

# Robust Optimization-Based Iterative Learning Control for Nonlinear Systems With Nonrepetitive Uncertainties

Deyuan Meng, *Senior Member, IEEE* and Jingyao Zhang

**Abstract**—This paper aims to solve the robust iterative learning control (ILC) problems for nonlinear time-varying systems in the presence of nonrepetitive uncertainties. A new optimization-based method is proposed to design and analyze adaptive ILC, for which robust convergence analysis via a contraction mapping approach is realized by leveraging properties of substochastic matrices. It is shown that robust tracking tasks can be realized for optimization-based adaptive ILC, where the boundedness of system trajectories and estimated parameters can be ensured, regardless of unknown time-varying nonlinearities and nonrepetitive uncertainties. Two simulation tests, especially implemented for an injection molding process, demonstrate the effectiveness of our robust optimization-based ILC results.

**Index Terms**—Adaptive iterative learning control (ILC), nonlinear time-varying system, robust convergence, substochastic matrix.

## I. INTRODUCTION

INTELLIGENT control has generated considerable research interest in both theories and applications of linear/nonlinear systems, where of particular note are the learning-based design methods (see, e.g., [1]–[5]). As a class of effective intelligent control methods with the specific focus on realizing the perfect tracking tasks for the systems that are repetitively executed, iterative learning control (ILC) has been considered as one of the most practically important learning-based design methods in many application fields, see, e.g., [6] for multi-axis robots, [7] for micro aerial vehicles, [8] for linear motor positioning systems, [9] for high-speed trains, and [10] for Chylla-Haase reactors. The readers are referred to the detailed explanations that have been introduced for characteristics and applications of ILC in the surveys of, e.g., [11]–[13]. In particular, ILC has been regarded as one of the most famous and applicable data-driven control methods [14]–[18] that may be alternatively called the model-free control methods [19]–[21], where the accurate models are generally not needed for the ILC algorithm design as well as convergence analysis.

Manuscript received December 15, 2020; accepted January 19, 2021. This work was supported by the National Natural Science Foundation of China (61873013, 61922007). Recommended by Associate Editor Qinglai Wei. (*Corresponding author: Deyuan Meng*)

Citation: D. Y. Meng and J. Y. Zhang, “Robust optimization-based iterative learning control for nonlinear systems with nonrepetitive uncertainties,” *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 5, pp. 1001–1014, May 2021.

The authors are with the Seventh Research Division, Beihang University (BUAA), Beijing 100191, and also with the School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100191, China (e-mail: dymeng@buaa.edu.cn; zhangjingyao@buaa.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/JAS.2021.1003973

Typically, this class of data-driven ILC methods are explored based on an optimization issue that focuses directly on nonlinear systems, regardless of unknown nonlinearities.

In addition to the tight relation to optimization-based design, the data-driven ILC employs a dynamical linearization approach to overcome unknown nonlinearities and an adaptive approach to estimate linearization parameters [15]–[21]. This yields a class of optimization-based adaptive ILC methods that are capable of accommodating unknown dynamics in both nonlinear systems and their dynamical linearization models. For the convergence analysis, optimization-based adaptive ILC adopts a contraction mapping (CM)-based approach that is usually implemented via the eigenvalue analysis. Though it is popularly applied in ILC, the eigenvalue-based CM approach requires ILC processes to have iteration-independent parameters from the perspective of standard linear system theory [22], [23].

It is worth emphasizing that for nonlinear control plants, the dynamical linearization inevitably leads to iteration-dependent model parameters [15]–[21]. This renders the eigenvalue-based CM approach no longer effective in implementing convergence analysis of optimization-based adaptive ILC. Another issue left to settle for optimization-based adaptive ILC is robustness with regard to iteration-dependent uncertainties that are considered to be practically important for ILC [24]–[31]. Actually, the robust issue has not been well studied for optimization-based adaptive ILC (see, e.g., [15]–[21]). It is mainly due to that the iteration-dependent uncertainties may bring challenging difficulties into ILC convergence in the presence of nonrepetitiveness created by iteration-dependent model parameters. To accommodate the effects arising from nonrepetitiveness, new design and analysis approaches for ILC usually need to be explored, see, e.g., [18] for an extended state observer-based design approach and [24], [28] for a double-dynamics analysis (DDA) approach. Despite these new approaches, the eigenvalue analysis is still leveraged in [18], and linear systems are only considered in [24], [28].

In this paper, we are devoted to exploring the robust problem of optimization-based adaptive ILC that is subject to unknown time-varying nonlinearities and nonrepetitive uncertainties due to iteration-dependent initial shifts and disturbances. The main contributions of our established design and analysis results are specified as follows.

1) We propose a new optimization-based design method for

adaptive ILC. This new design method makes it feasible to directly apply the CM-based analysis approach of ILC to develop the boundedness of estimated parameters that are used in our adaptive updating law for the estimation of unknown time-varying nonlinearities.

2) We introduce a new robust convergence analysis method for optimization-based adaptive ILC by implementing a DDA approach and resorting to the use of the properties of the substochastic matrices. This makes it possible to not only accomplish the robust convergence analysis of optimization-based adaptive ILC, but also guarantee the boundedness of all the system trajectories.

3) Our design methods and analysis results of optimization-based adaptive ILC can effectively work, regardless of the presence of nonrepetitive uncertainties. This particularly helps to overcome the drawbacks of those methods and results for optimization-based adaptive ILC established through applying the eigenvalue-based CM approach in, e.g., [16], [17].

In addition, we demonstrate the robust performance of our proposed optimization-based adaptive ILC with two simulation examples, regardless of the initial shifts and disturbances that are varying with respect to both iteration and time.

The rest of our paper is organized as follows. In Section II, a robust tracking problem of optimization-based ILC is given. In Section III, an optimization-based adaptive ILC is accordingly proposed, and the main design and analysis results are derived. Simulation tests and concluding remarks are made in Sections IV and V, respectively.

*Notations:* Let  $\mathbf{1}_n = [1, 1, \dots, 1]^T \in \mathbb{R}^n$ ,  $\mathbb{Z} = \{1, 2, 3, \dots\}$ , and  $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ . Denote  $\mathbb{Z}_T = \{0, 1, \dots, T\}$ ,  $\forall T \in \mathbb{Z}$ , and  $\Delta : l_k(t) \rightarrow \Delta l_k(t) = l_k(t) - l_{k-1}(t)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_+$  as a difference operator of  $l_k(t) \in \mathbb{R}^n$ . For  $A = [a_{ij}] \in \mathbb{R}^{m \times n}$ ,  $\|A\|_2$  and  $\|A\|_\infty$  are the spectral norm and the maximum row sum matrix norm of  $A$ , respectively. If  $a_{ij} \geq 0$ ,  $\forall i = 1, 2, \dots, m$ ,  $\forall j = 1, 2, \dots, n$ ,  $A$  is called a nonnegative matrix and denoted by  $A \geq 0$ , based on which  $B \geq C$  means  $B - C \geq 0$ ,  $\forall B, C \in \mathbb{R}^{m \times n}$ . Define  $|A| = [\|a_{ij}\|]$  from  $A$ , and  $|A| \geq 0$  is a trivial nonnegative matrix. For any  $A \geq 0$ ,  $A$  is called a substochastic matrix if  $A\mathbf{1}_n \leq \mathbf{1}_n$ . Given  $\{D_j \in \mathbb{R}^{n \times m} : j \in \mathbb{Z}_+\}$ , let  $\sum_{j=h}^i D_j = 0$  (that is, the null matrix) if  $h = 0$  and  $i = -1$ ; and when  $m = n$ ,  $\prod_{j=h}^i D_j = D_i D_{i-1} \cdots D_h$  if  $i \geq h$  and  $\prod_{j=h}^i D_j = I$  (that is, the identity matrix) if  $i < h$ .

## II. PROBLEM STATEMENT

Let  $k \in \mathbb{Z}_+$  and  $t \in \mathbb{Z}_T$  denote the iteration and time indices, respectively, and then we consider a nonlinear system with the following input-output dynamics:

$$y_k(t+1) = f(y_k(t), \dots, y_k(t-l), u_k(t), \dots, u_k(t-n), t) + w_k(t)$$

$$\text{with } y_k(i) = \begin{cases} 0, & i < 0 \\ y_0 + \delta_k, & i = 0 \end{cases} \quad \text{and } u_k(i) = 0, \quad i < 0 \quad (1)$$

where  $y_k(t) \in \mathbb{R}$  is the output (with an order  $l \in \mathbb{Z}_+$ ),  $u_k(t) \in \mathbb{R}$  is the input (with an order  $n \in \mathbb{Z}_+$ ),  $w_k(t)$  is the nonrepetitive disturbance,  $\delta_k$  is the nonrepetitive initial shift from a constant output  $y_0$ , and  $f$  is an unknown nonlinear function such that

$$f : \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{l+n+3} \rightarrow \mathbb{R}.$$

In the following analysis, we will write this nonlinear function as  $f$  or  $f(x_1, x_2, \dots, x_{l+n+3})$ , where  $x_i \in \mathbb{R}$  is the  $i$ th independent variable of  $f$  for each  $i = 1, 2, \dots, l+n+3$ .

*Problem Statement:* Let  $e_k(t) = y_d(t) - y_k(t)$  be the tracking error between the output  $y_k(t)$  and the desired output  $y_d(t) \in \mathbb{R}$  over  $k \in \mathbb{Z}_+$  and  $t \in \mathbb{Z}_T$ , and assume  $e_i(t) = 0$ ,  $\forall t \in \mathbb{Z}_T$  if  $i < 0$ . Then, we target at designing an ILC algorithm for the nonlinear system (1) to achieve a perfect tracking task as

$$\lim_{k \rightarrow \infty} e_k(t+1) = 0, \quad \forall t \in \mathbb{Z}_{T-1} \quad (2)$$

via the solving of an optimization problem with the following index over  $k \in \mathbb{Z}$  and  $t \in \mathbb{Z}_{T-1}$  (see also [16], [17]):

$$J(u_k(t)) = \left[ \sum_{i=1}^m \gamma_i e_{k-i+1}(t+1) \right]^2 + \lambda [\Delta u_k(t)]^2 \quad (3)$$

where  $\gamma_i > 0$ ,  $i = 1, 2, \dots, m$  and  $\lambda > 0$  are learning parameters to be selected, and  $m \in \mathbb{Z}$  is the order to denote tracking errors considered in (3) over iterations. However, due to the presence of nonrepetitive uncertainties, the perfect tracking task (2) may no longer be achieved in general, and instead a robust tracking task is usually considered of practical importance such that the tracking error can be decreased to a small neighborhood of the origin with increasing iterations, namely,

$$\limsup_{k \rightarrow \infty} |e_k(t+1)| \leq \beta_{e_{\sup}}, \quad \forall t \in \mathbb{Z}_{T-1} \quad (4)$$

where  $\beta_{e_{\sup}}(t) > 0$  is a small bound that depends continuously on the variation bounds of the nonrepetitive uncertainties.

To perform the abovementioned robust tracking tasks for the nonlinear system (1), we need two basic assumptions about the continuous differentiability of the unknown nonlinear function and the boundedness of nonrepetitive uncertainties.

*Assumption 1:* Let  $f$  be a continuously differentiable nonlinear function such that 1)  $\partial f / \partial x_i$ ,  $\forall i = 1, 2, \dots, l+n+2$  is bounded, and 2)  $\partial f / \partial x_{l+2}$  is sign-fixed. To be specific, we assume some finite upper bound  $\beta_f^- > 0$  such that

$$\left| \frac{\partial f}{\partial x_i} (x_1, x_2, \dots, x_{l+n+2}, t) \right| \leq \beta_f^- \quad \forall x_i \in \mathbb{R}, \quad i = 1, 2, \dots, l+n+2, \quad \forall t \in \mathbb{Z}_{T-1} \quad (5)$$

and without loss of generality, assume that  $\partial f / \partial x_{l+2}$  is positive and has some finite lower bound  $\beta_f^+ > 0$ , namely,

$$\begin{aligned} \frac{\partial f}{\partial x_{l+2}} (x_1, x_2, \dots, x_{l+n+2}, t) &\geq \beta_f^+ \\ \forall x_i \in \mathbb{R}, \quad i &= 1, 2, \dots, l+n+2, \quad \forall t \in \mathbb{Z}_{T-1}. \end{aligned} \quad (6)$$

*Assumption 2:* Let  $w_k(t)$  and  $\delta_k$  be bounded nonrepetitive uncertainties such that

$$\begin{aligned} |w_k(t)| &\leq \beta_w(t), \quad \forall k \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_{T-1} \\ |\delta_k| &\leq \beta_\delta, \quad \forall k \in \mathbb{Z}_+ \end{aligned} \quad (7)$$

for some finite bounds  $\beta_w(t) > 0$  and  $\beta_\delta > 0$ .

