p \le \infty$$

Additionally, for a continuous functional matrix $U(t) = [u_{ij}]_{m\times n} : [0, T] \to \mathbf{R}^{m\times n}$, its induced Lebesgue-$p$ norm is written as follows

$$\||U(\cdot)|\|_p = [\int_0^T (\max_{i=1,2,\cdots,m} \sum_{j=1}^n |u^{(ij)}(t)|)^p \mathrm{d}t]^{\frac{1}{p}}, 1 \le p \le \infty$$

The convergence issue of the algorithm in the sense of Lebesgue-p is deduced in the next section.

2 Convergence analysis

We now focus on the convergence issue.

Theorem 1. Assume that system (8) is stable and the matrix $C\exp(F(t))B$ is nonsingular. Then, the iterative learning control algorithm (2) is bounded with respect to (8) in the sense of Lebesgue-p norm if the system parameters and learning gains satisfy the following condition:

$$\begin{aligned} \rho = \ & \|(F_1 Q_k)^{-1}(I - CB\Gamma_d)F_1 Q_k\| + \\ & \||(F_1 Q_k)^{-1} C\exp(F(\cdot))(FB\Gamma_d + B\Gamma_p)F_1 Q_k|\|_1 < 1 \end{aligned}$$

Proof. We shall consider the output error modified as $S_k^{(-1)}\boldsymbol{e}_k(t)$ by the steady-state value matrix S_k only if the matrix S_k is nonsingular. Due to the nonlinearity, it is rather difficult to get the explicit solution to the differential equations (8). However, from the viewpoint of mathematics, we can express its solution implicitly as

$$\boldsymbol{y}(t) = C\int_0^t \exp(F(t-\tau))(B\boldsymbol{r}(\tau) + h(\boldsymbol{x}(\tau)))\mathrm{d}\tau$$

Similarly, the solution to the differential equation (9) can be formulated as

$$\boldsymbol{y}_k(t) = C\int_0^t \exp(F(t-\tau))(B\boldsymbol{r}_k(\tau) + h(\boldsymbol{x}_k(\tau)))\mathrm{d}\tau \quad (10)$$

By considering (10), we get

$$\begin{aligned}
&S_k^{-1}\boldsymbol{e}_k(t) = S_k^{-1}\boldsymbol{y}_{d_k}(t) - S_k^{-1}\boldsymbol{y}_k(t) = \\
&\quad S_{k-1}^{-1}\boldsymbol{e}_{k-1}(t) - [S_k^{-1}C\boldsymbol{x}_k(t) - S_{k-1}^{-1}C\boldsymbol{x}_{k-1}(t)] = \\
&\quad S_{k-1}^{-1}\boldsymbol{e}_{k-1}(t) - [S_k^{-1}C\int_0^t \exp(F(t-\tau))B\boldsymbol{r}_k(\tau)\mathrm{d}\tau - \\
&\quad S_{k-1}^{-1}C\int_0^t \exp(F(t-\tau))B\boldsymbol{r}_{k-1}(\tau)\mathrm{d}\tau] - \\
&\quad [S_k^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_k(\tau))\mathrm{d}\tau - \\
&\quad S_{k-1}^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_{k-1}(\tau))\mathrm{d}\tau] = \\
&\quad S_{k-1}^{-1}\boldsymbol{e}_{k-1}(t) - S_k^{-1}\int_0^t \exp(F(t-\tau))BQ_kQ_{k-1}^{-1}\times \\
&\quad [\boldsymbol{r}_{k-1}(\tau) + \Gamma_p\boldsymbol{e}_{k-1}(\tau) + \Gamma_d\dot{\boldsymbol{e}}_{k-1}(\tau)]\mathrm{d}\tau - \\
&\quad S_{k-1}^{-1}C\int_0^t \exp(F(t-\tau))B\boldsymbol{r}_{k-1}(\tau)\mathrm{d}\tau] - \\
&\quad [S_k^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_k(\tau))\mathrm{d}\tau - \\
&\quad S_{k-1}^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_{k-1}(\tau))\mathrm{d}\tau] = \\
&\quad S_{k-1}^{-1}\boldsymbol{e}_{k-1}(t) - \\
&\quad S_k^{-1}C\exp(F(t-\tau))BQ_kQ_{k-1}^{-1}\Gamma_d\boldsymbol{e}_{k-1}(\tau)|_{\tau=0}^{\tau=t} - \\
&\quad [S_k^{-1}C\int_0^t \exp(F(t-\tau))FBQ_kQ_{k-1}^{-1}\Gamma_d\boldsymbol{e}_{k-1}(\tau)]\mathrm{d}\tau - \\
&\quad [S_k^{-1}C\int_0^t \exp(F(t-\tau))FQ_kQ_{k-1}^{-1}\Gamma_p\boldsymbol{e}_{k-1}(\tau)]\mathrm{d}\tau - \\
&\quad \int_0^t [S_k^{-1}C\exp(F(t-\tau))BQ_k - \\
&\quad S_{k-1}^{-1}C\exp(F(t-\tau))BQ_{k-1}]Q_{k-1}^{-1}\boldsymbol{r}_{k-1}(\tau)\mathrm{d}\tau - \\
&\quad [S_k^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_k(\tau))\mathrm{d}\tau - \\
&\quad S_{k-1}^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_{k-1}(\tau))\mathrm{d}\tau]
\end{aligned}$$

