This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Design and Analysis of Data-Driven Learning Control: An Optimization-Based Approach

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

***Abstract*— Learning to perform perfect tracking tasks based on measurement data is desirable in the controller design of systems operating repetitively. This motivates this article to seek an optimization-based design and analysis approach for data-driven learning control systems by focusing on iterative learning control (ILC) of repetitive systems with unknown nonlinear time-varying dynamics. It is shown that perfect output tracking can be realized with updating inputs, where no explicit model knowledge but only measured input–output data are leveraged. In particular, adaptive updating strategies are proposed to obtain parameter estimations of nonlinearities. A double-dynamics analysis approach is applied to establish ILC convergence, together with boundedness of input, output, and estimated parameters, which benefits from employing properties of nonnegative matrices. Simulations are implemented to verify the validity of our optimization-based adaptive ILC.**

***Index Terms*— Adaptive updating, data-driven learning control, iterative learning control (ILC), nonlinear system, nonnegative matrix, optimization-based design and analysis.**

I. Introduction

LEARNING from measurement data but with no or limited model knowledge has become one of the most practically important problems in many application fields, such as robots [1], [2], batch processes [3], rail transportation [4], and servo systems [5]. This motivates a class of learning control approaches designed by mainly resorting to the measurement data, rather than to the plant models of controlled systems. One of the most popular learning control approaches is proposed with a focus on acquiring the learning abilities of robots from repetitive executions (iterations or trials), leading to the so-called "iterative learning control (ILC)" that is simple and easy to implement even with limited plant knowledge [6], [7]. Due to the operation executed using only measurement data, ILC is considered as one of the natural data-driven control approaches [8]. Since ILC is motivated from the physical learning patterns of human beings [9], it is also cataloged as one of the typical intelligent control approaches [10]. In particular, ILC effectively applies to general nonlinear plants [11], [12] and robustly works with the capability of rejecting the external disturbances/noises and initial shifts (see [13]–[19]).

One of the salient characteristics of ILC is to provide design tools to overcome shortcomings of conventional control design approaches. In particular, the design of ILC can be leveraged to improve the transient response performances for the controlled systems such that the perfect tracking objectives can be derived even in the presence of uncertain or unknown system structures and nonlinearities [9]–[11]. This class of high-performance tasks can be achieved over finite time steps gradually with increasing iterations. As a consequence, the convergence problem for ILC generally refers to the stability with respect to iteration because of the finite duration of time, which is considered as one of the key problems of ILC. There have been many effective methods to deal with ILC convergence problems, especially those based on the contraction mapping (CM) principle. To gain additional convergence properties, the optimization-based design together with CM-based analysis has been used as a good alternative for ILC in, e.g., [20]–[23]. It has been reported that optimization-based ILC can be designed to improve the convergence rate or even accomplish the monotonic convergence to better transient learning behaviors.

In the literature, there have been introduced different classes of design approaches to optimization-based ILC. The first class is called norm-optimal approach that is developed by resorting to the lifted system representation of ILC (see [24]–[32]). The norm-optimal ILC has wide potential applications for, e.g., robotic systems [24], [25], overhead cranes [26], and permanent magnet linear motors [27], regarding which practical problems have also been discussed, such as robustness against repetitive model uncertainties [28], [29], improvement of computational efficiencies [30], [31], and extension to accommodate nonlinear dynamics [26], [32]. The second class is devoted to stochastic ILC such that the optimization-based design can be explored to overcome ill effects arising from random (iteration-dependent) disturbances and noises (see [33], [34]). It is worth noting that all aforementioned optimization-based ILC approaches are either focused directly on linear systems [24], [25], [27]–[31], [31], [33], [34] or extended from linear systems to nonlinear systems with known linearized models [26], [32]. By contrast, the third class of optimization-based ILC approaches has been exploited by directly dealing with nonlinear systems subject to unknown nonlinearities, which creates data-driven or model-free optimal ILC (see [35]–[42]).

Manuscript received March 25, 2020; revised February 2, 2021; accepted March 30, 2021. This work was supported by the National Natural Science Foundation of China under Grant 61922007 and Grant 61873013. *(Corresponding author: Deyuan Meng.)*

The authors are with the Seventh Research Division, Beihang University (BUAA), Beijing 100191, China, and also with the School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100191, China (e-mail: dymeng23@126.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2021.3070920.

Digital Object Identifier 10.1109/TNNLS.2021.3070920

The data-driven optimal ILC requires no explicit models for algorithms design and convergence analyses, which is achieved by combining a dynamical linearization approach for nonlinear systems with an adaptive estimation approach for linearization parameters [35]–[41]. This also leads to a type of optimization-based adaptive ILC that permits not only the nonlinear systems but also their dynamical linearization models to have unknown dynamics and model structures. Furthermore, the optimization-based adaptive ILC has a property that its convergence analysis can be developed through the CM approach, especially through the eigenvalue-based CM approach. The eigenvalue analysis is well known as an easy-to-implement and popular approach for ILC convergence. However, despite these good properties, the eigenvalue-based CM approach is restricted to ILC processes with iteration-independent parameters based on the basic linear system theory [43], [44]. As a consequence, those eigenvalue-based results of data-driven optimal ILC for nonlinear systems may no longer work effectively, which is the main motivation of our present study to propose new optimization-based design and analysis approaches for ILC in the presence of unknown nonlinear dynamics.

In this article, we contribute to exploiting optimization-based ILC for nonlinear systems, in which we particularly propose an adaptive updating law for estimation of unknown time-varying nonlinearities. It is shown that the boundedness of all estimated parameters can be ensured directly from an optimization-based design. Furthermore, the ILC convergence is achieved, together with the boundedness of system trajectories, for which we introduce a double-dynamics analysis (DDA) approach by leveraging the properties of nonnegative matrices. Based on comparisons with the relevant existing results, the following main contributions are summarized for our optimization-based adaptive ILC.

1) We give a design and analysis approach of optimization-based adaptive ILC for nonlinear time-varying systems. It can result in a data-driven optimal ILC algorithm that, however, differs from those of, e.g., [35]–[41], especially for the updating law of parameter estimation. An advantage of our newly given approach is that all the estimated parameters can naturally be ensured to be bounded under the selection of weighting factors in our algorithm.
2) We propose an analysis approach to address convergence problems of optimization-based adaptive ILC. It benefits from implementing a DDA-based approach to ILC based on properties of nonnegative matrices. A consequence of this is the exploration of selection conditions for learning parameters such that we not only exploit the boundedness of system trajectories but also achieve the perfect output tracking tasks. Furthermore, our ILC convergence results avoid performing the eigenvalue analysis that is required in, e.g., [35]–[39], and [41].

In addition, we carry out simulation tests to demonstrate the effectiveness of our algorithm that optimization-based adaptive ILC both guarantees the boundedness of all system trajectories and achieves the prescribed perfect tracking tasks.

We organize the remainder sections of this article as follows. In Section II, we present the optimization-based ILC problem, for which an algorithm of optimization-based adaptive ILC is designed in Section III. The main ILC convergence results are established in Section IV. Simulations are performed in Section V, and then, conclusions are made in Section VI. The proofs of all lemmas are given in appendices.

Notations: Let $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$, $\mathbb{Z} = \{1, 2, 3, \ldots\}$, $\mathbb{Z}_T = \{0, 1, \ldots, T\}$ with any $T \in \mathbb{Z}$, and $\mathbf{1}_n = [1, 1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^n$. For a matrix $A = [a_{ij}] \in \mathbb{R}^{n\times m}$, $\|A\|$ denotes any norm of A, where specifically, $\|A\|_\infty$ and $\|A\|_2$ are the maximum row sum matrix norm and the spectral norm of A, respectively. Let $m = 1$, and then, $\|A\|_\infty$ and $\|A\|_2$ become the l_∞ and l_2 norms of a vector $A \in \mathbb{R}^n$, respectively. When $m = n$, $\rho(A)$ denotes the spectral radius of a square matrix $A \in \mathbb{R}^{n\times n}$. We call A a nonnegative matrix if $a_{ij} \geq 0$, $\forall i = 1, 2, \ldots, n$, $\forall j = 1, 2, \ldots, m$, which is denoted by $A \geq 0$. A trivial nonnegative matrix induced by A is $|A| = [|a_{ij}|] \geq 0$, and for any two matrices $A, B \in \mathbb{R}^{n\times m}$, $A \geq B$ means $A - B \geq 0$. For a matrix sequence $\{A_i \in \mathbb{R}^{n\times m} : i \in \mathbb{Z}_+\}$, let $\sum_{i=h}^{j} A_i = 0$ (i.e., the null matrix of appropriate dimensions) for $j = -1$ and $h = 0$, and for $m = n$, let $\prod_{i=h}^{j} A_i = A_j A_{j-1} \cdots A_h$ if $j \geq h$ and $\prod_{i=h}^{j} A_i = I$ (i.e., the identity matrix of appropriate dimensions) if $j < h$. A difference operator of a vector $\tau_k(t) \in \mathbb{R}^n$ is defined as $\Delta : \tau_k(t) \to \Delta\tau_k(t) = \tau_k(t) - \tau_{k-1}(t)$, $\forall k \in \mathbb{Z}$ and $\forall t \in \mathbb{Z}_+$.

II. Problem Statement

Consider a class of nonlinear discrete-time-varying systems with input–output dynamics described by

$$y_k(t+1) = f(y_k(t), \ldots, y_k(t-l), u_k(t), \ldots, u_k(t-n), t)$$
$$\text{with } y_k(i) = \begin{cases} 0, & i < 0 \\ y_0, & i = 0 \end{cases} \text{ and } u_k(i) = 0, \quad i < 0 \tag{1}$$

where $t \in \mathbb{Z}_{T-1}$ and $k \in \mathbb{Z}_+$ are the time and iteration indexes, respectively, $y_k(t) \in \mathbb{R}$ and $u_k(t) \in \mathbb{R}$ are the output and input, respectively, $l \in \mathbb{Z}_+$ and $n \in \mathbb{Z}_+$ are nonnegative integers that represent the system output and input orders, respectively, and

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

is an unknown nonlinear function. For the sake of convenience, we write this nonlinear function as f or $f(x_1, x_2, \ldots, x_{l+n+3})$, where $x_i \in \mathbb{R}$, $i = 1, 2, \ldots, l+n+3$ denotes the ith independent variable of f.

Problem Statement: Given any desired reference trajectory $y_d(t) \in \mathbb{R}$ over $t \in \mathbb{Z}_T$, the objective of this article is to design an ILC algorithm based on solving an optimization problem such that the uncertain nonlinear system (1) achieves the following perfect tracking task:

$$\lim_{k\to\infty} y_k(t) = y_d(t) \quad \forall t \in \mathbb{Z}_T \setminus \{0\}. \tag{2}$$

Correspondingly, we are interested in the optimization problem by leveraging the following index over $t \in \mathbb{Z}_{T-1}$ and $k \in \mathbb{Z}$ (see also [37] and [38]):

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

where $e_k(t) = y_d(t) - y_k(t)$ denotes the (output) tracking error, $\Delta u_k(t) = u_k(t) - u_{k-1}(t)$ represents the input error between two sequential iterations, and $\lambda > 0$ and $\gamma_i > 0$, $i = 1, 2, \ldots, m$, are some positive learning parameters. In (3), we consider a high order $m \in \mathbb{Z}$ for the tracking errors of interest over iterations and adopt $e_i(t+1) = 0$, $\forall t \in \mathbb{Z}_{T-1}$ if $i < 0$.