*Remark 1:* The considered nonlinear system (1) includes, as a particular case, the class of deterministic autoregressive

moving average models, which are typically seen in the applications of ILC (see, e.g., an injection molding process [32]). Moreover, (1) can be modeled to represent the input-output relation for a class of affine nonlinear systems in the state-space description form of

$$x_k(t+1) = h(x_k(t), t) + b(t)u_k(t) + w_k^x(t)$$

$$y_k(t) = c(t)x_k(t) + w_k^y(t)$$

for some state  $x_k(t) \in \mathbb{R}^n$ , disturbances  $w_k^x(t) \in \mathbb{R}^n$  and  $w_k^y(t) \in \mathbb{R}$ , nonlinear function  $h: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ , and system matrices  $b(t) \in \mathbb{R}^n$  and  $c(t) \in \mathbb{R}^{1 \times n}$  fulfilling  $c(t+1)b(t) \neq 0$ ,  $\forall t \in \mathbb{Z}_{T-1}$ . It is worth noticing that this class of affine nonlinear systems, covered by (1), are popularly considered in nonlinear ILC owing to the wide applications [33].

*Remark 2:* Since the nonlinear time-varying dynamics of (1) are unknown, Assumption 1 is a commonly used condition that provides basic guarantees for both dynamical linearization of unknown nonlinearities and optimization-based design of adaptive ILC (see, e.g., [15]–[21]). In particular, it follows from (5) and (6) in Assumption 1 that  $\partial f / \partial x_{l+2}$  satisfies

$$\frac{\partial f}{\partial x_{l+2}}(x_1, x_2, \dots, x_{l+n+2}, t) \in [\beta_f, \bar{\beta}_f]$$

$$\forall x_i \in \mathbb{R}, i = 1, 2, \dots, l+n+2, \quad \forall t \in \mathbb{Z}_{T-1}. \quad (8)$$

*Remark 3:* It is worth noting that Assumption 2 is a commonly considered ILC condition on the class of nonrepetitive uncertainties since it is generally acceptable for many practical applications (see, e.g., [24], [28]). Of particular note is to use a further condition on the convergence of nonrepetitive uncertainties such that

$$\lim_{k \rightarrow \infty} [w_k(t) - w_{k-1}(t)] = 0, \quad \forall t \in \mathbb{Z}_{T-1}, \quad \lim_{k \rightarrow \infty} (\delta_k - \delta_{k-1}) = 0 \quad (9)$$

which may be considered as an additional requirement of Assumption 2 for the accomplishment of the perfect tracking task (2), despite the presence of nonrepetitive uncertainties. A trivial case most considered for (9) is the absence of nonrepetitive uncertainties in ILC, namely, (9) collapses into

$$w_k(t) \equiv w(t), \quad \forall k \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_{T-1}, \quad \delta_k \equiv \delta, \quad \forall k \in \mathbb{Z}_+ \quad (10)$$

for some iteration-independent quantities  $w(t)$  and  $\delta$ . It should be noticed that (10) can imply (9), but not vice versa. Actually, the objective of robust ILC is realized if (4) can be achieved in the presence of nonrepetitive uncertainties satisfying (7), and furthermore (2) can be achieved in the absence of nonrepetitive uncertainties (i.e., (9) or (10) holds).

### III. MAIN RESULTS

#### A. Design of Optimization-Based Adaptive ILC

We design optimization-based adaptive ILC to overcome the effect of unknown nonlinear time-varying dynamics on seeking output tracking objectives for (1), regardless of the presence of nonrepetitive uncertainties. Towards this end, we establish the following lemma to derive an extended dynamical linearization for the unknown nonlinear time-varying dynamics of (1).

*Lemma 1:* If Assumption 1 is satisfied for the nonlinear system (1), then an extended dynamical linearization can be

given for (1) as

$$\begin{aligned} \begin{bmatrix} y_i(1) \\ y_i(2) \\ \vdots \\ y_i(T) \end{bmatrix} - \begin{bmatrix} y_j(1) \\ y_j(2) \\ \vdots \\ y_j(T) \end{bmatrix} &= \Theta_{i,j} \left( \begin{bmatrix} u_i(0) \\ u_i(1) \\ \vdots \\ u_i(T-1) \end{bmatrix} - \begin{bmatrix} u_j(0) \\ u_j(1) \\ \vdots \\ u_j(T-1) \end{bmatrix} \right) \\ &\quad + \Upsilon_{i,j} \left( \begin{bmatrix} w_i(0) \\ w_i(1) \\ \vdots \\ w_i(T-1) \end{bmatrix} - \begin{bmatrix} w_j(0) \\ w_j(1) \\ \vdots \\ w_j(T-1) \end{bmatrix} \right) \\ &\quad + \begin{bmatrix} \vartheta_{i,j,0} \\ \vartheta_{i,j,1} \\ \vdots \\ \vartheta_{i,j,T-1} \end{bmatrix} (\delta_i - \delta_j), \quad \forall i, j \in \mathbb{Z}_+ \end{aligned} \quad (11)$$

where for any  $i, j \in \mathbb{Z}_+$ ,  $\Theta_{i,j}$  and  $\Upsilon_{i,j}$  are some bounded lower triangular matrices in the form of

$$\Theta_{i,j} = \begin{bmatrix} \theta_{i,j,0}(0) & 0 & \cdots & 0 \\ \theta_{i,j,1}(0) & \theta_{i,j,1}(1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \theta_{i,j,T-1}(0) & \cdots & \theta_{i,j,T-1}(T-2) & \theta_{i,j,T-1}(T-1) \end{bmatrix}$$

$$\Upsilon_{i,j} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ v_{i,j,1}(0) & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ v_{i,j,T-1}(0) & \cdots & v_{i,j,T-1}(T-2) & 1 \end{bmatrix}$$

and  $\vartheta_{i,j,t}$ ,  $\forall t \in \mathbb{Z}_{T-1}$  is some bounded parameter. Namely, for some finite bound  $\beta_\theta > 0$ ,

$$\begin{aligned} |\theta_{i,j,t}(\xi)| &\leq \beta_\theta, \quad \forall \xi \in \mathbb{Z}_t, \forall t \in \mathbb{Z}_{T-1}, \forall i, j \in \mathbb{Z}_+ \\ |v_{i,j,t}(\xi)| &\leq \beta_\theta, \quad \forall \xi \in \mathbb{Z}_{t-1}, \forall t \in \mathbb{Z}_{T-1}, \forall i, j \in \mathbb{Z}_+ \\ |\vartheta_{i,j,t}| &\leq \beta_\theta, \quad \forall t \in \mathbb{Z}_{T-1}, \forall i, j \in \mathbb{Z}_+ \end{aligned} \quad (12)$$

where, more precisely,  $\theta_{i,j,t}(t)$  fulfills

$$\theta_{i,j,t}(t) \in [\beta_f, \bar{\beta}_f], \quad \forall t \in \mathbb{Z}_{T-1}, \forall i, j \in \mathbb{Z}_+. \quad (13)$$

*Proof:* With Assumption 1, we leverage the differential mean value theorem and can then develop this lemma by considering the derivation rules of compound functions. For clarity, we provide the proof details in the Appendix, which can also be found at <https://arxiv.org/abs/1908.02447>. ■

Of specific interest is the application of Lemma 1 to disclose the input-output relation between two sequential iterations for the nonlinear system (1). Namely, by taking  $i = k$  and  $j = k-1$  in (11), we have

$$\begin{aligned} \Delta y_k(t+1) &= \sum_{i=0}^t \theta_{k,k-1,t}(i) \Delta u_k(i) \\ &\quad + \Delta w_k(t) + \sum_{i=0}^{t-1} v_{k,k-1,t}(i) \Delta w_k(i) + \vartheta_{k,k-1,t} \Delta \delta_k \\ &\triangleq \overrightarrow{\Delta u_k}(t) \overrightarrow{\theta_{k,k-1,t}}(t) + \varphi_k(t), \quad \forall t \in \mathbb{Z}_{T-1}, \forall k \in \mathbb{Z} \end{aligned} \quad (14)$$

where  $\overrightarrow{\Delta u_k}(t)$ ,  $\overrightarrow{\theta_{k,k-1,t}}(t)$  and  $\varphi_k(t)$  are given by

$$\begin{aligned}\vec{u}_k(t) &= [u_k(0), u_k(1), \dots, u_k(t)]^T \\ \overrightarrow{\theta_{k,k-1,t}}(t) &= [\theta_{k,k-1,t}(0), \theta_{k,k-1,t}(1), \dots, \theta_{k,k-1,t}(t)]^T \\ \varphi_k(t) &= \Delta w_k(t) + \sum_{i=0}^{t-1} v_{k,k-1,t}(i) \Delta w_k(i) + \vartheta_{k,k-1,t} \Delta \delta_k.\end{aligned}$$

It is obvious from (14) that the linearization parameters help to describe the dynamic evolution of ILC along the iteration axis. However, Lemma 1 shows that the linearization parameters are unknown. In addition, the nonrepetitive uncertainties also play an important role in influencing the dynamic evolution of (14). These render it hard to find the optimal solution to (3). Toward this end, we select a bounded initial input  $u_0(t)$  over  $t \in \mathbb{Z}_{T-1}$ , and then calculate  $\partial J(u_k(t)) / \partial u_k(t) = 0$  and find an estimation  $\hat{\theta}_{k,k-1,t}(i)$  instead of employing  $\theta_{k,k-1,t}(i)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$ ,  $\forall i \in \mathbb{Z}_t$  such that we can present an updating law with respect to the input as

$$\begin{aligned}u_k(t) &= u_{k-1}(t) - \frac{\gamma_1^2 \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \sum_{i=0}^{t-1} \hat{\theta}_{k,k-1,t}(i) [u_k(i) \\ &\quad - u_{k-1}(i)] + \frac{\gamma_1 \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \left[ \gamma_1 e_{k-1}(t+1) \right. \\ &\quad \left. + \sum_{i=2}^m \gamma_i e_{k-i+1}(t+1) \right], \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1}. \quad (15)\end{aligned}$$

To implement (15), we need to determine  $\hat{\theta}_{k,k-1,t}(i)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$ ,  $\forall i \in \mathbb{Z}_t$ , for which we also resort to an optimization-based approach. Let

$$\overrightarrow{\hat{\theta}_{k,k-1,t}}(t) = [\hat{\theta}_{k,k-1,t}(0), \hat{\theta}_{k,k-1,t}(1), \dots, \hat{\theta}_{k,k-1,t}(t)]^T$$

and then we leverage an optimization index as

$$\begin{aligned}H\left(\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right) &= \left[ \Delta y_{k-1}(t+1) - \Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k,k-1,t}}(t) \right]^2 \\ &\quad + \mu_1 \left\| \overrightarrow{\hat{\theta}_{k,k-1,t}}(t) - \overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t) \right\|_2^2 \\ &\quad + \mu_2 \left\| \overrightarrow{\hat{\theta}_{k,k-1,t}}(t) \right\|_2^2, \quad \forall t \in \mathbb{Z}_{T-1}, \forall k \geq 2 \quad (16)\end{aligned}$$

where  $\mu_1 \geq 0$  and  $\mu_2 \geq 0$  are two weighting factors. In contrast to, e.g., [15]–[18], we introduce a novel optimization index in (16) to estimate unknown system parameters for the optimization-based adaptive ILC by adding the third term in (16). This, however, is crucial in exploring the properties of the estimated parameters, especially the boundedness and robustness of them in the presence of nonrepetitive uncertainties. To proceed, we give an adaptive updating scheme for the parameter estimation based on solving  $\partial H\left(\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right) / \partial \overrightarrow{\hat{\theta}_{k,k-1,t}}(t) = 0$ , where the following two separate steps are involved.

*S1) Initialization:* Select a bounded estimation  $\hat{\theta}_{1,0,t}(i)$ ,  $\forall t \in \mathbb{Z}_{T-1}$ ,  $\forall i \in \mathbb{Z}_t$ , where for any (small) scalar  $\varepsilon > 0$ , choose  $\hat{\theta}_{1,0,t}(t)$  to satisfy

$$\hat{\theta}_{1,0,t}(t) \geq \varepsilon, \quad \forall t \in \mathbb{Z}_{T-1}. \quad (17)$$

*S2) Update:* Implement an updating law about the parameter estimation, such that

$$\begin{aligned}\hat{\theta}_{k,k-1,t}(i) &= \frac{\mu_1}{\mu_1 + \mu_2} \hat{\theta}_{k-1,k-2,t}(i) \\ &\quad + \frac{\Delta u_{k-1}(i)}{\mu_1 + \mu_2 + \sum_{j=0}^t \Delta u_{k-1}^2(j)} \left[ \Delta y_{k-1}(t+1) \right. \\ &\quad \left. - \frac{\mu_1}{\mu_1 + \mu_2} \sum_{j=0}^t \hat{\theta}_{k-1,k-2,t}(j) \Delta u_{k-1}(j) \right] \\ &\quad \forall k \geq 2, \quad \forall t \in \mathbb{Z}_{T-1}, \quad \forall i \in \mathbb{Z}_t \quad (18)\end{aligned}$$

where, if  $\hat{\theta}_{k,k-1,t}(t) < \varepsilon$  occurs, then let us directly set

$$\hat{\theta}_{k,k-1,t}(t) = \hat{\theta}_{1,0,t}(t), \quad \forall k \geq 2, \forall t \in \mathbb{Z}_{T-1}. \quad (19)$$

It is worth highlighting that in the above optimization-based design of adaptive ILC, the effect of nonrepetitive uncertainties is not directly reflected. In spite of this fact, it will be disclosed that our optimization-based adaptive ILC is of robustness with respect to nonrepetitive uncertainties.