Because $\boldsymbol{e}_{k-1}(\mathbf{0}) = \mathbf{0}$, the above expression becomes

$$\begin{aligned}
&S_k^{-1}\boldsymbol{e}_k(t) = [I - S_k^{-1}CB\Gamma_dQ_kQ_{k-1}^{-1}S_{k-1}]S_{k-1}^{-1}\boldsymbol{e}_{k-1}(t) - \\
&\int_0^t [S_k^{-1}C\exp(F(t-\tau))(FB\Gamma_d + B\Gamma_p)\times \\
&Q_kQ_{k-1}^{-1}S_{k-1}S_{k-1}^{-1}\boldsymbol{e}_{k-1}(\tau)]\mathrm{d}\tau - \\
&\int_0^t [S_k^{-1}C\exp(F(t-\tau))BQ_k - S_{k-1}^{-1}C\times \\
&\exp(F(t-\tau))BQ_{k-1}]Q_{k-1}^{-1}\boldsymbol{r}_{k-1}(\tau)\mathrm{d}\tau - \\
&[S_k^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_k(\tau))\mathrm{d}\tau - S_{k-1}^{-1}C\times \\
&\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_{k-1}(\tau))\mathrm{d}\tau] \quad (11)
\end{aligned}$$

Calculating Lp-norm for both sides of (11) and using the Hausdorff-Young inequality of convolution integral yields

$$\begin{aligned}
&|||\, S_k^{-1}\boldsymbol{e}_k(\cdot)\, |||_p \leq \\
&\|\, I - S_k^{-1}CB\Gamma_dQ_kQ_{k-1}^{-1}S_{k-1}\, \|||||\, S_{k-1}^{-1}\boldsymbol{e}_{k-1}(\cdot)\, |||_p + \\
&|||\, S_k^{-1}C\exp(F(\cdot))(FB\Gamma_d + B\Gamma_p)Q_kQ_{k-1}^{-1}S_{k-1}\, |||_1 \times \\
&|||\, S_{k-1}^{-1}\boldsymbol{e}_{k-1}(\cdot)\, |||_p + \\
&|||\, S_k^{-1}C\exp(F(\cdot))BQ_k - S_{k-1}^{-1}C\exp(F(\cdot))BQ_{k-1}\, |||_1 \times \\
&|||\, Q_{k-1}^{-1}\boldsymbol{r}_{k-1}(\cdot)\, |||_p + \\
&|||\, S_k^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_k(\tau))\mathrm{d}\tau - S_{k-1}^{-1}C\times \\
&\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_{k-1}(\tau))\mathrm{d}\tau\, |||_1 \quad (12)
\end{aligned}$$