To address the abovementioned ILC problem, we introduce a fundamental assumption for the continuous differentiability of the unknown nonlinear function f.

(A1) Let f be continuously differentiable such that the partial derivatives with respect to the first $l+n+2$ independent variables are bounded, namely

$$\left|\frac{\partial f}{\partial x_i}(x_1, x_2, \ldots, x_{l+n+2}, t)\right| \le \beta_{\overline{f}} \quad \forall x_i \in \mathbb{R}, i = 1, 2, \ldots, l+n+2 \quad \forall t \in \mathbb{Z}_{T-1} \tag{4}$$

where $\beta_{\overline{f}} > 0$ is some finite bound. Furthermore, let the input–output coupling function, defined by $\partial f / \partial x_{l+2}$, be sign-fixed, which without any loss of generality is considered to be positive, namely

$$\frac{\partial f}{\partial x_{l+2}}(x_1, x_2, \ldots, x_{l+n+2}, t) \ge \beta_{\underline{f}} \quad \forall x_i \in \mathbb{R}, i = 1, 2, \ldots, l+n+2 \quad \forall t \in \mathbb{Z}_{T-1} \tag{5}$$

for some finite bound $\beta_{\underline{f}} > 0$.

Remark 1: In general, the globally Lipschitz condition is one of the basic requirements of nonlinear ILC [8]–[11], which can be satisfied for the nonlinear system (1) under Assumption (A1). To be specific, if we apply the mean value theorem (see [45, p. 651]), then for any x_i and $\bar{x}_i$, $i = 1, 2, \ldots, l+n+2$, there exists some $z_i = \sigma x_i + (1-\sigma)\bar{x}_i$ with $\sigma \in [0, 1]$ such that

$$f(x_1, x_2, \ldots, x_{l+n+2}, t) - f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{l+n+2}, t) = \sum_{i=1}^{l+n+2} \left.\frac{\partial f}{\partial x_i}\right|_{(z_1, z_2, \ldots, z_{l+n+2}, t)} (x_i - \bar{x}_i) \quad \forall t \in \mathbb{Z}_{T-1}$$

which, together with (4), leads to

$$|f(x_1, x_2, \ldots, x_{l+n+2}, t) - f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{l+n+2}, t)| \le \beta_{\overline{f}} \sum_{i=1}^{l+n+2} |x_i - \bar{x}_i| \quad \forall t \in \mathbb{Z}_{T-1}.$$

Since less plant information on the uncertain nonlinear system (1) is known, an adaptive ILC law is generally needed to reach the perfect tracking objective (2) via handling the optimization problem with the index (3) (see [35]–[38]). This requires the sign of the system input–output coupling function $\partial f / \partial x_{l+2}$ to be fixed in the optimization-based design of adaptive ILC, as made in Assumption (A1). In particular, we can see from (4) and (5) that $\partial f / \partial x_{l+2}$ is not only sign-fixed but also bounded. Namely, we have

$$\frac{\partial f}{\partial x_{l+2}}(x_1, x_2, \ldots, x_{l+n+2}, t) \in \left[\beta_{\underline{f}}, \beta_{\overline{f}}\right] \quad \forall x_i \in \mathbb{R}, i = 1, 2, \ldots, l+n+2 \quad \forall t \in \mathbb{Z}_{T-1}. \tag{6}$$

III. Optimization-Based Adaptive ILC

In this section, we present a design method for optimization-based adaptive ILC, regardless of controlled systems subject to unknown nonlinear time-varying dynamics. We thus propose a helpful lemma to develop an extended dynamical linearization for the unknown nonlinear time-varying dynamics such that we may realize an adaptive ILC design by solving the optimization problem with the index (3).

Lemma 1: For the nonlinear system (1) under Assumption (A1), an extended dynamical linearization can be given by

$$\begin{bmatrix} y_i(1) \\ y_i(2) \\ \vdots \\ y_i(T) \end{bmatrix} - \begin{bmatrix} y_j(1) \\ y_j(2) \\ \vdots \\ y_j(T) \end{bmatrix} = \Theta_{i,j} \left(\begin{bmatrix} u_i(0) \\ u_i(1) \\ \vdots \\ u_i(T-1) \end{bmatrix} - \begin{bmatrix} u_j(0) \\ u_j(1) \\ \vdots \\ u_j(T-1) \end{bmatrix} \right) \tag{7}$$

together with $\Theta_{i,j}$, $\forall i, j \in \mathbb{Z}_+$ being given in a lower triangular matrix form of

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

of which all nonzero entries can be guaranteed to be bounded, namely, for some finite bound $\beta_\theta > 0$

$$\left|\theta_{i,j,t}(\xi)\right| \le \beta_\theta \quad \forall \xi \in \mathbb{Z}_t \ \forall t \in \mathbb{Z}_{T-1} \ \forall i, j \in \mathbb{Z}_+. \tag{8}$$

In particular, the diagonal entries of $\Theta_{i,j}$ satisfy

$$\theta_{i,j,t}(t) \in \left[\beta_{\underline{f}}, \beta_{\overline{f}}\right] \quad \forall t \in \mathbb{Z}_{T-1} \ \forall i, j \in \mathbb{Z}_+. \tag{9}$$

Proof: This lemma can be established based on the differential mean value theorem and with the derivation rules of compound functions, where the facts of (4) and (6) should be noticed. The proof is easy to be developed, which is hence omitted here for clarity (for the proof details, see the proof of Lemma 1 given at https://arxiv.org/abs/1908.02447). □

Note that to derive (8), we generally have $\beta_\theta \ge \beta_{\overline{f}}$ in Lemma 1. This leads to the estimation of $\theta_{i,j,t}(t)$ in a more reasonable form (9), rather than $\theta_{i,j,t}(t) \in [\beta_{\underline{f}}, \beta_\theta]$, $\forall t \in \mathbb{Z}_{T-1}$, $\forall i, j \in \mathbb{Z}_+$. As an application of Lemma 1, we focus upon the input–output relationship between two sequential iterations k and $k-1$ for the nonlinear system (1). Namely, by letting $i = k$ and $j = k-1$ in (7), we can obtain

$$\Delta y_k(t+1) = \sum_{i=0}^{t} \theta_{k,k-1,t}(i) \Delta u_k(i) \triangleq \Delta \overrightarrow{u_k}^{\mathrm{T}}(t) \overrightarrow{\theta_{k,k-1,t}}(t) \quad \forall t \in \mathbb{Z}_{T-1} \ \forall k \in \mathbb{Z} \tag{10}$$

where $\overrightarrow{u_k}(t)$ and $\overrightarrow{\theta_{k,k-1,t}}(t)$ are defined, respectively, from $u_k(t)$ and $\theta_{k,k-1,t}(t)$ as

$$\overrightarrow{u_k}(t) = [u_k(0), u_k(1), \ldots, u_k(t)]^{\mathrm{T}}$$
$$\overrightarrow{\theta_{k,k-1,t}}(t) = \left[\theta_{k,k-1,t}(0), \theta_{k,k-1,t}(1), \ldots, \theta_{k,k-1,t}(t)\right]^{\mathrm{T}}.$$

In view of this discussion, we proceed to explore the "optimal solution" for the nonlinear system (1) to solve the optimization problem with the index (3).

Lemma 2: For the nonlinear system (1) under Assumption (A1), the solution that optimizes the index (3) for $t \in \mathbb{Z}_{T-1}$ and $k \in \mathbb{Z}$ can be presented in an updating form of

$$\begin{aligned}u_k(t) &= u_{k-1}(t)\\&\quad-\frac{\gamma_1^2\theta_{k,k-1,t}(t)}{\lambda+\gamma_1^2\theta_{k,k-1,t}^2(t)}\sum_{i=0}^{t-1}\theta_{k,k-1,t}(i)\big[u_k(i)-u_{k-1}(i)\big]\\&\quad+\frac{\gamma_1\theta_{k,k-1,t}(t)}{\lambda+\gamma_1^2\theta_{k,k-1,t}^2(t)}\bigg[\gamma_1 e_{k-1}(t+1)\\&\qquad+\sum_{i=2}^{m}\gamma_i e_{k-i+1}(t+1)\bigg]. \quad (11)\end{aligned}$$

Proof: This lemma can be obtained by incorporating (10) into optimizing the index (3) for the nonlinear system (1), of which the proof details are given in Appendix A. □

By Lemma 2, if the exact information of $\theta_{k,k-1,t}(i)$, $\forall k \in \mathbb{Z}$, $\forall t \in \mathbb{Z}_{T-1}$, and $\forall i \in \mathbb{Z}_t$, is available, then the input $u_k(t)$ determined by (11) is capable of optimizing the index (3). However, these parameters are induced by the dynamical linearization of the unknown nonlinear time-varying dynamics involved in (1) (see Lemma 1), which thus are generally unavailable. To overcome this problem, we present a parameter estimation algorithm with respect to iteration such that we can develop an estimated value $\hat{\theta}_{k,k-1,t}(i)$ of $\theta_{k,k-1,t}(i)$, $\forall k \in \mathbb{Z}$, $\forall t \in \mathbb{Z}_{T-1}$, and $\forall i \in \mathbb{Z}_t$. We explore an optimization-based approach to calculating these estimation parameters based on the following index:

$$\begin{aligned}H\left(\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right) &= \left[\Delta y_{k-1}(t+1)-\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right]^2\\&\quad+\mu_1\left\|\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)-\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t)\right\|_2^2\\&\quad+\mu_2\left\|\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right\|_2^2 \quad \forall t\in\mathbb{Z}_{T-1}\ \forall k\geq 2\end{aligned} \quad (12)$$

where $\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)$ is the estimation of $\overrightarrow{\theta_{k,k-1,t}}(t)$, defined as

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

and $\mu_1 \geq 0$ and $\mu_2 \geq 0$ are two nonnegative weighting factors. It is worth emphasizing that (12) proposes a new optimization index to accomplish the estimation of unknown system parameters in optimization-based adaptive ILC, by contrast with the existing relevant results of, e.g., [35]–[41].

Remark 2: For three terms involved in (12), the first term is to provide $\overrightarrow{\theta_{k,k-1,t}}(t)$ with a reasonable estimation $\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)$, the second term is to render $\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)$ a slowly varying estimation along the iteration axis, and the third term is to guarantee the boundedness of $\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)$. It is further expected that $\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)$ does not converge to zero with the increasing of iterations. This can be realized through the selections of two weighting factors in (12). In fact, the bigger each of the weighting factors is, the more important its weighted term is in (12). As the candidates, $\mu_1 = 1$ may be directly employed without any loss of generality, whereas a relatively small μ_2 should be adopted to both ensure $\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)$ to be bounded and avoid it to converge to zero (e.g., $\mu_2 = 0.001$ may be used when $\mu_1 = 1$, like that adopted in the examples of Section V).