### B. Robust Convergence Analyses of ILC

Next, the robust convergence analysis of optimization-based adaptive ILC for the nonlinear system (1) is explored. Towards this end, the dynamics of the tracking error are considered, and by combining (15) with (14), it can be verified that

$$\begin{aligned}e_k(t+1) &= e_{k-1}(t+1) - \Delta y_k(t+1) \\ &= e_{k-1}(t+1) - \sum_{i=0}^t \theta_{k,k-1,t}(i) \Delta u_k(i) \\ &\quad - \Delta w_k(t) - \sum_{i=0}^{t-1} v_{k,k-1,t}(i) \Delta w_k(i) - \vartheta_{k,k-1,t} \Delta \delta_k \\ &= \left[ 1 - \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \theta_{k,k-1,t}(t) \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \right] e_{k-1}(t+1) \\ &\quad - \sum_{i=3}^m \frac{\gamma_1 \gamma_i \theta_{k,k-1,t}(i) \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} e_{k-i+1}(t+1) \\ &\quad + \kappa_k(t), \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1} \quad (20)\end{aligned}$$

where  $\kappa_k(t)$  is given by

$$\begin{aligned}\kappa_k(t) &= \frac{\gamma_1^2 \theta_{k,k-1,t}(t) \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \sum_{i=0}^{t-1} \hat{\theta}_{k,k-1,t}(i) \Delta u_k(i) \\ &\quad - \sum_{i=0}^{t-1} \theta_{k,k-1,t}(i) \Delta u_k(i) - \Delta w_k(t) \\ &\quad - \sum_{i=0}^{t-1} v_{k,k-1,t}(i) \Delta w_k(i) - \vartheta_{k,k-1,t} \Delta \delta_k. \quad (21)\end{aligned}$$

From (20), it is clear that the evolution process of the tracking error is iteration-dependent owing to its explicit dependence on  $\theta_{k,k-1,t}(t)$  and  $\hat{\theta}_{k,k-1,t}(t)$ . This renders the eigenvalue

analysis not applicable any longer for developing the ILC convergence based on (20). By such observation, we leverage the properties of substochastic matrices to develop a CM-based approach for the ILC convergence analysis, together with combining a DDA approach to the optimization-based adaptive ILC.

To proceed with the ILC convergence analysis based on (20), we explore the boundedness property of the uncertain parameter  $\theta_{k,k-1,t}(i)$  and its estimation  $\hat{\theta}_{k,k-1,t}(i)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$ ,  $\forall i \in \mathbb{Z}_t$ . Since Lemma 1 shows the boundedness of each parameter  $\theta_{k,k-1,t}(i)$ , we only need to develop that of the estimation  $\hat{\theta}_{k,k-1,t}(i)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$ ,  $\forall i \in \mathbb{Z}_t$ , as shown in the following theorem.

**Theorem 1:** For the nonlinear system (1) with Assumptions 1 and 2, let the updating law (15) for the input and the adaptive updating schemes S1) and S2) for the parameter estimation be applied. If  $\mu_1 > 0$  and  $\mu_2 > 0$ , then the boundedness of the estimation  $\hat{\theta}_{k,k-1,t}(i)$  can be ensured such that for some finite bound  $\beta_{\hat{\theta}} > 0$ ,

$$|\hat{\theta}_{k,k-1,t}(i)| \leq \beta_{\hat{\theta}}, \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1}, \forall i \in \mathbb{Z}_t \quad (22)$$

where  $\hat{\theta}_{k,k-1,t}(t)$  particularly satisfies

$$\hat{\theta}_{k,k-1,t}(t) \in [\varepsilon, \beta_{\hat{\theta}}], \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1}. \quad (23)$$

*Proof:* From (18), we can equivalently derive

$$\begin{aligned} \overrightarrow{\hat{\theta}_{k,k-1,t}(t)} &= \frac{\mu_1}{\mu_1 + \mu_2} \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \\ &+ \frac{1}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \left[ \Delta y_{k-1}(t+1) \right. \\ &\quad \left. - \frac{\mu_1}{\mu_1 + \mu_2} \Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \right] \Delta \overrightarrow{u_{k-1}}(t). \end{aligned} \quad (24)$$

If we define a symmetric matrix as

$$Q(\Delta \overrightarrow{u_{k-1}}(t)) = I - \frac{\Delta \overrightarrow{u_{k-1}}(t) \Delta \overrightarrow{u_{k-1}}^T(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \quad (25)$$

then  $\|Q(\Delta \overrightarrow{u_{k-1}}(t))\|_2 \leq 1$  holds. Since  $\Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)}$  is a scalar, we can arrive at  $\Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \Delta \overrightarrow{u_{k-1}}(t) = \Delta \overrightarrow{u_{k-1}}(t) \Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)}$ . Then, with the help of (25), we can deduce further from (24) that

$$\begin{aligned} \overrightarrow{\hat{\theta}_{k,k-1,t}(t)} &= \frac{\mu_1}{\mu_1 + \mu_2} \left[ \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \right. \\ &\quad \left. - \frac{\Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \right] \\ &\quad + \frac{\Delta y_{k-1}(t+1) \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \\ &= \frac{\mu_1}{\mu_1 + \mu_2} \left[ \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \right. \\ &\quad \left. - \frac{\Delta \overrightarrow{u_{k-1}}(t) \Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)}}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \right] \end{aligned}$$

$$\begin{aligned} &+ \frac{\Delta y_{k-1}(t+1) \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \\ &= \frac{\mu_1}{\mu_1 + \mu_2} \left[ I - \frac{\Delta \overrightarrow{u_{k-1}}(t) \Delta \overrightarrow{u_{k-1}}^T(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \right] \\ &\quad \times \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} + \frac{\Delta y_{k-1}(t+1) \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \\ &= \left[ \frac{\mu_1}{\mu_1 + \mu_2} Q(\Delta \overrightarrow{u_{k-1}}(t)) \right] \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \\ &\quad + \frac{\Delta y_{k-1}(t+1) \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2}. \end{aligned} \quad (26)$$

By combining (7) and (12) with (14), we can validate

$$\begin{aligned} |\Delta y_{k-1}(t+1)| &\leq \left| \Delta \overrightarrow{u_{k-1}}^T(t) \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \right| + |\varphi_{k-1}(t)| \\ &\leq \|\Delta \overrightarrow{u_{k-1}}(t)\|_2 \sqrt{\sum_{i=0}^t \theta_{k-1,k-2,t}^2(i) + |\varphi_{k-1}(t)|} \\ &\leq \sqrt{t+1} \beta_{\theta} \|\Delta \overrightarrow{u_{k-1}}(t)\|_2 \\ &\quad + 2\beta_w(t) + 2 \sum_{i=0}^{t-1} \beta_{\theta} \beta_w(i) + 2\beta_{\theta} \beta_{\delta} \\ &\leq \sqrt{T} \beta_{\theta} \|\Delta \overrightarrow{u_{k-1}}(t)\|_2 \\ &\quad + 2\beta_w + 2T\beta_{\theta}\beta_w + 2\beta_{\theta}\beta_{\delta} \end{aligned} \quad (27)$$

where  $\beta_w = \max_{t \in \mathbb{Z}_{T-1}} \beta_w(t)$ . The use of (27) further yields

$$\begin{aligned} &\left\| \frac{\Delta y_{k-1}(t+1) \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \right\|_2 \\ &\leq \frac{\sqrt{T} \beta_{\theta} \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \\ &\quad + \frac{(2\beta_w + 2T\beta_{\theta}\beta_w + 2\beta_{\theta}\beta_{\delta}) \|\Delta \overrightarrow{u_{k-1}}(t)\|_2}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \\ &\leq \sqrt{T} \beta_{\theta} + \frac{\beta_w + T\beta_{\theta}\beta_w + \beta_{\theta}\beta_{\delta}}{\sqrt{\mu_1 + \mu_2}}. \end{aligned} \quad (28)$$

Due to  $\|Q(\Delta \overrightarrow{u_{k-1}}(t))\|_2 \leq 1$ , we leverage (28) and can employ (26) to obtain

$$\begin{aligned} \left\| \overrightarrow{\hat{\theta}_{k,k-1,t}(t)} \right\|_2 &\leq \left\| \frac{\mu_1}{\mu_1 + \mu_2} Q(\Delta \overrightarrow{u_{k-1}}(t)) \right\|_2 \left\| \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \right\|_2 \\ &\quad + \left\| \frac{\Delta y_{k-1}(t+1) \Delta \overrightarrow{u_{k-1}}(t)}{\mu_1 + \mu_2 + \|\Delta \overrightarrow{u_{k-1}}(t)\|_2^2} \right\|_2 \\ &\leq \frac{\mu_1}{\mu_1 + \mu_2} \left\| \overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)} \right\|_2 \\ &\quad + \sqrt{T} \beta_{\theta} + \frac{\beta_w + T\beta_{\theta}\beta_w + \beta_{\theta}\beta_{\delta}}{\sqrt{\mu_1 + \mu_2}}. \end{aligned} \quad (29)$$

A straightforward consequence of (29) is

$$\begin{aligned} \left\| \overrightarrow{\hat{\theta}_{k,k-1,t}(t)} \right\|_2 &\leq \left( \frac{\mu_1}{\mu_1 + \mu_2} \right)^{k-1} \left\| \overrightarrow{\hat{\theta}_{1,0,t}(t)} \right\|_2 + \sum_{i=0}^{k-2} \left( \frac{\mu_1}{\mu_1 + \mu_2} \right)^i \\ &\quad \times \left( \sqrt{T} \beta_\theta + \frac{\beta_w + T \beta_\theta \beta_w + \beta_\theta \beta_\delta}{\sqrt{\mu_1 + \mu_2}} \right). \end{aligned} \quad (30)$$

Since  $\mu_1 > 0$  and  $\mu_2 > 0$  ensure  $\mu_1 / (\mu_1 + \mu_2) < 1$ , we directly use (30) to derive

$$\begin{aligned} \left\| \overrightarrow{\hat{\theta}_{k,k-1,t}(t)} \right\|_2 &\leq \left\| \overrightarrow{\hat{\theta}_{1,0,t}(t)} \right\|_2 \\ &\quad + \frac{\mu_1 + \mu_2}{\mu_2} \left( \sqrt{T} \beta_\theta + \frac{\beta_w + T \beta_\theta \beta_w + \beta_\theta \beta_\delta}{\sqrt{\mu_1 + \mu_2}} \right) \\ &\leq \beta_\theta, \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1} \end{aligned} \quad (31)$$

where

$$\beta_\theta = \max_{t \in \mathbb{Z}_{T-1}} \left\| \overrightarrow{\hat{\theta}_{1,0,t}(t)} \right\|_2 + \frac{\mu_1 + \mu_2}{\mu_2} \left( \sqrt{T} \beta_\theta + \frac{\beta_w + T \beta_\theta \beta_w + \beta_\theta \beta_\delta}{\sqrt{\mu_1 + \mu_2}} \right).$$

Owing to  $|\hat{\theta}_{k,k-1,t}(i)| \leq \left\| \overrightarrow{\hat{\theta}_{k,k-1,t}(t)} \right\|_2$ ,  $\forall i \in \mathbb{Z}_t$ , (31) guarantees (22). In particular, we can verify (23) because (19) also ensures  $\hat{\theta}_{k,k-1,t}(t) \geq \varepsilon$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$ . ■

In Theorem 1, it discloses that the boundedness of estimated parameters depends only on the nonnegative weighting factors  $\mu_1$  and  $\mu_2$  used in the parameter estimation updating law (18), and is independent of the positive learning parameters  $\lambda$  and  $\gamma_i$ ,  $i = 1, 2, \dots, m$  used in the input updating law (15). Actually, it requires  $\mu_1 > 0$  and  $\mu_2 > 0$  to accomplish Theorem 1, which is achieved through implementing the CM-based approach for the boundedness analysis of ILC.

*Remark 4:* The similar idea of CM-based approach has been applied to exploit the boundedness of estimated parameters for the optimization-based adaptive ILC in, e.g., [16], [17]. By contrast,  $\mu_2 = 0$  is however used in [16], [17]. In this special case, although a step-size factor  $\varsigma \in (0, 2)$  is added to modify  $Q(\Delta \overrightarrow{u_{k-1}}(t))$  as

$$Q(\Delta \overrightarrow{u_{k-1}}(t)) = I - \frac{\varsigma \Delta \overrightarrow{u_{k-1}}(t) \Delta \overrightarrow{u_{k-1}}^T(t)}{\mu_1 + \left\| \Delta \overrightarrow{u_{k-1}}(t) \right\|_2^2}$$

it can only be obtained that  $\left\| Q(\Delta \overrightarrow{u_{k-1}}(t)) \right\|_2 \leq 1$ , and thus, the CM-based approach can not be performed for the boundedness analysis of estimated parameters any longer. On the contrary, our new design overcomes such drawback by selecting  $\mu_2 > 0$ , and consequently makes an essential contribution to the design of optimization-based adaptive ILC.