According to the control theory, the steady-state value of the steered system is uniquely determined by the intrinsic mechanism of the system no matter how much internal information we acquire if the system is stable and the control decision is given. Therefore, it is reasonable to suppose that the relationship between the control decision and its corresponding steady-state value is as follows

$$S_{k-1} = F_1Q_{k-1} \quad (13)$$

We understand that matrix F_1 is diagonal because both S_{k-1} and Q_{k-1} are selected being diagonal. Hence, we can derive that

$$\begin{aligned}
&\|\, I - S_k^{-1}CB\Gamma_dQ_kQ_{k-1}^{-1}S_{k-1}\, \| = \\
&\|\, I - Q_k^{-1}F_1^{-1}CB\Gamma_dQ_kQ_{k-1}^{-1}F_1Q_{k-1}\, \| = \\
&\|\, (F_1Q_k)^{-1}(I - CB\Gamma_d)F_1Q_k\, \|
\end{aligned}$$

Simultaneously, we get

$$\begin{aligned}
&|||\, S_k^{-1}C\exp(F(\cdot))(FB\Gamma_d + B\Gamma_p)Q_kQ_{k-1}^{-1}S_{k-1}\, |||_1 = \\
&|||\, F_1Q_k{}^{-1}C\exp(F(\cdot))(FB\Gamma_d + B\Gamma_p)F_1Q_k\, |||_1, \\
&|||\, S_k^{-1}C\exp(F(\cdot))BQ_k - S_{k-1}^{-1}C\exp(F(\cdot))BQ_{k-1}\, |||_1 = \\
&|||\, Q_k^{-1}F_1^{-1}C\exp(F(\cdot))BQ_k - Q_{k-1}^{-1}F_1^{-1}\times \\
&C\exp(F(\cdot))BQ_{k-1}\, |||_1
\end{aligned}$$

Denote

$$\begin{aligned}
\Psi_k =& |||\, Q_k^{-1}F_1^{-1}C\exp(F(\cdot))BQ_k - Q_{k-1}^{-1}F_1^{-1}C\times \\
&\exp(F(\cdot))BQ_{k-1}\, |||_1 \\
E_k =& |||\, S_k^{-1}\boldsymbol{e}_k(\cdot)\, |||_p \\
R_k =& |||\, Q_k^{-1}\boldsymbol{r}_k(\cdot)\, |||_p \\
L_k =& |||\, S_k^{-1}C\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_k(\tau))\mathrm{d}\tau - S_{k-1}^{-1}C\times \\
&\int_0^t \exp(F(t-\tau))h(\boldsymbol{x}_{k-1}(\tau))\mathrm{d}\tau]\, |||_1 \\
\rho =& \|(F_1Q_k)^{-1}(I - CB\Gamma_d)F_1Q_k\| + |||(F_1Q_k)^{-1}C\times \\
&\exp(F(\cdot))(FB\Gamma_d + B\Gamma_p)F_1Q_k|||_1
\end{aligned}$$

The inequalition (12) is consequently reduced to

$$E_k \leq \rho E_{k-1} + \Psi_kR_k + L_k \quad (14)$$

It is reminded that we assume the system is stable, which leads to $\exp(F(\infty)) = 0$. Sequentially, it is reasonable to

assume that $\Psi_k \le \Psi$, $L_k \le L$, and $R_k \le R$. Hence, the inequalition (14) can be rewritten as

$$E_k \le \rho E_{k-1} + \Psi_k R_k + L_k \le \rho E_{k-1} + \Psi R + L \tag{15}$$