Remark 3: By comparison to the existing optimization-based nonlinear ILC of, e.g., [37] and [38], we add the third term in the optimization index (12). Although this seems a slight extension, it is worth emphasizing that a new optimization-based design of nonlinear ILC can be created to not only induce a parameter estimation algorithm but also yield a strict contraction process for the property analysis of estimated parameters. Our design via the adding of the third term in (12) is crucial for the use of CM-based methods to avoid the design drawbacks of the existing methods for optimization-based nonlinear ILC. For example, it makes it possible to avoid using the eigenvalue-based analysis in exploiting the boundedness of estimated parameters (see Theorem 1 as well as its proof for more details).

To proceed further with the index (12) by finding its optimal solution, the following lemma presents an updating law for the parameter estimation.

Lemma 3: For $t \in \mathbb{Z}_{T-1}$ and $k \geq 2$, the solution that optimizes the index (12) can be proposed in an updating form of

$$\begin{aligned}\overrightarrow{\hat{\theta}_{k,k-1,t}}(t) &= \frac{\mu_1}{\mu_1+\mu_2}\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t)\\&\quad+\frac{1}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\\&\quad\times\bigg[\Delta y_{k-1}(t+1)\\&\qquad-\frac{\mu_1}{\mu_1+\mu_2}\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t)\bigg]\Delta\overrightarrow{u_{k-1}}(t).\end{aligned} \quad (13)$$

Proof: This lemma can be established by optimizing the index (12) via a similar idea as the proof of Lemma 2 (see Appendix B for the proof details). □

We are in position to leverage the development of Lemmas 2 and 3 to propose an optimization-based adaptive ILC algorithm for the uncertain nonlinear system (1)

Algorithm 1 Optimization-Based Adaptive ILC
1) do step (S1);
2) let $k = 1$, and go to step 3) to start iteration;
3) apply $u_{k-1}(t)$ to operate the nonlinear system (1);
4) do step (S2) if $k \geq 2$; otherwise, go directly to step 5);
5) do step (S3);
6) let $k = k + 1$, and go back to step 3);

in which the steps (S1), (S2), and (S3) are presented as follows.

(S1) For any $t \in \mathbb{Z}_{T-1}$, choose any bounded initial input $u_0(t)$ and initial estimated value $\hat{\theta}_{1,0,t}(i)$ of $\theta_{1,0,t}(i)$, $\forall i \in \mathbb{Z}_t$. In particular, given any (small) scalar $\epsilon > 0$, choose $\hat{\theta}_{1,0,t}(t)$ such that

$$\hat{\theta}_{1,0,t}(t) \geq \epsilon \quad \forall t \in \mathbb{Z}_{T-1}. \quad (14)$$

(S2) For any $t \in \mathbb{Z}_{T-1}$, apply an updating law of the parameter estimation with respect to each iteration $k \geq 2$

and each time step $i \in \mathbb{Z}_t$ as

$$\begin{aligned}\hat{\theta}_{k,k-1,t}(i) = &\frac{\mu_1}{\mu_1+\mu_2}\hat{\theta}_{k-1,k-2,t}(i)\\&+\frac{\Delta u_{k-1}(i)}{\mu_1+\mu_2+\sum_{j=0}^{t}\Delta u_{k-1}^2(j)}\\&\times\Bigg[\Delta y_{k-1}(t+1)\\&\quad-\frac{\mu_1}{\mu_1+\mu_2}\sum_{j=0}^{t}\hat{\theta}_{k-1,k-2,t}(j)\Delta u_{k-1}(j)\Bigg].\end{aligned}\tag{15}$$

In particular, if $\hat{\theta}_{k,k-1,t}(t) < \epsilon$, then set

$$\hat{\theta}_{k,k-1,t}(t) = \hat{\theta}_{1,0,t}(t) \quad \forall t \in \mathbb{Z}_{T-1}.\tag{16}$$

(S3) For any $t \in \mathbb{Z}_{T-1}$, apply an updating law with respect to the input for each iteration $k \in \mathbb{Z}$ as

$$\begin{aligned}u_k(t) = &u_{k-1}(t)\\&-\frac{\gamma_1^2\hat{\theta}_{k,k-1,t}(t)}{\lambda+\gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\sum_{i=0}^{t-1}\hat{\theta}_{k,k-1,t}(i)\big[u_k(i)-u_{k-1}(i)\big]\\&+\frac{\gamma_1\hat{\theta}_{k,k-1,t}(t)}{\lambda+\gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\Bigg[\gamma_1 e_{k-1}(t+1)\\&\qquad+\sum_{i=2}^{m}\gamma_i e_{k-i+1}(t+1)\Bigg].\end{aligned}\tag{17}$$

Remark 4: From (14) and (16), we can obtain $\hat{\theta}_{k,k-1,t}(t) \geq \epsilon$, $\forall k \in \mathbb{Z}$ and $\forall t \in \mathbb{Z}_{T-1}$. This discloses that ϵ represents the smallest acceptable value of the estimation $\hat{\theta}_{k,k-1,t}(t)$ for the parameter $\theta_{k,k-1,t}(t)$, $\forall k \in \mathbb{Z}$ and $\forall t \in \mathbb{Z}_{T-1}$. The application of Algorithm 1 thus naturally avoids the zero convergence of $\hat{\theta}_{k,k-1,t}(t)$ along the iteration axis. In particular, when noting (5) or (6), we can also find that (14) guarantees $\hat{\theta}_{1,0,t}(t)$ to have the same sign as $\partial f/\partial x_{l+2}$. It is important for the practical implementation of adaptive ILC. In addition, Algorithm 1 greatly generalizes the existing optimization-based adaptive ILC algorithms in the literature. For example, if we set $\mu_2 = 0$, then Algorithm 1 collapses into the high-order adaptive ILC algorithm of, e.g., [37] and [38]; and furthermore, if we take $m = 1$, then it becomes the first-order adaptive ILC algorithm of, e.g., [35] and [36].

IV. Convergence Analysis Results

We next contribute to exploring the convergence analysis of the nonlinear system (1) that operates under Algorithm 1 of optimization-based adaptive ILC. Toward this end, we resort to the tracking error and can employ (10) to equivalently derive

$$\begin{aligned}e_k(t+1) &= e_{k-1}(t+1) - \Delta y_k(t+1)\\&= e_{k-1}(t+1)\\&\quad-\sum_{i=0}^{t}\theta_{k,k-1,t}(i)\Delta u_k(i) \quad \forall t \in \mathbb{Z}_{T-1}\ \forall k \in \mathbb{Z}\end{aligned}\tag{18}$$

in which nonrepetitive (namely, iteration-dependent) uncertain parameters $\theta_{k,k-1,t}(i)$, $\forall i \in \mathbb{Z}_t$, $\forall t \in \mathbb{Z}_{T-1}$, and $\forall k \in \mathbb{Z}$ are inevitably involved. It may result in challenging difficulties for exploiting robust convergence results of ILC. For example, the eigenvalue (or spectral radius) analysis is not applicable any longer when the system (matrix) parameters of the resulting ILC process are explicitly dependent upon iteration (see [44] for more detailed discussions). The traditional CM-based method of convergence analysis may even be not effective in ILC due to nonrepetitive uncertainties (see also [46]–[48]).

To make the abovementioned observations clearer to follow, we insert (17) into (18) and can further deduce

$$\begin{aligned}e_k(t+1) &= e_{k-1}(t+1) - \theta_{k,k-1,t}(t)\Delta u_k(t)\\&\quad-\sum_{i=0}^{t-1}\theta_{k,k-1,t}(i)\Delta u_k(i)\\&=\left[1-\frac{(\gamma_1^2+\gamma_1\gamma_2)\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda+\gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\right]e_{k-1}(t+1)\\&\quad-\sum_{i=3}^{m}\frac{\gamma_1\gamma_i\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda+\gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}e_{k-i+1}(t+1)\\&\quad+\kappa_k(t)\end{aligned}\tag{19}$$

where $\kappa_k(t)$ is a driving signal given by

$$\begin{aligned}\kappa_k(t) = &\frac{\gamma_1^2\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda+\gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\sum_{i=0}^{t-1}\hat{\theta}_{k,k-1,t}(i)\Delta u_k(i)\\&-\sum_{i=0}^{t-1}\theta_{k,k-1,t}(i)\Delta u_k(i).\end{aligned}\tag{20}$$

Obviously, we can see from (19) that the system parameters of the ILC process resulting from the nonlinear system (1) under Algorithm 1 depend explicitly on $\theta_{k,k-1,t}(t)$ and $\hat{\theta}_{k,k-1,t}(t)$ and, hence, are iteration-dependent. This renders the traditional CM-based method not applicable to ILC convergence analysis, especially those CM-based methods using eigenvalue analyses, to overcome which we apply a DDA approach to optimization-based adaptive ILC by leveraging the properties of nonnegative matrices (see [49, Ch. 8] for the detailed properties of nonnegative matrices).

A. Boundedness of Estimated System Parameters

As noted in (19) and (20), the uncertain parameter $\theta_{k,k-1,t}(i)$ and its estimation $\hat{\theta}_{k,k-1,t}(i)$, $\forall i \in \mathbb{Z}_t$, both play crucial roles in optimization-based adaptive ILC along the iteration axis. With Lemma 1, we can obtain a basic boundedness property of each parameter $\theta_{k,k-1,t}(i)$, $\forall i \in \mathbb{Z}_t$, whereas we gain the estimation $\hat{\theta}_{k,k-1,t}(i)$, $\forall i \in \mathbb{Z}_t$, in the process of applying Algorithm 1 to the nonlinear system (1), for which it is needed to determine whether the basic boundedness property holds. An affirmative answer to this question is provided in the following theorem.

Theorem 1: For the nonlinear system (1), let Assumption (A1) hold. If Algorithm 1 is applied with $\mu_1 > 0$ and $\mu_2 > 0$, then the boundedness of the estimation $\hat{\theta}_{k,k-1,t}(i)$ can be guaranteed such that

$$\left|\hat{\theta}_{k,k-1,t}(i)\right| \leq \beta_{\hat{\theta}} \quad \forall i \in \mathbb{Z}_t\ \forall t \in \mathbb{Z}_{T-1}\ \forall k \in \mathbb{Z}\tag{21}$$

for some finite bound $\beta_{\hat{\theta}} > 0$. In particular, $\hat{\theta}_{k,k-1,t}(t)$ satisfies

$$\hat{\theta}_{k,k-1,t}(t) \in \left[\epsilon, \beta_{\hat{\theta}}\right] \quad \forall t \in \mathbb{Z}_{T-1} \ \forall k \in \mathbb{Z}. \tag{22}$$

With Theorem 1, it is revealed that the estimated parameters of all the nonrepetitive uncertain parameters $\theta_{k,k-1,t}(i)$, $\forall i \in \mathbb{Z}_t$, $\forall t \in \mathbb{Z}_{T-1}$, and $\forall k \in \mathbb{Z}$ are bounded when employing Algorithm 1 for the nonlinear system (1). This boundedness result resorts to no conditions on the input updating law (17), which is even independent of the selections of the learning parameters λ and γ_i, $i = 1, 2, \ldots, m$. Furthermore, Theorem 1 is naturally ensured with $\mu_1 > 0$ and $\mu_2 > 0$ in the updating law (15) for parameter estimation but without adding a step-size factor in (15), which is different from, e.g., [37] and [38].