With Theorem 1, we proceed to develop robust convergence of ILC by achieving the boundedness of the system trajectories and the convergence of the tracking error, which is established in the following theorem by leveraging a DDA approach.

*Theorem 2:* Consider the nonlinear system (1) satisfying Assumptions 1 and 2, and let the updating law (15) for the input and the adaptive updating schemes S1 and S2 for the parameter estimation be applied. If

$$\gamma_1 + \gamma_2 > \sum_{i=3}^m \gamma_i, \quad \lambda > (\gamma_1^2 + \gamma_1 \gamma_2) \beta_{\bar{f}} \beta_\theta, \quad \mu_1 > 0, \quad \mu_2 > 0 \quad (32)$$

then the following results for boundedness and convergence of the optimization-based adaptive ILC hold.

1) The boundedness can be guaranteed for both input and output trajectories such that

$$\begin{cases} |u_k(t)| \leq \beta_u, & \forall k \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_{T-1} \\ |y_k(t)| \leq \beta_y, & \forall k \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_T \end{cases} \quad (33)$$

for some finite bounds  $\beta_u > 0$  and  $\beta_y > 0$ .

2) The robust tracking objective (4) of ILC can be realized. Further, the perfect tracking objective (2) of ILC can be achieved (with its limit being approached exponentially fast), provided that (9) is ensured (with an exponentially fast speed).

*Remark 5:* In view of Theorem 2, it can be obviously stated that our approach for the optimization-based adaptive ILC is not only applicable for accommodating unknown nonlinear time-varying dynamics but also effective in overcoming ill effects of nonrepetitive uncertainties. This benefits from our new design of the optimization-based adaptive ILC, together with the use of a DDA approach to its convergence analysis. Further, it is worth emphasizing that it is generally difficult to obtain robustness of data-driven ILC in the presence of nonrepetitive uncertainties, e.g., see [15]–[21]. By contrast, Theorem 2 successfully shows the robust analysis of data-driven ILC, in spite of nonrepetitive uncertainties arising from disturbances and initial shifts.

*Remark 6:* For the optimization-based adaptive ILC of nonlinear systems, the parameters  $\gamma_i$ ,  $i = 1, 2, \dots, m$ ,  $\lambda$ ,  $\mu_1$ , and  $\mu_2$  are generally selected according to the condition (32). There are also some specific selections of the learning parameters  $\gamma_i$ ,  $i = 1, 2, \dots, m$  and  $\lambda$ , such as the first-order case that is most considered for optimization-based adaptive ILC (i.e., by setting  $m = 1$  in (3)). Our established results particularly work for this popular case, where our design and analysis only need to take  $\gamma_i = 0$ ,  $i = 2, 3, \dots, m$ . More specifically, when  $m = 1$  holds, the only modification of our optimization-based adaptive ILC is that (15) collapses into

$$\begin{aligned} u_k(t) &= u_{k-1}(t) + \frac{\gamma_1^2 \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \left\{ e_{k-1}(t+1) \right. \\ &\quad \left. - \sum_{i=0}^{t-1} \hat{\theta}_{k,k-1,t}(i) [u_k(i) - u_{k-1}(i)] \right\} \end{aligned}$$

and accordingly, the selection conditions become  $\lambda > \gamma_1^2 \beta_{\bar{f}} \beta_\theta$ ,  $\mu_1 > 0$  and  $\mu_2 > 0$  such that the robust ILC results of Theorems 1 and 2 are still effective. For the two weighting factors  $\mu_1$  and  $\mu_2$ , we generally only require them to be positive, as shown in Theorems 1 and 2. From the implementation of (15) (see also the boundedness result (23)),  $\hat{\theta}_{k,k-1,t}(t)$  should not be zero, due to which a resetting condition (19) is involved in the adaptive updating schemes S1 and S2 for the parameter estimation. This performance actually depends on the selections of  $\mu_1$  and  $\mu_2$ , where we may directly set  $\mu_1 = 1$  without loss of generality and make  $\mu_2$  be relatively smaller than  $\mu_1$ . It is thus possible to simultaneously guarantee the boundedness of the estimated parameters  $\hat{\theta}_{k,k-1,t}(i)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$ ,  $\forall i \in \mathbb{Z}_t$  and avoid leading to a fast decreasing process of  $\hat{\theta}_{k,k-1,t}(t)$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$  along the

iteration axis according to (26).

To prove Theorem 2, we give some helpful properties for the robust convergence of the tracking error and the boundedness of the input, where we resort to the properties of substochastic matrices and nonnegative matrices. For clarity of our analysis, we introduce some basic facts related to nonnegative matrices.

*Lemma 2:* For any matrices  $A = [a_{ij}] \in \mathbb{R}^{n \times n}$  and  $B = [b_{ij}] \in \mathbb{R}^{n \times n}$ , two properties hold as follows:

- 1)  $|AB| \leq |A||B|$ ;
- 2)  $\|A\|_\infty = \| |A| \|_\infty$ .

Further, if  $A \geq 0$  and  $B \geq 0$ , then the following properties hold:

- 3)  $A \leq B$  ensures  $\|A\|_\infty \leq \|B\|_\infty$ ;
- 4)  $\|A\|_\infty = \|A\mathbf{1}_n\|_\infty$ .

*Proof:* Step 1: See the result (8.1.9) of [34, Chapter 8, p. 491].

Step 2: Thanks to  $|A| = [[a_{ij}]]$ , it follows straightforwardly from the definition of the maximum row sum matrix norm that

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \| |A| \|_\infty.$$

Step 3: Due to  $A \geq 0$  (i.e.,  $a_{ij} \geq 0$ ,  $\forall i, j$ ) and  $B \geq 0$  (i.e.,  $b_{ij} \geq 0$ ,  $\forall i, j$ ), we can derive

$$\begin{aligned} \|A\|_\infty &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} \\ \|B\|_\infty &= \max_{1 \leq i \leq n} \sum_{j=1}^n |b_{ij}| = \max_{1 \leq i \leq n} \sum_{j=1}^n b_{ij} \end{aligned}$$

and then with  $A \leq B$  (i.e.,  $a_{ij} \leq b_{ij}$ ,  $\forall i, j$ ), we can arrive at

$$\sum_{j=1}^n a_{ij} \leq \sum_{j=1}^n b_{ij} \leq \max_{1 \leq i \leq n} \sum_{j=1}^n b_{ij}, \quad \forall i = 1, 2, \dots, n$$

which further yields

$$\max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} \leq \max_{1 \leq i \leq n} \sum_{j=1}^n b_{ij}.$$

From the above facts,  $\|A\|_\infty \leq \|B\|_\infty$  is immediate.

Step 4: Owing to  $A\mathbf{1}_n = [\sum_{j=1}^n a_{1j}, \sum_{j=1}^n a_{2j}, \dots, \sum_{j=1}^n a_{nj}]^T$ , the use of  $A \geq 0$  leads to  $A\mathbf{1}_n \geq 0$ , and consequently, we have

$$\|A\mathbf{1}_n\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij}$$

from which  $\|A\mathbf{1}_n\|_\infty = \|A\|_\infty$  follows directly.  $\blacksquare$

Based on Lemma 2, we revisit (20) and develop two helpful convergence results for the tracking error in the lemma below.

*Lemma 3:* Consider the iterative process (20) for the tracking error over any  $k \in \mathbb{Z}$  and  $t \in \mathbb{Z}_{T-1}$ , and then the following two convergence results hold under the condition (32).

1) If  $\kappa_k(t)$  is bounded (i.e.,  $\sup_{k \in \mathbb{Z}} |\kappa_k(t)| \leq \beta_k(t)$  for some finite bound  $\beta_k(t) > 0$ ), then  $\sup_{k \in \mathbb{Z}_+} |e_k(t+1)| \leq \beta_e(t)$  and  $\limsup_{k \rightarrow \infty} |e_k(t+1)| \leq \beta_{e_{\text{sup}}}(t)$  hold for some finite bounds  $\beta_e(t)$  and  $\beta_{e_{\text{sup}}}(t)$  such that  $\beta_e(t) > \beta_{e_{\text{sup}}}(t) > 0$ .

2) If, moreover,  $\lim_{k \rightarrow \infty} \kappa_k(t) = 0$  (exponentially fast), then  $\lim_{k \rightarrow \infty} e_k(t+1) = 0$  (exponentially fast).

*Proof:* Let us denote

$$\begin{aligned} \vec{e}_k(t+1) &= [e_k(t+1), e_{k-1}(t+1), \dots, e_{k-m+2}(t+1)]^T \in \mathbb{R}^{m-1} \\ \vec{\kappa}_k(t) &= [\kappa_k(t), 0, \dots, 0]^T \in \mathbb{R}^{m-1} \\ p_{1,k}(t) &= 1 - \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \theta_{k,k-1,t} \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \\ p_{i,k}(t) &= -\frac{\gamma_1 \gamma_{i+1} \theta_{k,k-1,t} \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)}, \quad i = 2, 3, \dots, m-1 \end{aligned} \quad (34)$$

and then we can rewrite (20) as

$$\vec{e}_k(t+1) = P_k(t) \vec{e}_{k-1}(t+1) + \vec{\kappa}_k(t), \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1} \quad (35)$$

where  $P_k(t) \in \mathbb{R}^{(m-1) \times (m-1)}$  takes the form of

$$P_k(t) = \begin{bmatrix} p_{1,k}(t) & p_{2,k}(t) & \cdots & \cdots & p_{m-1,k}(t) \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}. \quad (36)$$

With (32), we can combine (13) and (23) to deduce

$$\begin{aligned} \sum_{i=1}^{m-1} |p_{i,k}(t)| &= \left| 1 - \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \theta_{k,k-1,t} \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \right| \\ &\quad + \sum_{i=3}^m \left| \frac{\gamma_1 \gamma_i \theta_{k,k-1,t} \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \right| \\ &\leq 1 - \frac{\gamma_1 \left( \gamma_1 + \gamma_2 - \sum_{i=3}^m \gamma_i \right) \beta_f \varepsilon}{\lambda + \gamma_1^2 \beta_{\hat{\theta}}^2} \\ &\triangleq \zeta < 1, \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1}. \end{aligned} \quad (37)$$

To proceed, we resort to the nonnegative matrix  $|P_k(t)|$  and consider it over a sequence  $\{(m-1)s+1 : s \in \mathbb{Z}_+\}$ . Based on (37), we can further deduce that  $|P_k(t)|$ ,  $\forall k \in \mathbb{Z}$ ,  $\forall t \in \mathbb{Z}_{T-1}$  is a substochastic matrix owing to  $|P_k(t)|\mathbf{1}_{m-1} \leq \mathbf{1}_{m-1}$ . Because the set of substochastic matrices is closed under the operation of matrix multiplication, we can obtain

$$\prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \mathbf{1}_{m-1} \leq \mathbf{1}_{m-1}, \quad \forall s \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_{T-1}$$

which, together with (37) describing a condition that is strictly less than one, further leads to (see also the proof of (71) given at <https://arxiv.org/abs/1908.02447>)

$$\prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \mathbf{1}_{m-1} \leq \zeta \mathbf{1}_{m-1}, \quad \forall s \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_{T-1}. \quad (38)$$

Since the use of the property 1) of Lemma 2 leads to

$$\left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} P_j(t) \right| \leq \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \right|$$

we leverage this fact and incorporate the properties 3) and 4) of Lemma 2 to deduce

$$\begin{aligned} \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} P_j(t) \right| \right\|_{\infty} &\leq \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \right| \right\|_{\infty} \\ &= \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \mathbf{1}_{m-1} \right| \right\|_{\infty} \end{aligned}$$

which, together with using the property 2) of Lemma 2, further results in

$$\begin{aligned} \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} P_j(t) \right| \right\|_{\infty} &= \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} P_j(t) \right| \right\|_{\infty} \\ &\leq \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \mathbf{1}_{m-1} \right| \right\|_{\infty}. \end{aligned}$$

As a consequence of this property and again with the property 3) of Lemma 2, we can explore (38) to obtain

$$\begin{aligned} \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} P_j(t) \right| \right\|_{\infty} &\leq \left\| \left| \prod_{j=(m-1)s+1}^{(m-1)s+m-1} |P_j(t)| \mathbf{1}_{m-1} \right| \right\|_{\infty} \\ &\leq \|\zeta \mathbf{1}_{m-1}\|_{\infty} \\ &= \zeta \\ &< 1, \quad \forall s \in \mathbb{Z}_+, \forall t \in \mathbb{Z}_{T-1}. \end{aligned} \quad (39)$$

For the iterative process (35) under the condition (39), we can exploit the idea of [24, Lemma 2] to develop that:

1) There can be determined some finite bounds satisfying  $\beta_{\vec{e}}(t) > \beta_{\vec{e}_{\text{sup}}}(t) > 0$  such that  $\sup_{k \in \mathbb{Z}_+} \|\vec{e}_k(t+1)\|_{\infty} \leq \beta_{\vec{e}}(t)$  and  $\limsup_{k \rightarrow \infty} \|\vec{e}_k(t+1)\|_{\infty} \leq \beta_{\vec{e}_{\text{sup}}}(t)$ , provided that  $\sup_{k \in \mathbb{Z}_+} \|\vec{k}_k(t)\|_{\infty} \leq \beta_{\vec{k}}(t)$  is ensured for some finite bound  $\beta_{\vec{k}}(t) > 0$ ;

2)  $\lim_{k \rightarrow \infty} \vec{e}_k(t+1) = 0$  (exponentially fast), provided that  $\lim_{k \rightarrow \infty} \vec{k}_k(t) = 0$  (exponentially fast).