Taking into account the assumption that $\rho < 1$ yields

$$\lim_{k\to\infty} \le \frac{\Psi R + L}{1-\rho} \tag{16}$$

□

Remark 1. From (16), we observe that the bound of output error contains two parts, that is, ΨR and L. The influence intensity of the term ΨR on the convergence is primarily dominated by the distinct scales of the control decisions and the coupling feature of the system. Due to the permutable limitation of matrix multiplication and the distinct scales of diagonal elements of the matrix Q_k, as well as the matrix Q_{k-1}, the term

$$\Psi_k = \|| \, Q_k^{-1} F_1^{-1} C \exp(F(\cdot)) B Q_k - Q_{k-1}^{-1} F_1^{-1} C \times \exp(F(\cdot)) B Q_{k-1} \, ||_1$$

which shows the bias of two similarity transformations of the normalized transmission matrix $F_1^{-1} C \exp(F(\cdot)) B$, may not vanish if the direct transmission matrix $C\exp(F(\cdot))B$ is nonsingular, unless either matrix $C\exp(F(\cdot))B$ is diagonal or matrix Q_k is proportional to the matrix Q_{k-1}. It is worth pointing out that $C\exp(F(\cdot))B$ being diagonal means that the large-scale system, as well as each of MIMO subsystem, is decoupled. In this case, the iterative learning control strategy is equivalent to that for a set of SISO subsystems. This phenomenon may not happen because of, at least, the existence of the interactions. The term may disappear if the large-scale system is decoupled or if matrix Q_k is proportional to matrix Q_{k-1}. In addition, the term concerning the upper bound L of L_k may not vanish since it reflects the limitation of the influence caused by the system nonlinear components and formulated as the discrepancy between two adjacent operations as above-mentioned.

Despite the distinguishable conclusion from the existing ILC for repetitive tracking task, the above formulation implies that the milder the nonlinearity, the less the output error. From the mathematics point of view, the conclusion is convincing because the scales of both the constrained control inputs and the specified desired trajectories are linearly amplified in agreement with the magnitudes of the control decisions and its corresponding steady-state values as the iteration index increases, whereas the system characteristics is featured to be nonlinear. Certainly, this will result in accumulation of system nonlinear components between two consecutive trials.

Remark 2. It is clear that, mathematically, the sup norm is the case when we choose $p = \infty$ in the concept of Lebesgue-p norm, while the Euclidean inner product induced norm is the case of $p = 2$. Thus, the existing ILC algorithm for tracking a unique desired trajectory is convergent in the sense of Lebesgue-p norm. This coincides with the result with respect to the sup norm shown in [6−7]. More significantly, the upper bound of the convergent interval of the present result does not depend on the system parameters and the learning gains.

Remark 3. Some components of the control decision may vanish, and in this case, the computational overflowing may emerge if the iterative learning control strategy (7) is still used. To avoid such emergence, we shall borrow the information of the previous implementation, which is conducted by the latest nonzero components of the control decision, in order to replace the term produced by the zero components.

3 Simulation result

In order to compare the simulation result of the nonlinear system with that of the linear system, we consider a large-scale system consisting of two subsystems described as

subsystem i:

$$\begin{cases} \dot{\boldsymbol{x}}^{(i)}(t) = A^{(i)} f^{(i)}(\boldsymbol{x}^{(i)}(t)) + B^{(i)}(t)\boldsymbol{r}^{(i)}(t) + \\ \qquad \varepsilon^{(i)} D^{(ij)} \boldsymbol{x}^{(j)}(t) \\ \boldsymbol{y}^{(i)}(t) = C^{(i)}\boldsymbol{x}^{(i)}(t), \quad \boldsymbol{x}^{(i)}(0) = \boldsymbol{0} \quad (i,j=1,2; i \ne j) \end{cases}$$

where

$$\boldsymbol{x}^{(1)}(t) = [x^{(11)}(t), x^{(12)}(t), x^{(13)}(t)]^{\mathrm{T}}$$