To prove Theorem 1, a useful lemma on the norm estimation of an iteration-dependent matrix operator is given as follows.

Lemma 4: For any $t \ge 0$ and $k \ge 2$, $\|Q(\Delta\overrightarrow{u_{k-1}}(t))\|_2 \le 1$ holds for a square matrix $Q(\Delta\overrightarrow{u_{k-1}}(t))$ defined as

$$Q\left(\Delta\overrightarrow{u_{k-1}}(t)\right) = I - \frac{\Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)}{\mu_1 + \mu_2 + \left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}. \tag{23}$$

Proof: The proof of this lemma can be given by exploiting the specific symmetric structure of $Q(\Delta\overrightarrow{u_{k-1}}(t))$, where the details are given in Appendix C. □

With Lemma 4, we show the proof of Theorem 1 as follows.

Proof of Theorem 1: It can be seen that (15) in Algorithm 1 is equivalently derived from (13) in Lemma 3. We thus revisit (13) and can employ (23) to deduce

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

where we also insert

$$\begin{aligned}
&\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t)\Delta\overrightarrow{u_{k-1}}(t) \\
&\qquad = \Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t).
\end{aligned}$$

By combining (8) and (10), we can validate

$$\begin{aligned}
|\Delta y_{k-1}(t+1)| &= \left|\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)\overrightarrow{\theta_{k-1,k-2,t}}(t)\right| \\
&\le \left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2\left\|\overrightarrow{\theta_{k-1,k-2,t}}(t)\right\|_2 \\
&= \left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2\sqrt{\sum_{i=0}^{t}\theta_{k-1,k-2,t}^2(i)} \\
&\le \sqrt{t+1}\beta_\theta\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2 \\
&\le \sqrt{T}\beta_\theta\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2.
\end{aligned} \tag{25}$$

We further explore (25) to derive

$$\begin{aligned}
\left\|\frac{\Delta y_{k-1}(t+1)\Delta\overrightarrow{u_{k-1}}(t)}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\right\|_2 &\le \frac{\sqrt{T}\beta_\theta\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2} \\
&\le \sqrt{T}\beta_\theta.
\end{aligned} \tag{26}$$

With Lemma 4, we consider (26) for (24) and can obtain

$$\begin{aligned}
\left\|\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right\|_2 &\le \left\|\frac{\mu_1}{\mu_1+\mu_2}Q\left(\Delta\overrightarrow{u_{k-1}}(t)\right)\right\|_2\left\|\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t)\right\|_2 \\
&\quad + \left\|\frac{\Delta y_{k-1}(t+1)\Delta\overrightarrow{u_{k-1}}(t)}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\right\|_2 \\
&\le \frac{\mu_1}{\mu_1+\mu_2}\left\|\overrightarrow{\hat{\theta}_{k-1,k-2,t}}(t)\right\|_2 + \sqrt{T}\beta_\theta
\end{aligned} \tag{27}$$

which can be adopted to yield

$$\begin{aligned}
\left\|\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right\|_2 &\le \left(\frac{\mu_1}{\mu_1+\mu_2}\right)^{k-1}\left\|\overrightarrow{\hat{\theta}_{1,0,t}}(t)\right\|_2 \\
&\quad + \sum_{i=0}^{k-2}\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^{i}\sqrt{T}\beta_\theta.
\end{aligned} \tag{28}$$

Due to $\mu_1/(\mu_1+\mu_2) < 1$, we can verify with (28) that

$$\begin{aligned}
\left\|\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\right\|_2 &\le \left\|\overrightarrow{\hat{\theta}_{1,0,t}}(t)\right\|_2 + \frac{\mu_1+\mu_2}{\mu_2}\sqrt{T}\beta_\theta \\
&\le \beta_{\hat{\theta}} \quad \forall t \in \mathbb{Z}_{T-1} \ \forall k \in \mathbb{Z}
\end{aligned} \tag{29}$$

where

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

Since $|\hat{\theta}_{k,k-1,t}(i)| \le \|\overrightarrow{\hat{\theta}_{k,k-1,t}}(t)\|_2$, $\forall i \in \mathbb{Z}_t$ holds, (21) follows as a direct consequence of (29). In particular, we can develop (22) by also considering that (16) ensures $\hat{\theta}_{k,k-1,t}(t) \ge \epsilon$. □

Remark 5: It is worth emphasizing that (24) essentially gives a nonrepetitive system with respect to iteration because of the system matrix $\mu_1 Q(\Delta\overrightarrow{u_{k-1}}(t))/(\mu_1+\mu_2)$ (see also [46]–[48]). Despite this issue, we can develop a strict CM condition as

$$\begin{aligned}
\left\|\frac{\mu_1}{\mu_1+\mu_2}Q\left(\Delta\overrightarrow{u_{k-1}}(t)\right)\right\|_2 &\le \frac{\mu_1}{\mu_1+\mu_2} \\
&< 1 \quad \forall t \in \mathbb{Z}_{T-1} \ \forall k \ge 2
\end{aligned} \tag{30}$$

and, hence, we can directly implement the CM-based approach to the boundedness analysis for the estimation $\hat{\theta}_{k,k-1,t}(i)$ of the uncertain parameter $\theta_{k,k-1,t}(i)$, $\forall i \in \mathbb{Z}_t$, $\forall t \in \mathbb{Z}_{T-1}$, and $\forall k \in \mathbb{Z}$. In fact, such a benefit is because of the optimization-based design result of Lemma 3, which can no longer be gained for $\mu_2 = 0$. We can consequently see that we may improve the boundedness analysis method used in, e.g., [37] and [38].

B. Convergence of Optimization-Based Adaptive ILC

We proceed to explore the system performances of (1) under Algorithm 1 of optimization-based adaptive ILC, including the boundedness of the system trajectories and the convergence of the tracking error. We thus revisit (19) that essentially shows a nonrepetitive higher order ILC process regarding the tracking error. To overcome the effect of higher order dynamics on ILC, we resort to a lifting technique to reformulate (19) as

$$\overrightarrow{e_k}(t+1) = P_k(t)\overrightarrow{e_{k-1}}(t+1) + \overrightarrow{\kappa_k}(t) \tag{31}$$

where $\overrightarrow{e_k}(t+1)$ and $\overrightarrow{\kappa_k}(t)$ are two vectors defined by

$$\begin{aligned}\overrightarrow{e_k}(t+1) &= \left[e_k(t+1), e_{k-1}(t+1), \ldots, e_{k-m+2}(t+1)\right]^{\mathrm{T}} \in \mathbb{R}^{m-1}\\ \overrightarrow{\kappa_k}(t) &= [\kappa_k(t), 0, \ldots, 0]^{\mathrm{T}} \in \mathbb{R}^{m-1}\end{aligned} \tag{32}$$

and $P_k(t) \in \mathbb{R}^{(m-1)\times(m-1)}$ is a correspondingly induced matrix in the form of

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

$$\begin{aligned} p_{1,k}(t) &= 1 - \frac{\left(\gamma_1^2 + \gamma_1\gamma_2\right)\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\\ p_{i,k}(t) &= -\frac{\gamma_1\gamma_{i+1}\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}, \quad i = 2, 3, \ldots, m-1. \end{aligned} \tag{33}$$

For (31), we can develop a convergence result by leveraging the nonrepetitive ILC results of, e.g., [46].

Lemma 5: For (31) over any $t \in \mathbb{Z}_{T-1}$, if

(C) there exist some iteration sequence $\{\omega_s(t)\}_{s=0}^{\infty}$ and some finite positive integer $\chi(t)$, satisfying $\omega_0(t) = 1$ and $0 < \omega_{s+1}(t) - \omega_s(t) \le \chi(t)$, such that

$$\left\|\prod_{k=\omega_s(t)}^{\omega_{s+1}(t)-1} P_k(t)\right\|_\infty \le \eta < 1 \quad \forall s \in \mathbb{Z}_+$$

then the following two results hold.

1) $\overrightarrow{e_k}(t+1)$ is bounded (that is, $\sup_{k\in\mathbb{Z}_+}\|\overrightarrow{e_k}(t+1)\|_\infty \le \beta_{\vec{e}}(t)$ for some finite bound $\beta_{\vec{e}}(t) > 0$), provided that $\overrightarrow{\kappa_k}(t)$ is bounded (that is, $\sup_{k\in\mathbb{Z}_+}\|\overrightarrow{\kappa_k}(t)\|_\infty \le \beta_{\vec{\kappa}}(t)$ for some finite bound $\beta_{\vec{\kappa}}(t) > 0$).
2) $\lim_{k\to\infty}\overrightarrow{e_k}(t+1) = 0$ (exponentially fast), provided that $\lim_{k\to\infty}\overrightarrow{\kappa_k}(t) = 0$ (exponentially fast).

Proof: The two results of this lemma can be established based on (31) by following the same steps as the establishments of the results 1) and 2) of [46, Lemma 2], respectively, for which the details are thus omitted here for simplicity. □

For the condition (C) in Lemma 5, we can verify

$$\left\|\prod_{k=\omega_s(t)}^{\omega_{s+1}(t)-1} P_k(t)\right\|_\infty \le \left\|\prod_{k=\omega_s(t)}^{\omega_{s+1}(t)-1} |P_k(t)|\right\|_\infty \quad \forall t \in \mathbb{Z}_+\ \forall k \in \mathbb{Z} \tag{34}$$

in which $|P_k(t)|$, compared with $P_k(t)$, becomes a nonnegative matrix given by

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

We explore the fact (34) based on the properties of nonnegative matrices, together with using the convergence result of Lemma 5, to establish a convergence result with respect to the tracking error satisfying (19).

Lemma 6: For (19) over any $t \in \mathbb{Z}_{T-1}$, if

$$\left|1 - \frac{\left(\gamma_1^2 + \gamma_1\gamma_2\right)\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\right| + \sum_{i=3}^{m}\left|\frac{\gamma_1\gamma_i\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\right| \le \zeta < 1 \quad \forall k \in \mathbb{Z} \tag{35}$$

then the following two results hold.