Based on these two facts, we notice the definitions of  $\vec{e}_k(t+1)$  and  $\vec{k}_k(t)$  in (34), and can easily accomplish two results of this lemma, of which the details are omitted for simplicity. ■

Next, we explore the boundedness of the system trajectories, for which the input dynamics along the iteration axis are used. To this end, we redescribe (15) as

$$\begin{aligned} u_k(t) &= u_{k-1}(t) + \frac{\gamma_1 \hat{\theta}_{k,k-1,t}(t) \sum_{i=3}^m \gamma_i e_{k-i+1}(t+1)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \\ &\quad - \frac{\gamma_1^2 \hat{\theta}_{k,k-1,t}(t) \sum_{i=0}^{t-1} \hat{\theta}_{k,k-1,t}(i) \Delta u_k(i)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \\ &\quad + \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \hat{\theta}_{k,k-1,t}(t) [y_d(t+1) - y_{k-1}(t+1)]}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)}. \end{aligned} \quad (40)$$

By considering the initial iteration (i.e.,  $j = 0$ ) and the  $(k-1)$ th

iteration (i.e.,  $i = k-1$ ) in (11), we can verify

$$\begin{aligned} y_{k-1}(t+1) &= \theta_{k-1,0,t}(t) u_{k-1}(t) + y_0(t+1) \\ &\quad + \sum_{i=0}^{t-1} \theta_{k-1,0,t}(i) u_{k-1}(i) - \sum_{i=0}^t \theta_{k-1,0,t}(i) u_0(i) \\ &\quad + [w_{k-1}(t) - w_0(t)] + \sum_{i=0}^{t-1} v_{k-1,0,t}(i) [w_{k-1}(i) \\ &\quad - w_0(i)] + \vartheta_{k-1,0,t} (\delta_{k-1} - \delta_0). \end{aligned} \quad (41)$$

As a consequence of substituting (41) into (40), the dynamics of the input can be formulated as

$$\begin{aligned} u_k(t) &= \left[ 1 - \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \hat{\theta}_{k,k-1,t}(t) \theta_{k-1,0,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \right] u_{k-1}(t) \\ &\quad + \psi_k(t), \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1} \end{aligned} \quad (42)$$

where  $\psi_k(t)$  is given by

$$\begin{aligned} \psi_k(t) &= \frac{\gamma_1 \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \left\{ (\gamma_1 + \gamma_2) \sum_{i=0}^t \theta_{k-1,0,t}(i) u_0(i) - (\gamma_1 \right. \\ &\quad \left. + \gamma_2) \sum_{i=0}^{t-1} \theta_{k-1,0,t}(i) u_{k-1}(i) - \gamma_1 \sum_{i=0}^{t-1} \hat{\theta}_{k,k-1,t}(i) \Delta u_k(i) \right. \\ &\quad \left. + (\gamma_1 + \gamma_2) e_0(t+1) + \sum_{i=3}^m \gamma_i e_{k-i+1}(t+1) \right. \\ &\quad \left. - (\gamma_1 + \gamma_2) [w_{k-1}(t) - w_0(t)] \right. \\ &\quad \left. - (\gamma_1 + \gamma_2) \sum_{i=0}^{t-1} v_{k-1,0,t}(i) [w_{k-1}(i) - w_0(i)] \right. \\ &\quad \left. - (\gamma_1 + \gamma_2) \vartheta_{k-1,0,t} (\delta_{k-1} - \delta_0) \right\}. \end{aligned} \quad (43)$$

Now, with (42), we present a boundedness result of the input in the following lemma.

**Lemma 4:** For the iterative process (42) of the input over any  $k \in \mathbb{Z}$  and  $t \in \mathbb{Z}_{T-1}$ , if the condition (32) is satisfied, then the boundedness of  $u_k(t)$  can be ensured (that is,  $\sup_{k \in \mathbb{Z}_+} |u_k(t)| \leq \beta_u(t)$  for some finite bound  $\beta_u(t) > 0$ ), provided that  $\psi_k(t)$  is bounded (that is,  $\sup_{k \in \mathbb{Z}_+} |\psi_k(t)| \leq \beta_{\psi}(t)$  for some finite bound  $\beta_{\psi}(t) > 0$ ).

*Proof:* Based on (32), the application of (13) and (23) results in

$$\begin{aligned} &\left| 1 - \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \hat{\theta}_{k,k-1,t}(t) \theta_{k-1,0,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \right| \\ &\leq 1 - \frac{(\gamma_1^2 + \gamma_1 \gamma_2) \beta_f \varepsilon}{\lambda + \gamma_1^2 \beta_{\hat{\theta}}^2} \triangleq \phi < 1, \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}_{T-1}. \end{aligned} \quad (44)$$

By considering (44) for (42), we can develop this lemma based on the result i) in [24, Lemma 2]. ■

Based on Lemmas 3 and 4, we present the proof of Theorem 2 by resorting to a DDA approach to ILC, instead of applying the eigenvalue-based analysis approach.

*Proof of Theorem 2:* This proof is obtained by induction over  $t \in \mathbb{Z}_{T-1}$ , where three steps are included as follows.

*Step 1:* Initialization results for  $t = 0$ . From (21), we have

$$\kappa_k(0) = -\Delta w_k(0) - \vartheta_{k,k-1,0} \Delta \delta_k, \quad \forall k \in \mathbb{Z} \quad (45)$$

which is guaranteed to be bounded under (7) and (12), namely,  $\sup_{k \in \mathbb{Z}} |\kappa_k(0)| \leq \beta_\kappa(0)$  for  $\beta_\kappa(0) = 2\beta_w(0) + 2\beta_\theta\beta_\delta$ . Then, as a consequence of the result 1) of Lemma 3, we can arrive at

$$\sup_{k \in \mathbb{Z}_+} |e_k(1)| \leq \beta_e(0) \text{ and } \limsup_{k \rightarrow \infty} |e_k(1)| \leq \beta_{e_{\sup}}(0) \quad (46)$$

for some finite bounds  $\beta_e(0) > \beta_{e_{\sup}}(0) > 0$ . In addition, using (43) yields

$$\begin{aligned} \psi_k(0) = & \frac{\gamma_1 \hat{\theta}_{k,k-1,0}(0)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,0}^2(0)} \left\{ (\gamma_1 + \gamma_2) \theta_{k-1,0,0}(0) u_0(0) \right. \\ & + (\gamma_1 + \gamma_2) e_0(1) + \sum_{i=3}^m \gamma_i e_{k-i+1}(1) \\ & - (\gamma_1 + \gamma_2) [w_{k-1}(0) - w_0(0)] \\ & \left. - (\gamma_1 + \gamma_2) \vartheta_{k-1,0,0}(\delta_{k-1} - \delta_0) \right\} \end{aligned}$$

which, together with (7), (12), (23), and (46), ensures

$$\begin{aligned} |\psi_k(0)| \leq & \frac{\gamma_1 \beta_{\hat{\theta}}}{\lambda + \gamma_1^2 \varepsilon^2} \left\{ (\gamma_1 + \gamma_2) [\beta_\theta |u_0(0)| + 2\beta_w(0) \right. \\ & + 2\beta_\theta\beta_\delta] + \sum_{i=1}^m \gamma_i \beta_e(0) \Big\} \\ & \triangleq \beta_\psi(0). \end{aligned} \quad (47)$$

With (47) and by considering Lemma 4 for the input process (42) at the initial time step  $t = 0$ ,  $\sup_{k \in \mathbb{Z}_+} |u_k(0)| \leq \beta_u(0)$  can be deduced for some finite bound  $\beta_u(0) > 0$ .

*Step 2:* Let  $N \in \mathbb{Z}_{T-1}$  be given. Assume that for all  $t = 0, 1, \dots, N-1$ ,  $\sup_{k \in \mathbb{Z}_+} |e_k(t+1)| \leq \beta_e(t)$ ,  $\limsup_{k \rightarrow \infty} |e_k(t+1)| \leq \beta_{e_{\sup}}(t)$ , and  $\sup_{k \in \mathbb{Z}_+} |u_k(t)| \leq \beta_u(t)$  are ensured for some finite bounds  $\beta_e(t) > 0$ ,  $\beta_{e_{\sup}}(t) > 0$ , and  $\beta_u(t) > 0$ , where  $\beta_e(t) > \beta_{e_{\sup}}(t)$ . Then, the same results will be proved for  $t = N$ .

In view of Lemma 1 and Theorem 1, we notice (21), employ the hypothesis presented for the time instants  $0, 1, \dots, N-1$  in this step, and can verify under Assumption 2 that

$$\begin{aligned} |\kappa_k(N)| = & \left| \frac{\gamma_1^2 \theta_{k,k-1,N}(N) \hat{\theta}_{k,k-1,N}(N)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,N}^2(N)} \sum_{i=0}^{N-1} \hat{\theta}_{k,k-1,N}(i) \Delta u_k(i) \right. \\ & - \sum_{i=0}^{N-1} \theta_{k,k-1,N}(i) \Delta u_k(i) - \Delta w_k(N) \\ & \left. - \sum_{i=0}^{N-1} v_{k,k-1,N}(i) \Delta w_k(i) - \vartheta_{k,k-1,N} \Delta \delta_k \right| \\ \leq & 2\beta_\theta \left( 1 + \frac{\gamma_1^2 \beta_{\hat{\theta}}^2}{\lambda + \gamma_1^2 \varepsilon^2} \right) \sum_{i=0}^{N-1} \beta_u(i) + 2\beta_w(N) \\ & + 2\beta_\theta \sum_{i=0}^{N-1} \beta_w(i) + 2\beta_\theta\beta_\delta \\ \triangleq & \beta_\kappa(N), \quad \forall k \in \mathbb{Z}. \end{aligned} \quad (48)$$

With (48) and based on the result 1) of Lemma 3, we can get

$$\sup_{k \in \mathbb{Z}_+} |e_k(N+1)| \leq \beta_e(N), \quad \limsup_{k \rightarrow \infty} |e_k(N+1)| \leq \beta_{e_{\sup}}(N) \quad (49)$$

where  $\beta_e(N) > \beta_{e_{\sup}}(N) > 0$  are some finite bounds. Similarly, the use of (43) yields

$$\begin{aligned} \psi_k(N) = & \frac{\gamma_1 \hat{\theta}_{k,k-1,N}(N)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,N}^2(N)} \left[ (\gamma_1 + \gamma_2) \sum_{i=0}^N \theta_{k-1,0,N}(i) u_0(i) - (\gamma_1 \right. \\ & + \gamma_2) \sum_{i=0}^{N-1} \theta_{k-1,0,N}(i) u_{k-1}(i) - \gamma_1 \sum_{i=0}^{N-1} \hat{\theta}_{k,k-1,N}(i) \Delta u_k(i) \\ & + (\gamma_1 + \gamma_2) e_0(N+1) + \sum_{i=3}^m \gamma_i e_{k-i+1}(N+1) \\ & - (\gamma_1 + \gamma_2) [w_{k-1}(N) - w_0(N)] \\ & - (\gamma_1 + \gamma_2) \sum_{i=0}^{N-1} v_{k-1,0,N}(i) [w_{k-1}(i) - w_0(i)] \\ & \left. - (\gamma_1 + \gamma_2) \vartheta_{k-1,0,N}(\delta_{k-1} - \delta_0) \right] \end{aligned}$$

for which we insert (49), together with the boundedness results of Lemma 1, Theorem 1, Assumption 2, and the hypothesis made in this step, to obtain

$$\begin{aligned} |\psi_k(N)| \leq & \frac{\gamma_1 |\hat{\theta}_{k,k-1,N}(N)|}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,N}^2(N)} \left[ (\gamma_1 + \gamma_2) \sum_{i=0}^N |\theta_{k-1,0,N}(i)| |u_0(i)| \right. \\ & + (\gamma_1 + \gamma_2) \sum_{i=0}^{N-1} |\theta_{k-1,0,N}(i)| |u_{k-1}(i)| \\ & + \gamma_1 \sum_{i=0}^{N-1} |\hat{\theta}_{k,k-1,N}(i)| |\Delta u_k(i)| \\ & + (\gamma_1 + \gamma_2) |e_0(N+1)| + \sum_{i=3}^m \gamma_i |e_{k-i+1}(N+1)| \\ & + (\gamma_1 + \gamma_2) [|w_{k-1}(N)| + |w_0(N)|] \\ & + (\gamma_1 + \gamma_2) \sum_{i=0}^{N-1} |v_{k-1,0,N}(i)| [|w_{k-1}(i)| + |w_0(i)|] \\ & \left. + (\gamma_1 + \gamma_2) |\vartheta_{k-1,0,N}| (|\delta_{k-1}| + |\delta_0|) \right] \\ \leq & \frac{\gamma_1 \beta_{\hat{\theta}}}{\lambda + \gamma_1^2 \varepsilon^2} \left\{ (\gamma_1 + \gamma_2) \beta_\theta \sum_{i=0}^N |u_0(i)| \right. \\ & + \left( \gamma_1 \beta_\theta + \gamma_2 \beta_\theta + 2\gamma_1 \beta_{\hat{\theta}} \right) \sum_{i=0}^{N-1} \beta_u(i) + \sum_{i=1}^m \gamma_i \beta_e(N) \\ & + 2(\gamma_1 + \gamma_2) \left[ \beta_w(N) + \beta_\theta \sum_{i=0}^{N-1} \beta_w(i) + \beta_\theta \beta_\delta \right] \left. \right\} \\ \triangleq & \beta_\psi(N). \end{aligned} \quad (50)$$

Based on the application of Lemma 4 to the input process (42)

for the time step  $t = N$  and with (50), we can thus conclude that  $\sup_{k \in \mathbb{Z}_+} |u_k(N)| \leq \beta_u(N)$  holds for some finite bound  $\beta_u(N) > 0$ . This, together with (49), implies that all the results made in the hypothesis for  $t = 0, 1, \dots, N-1$  also hold for  $t = N$ .