$$\boldsymbol{x}^{(2)}(t) = [x^{(21)}(t), x^{(22)}(t), x^{(23)}(t)]^{\mathrm{T}}$$

$$f^{(1)}(\boldsymbol{x}^{(1)}(t)) = [x^{(11)}(t), x^{(12)}(t), \arctan(x^{(13)}(t))]^{\mathrm{T}}$$

$$f^{(2)}(\boldsymbol{x}^{(2)}(t)) = [x^{(21)}(t), x^{(22)}(t), f^{(23)}(x^{(23)}(t))]^{\mathrm{T}}$$

$$\boldsymbol{r}^{(1)}(t) = [r^{(11)}(t), r^{(12)}(t)]^{\mathrm{T}}$$

$$\boldsymbol{y}^{(1)}(t) = [y^{(11)}(t), y^{(12)}(t)]^{\mathrm{T}}$$

$$A^{(1)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.04 & -0.2 & -0.86 \end{bmatrix}$$

$$A^{(2)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & -1.3 \end{bmatrix}$$

$$B^{(1)} = \begin{bmatrix} 0 & 0.2 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{B}^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$C^{(1)} = \begin{bmatrix} 0.05 & 0.1 & 0.02 \\ 0.04 & 0.05 & 0.03 \end{bmatrix}$$

$$\boldsymbol{C}^{(2)} = \begin{bmatrix} 0 & 1 & 0.8 \end{bmatrix}$$

$$D^{(12)} = A^{(1)}, \quad D^{(21)} = A^{(2)}$$

where $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$ are intensities of the interactions of subsystems each other, and $\arctan(x^{13}(t))$ can be considered as a kind of saturation nonlinearity, and

$$f^{(23)}(x^{(23)}) = \begin{cases} \dfrac{1}{3}x^{(23)}, & \text{if } |\, x^{(23)} \,| > 1 \\ x^{(23)} - \dfrac{2}{3}(x^{(23)})^3, & \text{if } |\, x^{(23)} \,| \le 1 \end{cases}$$

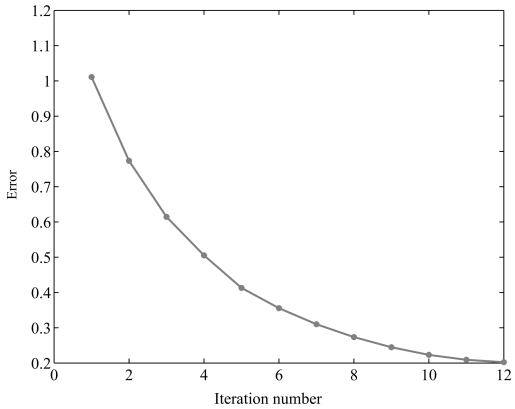
can be viewed as a kind of nonlinearity.

$$\boldsymbol{y}_{d_0}^{(1)}(t) = \begin{bmatrix} y_{d_0}^{(11)}(t) \\ y_{d_0}^{(12)}(t) \end{bmatrix} = \begin{bmatrix} 1-\exp(-0.3t) \\ 1-\exp(-0.2t) \end{bmatrix}$$

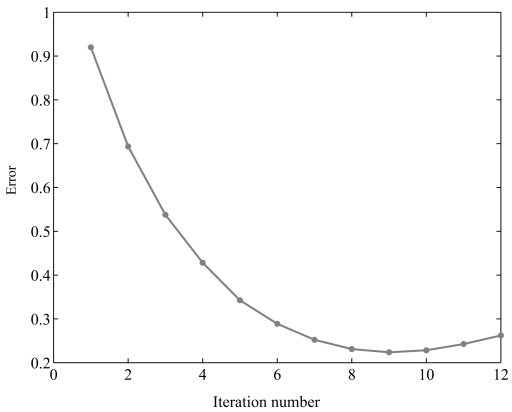
$$y_{d_0}^{(2)}(t) = 1 - \exp(-0.8t)$$

$$\Gamma_p^{(1)} = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix}, \quad \Gamma_d^{(1)} = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.8 \end{bmatrix}$$