1) $e_k(t+1)$ is bounded (namely, $\sup_{k\in\mathbb{Z}_+}|e_k(t+1)| \le \beta_e(t)$ for some finite bound $\beta_e(t) > 0$), provided that $\kappa_k(t)$ is bounded (namely, $\sup_{k\in\mathbb{Z}_+}|\kappa_k(t)| \le \beta_\kappa(t)$ for some finite bound $\beta_\kappa(t) > 0$).
2) $\lim_{k\to\infty} e_k(t+1) = 0$ (exponentially fast), provided that $\lim_{k\to\infty}\kappa_k(t) = 0$ (exponentially fast).

Proof: This lemma can be developed via a nonnegative matrix-based analysis approach and by noting the definitions (32) and (33). For the proof details, see Appendix D. □

Even though Lemma 6 may help to achieve the convergence analysis of the tracking error, it is no longer applicable for the boundedness analysis of the system trajectories. We thus resort to the DDA approach to ILC and exploit the dynamic evolution of input along the iteration axis to implement the boundedness analysis in the presence of nonrepetitive uncertainties (see also [46], [47]). Toward this end, we rewrite (17) as

$$\begin{aligned} u_k(t) = u_{k-1}(t) &+ \frac{\gamma_1\hat{\theta}_{k,k-1,t}(t)\sum_{i=3}^{m}\gamma_i e_{k-i+1}(t+1)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\\ &- \frac{\gamma_1^2\hat{\theta}_{k,k-1,t}(t)\sum_{i=0}^{t-1}\hat{\theta}_{k,k-1,t}(i)\Delta u_k(i)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\\ &+ \frac{\left(\gamma_1^2 + \gamma_1\gamma_2\right)\hat{\theta}_{k,k-1,t}(t)\left[y_d(t+1) - y_{k-1}(t+1)\right]}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}. \end{aligned} \tag{36}$$

As a direct application of (7) for the initial iteration (i.e., $j = 0$) and the $(k-1)$th iteration (i.e., $i = k-1$), we can derive

$$\begin{aligned} y_{k-1}(t+1) &= y_0(t+1) + \sum_{i=0}^{t}\theta_{k-1,0,t}(i)\left[u_{k-1}(i) - u_0(i)\right]\\ &= \theta_{k-1,0,t}(t)u_{k-1}(t) + y_0(t+1)\\ &\quad + \sum_{i=0}^{t-1}\theta_{k-1,0,t}(i)u_{k-1}(i) - \sum_{i=0}^{t}\theta_{k-1,0,t}(i)u_0(i). \end{aligned} \tag{37}$$

By substituting (37) into (36), we can obtain

$$u_k(t) = \left[1 - \frac{\left(\gamma_1^2 + \gamma_1\gamma_2\right)\hat{\theta}_{k,k-1,t}(t)\theta_{k-1,0,t}(t)}{\lambda + \gamma_1^2\hat{\theta}_{k,k-1,t}^2(t)}\right]u_{k-1}(t) + \psi_k(t) \tag{38}$$

where $\psi_k(t)$ is a driving signal given by

$$\begin{aligned}\psi_k(t) = \frac{\gamma_1 \hat{\theta}_{k,k-1,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \Bigg[&(\gamma_1+\gamma_2)\sum_{i=0}^{t} \theta_{k-1,0,t}(i) u_0(i) \\ &- (\gamma_1+\gamma_2)\sum_{i=0}^{t-1} \theta_{k-1,0,t}(i) u_{k-1}(i) \\ &- \gamma_1 \sum_{i=0}^{t-1} \hat{\theta}_{k,k-1,t}(i) \Delta u_k(i) \\ &+ (\gamma_1+\gamma_2) e_0(t+1) \\ &+ \sum_{i=3}^{m} \gamma_i e_{k-i+1}(t+1) \Bigg]. \end{aligned} \tag{39}$$

With (38), we propose a lemma for the boundedness of the input $u_k(t)$ with respect to any bounded driving signal $\psi_k(t)$.

Lemma 7: For (38), over any $t \in \mathbb{Z}_{T-1}$, if

$$\left| 1 - \frac{\left(\gamma_1^2 + \gamma_1\gamma_2\right)\hat{\theta}_{k,k-1,t}(t)\theta_{k-1,0,t}(t)}{\lambda + \gamma_1^2 \hat{\theta}_{k,k-1,t}^2(t)} \right| \le \phi < 1 \quad \forall k \in \mathbb{Z} \tag{40}$$

then $u_k(t)$ is ensured to be bounded such that $\sup_{k\in\mathbb{Z}_+} |u_k(t)| \le \beta_u(t)$ holds for some finite bound $\beta_u(t) > 0$, provided that $\psi_k(t)$ is bounded (i.e., $\sup_{k\in\mathbb{Z}_+} |\psi_k(t)| \le \beta_\psi(t)$ for some finite bound $\beta_\psi(t) > 0$).

Proof: A consequence of the result 1) of [46, Lemma 2]. □

Based on the analysis results of Lemmas 6 and 7, we present the following theorem to develop tracking results for uncertain nonlinear systems using Algorithm 1 of optimization-based adaptive ILC.

Theorem 2: For the nonlinear system (1), let Assumption (A1) be satisfied. If Algorithm 1 is applied with $\mu_1 > 0$, $\mu_2 > 0$, and

$$\gamma_1 + \gamma_2 > \sum_{i=3}^{m} \gamma_i, \quad \lambda > \left(\gamma_1^2 + \gamma_1\gamma_2\right)\beta_{\overline{f}}\beta_{\hat{\theta}} \tag{41}$$

then both boundedness and convergence of optimization-based adaptive ILC can be ensured, namely, the following conditions hold.

1) The boundedness of the input and output trajectories can be guaranteed during the ILC process such that

$$\begin{aligned} |u_k(t)| &\le \beta_u \quad \forall t \in \mathbb{Z}_{T-1}\ \forall k \in \mathbb{Z}_+ \\ |y_k(t)| &\le \beta_y \quad \forall t \in \mathbb{Z}_T\ \forall k \in \mathbb{Z}_+ \end{aligned} \tag{42}$$

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

2) The perfect tracking objective (2) of ILC can be achieved with its limit being approached exponentially fast.

Remark 6: From Theorem 2, it can be seen that Algorithm 1 is effective in accomplishing the perfect output tracking tasks of ILC, together with guaranteeing the boundedness of all the system trajectories, even in the presence of unknown nonlinear time-varying dynamics. Moreover, it is worth highlighting that Theorem 2 actually provides a class of data-driven ILC results because the implementation of the nonlinear system (1) under Algorithm 1 leverages only the input and output data. Also, we need quite limited estimation knowledge of (1) to establish Theorem 2, as well as to gain Theorem 1. Similar contributions have been made in, e.g., [37] and [38], which, however, employ the eigenvalue-based CM approach to the convergence analysis of ILC and can no longer apply to the analysis of our results. This can be clearly seen from (19) and (38) that yield nonrepetitive ILC processes and make the eigenvalue analysis not applicable for their convergence analysis any longer (see also [44]).

Remark 7: Another issue worth noticing is the interdependent relation between (19) and (20) for the tracking error dynamics and (38) and (39) for the input dynamics. This naturally leads to that the convergence analysis of the tracking error and the boundedness analysis of the input are interdependent with each other. Thus, the CM-based approach to ILC can not be adopted to develop Theorem 2, which motivates us to implement a DDA approach. A benefit of employing a DDA approach is to make the selection condition (41) independent of the length T of the learning time interval. It consequently improves the selection condition used in, e.g., [38] that depends heavily on the length of the learning time interval. Another benefit is the exponential convergence rate ensured in Theorem 2 for optimization-based ILC in the presence of nonlinear time-varying systems.

Remark 8: To implement Algorithm 1 of optimization-based adaptive ILC, the parameters γ_i, $i = 1, 2, \ldots, m$, and λ need to be selected under the condition (41). By the selection condition $\gamma_1 + \gamma_2 > \sum_{i\ge 3} \gamma_i$ in (41), the importance of $e_k(t)$ and $e_{k-1}(t)$ is higher than that of the tracking errors from the previous iterations less than $k-1$ in the optimization index (3). For any bounds of nonlinearities and estimated parameters, there always exists some sufficiently large λ satisfying $\lambda > (\gamma_1^2 + \gamma_1\gamma_2)\beta_{\overline{f}}\beta_{\hat{\theta}}$ in (41). However, by (17), the larger λ is, the smaller $u_k(t) - u_{k-1}(t)$ is, which leads to a more conservative update process of ILC and, thus, a relatively slow convergence speed of ILC. The two weighting factors $\mu_1 > 0$ and $\mu_2 > 0$ required in Theorems 1 and 2 actually affect the parameter estimation process. As shown in Remark 2, $\mu_1 = 1$ and $\mu_2 = 0.001$ can be specifically adopted such that the boundedness of the estimated parameters $\hat{\theta}_{k,k-1,t}(i), k \in \mathbb{Z}$ and $i \in \mathbb{Z}_t$, can be ensured, and a fast decreasing process of $\hat{\theta}_{k,k-1,t}(i), k \in \mathbb{Z}$ and $i \in \mathbb{Z}_t$, can be simultaneously avoided along the iteration axis.

Although Lemmas 6 and 7 show preliminary analysis results for the development of Theorem 2, they resort to two different conditions (35) and (40). To overcome this issue, we introduce a helpful lemma to disclose the relations among the conditions (35), (40), and (41).

Lemma 8: For the nonlinear system (1) under Assumption (A1), if the condition (41) is satisfied, then both conditions (35) and (40) can be simultaneously guaranteed.

Proof: This lemma can be proved with the boundedness results of Lemma 1 and Theorem 1, where the proof details are given in Appendix E. □

Now, by utilizing Lemmas 6–8, we are in position to present the proof of Theorem 2, for which a DDA approach instead of the eigenvalue-based analysis approach to ILC is implemented.

Proof of Theorem 2: It follows by Lemma 8 that the selection condition (41) in this theorem ensures the validity of Lemmas 6 and 7. Then, we perform induction over $t \in \mathbb{Z}_{T-1}$ to complete this proof with two steps.

Step i): Let $t = 0$, and then, we prove that $\lim_{k\to\infty} e_k(1) = 0$ exponentially fast, $\lim_{k\to\infty} \Delta u_k(0) = 0$ exponentially fast, and $\sup_{k\in\mathbb{Z}_+} |u_k(0)| \le \beta_u(0)$ for some finite bound $\beta_u(0) > 0$.