From the analysis of Steps 1 and 2, we perform induction and can easily deduce the boundedness result 1) and the robust tracking result (4) of Theorem 2.

*Step 3:* If (9) additionally holds (with an exponentially fast speed), then for  $t = 0$ , the use of (45) leads to  $\lim_{k \rightarrow \infty} \kappa_k(0) = 0$  (exponentially fast), with which we can obtain  $\lim_{k \rightarrow \infty} e_k(1) = 0$  (exponentially fast) based on the result 2) of Lemma 3. Since (15) ensures

$$\Delta u_k(0) = \frac{\gamma_1 \hat{\theta}_{k,k-1,0}(0)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,0}^2(0)} \left[ \gamma_1 e_{k-1}(1) + \sum_{i=2}^m \gamma_i e_{k-i+1}(1) \right]$$

we further get  $\lim_{k \rightarrow \infty} \Delta u_k(0) = 0$  (exponentially fast). Thus, if we make a hypothesis that for  $N \in \mathbb{Z}_{T-1}$ ,  $\lim_{k \rightarrow \infty} e_k(t+1) = 0$  (exponentially fast) and  $\lim_{k \rightarrow \infty} \Delta u_k(t) = 0$  (exponentially fast) hold for all  $t = 0, 1, \dots, N-1$ , then for  $t = N$ , by noting

$$\begin{aligned} |\kappa_k(N)| &\leq \beta_\theta \left( 1 + \frac{\gamma_1^2 \beta_{\hat{\theta}}^2}{\lambda + \gamma_1^2 \varepsilon^2} \right) \sum_{i=0}^{N-1} |\Delta u_k(i)| \\ &\quad + |\Delta w_k(N)| + \beta_\theta \sum_{i=0}^{N-1} |\Delta w_k(i)| + \beta_\theta |\Delta \delta_k| \end{aligned}$$

$$\rightarrow 0 \text{ (exponentially fast), as } k \rightarrow \infty \quad (51)$$

we can apply the result 2) of Lemma 3 to derive  $\lim_{k \rightarrow \infty} e_k(N+1) = 0$  (exponentially fast), which together with

$$\begin{aligned} |\Delta u_k(N)| &= \left| -\frac{\gamma_1^2 \hat{\theta}_{k,k-1,N}(N)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,N}^2(N)} \sum_{i=0}^{N-1} \hat{\theta}_{k,k-1,N}(i) \Delta u_k(i) \right. \\ &\quad \left. + \frac{\gamma_1 \hat{\theta}_{k,k-1,N}(N)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,N}^2(N)} \left[ \gamma_1 e_{k-1}(N+1) \right. \right. \\ &\quad \left. \left. + \sum_{i=2}^m \gamma_i e_{k-i+1}(N+1) \right] \right| \\ &\leq \frac{\gamma_1^2 \beta_{\hat{\theta}}^2}{\lambda + \gamma_1^2 \varepsilon^2} \sum_{i=0}^{N-1} |\Delta u_k(i)| + \frac{\gamma_1 \beta_{\hat{\theta}}}{\lambda + \gamma_1^2 \varepsilon^2} \left[ \gamma_1 |e_{k-1}(N+1)| \right. \\ &\quad \left. + \sum_{i=2}^m \gamma_i |e_{k-i+1}(N+1)| \right] \end{aligned}$$

further yields  $\lim_{k \rightarrow \infty} \Delta u_k(N) = 0$  (exponentially fast). Namely, we can conclude by induction that for  $t \in \mathbb{Z}_{T-1}$ ,  $\lim_{k \rightarrow \infty} e_k(t+1) = 0$  (exponentially fast) and  $\lim_{k \rightarrow \infty} \Delta u_k(t) = 0$  (exponentially fast). The perfect tracking result (2) can thus be obtained for Theorem 2. ■

#### IV. SIMULATION TESTS

To illustrate the effectiveness of the proposed optimization-based adaptive ILC, we perform simulation tests by considering a numerical example and an injection molding process.

*Example 1:* Consider the nonlinear system (1) with

nonlinear dynamics described in a specific form of

$$\begin{aligned} f(y_k(t), y_k(t-1), u_k(t), u_k(t-1), t) &= \sin(y_k(t)) \\ &\quad + \cos(y_k(t-1)) + \frac{t+1}{t+2} u_k(t) + \cos(y_k(t)) \sin(u_k(t-1)) \end{aligned}$$

and with nonrepetitive uncertainties given by

$$w_k(t) = 0.01 \chi_w(k, t), \quad y_k(0) = 1.5 + 0.01 \chi_y(k)$$

where  $\chi_w(k, t)$  and  $\chi_y(k)$  vary within  $[-1, 1]$  arbitrarily for all  $k$  and  $t$ . For the robust tracking task (4), we adopt the desired reference trajectory as

$$y_d(t) = 5 \sin\left(\frac{2\pi t}{50}\right) + \frac{t(50-t)}{375}, \quad \forall t \in \mathbb{Z}_{50}.$$

To implement the updating law (15) and the adaptive updating schemes S1) and S2), we employ the parameters shown in Table I, and select the initial value  $\hat{\theta}_{0,-1,t}(i)$  of estimated parameters as  $\hat{\theta}_{0,-1,t}(i) = 0.9$ ,  $\forall i \in \mathbb{Z}_t$ ,  $\forall t \in \mathbb{Z}_{T-1}$  and the initial input  $u_0(t)$  as the zero input, i.e.,  $u_0(t) = 0$ ,  $\forall t \in \mathbb{Z}_{T-1}$ . It can be verified that the selection condition (32) is guaranteed under the parameters presented in Table I.

TABLE I  
PARAMETERS USED IN (15), S1) AND S2)

$\lambda$	$\gamma_1$	$\gamma_2$	$\gamma_3$	$\mu_1$	$\mu_2$	$\varepsilon$
1	0.8	0.14	0.06	1	0.001	0.01

In Fig. 1, we describe the curve of the iteration evolution of the input, which is evaluated with  $\max_{t \in \mathbb{Z}_{50}} |u_k(t)|$ , for the first 1000 iterations. It is clear from Fig. 1 that the input is bounded for all time steps and all iterations. For illustration of the robust output tracking performances, we depict the iteration evolution of the tracking error evaluated with  $\max_{t \in \mathbb{Z}_{49}} |e_k(t+1)|$  for the first 1000 iterations in Fig. 2. This figure clearly demonstrates that the robust tracking objective (4) can be realized, regardless of a small remaining tracking error due to the presence of the nonrepetitive uncertainties. In particular, we depict the tracking performance with respect to the desired reference trajectory in Fig. 3, which is achieved for the system output refined through the optimization-based adaptive ILC after the 400th iteration. We can clearly observe from Fig. 3 that the learned system output almost perfectly tracks the desired reference trajectory (except the initial time step) despite the nonrepetitive initial shifts and disturbances. The simulation results of Figs. 1–3 validate that our proposed optimization-based adaptive ILC works robustly and effectively, regardless of nonrepetitive uncertainties.

*Example 2:* Consider an injection molding process, devoting to the dynamics between the nozzle pressure and the hydraulic control valve opening, described by (see also [32])

$$\begin{aligned} y_k(t+1) &= 1.607 y_k(t) - 0.6086 y_k(t-1) \\ &\quad + 1.239 u_k(t) - 0.9282 u_k(t-1) + w_k(t) \end{aligned}$$

where  $u_k(t)$  and  $y_k(t)$  now represent the hydraulic control valve opening and the nozzle pressure, respectively, and  $w_k(t)$  results from the external disturbances and/or unmodeled uncertainties. Let the nozzle pressure control process be

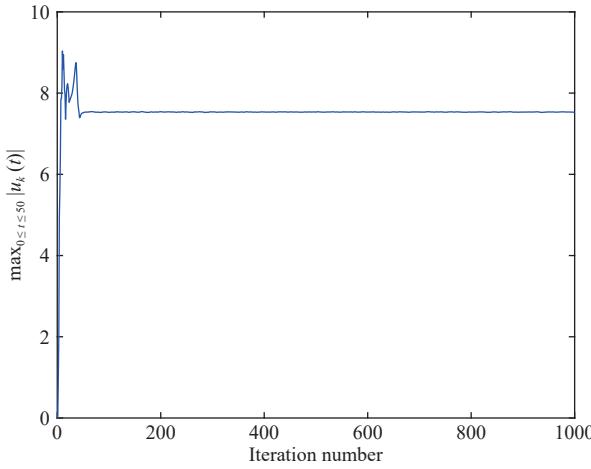


Fig. 1. Example 1: The input evolution versus iteration.

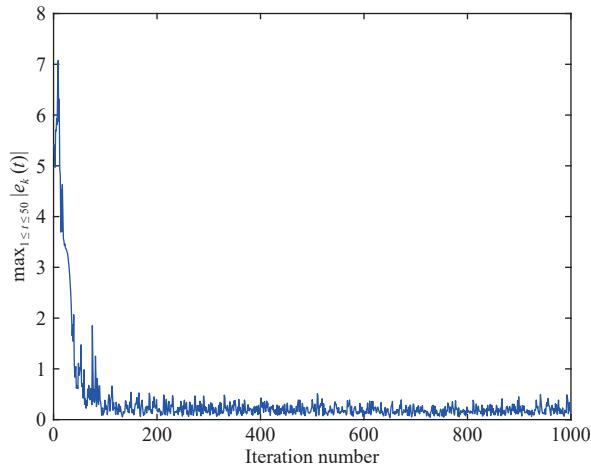


Fig. 2. Example 1: The convergence performance of the tracking error over the first 1000 iterations.

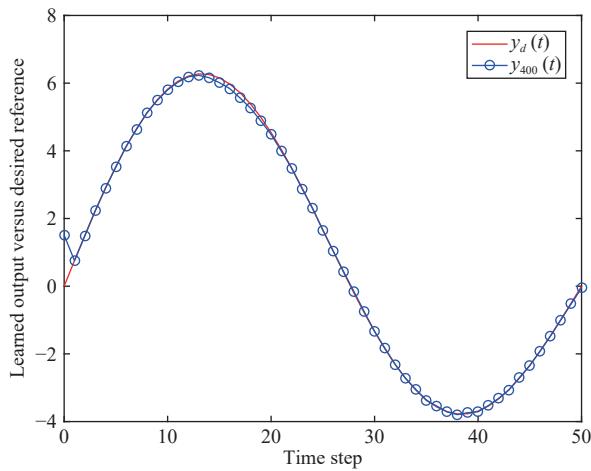


Fig. 3. Example 1: The tracking performance between the output learned after the 400th iteration and the desired reference.

subject to nonrepetitive uncertainties arising from the initial shifts and the external disturbances in the form of

$$y_k(0) = 10 + \bar{\chi}_y(k), \quad w_k(t) = \bar{\chi}_w(k, t)$$

for some  $\bar{\chi}_y(k)$  and  $\bar{\chi}_w(k, t)$  varying within  $[-1, 1]$  arbitrarily for all  $k$  and  $t$ .

To implement the robust tracking task, the desired reference trajectory is given by

$$y_d(t) = \begin{cases} 150, & 0 \leq t \leq 50 \\ 300, & 51 \leq t \leq 100 \end{cases}$$

and our optimization-based adaptive ILC, which is comprised of the updating law (15) and the adaptive updating schemes S1) and S2), is applied by adopting the same parameters as shown in Table I and the same initial settings as used in Example 1. It can be validated that the needed robust convergence conditions of ILC in Theorems 1 and 2 are satisfied.