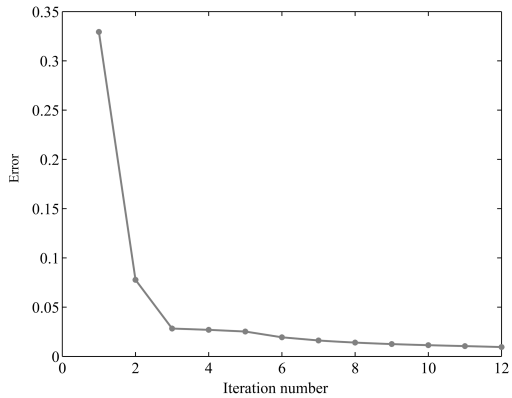
$$\Gamma_p^{(2)} = 0.6, \quad \Gamma_d^{(2)} = 0.7$$



(a) Normalized output error of $y^{(11)}(t)$



(b) Normalized output error of $y^{(12)}(t)$



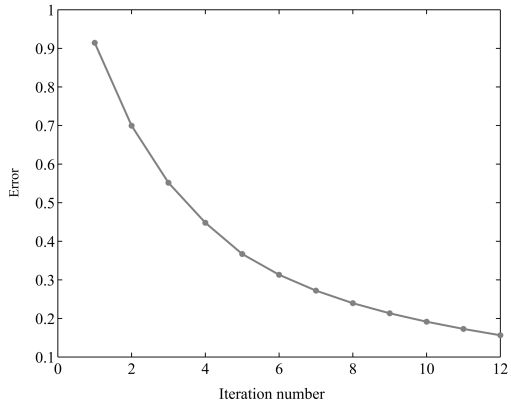
(c) Normalized output error of $y^{(2)}(t)$

Fig. 2 The Lebesgue-2 norm of the normalized out error for nonlinear system

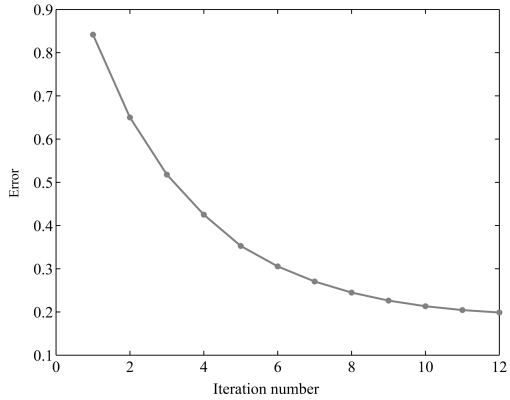
The overall control decisions are proportional, that is,

$$[c_1^{(11)}, c_2^{(11)}, \cdots, c_{12}^{(11)}] = [1.5, 1.4, 1.3, 1.2, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3]$$

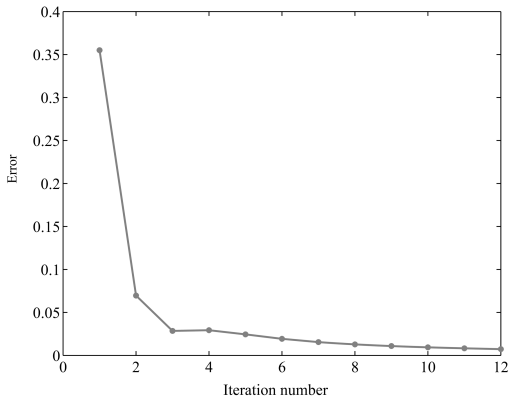
$$[c_1^{(12)}, c_2^{(12)}, \cdots, c_{12}^{(12)}] = 0.8 \times [1.5, 1.4, 1.3, 1.2, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3]$$