From (20), it follows $\kappa_k(0) = 0$, $\forall k \in \mathbb{Z}_+$. We hence consider Lemma 6 for (19) and can obtain

$$\sup_{k\in\mathbb{Z}_+} |e_k(1)| \le \beta_e(0) \text{ and } \lim_{k\to\infty} e_k(1) = 0 \text{ exponentially fast} \tag{43}$$

for some finite bound $\beta_e(0) > 0$. From (17), we can deduce

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

which, together with (22) and (43), leads to $\lim_{k\to\infty} \Delta u_k(0) = 0$ exponentially fast. In addition, we know from (39) that

$$\psi_k(0) = \frac{\gamma_1 \hat{\theta}_{k,k-1,0}(0)}{\lambda + \gamma_1^2 \hat{\theta}^2_{k,k-1,0}(0)} \Bigg[(\gamma_1 + \gamma_2)\theta_{k-1,0,0}(0) u_0(0) + (\gamma_1 + \gamma_2) e_0(1) + \sum_{i=3}^{m} \gamma_i e_{k-i+1}(1) \Bigg]$$

and then by the boundedness results of Lemma 1 and Theorem 1, we can derive

$$\begin{aligned} |\psi_k(0)| &\le \frac{\gamma_1 \beta_{\hat{\theta}}}{\lambda + \gamma_1^2 \varepsilon^2} \left[(\gamma_1 + \gamma_2) \beta_\theta |u_0(0)| + \sum_{i=1}^{m} \gamma_i \beta_e(0) \right] \\ &\triangleq \beta_\psi(0). \end{aligned} \tag{44}$$

With (44), the use of Lemma 7 yields $\sup_{k\in\mathbb{Z}_+} |u_k(0)| \le \beta_u(0)$ for some finite bound $\beta_u(0) > 0$.

Step ii): For all $t = 0, 1, \ldots, N-1$ with any given $N \in \mathbb{Z}_{T-1}$, let $\lim_{k\to\infty} e_k(t+1) = 0$ exponentially fast, $\lim_{k\to\infty} \Delta u_k(t) = 0$ exponentially fast, and $\sup_{k\in\mathbb{Z}_+} |u_k(t)| \le \beta_u(t)$ for some finite bound $\beta_u(t) > 0$. Then, for $t = N$, we will prove that the hypothesis made for the two exponential convergence results and one boundedness result also holds.

With the hypothesis made for the time steps $0, 1, \ldots, N-1$ in Step ii) and by applying the boundedness results of Lemma 1 and Theorem 1, we can employ (20) to verify

$$\begin{aligned} |\kappa_k(N)| &= \Bigg| \frac{\gamma_1^2 \theta_{k,k-1,N}(N) \hat{\theta}_{k,k-1,N}(N)}{\lambda + \gamma_1^2 \hat{\theta}^2_{k,k-1,N}(N)} \sum_{i=0}^{N-1} \hat{\theta}_{k,k-1,N}(i) \Delta u_k(i) \\ &\quad - \sum_{i=0}^{N-1} \theta_{k,k-1,N}(i) \Delta u_k(i) \Bigg| \\ &\le 2\beta_\theta \left(1 + \frac{\gamma_1^2 \beta_{\hat{\theta}}^2}{\lambda + \gamma_1^2 \varepsilon^2} \right) \sum_{i=0}^{N-1} \beta_u(i) \\ &\triangleq \beta_\kappa(N) \quad \forall k \in \mathbb{Z}_+ \end{aligned} \tag{45}$$

and

$$\begin{aligned} |\kappa_k(N)| &\le \beta_\theta \left(1 + \frac{\gamma_1^2 \beta_{\hat{\theta}}^2}{\lambda + \gamma_1^2 \varepsilon^2} \right) \sum_{i=0}^{N-1} |\Delta u_k(i)| \\ &\to 0 \text{ exponentially fast}, \quad \text{as } k \to \infty. \end{aligned} \tag{46}$$

We then leverage (45) and (46) and apply Lemma 6 to deduce

$$\begin{aligned} &\sup_{k\in\mathbb{Z}_+} |e_k(N+1)| \le \beta_e(N) \\ &\lim_{k\to\infty} e_k(N+1) = 0 \text{ exponentially fast} \end{aligned} \tag{47}$$

for some finite bound $\beta_e(N) > 0$. Since we can use (17) to get

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

we follow the same lines as (46) and insert (47) to derive:

$$\begin{aligned} |\Delta u_k(N)| &\le \frac{\gamma_1^2 \beta_{\hat{\theta}}^2}{\lambda + \gamma_1^2 \varepsilon^2} \sum_{i=0}^{N-1} |\Delta u_k(i)| \\ &\quad + \frac{\gamma_1 \beta_{\hat{\theta}}}{\lambda + \gamma_1^2 \varepsilon^2} \Bigg[\gamma_1 |e_{k-1}(N+1)| \\ &\qquad + \sum_{i=2}^{m} \gamma_i |e_{k-i+1}(N+1)| \Bigg] \\ &\to 0 \text{ exponentially fast}, \quad \text{as } k \to \infty \end{aligned}$$

which implies that $\lim_{k\to\infty} \Delta u_k(N) = 0$ exponentially fast. From (39), we can obtain

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

with which we can validate

$$\begin{aligned} |\psi_k(N)| &\le \frac{\gamma_1 \beta_{\hat{\theta}}}{\lambda + \gamma_1^2 \varepsilon^2} \Bigg[(\gamma_1 + \gamma_2) \beta_\theta \sum_{i=0}^{N} |u_0(i)| \\ &\qquad + \left(\gamma_1 \beta_\theta + \gamma_2 \beta_\theta + 2\gamma_1 \beta_{\hat{\theta}} \right) \sum_{i=0}^{N-1} \beta_u(i) \\ &\qquad + \sum_{i=1}^{m} \gamma_i \beta_e(N) \Bigg] \\ &\triangleq \beta_\psi(N). \end{aligned} \tag{48}$$

Based on (48), we consider Lemma 7 for (38) and can develop $\sup_{k\in\mathbb{Z}_+} |u_k(N)| \le \beta_u(N)$ for some finite bound $\beta_u(N) > 0$. We can thus conclude that the hypothesis made for $t = 0, 1, \ldots, N-1$ in this step also holds for $t = N$.

TABLE I
LEARNING PARAMETERS USED IN EXAMPLE 1

λ	γ_1	γ_2	γ_3	μ_1	μ_2	ε
1	0.8	0.14	0.06	1	0.001	0.01

By induction based on the analysis of the above steps i) and ii), we can arrive at

$$\sup_{k\in\mathbb{Z}_+}|u_k(t)| \le \beta_u(t)$$
$$\lim_{k\to\infty} e_k(t+1) = 0 \text{ exponentially fast } \quad \forall t \in \mathbb{Z}_{T-1} \quad (49)$$

with which we can further employ Lemma 6 to get

$$\sup_{k\in\mathbb{Z}_+}|e_k(t+1)| \le \beta_e(t) \quad \forall t \in \mathbb{Z}_{T-1}. \quad (50)$$

The use of (50) yields $\sup_{k\in\mathbb{Z}_+}|y_k(t+1)| \le \beta_e(t)+|y_d(t+1)|$, $\forall t \in \mathbb{Z}_{T-1}$, which together with (49) leads to $\sup_{k\in\mathbb{Z}_+}|u_k(t)| \le \beta_u$, $\forall t \in \mathbb{Z}_{T-1}$, and $\sup_{k\in\mathbb{Z}_+}|y_k(t)| \le \beta_y$, $\forall t \in \mathbb{Z}_T$ by taking

$$\beta_u = \max_{t\in\mathbb{Z}_{T-1}} \beta_u(t)$$
$$\beta_y = \max\left\{|y_0|, \max_{t\in\mathbb{Z}_{T-1}}\{\beta_e(t)+|y_d(t+1)|\}\right\}.$$

We can also derive from (49) that the perfect tracking objective (2) holds with its limit being approached exponentially fast. The proof of Theorem 2 is complete. □

V. SIMULATION TESTS

For the illustration of our optimization-based adaptive ILC algorithm, we present a numerical example and also implement simulation tests on an injection molding process.

Example 1: Let $T = 50$, and consider the nonlinear system (1) given specifically as

$$y_k(t+1) = 2\sin(y_k(t)) + \cos(y_k(t-1)) + \frac{t+1}{t+2}u_k(t)$$

where the initial output is set as $y_0 = 1.5$. The perfect tracking task (2) is simulated with the desired reference trajectory as

$$y_d(t) = 3\cos\left(\frac{2\pi t}{50}\right) + \frac{t^2(50-t)}{1500} \quad \forall t \in \mathbb{Z}_T.$$

To perform simulations with Algorithm 1, the parameters in Table I are selected under the condition (41), and the initial estimated value $\hat{\theta}_{0,-1,t}(i)$ is chosen as $\hat{\theta}_{0,-1,t}(i) = 0.9$, $\forall i \in \mathbb{Z}_t$ and $\forall t \in \mathbb{Z}_{T-1}$. In addition, we without loss of generality use the zero initial input: $u_0(t) = 0, \forall t \in \mathbb{Z}_{T-1}$.

In Fig. 1, the iteration evolution of the input is plotted in the sense of $\sum_{t=0}^{49}|u_k(t)|$ for the first 300 iterations. It can be easily observed from Fig. 1 that the input is bounded for all time steps and all iterations. To demonstrate the output tracking performances, the iteration evolution of the tracking error is depicted in the sense of $\sum_{t=1}^{50}|e_k(t)|$ for the first 300 iterations in Fig. 2. It is clear from this figure that the output tracking error converges to zero exponentially fast along the iteration axis. Because the desired reference $y_d(t)$ is bounded, Fig. 2 also implies the boundedness of the output for all time steps and all iterations. In addition, Fig. 3 shows the tracking performance of the desired reference by the system output resulting from our optimization-based adaptive ILC after 300 iterations. It is obviously revealed in Fig. 3 that the perfect output tracking tasks of nonlinear systems can be achieved with our optimization-based adaptive ILC, regardless of unknown nonlinear time-varying dynamics.

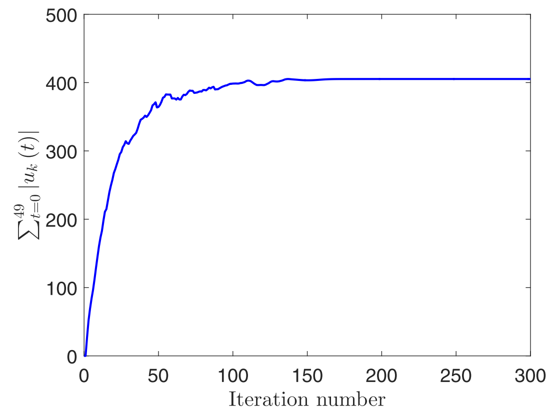

Fig. 1. (Example 1) Bounded evolution of the input along the iteration axis.

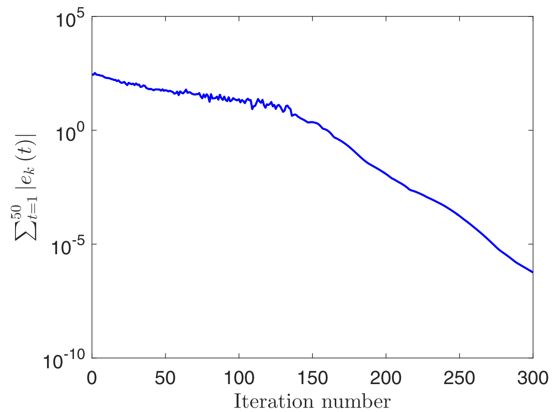

Fig. 2. (Example 1) Exponential convergence of the tracking error along the iteration axis.