Similarly to Figs. 1–3, Figs. 4–6 are depicted to demonstrate the system performance of the injection molding process when operating with the use of our optimization-based adaptive ILC. The input evolution versus iteration is depicted in Fig. 4, which illustrates the boundedness of system trajectories. In Fig. 5, the robust convergence performance is illustrated for the optimization-based adaptive ILC by describing the evolution of the tracking error versus iteration. The high-precision tracking performance is demonstrated in Fig. 6 that describes the output learned with our optimization-based adaptive ILC after the 100th iteration, as well as the desired reference trajectory. It can be obviously seen that the illustrations of Figs. 4–6 coincide with our robust optimization-based adaptive ILC results of nonlinear systems.

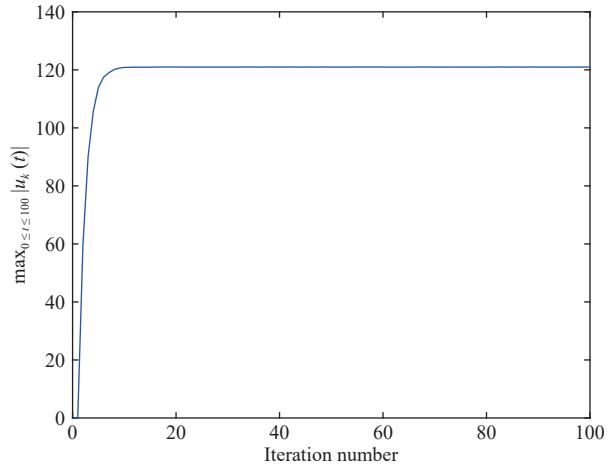


Fig. 4. Example 2: The bounded evolution of the input versus iteration.

**Discussions:** The simulation tests performed in Examples 1 and 2 validate the robustness and effectiveness of our presented optimization-based adaptive ILC for nonlinear systems in spite of unknown nonlinearities and nonrepetitive uncertainties. Due to the limited use of model information, they also demonstrate that our optimization-based adaptive ILC results may provide a feasible way to the design and analysis of data-driven methods. In particular, the illustrations of Figs. 3 and 6 can disclose that our design method of optimization-based adaptive ILC works effectively for accomplishing the high-precision tracking tasks of nonlinear systems

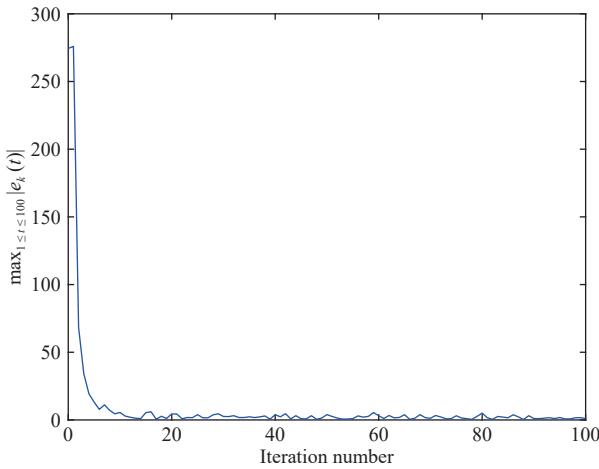


Fig. 5. Example 2: The robust convergence performance of the tracking error for the first 100 iterations.

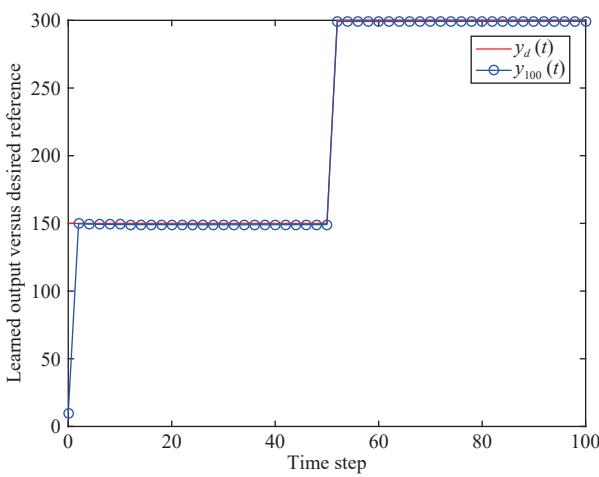


Fig. 6. Example 2: The high-precision tracking performance between the output learned after the 100th iteration and the desired reference trajectory.

subject to the nonrepetitive uncertainties, especially in comparison with those methods proposed in, e.g., [16], [17].

## V. CONCLUSIONS

In this paper, robust convergence problems for the optimization-based adaptive ILC of nonlinear time-varying systems subject to iteration-dependent initial shifts and disturbances have been discussed. A new design method has been introduced to bridge the gap between optimization-based and CM-based approaches for ILC. By incorporating properties of substochastic matrices, a DDA approach integrated with CM-based analyses has been explored to establish robust convergence results of ILC, which avoids performing the eigenvalue analysis to gain convergence of iterative processes subject to iteration-dependent parameters. These advantages make our design and analysis methods for the optimization-based adaptive ILC effective and robust in spite of nonrepetitive uncertainties. In addition, the simulation tests implemented through a numerical example and for an injection molding process have demonstrated the validity of our robust optimization-based adaptive ILC results.

## APPENDIX

### PROOF OF LEMMA 1

*Proof:* An inductive analysis on  $t$  is performed to prove this lemma, and the proof is separated into two steps as follows.

*Step 1:* Let  $t = 0$ . Then, the use of (1) gives

$$\begin{aligned} y_k(1) &= f(y_k(0), 0, \dots, 0, u_k(0), 0, \dots, 0) + w_k(0) \\ &\triangleq \bar{g}^0(y_k(0), u_k(0), w_k(0)) \end{aligned}$$

based on which we have

$$\begin{aligned} \frac{\partial \bar{g}^0}{\partial y_k(0)} &= \frac{\partial f}{\partial x_1} \Big|_{(y_k(0), 0, \dots, 0, u_k(0), 0, \dots, 0)} \\ \frac{\partial \bar{g}^0}{\partial u_k(0)} &= \frac{\partial f}{\partial x_{l+2}} \Big|_{(y_k(0), 0, \dots, 0, u_k(0), 0, \dots, 0)} \\ \frac{\partial \bar{g}^0}{\partial w_k(0)} &= 1. \end{aligned}$$

By employing (5) and (6), we can further derive

$$\left| \frac{\partial \bar{g}^0}{\partial y_k(0)} \right| \leq \beta_{\bar{f}} \triangleq \beta_\theta(0), \quad \left| \frac{\partial \bar{g}^0}{\partial u_k(0)} \right| \in [\beta_{\underline{f}}, \beta_{\bar{f}}], \quad \left| \frac{\partial \bar{g}^0}{\partial w_k(0)} \right| = 1.$$

*Step 2:* Let us consider any  $N \in \mathbb{Z}$ . For  $t = 0, 1, \dots, N-1$ , we assume  $y_k(t+1) = \bar{g}^t(y_0, u_k(0), \dots, u_k(t), w_k(0), \dots, w_k(t))$  and simultaneously that it satisfies

$$\begin{aligned} \left| \frac{\partial \bar{g}^t}{\partial y_k(0)} \right| &\leq \beta_\theta(t), \quad \left| \frac{\partial \bar{g}^t}{\partial u_k(t)} \right| \in [\beta_{\underline{f}}, \beta_{\bar{f}}], \quad \left| \frac{\partial \bar{g}^t}{\partial w_k(t)} \right| = 1 \\ \left| \frac{\partial \bar{g}^t}{\partial u_k(0)} \right| &\leq \beta_\theta(t), \dots, \left| \frac{\partial \bar{g}^t}{\partial u_k(t-1)} \right| \leq \beta_\theta(t) \\ \left| \frac{\partial \bar{g}^t}{\partial w_k(0)} \right| &\leq \beta_\theta(t), \dots, \left| \frac{\partial \bar{g}^t}{\partial w_k(t-1)} \right| \leq \beta_\theta(t) \end{aligned}$$

for some finite bound  $\beta_\theta(t) > 0$ . Next, we show that for  $t = N$ , we can deduce the same results.

When we consider (1) for  $t = N$ , the use of the hypothesis made for  $t = 0, 1, \dots, N-1$  leads to

$$\begin{aligned} y_k(N+1) &= f(y_k(N), \dots, y_k(N-l), u_k(N), \dots, u_k(N-n), N) \\ &\quad + w_k(N) \\ &= f(\bar{g}^{N-1}, \dots, \bar{g}^{N-1-l}, u_k(N), \dots, u_k(N-n), N) + w_k(N) \\ &\triangleq \bar{g}^N(y_k(0), u_k(0), \dots, u_k(N), w_k(0), \dots, w_k(N)). \end{aligned}$$

For  $\bar{g}^N$ , we employ the derivation rules of compound functions to deduce

$$\begin{aligned} \frac{\partial \bar{g}^N}{\partial y_k(0)} &= \sum_{i=0}^l \frac{\partial f}{\partial \bar{g}^{N-1-i}} \frac{\partial \bar{g}^{N-1-i}}{\partial y_k(0)} \\ \frac{\partial \bar{g}^N}{\partial u_k(0)} &= \sum_{i=0}^l \frac{\partial f}{\partial \bar{g}^{N-1-i}} \frac{\partial \bar{g}^{N-1-i}}{\partial u_k(0)} \\ &\vdots \\ \frac{\partial \bar{g}^N}{\partial u_k(N-1)} &= \frac{\partial f}{\partial \bar{g}^{N-1}} \frac{\partial \bar{g}^{N-1}}{\partial u_k(N-1)} + \frac{\partial f}{\partial u_k(N-1)} \\ \frac{\partial \bar{g}^N}{\partial u_k(N)} &= \frac{\partial f}{\partial u_k(N)} \end{aligned}$$

and

$$\begin{aligned}\frac{\partial \bar{g}^N}{\partial w_k(0)} &= \sum_{i=0}^l \frac{\partial f}{\partial \bar{g}^{N-1-i}} \frac{\partial \bar{g}^{N-1-i}}{\partial w_k(0)} \\ &\vdots \\ \frac{\partial \bar{g}^N}{\partial w_k(N-1)} &= \frac{\partial f}{\partial \bar{g}^{N-1}} \frac{\partial \bar{g}^{N-1}}{\partial w_k(N-1)} \\ \frac{\partial \bar{g}^N}{\partial w_k(N)} &= 1.\end{aligned}$$

Again with the made hypothesis and by inserting (5) and (8), we can obtain

$$\begin{aligned}\left| \frac{\partial \bar{g}^N}{\partial y_k(0)} \right| &\leq \sum_{i=0}^l \left| \frac{\partial f}{\partial \bar{g}^{N-1-i}} \right| \left| \frac{\partial \bar{g}^{N-1-i}}{\partial y_k(0)} \right| \leq \beta_\theta(N) \\ \left| \frac{\partial \bar{g}^N}{\partial u_k(0)} \right| &\leq \sum_{i=0}^l \left| \frac{\partial f}{\partial \bar{g}^{N-1-i}} \right| \left| \frac{\partial \bar{g}^{N-1-i}}{\partial u_k(0)} \right| \leq \beta_\theta(N) \\ &\vdots \\ \left| \frac{\partial \bar{g}^N}{\partial u_k(N-1)} \right| &\leq \left| \frac{\partial f}{\partial \bar{g}^{N-1}} \right| \left| \frac{\partial \bar{g}^{N-1}}{\partial u_k(N-1)} \right| + \left| \frac{\partial f}{\partial u_k(N-1)} \right| \\ &\leq \beta_\theta(N) \\ \frac{\partial \bar{g}^N}{\partial u_k(N)} &= \frac{\partial f}{\partial u_k(N)} \in [\beta_{\underline{f}}, \beta_{\bar{f}}]\end{aligned}$$

and

$$\begin{aligned}\left| \frac{\partial \bar{g}^N}{\partial w_k(0)} \right| &\leq \sum_{i=0}^l \left| \frac{\partial f}{\partial \bar{g}^{N-1-i}} \right| \left| \frac{\partial \bar{g}^{N-1-i}}{\partial w_k(0)} \right| \leq \beta_\theta(N) \\ &\vdots \\ \left| \frac{\partial \bar{g}^N}{\partial w_k(N-1)} \right| &\leq \left| \frac{\partial f}{\partial \bar{g}^{N-1}} \right| \left| \frac{\partial \bar{g}^{N-1}}{\partial w_k(N-1)} \right| \leq \beta_\theta(N) \\ \frac{\partial \bar{g}^N}{\partial w_k(N)} &= 1\end{aligned}$$

where  $\beta_\theta(N) = (l+1)\beta_{\bar{f}} \max_{t \in \mathbb{Z}_{N-1}} \beta_\theta(t) + \beta_{\bar{f}}$  can be adopted as a candidate.