(a) Normalized output error of $y^{(11)}(t)$



(b) Normalized output error of $y^{(12)}(t)$



(c) Normalized output error of $y^{(2)}(t)$

Fig. 3 The Lebesgue-2 norm of the normalized output error for the corresponding linearized system

$$[c_1^{(2)}, c_2^{(2)}, \cdots, c_{12}^{(2)}] = 0.5 \times [1.5, 1.4, 1.3, 1.2, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3]$$

The corresponding iterative learning control updating law is as follows:

$$\boldsymbol{r}_k^{(1)}(t) = \begin{bmatrix} r_k^{(11)}(t) \\ r_k^{(12)}(t) \end{bmatrix} = \begin{bmatrix} \dfrac{c_k^{(11)}}{c_{k-1}^{(11)}} & 0 \\ 0 & \dfrac{c_k^{(12)}}{c_{k-1}^{(12)}} \end{bmatrix} \times$$

$$\left(\begin{bmatrix} r_{k-1}^{(11)}(t) \\ r_{k-1}^{(12)}(t) \end{bmatrix} + \Gamma_p^{(1)} \begin{bmatrix} e_{k-1}^{(11)}(t) \\ e_{k-1}^{(12)}(t) \end{bmatrix} + \Gamma_d^{(1)} \begin{bmatrix} \dot{e}_{k-1}^{(11)}(t) \\ \dot{e}_{k-1}^{(12)}(t) \end{bmatrix} \right)$$

$$r_k^{(2)}(t) = \frac{c_k^{(2)}}{c_{k-1}^{(2)}}[r_{k-1}^{(2)}(t) + \Gamma_p^{(2)} e_{k-1}^{(2)}(t) + \Gamma_d^{(2)} \dot{e}_{k-1}^{(2)}(t)]$$

$$t \in [0, 40], k = 1, 2, \cdots, 12.$$

where

$$\begin{bmatrix} e_{k-1}^{(11)}(t) \\ e_{k-1}^{(12)}(t) \end{bmatrix} = \begin{bmatrix} c_{k-1}^{(11)} y_{d_0}^{(11)}(t) - y_{k-1}^{(11)}(t) \\ c_{k-1}^{(12)} y_{d_0}^{(12)}(t) - y_{k-1}^{(12)}(t) \end{bmatrix}$$

$$e_{k-1}^{(2)}(t) = c_{k-1}^{(2)} y_{d_0}^{(2)}(t) - y_{k-1}^{(2)}(t).$$

The interaction intensities are chosen to be $\varepsilon_1 = 0.02$ and $\varepsilon_2 = 0.01$. The Lebesgue-2 norm of the normalized output error for nonlinear system is illustrated in Figs. 2(a)∼(c). The Lebesgue-2 norm of the normalized output error for the corresponding linearized system is shown in Figs. 3(a)∼(c).

Comparing the Lebesgue-2 norm of the normalized output errors of the nonlinear system as shown in Fig. 2 with that of linearized system as exhibited in Fig. 3, we observe that the influence of the nonlinearity is obviously visible. This testifies the validity of the theoretical conclusion.

4 Conclusion

In this paper, an iterative learning controller is imbedded into the procedure of steady-state hierarchical optimization utilizing feedback information. The distinct scales of the control decision sequence are considered by introducing some appropriate amplified gain matrices not only to the formation of the updating law but also to the specification of the desired reference trajectories. The theoretical analysis in the sense of Lebesgue-p norm indicates that the influence factors on the convergence are mainly incurred by the multidimensionality and the nonlinearity. The conclusion is discriminated from the existing result for repetitive tracking task. This phenomenon seems rational since the generalization of the universe of discourse to the nonrepetitive tracking tasks would give rise to the degeneration of the convergence properties. Despite the unavoidable interventions, the proposed iterative learning control scheme can effectively and efficiently improve the dynamic performance of the transient response.