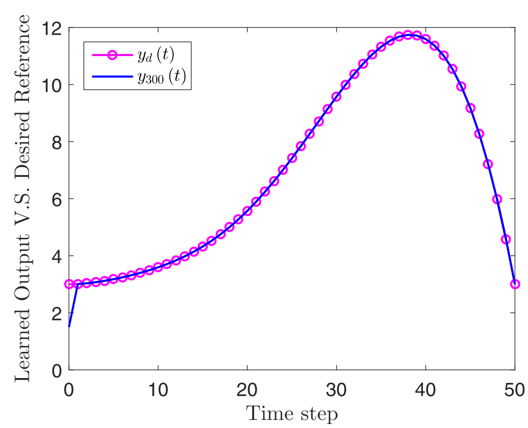

Fig. 3. (Example 1) Output tracking performance of optimization-based adaptive ILC after 300 iterations.

Example 2: Consider an injection molding process with the input–output dynamics fulfilling (1) such that (see also [3])

$$y_k(t+1) = 1.607y_k(t) - 0.6086y_k(t-1) + 1.239u_k(t) - 0.9282u_k(t-1), \quad t \in \mathbb{Z}_{200}.$$

TABLE II
LEARNING PARAMETERS USED IN EXAMPLE 2

λ	γ_1	γ_2	μ_1	μ_2	ε
1	0.95	0.05	1	0.001	0.01

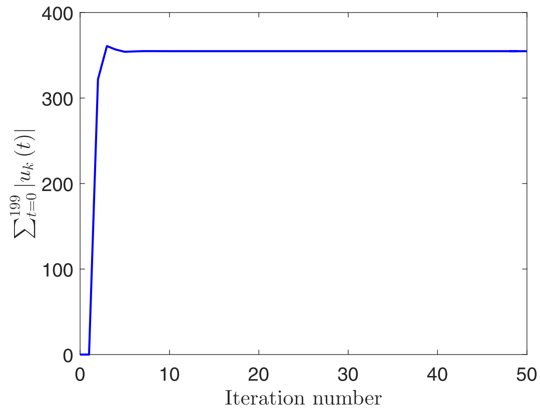

Fig. 4. (Example 2) Bounded evolution of the input along the iteration axis.

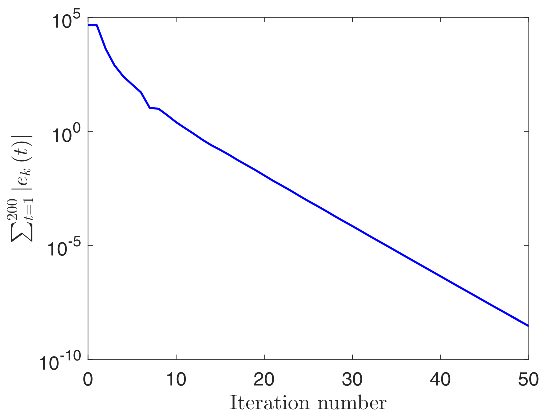

Fig. 5. (Example 2) Monotonic convergence of the tracking error along the iteration axis.

By also following [3], we adopt the desired reference trajectory specified by:

$$y_d(t) = \begin{cases} 200, & 0 \le t \le 100 \\ 300, & 101 \le t \le 200. \end{cases}$$

We take the initial output as $y_0 = 10$ that is obviously different from $y_d(0)$ and, thus, results in a constant initial shift. To apply Algorithm 1, we utilize the parameters presented in Table II, under which the selection condition (41) can be satisfied. We also employ the initial estimated value $\hat{\theta}_{0,-1,t}(i)$ and the initial input $u_0(t)$ as used in Example 1.

In Figs. 4–6, the validity of our optimization-based adaptive ILC applied to the injection molding process is demonstrated. To be more specific, the iteration evolution of the input in Fig. 4 indicates that we can guarantee the boundedness of all the system trajectories, and the iteration evolution of the tracking error in Fig. 5 discloses that we can ensure the tracking error to monotonically and exponentially decrease to zero. In particular, Fig. 6 shows that the perfect tracking of the desired reference can be achieved by our optimization-based adaptive ILC only after 50 iterations.

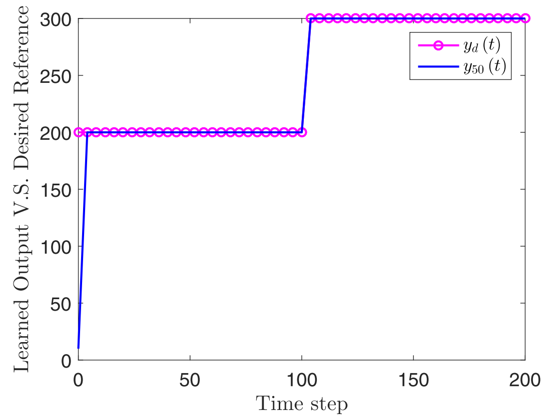

Fig. 6. (Example 2) Output tracking performance of optimization-based adaptive ILC after 50 iterations.

Discussions on the Simulation Results of Examples 1 and 2: The illustrations of Figs. 1–6 are consistent with our convergence results of optimization-based adaptive ILC established in Theorems 1 and 2. This verifies the validity of our proposed Algorithm 1 for the perfect tracking of nonlinear time-varying systems. In addition, Figs. 1–6 show that our design and analysis of optimization-based adaptive ILC not only generalize the relevant existing results of data-driven ILC in, e.g., [37] and [38] by proposing a new algorithm but also proceed further to make improvements of them by exploring a new convergence analysis approach.

VI. CONCLUSION

In this article, the algorithm design and convergence analysis of optimization-based adaptive ILC for nonlinear systems have been discussed despite unknown time-varying uncertainties. A new design approach has been given to exploit optimization-based adaptive ILC, especially through presenting an improved optimization index to obtain an updating law for the parameter estimation. Simultaneously, a new analysis approach has been proposed to cope with convergence problems for optimization-based adaptive ILC, which resorts to the DDA-based approach to ILC convergence and takes advantage of the good properties of nonnegative matrices. It has been shown that our established results may proceed further with the data-driven ILC problems investigated in, e.g., [35]–[41]. Simulation test results have also been provided to demonstrate the effectiveness of our obtained results for optimization-based adaptive ILC.

APPENDIX A
PROOF OF LEMMA 2

By noting (18), we first reformulate (3) as

$$\begin{aligned} J(u_k(t)) = {} & \lambda[\Delta u_k(t)]^2 \\ & + \Bigg\{ \gamma_1 \Bigg[-\theta_{k,k-1,t}(t) u_k(t) \\ & \quad + \theta_{k,k-1,t}(t) u_{k-1}(t) + e_{k-1}(t+1) \\ & \quad - \sum_{i=0}^{t-1} \theta_{k,k-1,t}(i) \Delta u_k(i) \Bigg] \\ & + \sum_{i=2}^{m} \gamma_i e_{k-i+1}(t+1) \Bigg\}^2. \end{aligned} \tag{51}$$

Then, to determine $u_k(t)$ that can optimize (3), it may generally resort to the condition (see also [37]–[39])

$$\frac{\partial J(u_k(t))}{\partial u_k(t)} = 0 \tag{52}$$

for which we can benefit from (51) to deduce

$$\begin{aligned}\frac{\partial J(u_k(t))}{\partial u_k(t)} &= 2\lambda\big[u_k(t) - u_{k-1}(t))\big]\\ &\quad - 2\gamma_1\theta_{k,k-1,t}(t)\Bigg\{\gamma_1\Bigg[-\theta_{k,k-1,t}(t)u_k(t)\\ &\qquad + \theta_{k,k-1,t}(t)u_{k-1}(t) + e_{k-1}(t+1)\\ &\qquad - \sum_{i=0}^{t-1}\theta_{k,k-1,t}(i)\Delta u_k(i)\Bigg]\\ &\qquad + \sum_{i=2}^{m}\gamma_i e_{k-i+1}(t+1)\Bigg\}.\end{aligned} \tag{53}$$

A straightforward consequence of inserting (53) into (52) is an optimal ILC law gained for $t \in \mathbb{Z}_{T-1}$ and $k \in \mathbb{Z}$ in the updating form of (11).

Appendix B
Proof of Lemma 3

From (12), we can derive

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

We take $\partial H(\overrightarrow{\hat{\theta}_{k,k-1,t}(t)})/\partial\overrightarrow{\hat{\theta}_{k,k-1,t}(t)} = 0$ and, consequently, can deduce

$$\begin{aligned}&\left[\mu_1 I + \mu_2 I + \Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)\right]\overrightarrow{\hat{\theta}_{k,k-1,t}(t)}\\ &\quad = \Delta y_{k-1}(t+1)\Delta\overrightarrow{u_{k-1}}(t) + \mu_1\overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)}.\end{aligned} \tag{54}$$

In addition, we can verify

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

which, together with (54), leads to

$$\begin{aligned}&\overrightarrow{\hat{\theta}_{k,k-1,t}(t)}\\ &= \frac{\Delta y_{k-1}(t+1)}{\mu_1+\mu_2}\left[I - \frac{\Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\right]\Delta\overrightarrow{u_{k-1}}(t)\\ &\quad + \frac{\mu_1}{\mu_1+\mu_2}\left[I - \frac{\Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\right]\overrightarrow{\hat{\theta}_{k-1,k-2,t}(t)}.\end{aligned} \tag{55}$$

By noting that

$$\begin{aligned}&\left[I - \frac{\Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\right]\Delta\overrightarrow{u_{k-1}}(t)\\ &\qquad = \frac{\mu_1+\mu_2}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}\Delta\overrightarrow{u_{k-1}}(t)\end{aligned}$$

and that

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

we can employ (55) to further obtain

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

namely, (13) holds.

Appendix C
Proof of Lemma 4

We can easily see from (23) that $Q\big(\Delta\overrightarrow{u_{k-1}}(t)\big) \in \mathbb{R}^{(t+1)\times(t+1)}$ is a real symmetric matrix and can thus obtain

$$\left\|Q\big(\Delta\overrightarrow{u_{k-1}}(t)\big)\right\|_2 = \rho\Big(Q\big(\Delta\overrightarrow{u_{k-1}}(t)\big)\Big). \tag{56}$$

We can also verify that for $\Delta\overrightarrow{u_{k-1}}(t)\Delta\overrightarrow{u_{k-1}}^{\mathrm{T}}(t)$, the eigenvalues are either $\|\Delta\overrightarrow{u_{k-1}}(t)\|_2^2$ or 0 (with a multiplicity of t). By noting this fact, we can further get from (23) that for $Q(\Delta\overrightarrow{u_{k-1}}(t))$, the eigenvalues are either $(\mu_1+\mu_2)/(\mu_1+\mu_2+\|\Delta\overrightarrow{u_{k-1}}(t)\|_2^2)$ or 1 (with a multiplicity of t). As a direct consequence, we have

$$\rho\Big(Q\big(\Delta\overrightarrow{u_{k-1}}(t)\big)\Big) = \begin{cases}1, & t > 0\\ \dfrac{\mu_1+\mu_2}{\mu_1+\mu_2+\left\|\Delta\overrightarrow{u_{k-1}}(t)\right\|_2^2}, & t = 0\end{cases} \quad \forall k \geq 2$$

which, together with (56), yields

$$\left\|Q\big(\Delta\overrightarrow{u_{k-1}}(t)\big)\right\|_2 = \rho\Big(Q\big(\Delta\overrightarrow{u_{k-1}}(t)\big)\Big) \leq 1.$$

Namely, Lemma 4 is developed.