Based on the analysis of the above Steps 1 and 2, we can conclude by induction that for any  $t \in \mathbb{Z}_{T-1}$  and  $k \in \mathbb{Z}_+$ ,

$$\begin{aligned}y_k(t+1) &= \bar{g}^t(y_k(0), u_k(0), \dots, u_k(t), w_k(0), \dots, w_k(t)) \text{ with} \\ &\left\{ \begin{array}{l} \left| \frac{\partial \bar{g}^t}{\partial y_k(0)} \right| \leq \beta_\theta(t), \quad \frac{\partial \bar{g}^t}{\partial u_k(t)} \in [\beta_{\underline{f}}, \beta_{\bar{f}}], \quad \frac{\partial \bar{g}^t}{\partial w_k(t)} = 1 \\ \left| \frac{\partial \bar{g}^t}{\partial u_k(0)} \right| \leq \beta_\theta(t), \dots, \left| \frac{\partial \bar{g}^t}{\partial u_k(t-1)} \right| \leq \beta_\theta(t) \\ \left| \frac{\partial \bar{g}^t}{\partial w_k(0)} \right| \leq \beta_\theta(t), \dots, \left| \frac{\partial \bar{g}^t}{\partial w_k(t-1)} \right| \leq \beta_\theta(t) \end{array} \right.\end{aligned}$$

where  $\bar{g}^t$  is some continuously differentiable function such that

$$\bar{g}^t : \underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{2t+3} \rightarrow \mathbb{R}$$

and  $\beta_\theta(t) > 0$  is some finite bound. For convenience, we write

$\bar{g}^t$  in terms of  $\bar{g}^t(z_1, z_2, \dots, z_{2t+3})$ , where  $z_i \in \mathbb{R}$ ,  $i = 1, 2, \dots, 2t+3$  represents the  $i$ th independent variable of  $\bar{g}^t$ . Then based on the use of the mean value theorem (see, e.g., [35], p. 651]), we can validate

$$\begin{aligned}&y_i(t+1) - y_j(t+1) \\ &= \left[ \frac{\partial \bar{g}^t}{\partial z_1}, \frac{\partial \bar{g}^t}{\partial z_2}, \dots, \frac{\partial \bar{g}^t}{\partial z_{2t+3}} \right] \Big|_{(z_1, z_2, \dots, z_{2t+3})=(z_1^*, z_2^*, \dots, z_{2t+3}^*)} \\ &\times \begin{pmatrix} y_i(0) & y_j(0) \\ u_i(0) & u_j(0) \\ \vdots & \vdots \\ u_i(t) & u_j(t) \\ w_i(0) & w_j(0) \\ \vdots & \vdots \\ w_i(t) & w_j(t) \end{pmatrix} \\ &= \left[ \frac{\partial \bar{g}^t}{\partial z_2}, \frac{\partial \bar{g}^t}{\partial z_3}, \dots, \frac{\partial \bar{g}^t}{\partial z_{t+2}} \right] \Big|_{(z_1, z_2, \dots, z_{2t+3})=(z_1^*, z_2^*, \dots, z_{2t+3}^*)} \\ &\times \begin{pmatrix} u_i(0) & u_j(0) \\ u_i(1) & u_j(1) \\ \vdots & \vdots \\ u_i(t) & u_j(t) \end{pmatrix} \\ &+ \left[ \frac{\partial \bar{g}^t}{\partial z_{t+3}}, \frac{\partial \bar{g}^t}{\partial z_{t+4}}, \dots, \frac{\partial \bar{g}^t}{\partial z_{2t+3}} \right] \Big|_{(z_1, z_2, \dots, z_{2t+3})=(z_1^*, z_2^*, \dots, z_{2t+3}^*)} \\ &\times \begin{pmatrix} w_i(0) & w_j(0) \\ w_i(1) & w_j(1) \\ \vdots & \vdots \\ w_i(t) & w_j(t) \end{pmatrix} \\ &+ \frac{\partial \bar{g}^t}{\partial z_1} \Big|_{(z_1, z_2, \dots, z_{2t+3})=(z_1^*, z_2^*, \dots, z_{2t+3}^*)} (\delta_i - \delta_j) \quad (52)\end{aligned}$$

where

$$\begin{aligned}(z_1^*, z_2^*, \dots, z_{2t+3}^*) &= \bar{\omega}(y_i(0), u_i(0), \dots, u_i(t), w_i(0), \dots, w_i(t)) \\ &+ (1 - \bar{\omega})(y_j(0), u_j(0), \dots, u_j(t), w_j(0), \dots, w_j(t))\end{aligned}$$

for some  $\bar{\omega} \in [0, 1]$ . Clearly, (52) can be rewritten in a compact form of (11). Moreover, by setting  $\beta_\theta = \max_{t \in \mathbb{Z}_{T-1}} \beta_\theta(t)$ , the boundedness results of (12) and (13) can also be obtained. ■

## REFERENCES

- [1] H. Zhang, Q. Wei, and Y. Luo, "A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy HDP iteration algorithm," *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 38, no. 4, pp. 937–942, Aug. 2008.
- [2] S. John and J. O. Pedro, "Neural network-based adaptive feedback linearization control of antilock braking system," *Int. Journal of Artificial Intelligence*, vol. 10, no. S13, pp. 21–40, Mar. 2013.

- [3] S. Preitl, R.-E. Precup, Z. Preitl, S. Vaivoda, S. Kilyeni, and J. K. Tar, "Iterative feedback and learning control. Servo systems applications," *IFAC Proceedings Volumes*, vol. 40, no. 8, pp. 16–27, Jul. 2007.
- [4] J. Zhang and D. Meng, "Convergence analysis of saturated iterative learning control systems with locally Lipschitz nonlinearities," *IEEE Trans. Neural Networks and Learning Systems*, vol. 31, no. 10, pp. 4025–4035, Oct. 2020.
- [5] D. Meng and K. L. Moore, "Contraction mapping-based robust convergence of iterative learning control with uncertain, locally-Lipschitz nonlinearity," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 50, no. 2, pp. 442–454, Feb. 2020.
- [6] K. L. Barton and A. G. Alleyne, "A norm optimal approach to time-varying ILC with application to a multi-axis robotic testbed," *IEEE Trans. Control Systems Technology*, vol. 19, no. 1, pp. 166–180, Jan. 2011.
- [7] W. He, T. Meng, X. He, and C. Sun, "Iterative learning control for a flapping wing micro aerial vehicle under distributed disturbances," *IEEE Trans. Cybernetics*, vol. 49, no. 4, pp. 1524–1535, Apr. 2019.
- [8] P. Janssens, G. Pipeleers, and J. Swevers, "A data-driven constrained norm-optimal iterative learning control framework for LTI systems," *IEEE Trans. Control Systems Technology*, vol. 21, no. 2, pp. 546–551, Mar. 2013.
- [9] Q. Yu, Z. Hou, and J. Xu, "D-type ILC based dynamic modeling and norm optimal ILC for high-speed trains," *IEEE Trans. Control Systems Technology*, vol. 26, no. 2, pp. 652–663, Mar. 2018.
- [10] J. Wang, Y. Wang, L. Cao, and Q. Jin, "Adaptive iterative learning control based on unfalsified strategy for Chyba-Haase reactor," *IEEE/CAA Journal of Automatica Sinica*, vol. 1, no. 4, pp. 347–360, Oct. 2014.
- [11] D. Shen, "Iterative learning control with incomplete information: A survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 5, pp. 885–901, Sept. 2018.
- [12] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control: A learning-based method for high-performance tracking control," *IEEE Control Systems Magazine*, vol. 26, no. 3, pp. 96–114, Jun. 2006.
- [13] H.-S. Ahn, Y. Chen, and K. L. Moore, "Iterative learning control: Brief survey and categorization," *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1099–1121, Nov. 2007.
- [14] M. Minakais, S. Mishra, and J. T. Wen, "Database-driven iterative learning for building temperature control," *IEEE Trans. Automation Science and Engineering*, vol. 16, no. 4, pp. 1896–1906, Oct. 2019.
- [15] R.-H. Chi and Z.-S. Hou, "Dual-stage optimal iterative learning control for nonlinear non-affine discrete-time systems," *Acta Automatica Sinica*, vol. 33, no. 10, pp. 1061–1065, Oct. 2007.
- [16] R. Chi, Z. Hou, B. Huang, and S. Jin, "A unified data-driven design framework of optimality-based generalized iterative learning control," *Computers and Chemical Engineering*, vol. 77, pp. 10–23, Jun. 2015.
- [17] R. Chi, Z. Hou, S. Jin, and B. Huang, "Computationally efficient data-driven higher order optimal iterative learning control," *IEEE Trans. Neural Networks and Learning Systems*, vol. 29, no. 12, pp. 5971–5980, Dec. 2018.
- [18] Y. Hui, R. Chi, B. Huang, and Z. Hou, "Extended state observer-based data-driven iterative learning control for permanent magnet linear motor with initial shifts and disturbances," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 51, no. 3, pp. 1881–1891, Mar. 2021.
- [19] Z. Hou and S. Jin, *Model Free Adaptive Control: Theory and Applications*. Boca Raton, USA: CRC Press, 2013.
- [20] X. Bu, S. Wang, Z. Hou, and W. Liu, "Model free adaptive iterative learning control for a class of nonlinear systems with randomly varying iteration lengths," *Journal of the Franklin Institute*, vol. 356, no. 5, pp. 2491–2504, Mar. 2019.
- [21] X. Bu, Q. Yu, Z. Hou, and W. Qian, "Model free adaptive iterative learning consensus tracking control for a class of nonlinear multiagent systems," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 49, no. 4, pp. 677–686, Apr. 2019.
- [22] W. J. Rugh, *Linear System Theory*. Upper Saddle River, NJ, USA: Prentice Hall, 1996.
- [23] D. Meng, Y. Jia, and J. Du, "Stability of varying two-dimensional roesser systems and its application to iterative learning control convergence analysis," *IET Control Theory and Applications*, vol. 9, no. 8, pp. 1221–1228, May 2015.
- [24] D. Meng and K. L. Moore, "Robust iterative learning control for nonrepetitive uncertain systems," *IEEE Trans. Automatic Control*, vol. 62, no. 2, pp. 907–913, Feb. 2017.
- [25] Q. Zhu, J.-X. Xu, D. Huang, and G.-D. Hu, "Iterative learning control design for linear discrete-time systems with multiple high-order internal models," *Automatica*, vol. 62, pp. 65–76, Dec. 2015.
- [26] D. Meng, "Convergence conditions for solving robust iterative learning control problems under nonrepetitive model uncertainties," *IEEE Trans. Neural Networks and Learning Systems*, vol. 30, no. 6, pp. 1908–1919, Jun. 2019.
- [27] J. Zhang, B. Cui, X. Dai, and Z. Jiang, "Iterative learning control for distributed parameter systems based on non-collocated sensors and actuators," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 3, pp. 865–871, May 2020.
- [28] D. Meng and J. Zhang, "Robust tracking of nonrepetitive learning control systems with iteration-dependent references," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 51, no. 2, pp. 842–852, Feb. 2021.
- [29] D. Meng and J. Zhang, "Convergence analysis of robust iterative learning control against nonrepetitive uncertainties: System equivalence transformation," *IEEE Trans. Neural Networks and Learning Systems*, to be published, DOI: 10.1109/TNNLS.2020.3016057.
- [30] X. Jin, "Fault-tolerant iterative learning control for mobile robots non-repetitive trajectory tracking with output constraints," *Automatica*, vol. 94, pp. 63–71, Aug. 2018.
- [31] M. Yu, D. Huang, and W. He, "Robust adaptive iterative learning control for discrete-time nonlinear systems with both parametric and nonparametric uncertainties," *Int. Journal of Adaptive Control and Signal Processing*, vol. 30, no. 7, pp. 972–985, Jul. 2016.
- [32] J. Shi, F. Gao, and T.-J. Wu, "Integrated design and structure analysis of robust iterative learning control system based on a two-dimensional model," *Industrial and Engineering Chemistry Research*, vol. 44, no. 21, pp. 8095–8105, Sept. 2005.
- [33] J.-X. Xu, "A survey on iterative learning control for nonlinear systems," *Int. Journal of Control*, vol. 84, no. 7, pp. 1275–1294, Jul. 2011.
- [34] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [35] H. K. Khalil, *Nonlinear Systems*. Upper Saddle River, New Jersey, USA: Prentice Hall, 2002.



**Deyuan Meng** (M'12-SM'18) received the B.S. degree in mathematics and applied mathematics from the Ocean University of China, Qingdao, China, in 2005, and the Ph.D. degree in control theory and control engineering from Beihang University (BUAA), Beijing, China, in 2010. From 2012 to 2013, he was a Visiting Scholar with the Department of Electrical Engineering and Computer Science, Colorado School of Mines, Golden, CO, USA. He is currently a Full Professor with the Seventh Research

Division and the School of Automation Science and Electrical Engineering, Beihang University (BUAA), China.

His current research interests include iterative learning control, data-driven control, and multi-agent systems.



**Jingyao Zhang** received the B.S. degree in information and computing science and the M.S. degree in mathematics from Beihang University (BUAA), Beijing, China, in 2015 and 2018, respectively. He is currently a Ph.D. candidate at the School of Automation Science and Electrical Engineering, Beihang University (BUAA), China.

His current research interests include iterative learning control and nonlinear control.