References

1 Bien Z, Xu J. *Iterative Learning Control: Analysis, Design, Integration and Applications*. Boston: Kluwer Academic Publishers, 1998

2 Xie Sheng-Li, Tian Sen-Ping, Xie Zhen-Dong. *Theory and Applications of Iterative Learning Control*. Beijing: Science Press, 2005 (in Chinses)

3 Arif M, Ishihara T, Inooka H. Prediction-based iterative learning control (PILC) for uncertain dynamic nonlinear systems using identification technique. *Journal of Intelligent and Robotic Systems*, 2002, **27**(3): 291−304

4 Sun Ming-Xuan, Wang Dan-Wei. Closed-loop iterative learning control for nonlinear systems with initial shifts. *International Journal of Adaptive Control and Signal Processing*, 2002, **16**: 515−538

5 Tayebi A, Xu J X. Observer-based iterative learning for a class of time-varying nonlinear systems. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, 2003, **50**(3): 452−455

6 Park K H, Bien Z. A study on iterative learning control with adjustment of learning interval for monotone convergence in the sense of sup-norm. *Asian Journal of Control*, 2002, **4**(1): 111−118

7 Park K H, Bien Z. Intervalized iterative learning control for monotonic convergence in the sense of sup-norm. *International Journal of Control*, 2005, **78**(15): 1218−1227

8 Xie Sheng-Li, Tian Sen-Ping, Xie Zhen-Dong. A new algorithm of iterative learning control based on geometric vector plots. *Acta Automatica Sinica*, 2004, **30**(2): 161−168 (in Chinese)

9 Xu J X. Direct learning of control efforts for trajectories with different magnitude scales. *Automatica*, 1997, **33**(12): 2191−2195

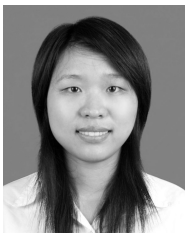
10 Saab S S, Vogt W G, Mickle M H. Learning control algorithm for tracking "slowly" varying trajectories. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 1997, **27**(4): 657−670

11 Xu J X, Xu J. On iterative learning from different tracking tasks in the presence of time-varying uncertainties. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 2004, **34**(1): 589−597

12 Ruan X E, Bien Z Z, Park K H. Decentralized iterative learning control to large-scale industrial processes for nonrepetitive trajectories tracking. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 2008, **38**(1): 238−252

13 Khalil H K. *Nonlinear Systems*. New York: Prentice Hall, 2002

RUAN Xiao-E Received her bachelor and master degrees in mathematics from Shaanxi Normal University in 1988 and 1995, respectively, and Ph. D. degree in control science and engineering from the Institute of Systems Engineering in Xi′an Jiaotong University in 2002. Since 1995, she has been in the Department of Mathematics, Faculty of Science, Xi′an Jiaotong University. From March 2003 to August 2004, she was a postdoctoral researcher with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology. Her research interest covers steady-state hierarchical optimization of large-scale systems, and iterative leaning control. E-mail: wruanxe@mail.xjtu.edu.cn



CHEN Feng-Min Received her bachelor degree from the Department of Applied Mathematics, Liaocheng University, Shandong Province, in 2005. She is now a graduate student in the Department of Mathematics, Faculty of Science, Xi′an Jiaotong University, Xi′an. Her research interest covers iterative learning control, and optimization control.
E-mail: fengminc@gmail.com

WAN Bai-Wu Graduated from the Institute of Communication, Jiao Tong University, Shanghai, in 1951. He has been a professor of control and systems engineering with the Institute of Systems Engineering, Xi′an Jiaotong University. He is an honorary member of the Council of Chinese Association of Automation. He was a member of editorial boards for three top Chinese journals of automatic control, and for the *Proceedings of Institution of Mechanical Engineers, Part I, Journal of Systems and Control Engineering*, United Kingdom, and is a member of Technical Committee of Large Scale Systems of IFAC. His research interest covers large scale systems theory and application, product quality control, and intelligent control. Corresponding author of this paper. E-mail:wanbw@mail.xjtu.edu.cn