APPENDIX D
PROOF OF LEMMA 6

To prove Lemma 6, we exploit the properties of nonnegative matrices to disclose the relationship between the condition (35) in Lemma 6 and the condition (C) in Lemma 5.

Lemma 9: For any $t \in \mathbb{Z}_{T-1}$, the condition (C) in Lemma 5 holds as a consequence of the condition (35) in Lemma 6.

Proof: Let us take $\omega_s(t) = (m-1)s+1$, $\forall t \in \mathbb{Z}_{T-1}$ in (34), and we can use the properties of nonnegative matrices to deduce

$$\begin{aligned}\left\|\prod_{k=(m-1)s+1}^{(m-1)s+m-1} P_k(t)\right\|_\infty &\le \left\|\prod_{k=(m-1)s+1}^{(m-1)s+m-1} |P_k(t)|\right\|_\infty \\ &= \left\|\prod_{k=(m-1)s+1}^{(m-1)s+m-1} |P_k(t)|\mathbf{1}_{m-1}\right\|_\infty. \end{aligned} \tag{57}$$

To proceed with (57), we adopt an inductive analysis approach to show a property that if (35) holds, then

$$\prod_{k=(m-1)s+1}^{(m-1)s+m-1} |P_k(t)|\mathbf{1}_{m-1} \le \zeta\mathbf{1}_{m-1}. \tag{58}$$

Step 1): For $i=1$, we consider

$$\prod_{k=(m-1)s+1}^{(m-1)s+i} |P_k(t)|\mathbf{1}_{m-1} = \left|P_{(m-1)s+1}(t)\right|\mathbf{1}_{m-1}$$

and then by employing the definition of the nonnegative matrix $|P_{(m-1)s+1}(t)|$, we can gain that its induced nonnegative vector $|P_{(m-1)s+1}(t)|\mathbf{1}_{m-1}$ satisfies

$$\begin{aligned}\left|P_{(m-1)s+1}(t)\right|\mathbf{1}_{m-1} &= \begin{bmatrix} \sum_{j=1}^{m-1}\left|p_{j,(m-1)s+1}(t)\right| \\ \mathbf{1}_{m-2}\end{bmatrix} \\ &\le \begin{bmatrix}\zeta \\ \mathbf{1}_{m-2}\end{bmatrix} \\ &\triangleq \begin{bmatrix}\zeta\mathbf{1}_i \\ \mathbf{1}_{m-i-1}\end{bmatrix}. \end{aligned} \tag{59}$$

Step 2): For any $i \ge 1$, we explore the fact (59) to make the following hypothesis:

$$\prod_{k=(m-1)s+1}^{(m-1)s+i} |P_k(t)|\mathbf{1}_{m-1} \le \begin{bmatrix}\zeta\mathbf{1}_i \\ \mathbf{1}_{m-i-1}\end{bmatrix}. \tag{60}$$

Then, for the next step $i+1$, we insert (60) and can again apply the properties of nonnegative matrices to derive

$$\begin{aligned}&\prod_{k=(m-1)s+1}^{(m-1)s+i+1} |P_k(t)|\mathbf{1}_{m-1} \\ &= \left|P_{(m-1)s+i+1}(t)\right|\left[\prod_{l=(m-1)s+1}^{(m-1)s+i} |P_k(t)|\mathbf{1}_{m-1}\right] \\ &\le \left|P_{(m-1)s+i+1}(t)\right|\begin{bmatrix}\zeta\mathbf{1}_i \\ \mathbf{1}_{m-i-1}\end{bmatrix}\end{aligned}$$

which, together with the definition of $|P_k(t)|$ and the condition (35), leads to

$$\begin{aligned}\prod_{k=(m-1)s+1}^{(m-1)s+i+1} |P_k(t)|\mathbf{1}_{m-1} &\le \begin{bmatrix}\left(\begin{array}{c}\zeta\sum_{j=1}^{i}\left|p_{j,(m-1)s+i+1}(t)\right| \\ +\sum_{j=i+1}^{m-1}\left|p_{j,(m-1)s+i+1}(t)\right|\end{array}\right) \\ \zeta\mathbf{1}_i \\ \mathbf{1}_{m-i-2}\end{bmatrix} \\ &\le \begin{bmatrix}\sum_{j=1}^{m-1}\left|p_{j,(m-1)s+i+1}(t)\right| \\ \zeta\mathbf{1}_i \\ \mathbf{1}_{m-i-2}\end{bmatrix} \\ &\le \begin{bmatrix}\zeta\mathbf{1}_{i+1} \\ \mathbf{1}_{m-i-2}\end{bmatrix}. \end{aligned} \tag{61}$$

Clearly, (61) implies that the hypothesis made in (60) can also hold by updating i with $i+1$.

With the above analysis of steps 1) and 2) and by induction, we can conclude that (58) holds. By combining (57) and (58), we can further deduce

$$\begin{aligned}\left\|\prod_{k=(m-1)s+1}^{(m-1)s+m-1} P_k(t)\right\|_\infty &\le \left\|\prod_{k=(m-1)s+1}^{(m-1)s+m-1} |P_k(t)|\mathbf{1}_{m-1}\right\|_\infty \\ &\le \|\zeta\mathbf{1}_{m-1}\|_\infty \\ &= \zeta < 1. \end{aligned} \tag{62}$$

Consequently, (62) guarantees that the condition (C) in Lemma 5 can be developed by particularly setting $\omega_s(t) = (m-1)s+1$, $\forall t \in \mathbb{Z}_{T-1}$, and $\eta = \zeta$. □

In addition, for the relationship between (19) and (31), we clearly have two equivalent results in the following lemma.

Lemma 10: For any $t \in \mathbb{Z}_{T-1}$, the following two results hold.

1) $\overrightarrow{e_k}(t+1)$ is bounded (respectively, $\lim_{k\to\infty}\overrightarrow{e_k}(t+1) = 0$) if and only if $e_k(t+1)$ is bounded (respectively, $\lim_{k\to\infty} e_k(t+1) = 0$);
2) $\overrightarrow{\kappa_k}(t)$ is bounded (respectively, $\lim_{k\to\infty}\overrightarrow{\kappa_k}(t) = 0$) if and only if $\kappa_k(t)$ is bounded (respectively, $\lim_{k\to\infty}\kappa_k(t) = 0$).

Proof: A consequence of the definitions for $\overrightarrow{e_k}(t+1)$ and $\overrightarrow{\kappa_k}(t)$ in (32). □

Note that the equivalence between two convergence results in Lemma 10 also holds when they are exponentially fast. With Lemmas 9 and 10, we can prove Lemma 6 as follows.

Proof of Lemma 6: Based on Lemmas 5 and 9, we know that if the condition (35) holds, then $\lim_{k\to\infty}\overrightarrow{e_k}(t+1) = 0$ (exponentially fast), provided that $\lim_{k\to\infty}\overrightarrow{\kappa_k}(t) = 0$ (exponentially fast). By the two equivalent results of Lemma 10, we can further conclude $\lim_{k\to\infty} e_k(t+1) = 0$ (exponentially fast), provided that $\lim_{k\to\infty}\kappa_k(t) = 0$ (exponentially fast). In the same way, we can prove that $e_k(t+1)$ is bounded, provided that $\kappa_k(t)$ is bounded. Namely, Lemma 6 is obtained. □

Appendix E
Proof of Lemma 8

With Lemma 1 and Theorem 1, we can obtain

$$\hat{\theta}^2_{k,k-1,t}(t) \leq \beta^2_{\hat{\theta}}$$

$$\beta_{\underline{f}}\epsilon \leq \theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)$$

$$\frac{(\gamma_1^2+\gamma_1\gamma_2)\theta_{k,k-1,t}(t)\hat{\theta}_{k,k-1,t}(t)}{\lambda+\gamma_1^2\hat{\theta}^2_{k,k-1,t}(t)} \leq \frac{(\gamma_1^2+\gamma_1\gamma_2)\beta_{\overline{f}}\beta_{\hat{\theta}}}{\lambda} \tag{63}$$

which, together with (41), leads to

$$\frac{\gamma_1\left(\gamma_1+\gamma_2-\sum_{i=3}^{m}\gamma_i\right)\beta_{\underline{f}}\epsilon}{\lambda+\gamma_1^2\beta^2_{\hat{\theta}}} \leq \frac{(\gamma_1^2+\gamma_1\gamma_2)\beta_{\overline{f}}\beta_{\hat{\theta}}}{\lambda} < 1. \tag{64}$$

By inserting (63) and (64) into (35), we can verify

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

that is, the condition (35) holds.

In the same way as (63), we can apply (9) and (22) to derive

$$\beta_{\underline{f}}\epsilon \leq \hat{\theta}_{k,k-1,t}(t)\theta_{k-1,0,t}(t) \leq \beta_{\overline{f}}\beta_{\hat{\theta}}$$

and further by (41), we can deduce

$$\begin{aligned}
\frac{(\gamma_1^2+\gamma_1\gamma_2)\beta_{\underline{f}}\epsilon}{\lambda+\gamma_1^2\beta^2_{\hat{\theta}}} &\leq \frac{(\gamma_1^2+\gamma_1\gamma_2)\hat{\theta}_{k,k-1,t}(t)\theta_{k-1,0,t}(t)}{\lambda+\gamma_1^2\hat{\theta}^2_{k,k-1,t}(t)} \\
&\leq \frac{(\gamma_1^2+\gamma_1\gamma_2)\beta_{\overline{f}}\beta_{\hat{\theta}}}{\lambda} \\
&< 1
\end{aligned}$$

which, together with (40), results in

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

Namely, the condition (40) holds.

Acknowledgment

The authors would like to thank all the editors and reviewers for their helpful comments and suggestions.

References

[1] Z. Zhang, Z. Li, Y. Zhang, Y. Luo, and Y. Li, "Neural-dynamic-method-based dual-arm CMG scheme with time-varying constraints applied to humanoid robots," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3251–3262, Dec. 2015.

[2] W. He, Y. Sun, Z. Yan, C. Yang, Z. Li, and O. Kaynak, "Disturbance observer-based neural network control of cooperative multiple manipulators with input saturation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 5, pp. 1735–1746, May 2020.

[3] 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," *Ind. Eng. Chem. Res.*, vol. 44, no. 21, pp. 8095–8105, Oct. 2005.

[4] H. Ji, Z. Hou, and R. Zhang, "Adaptive iterative learning control for high-speed trains with unknown speed delays and input saturations," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 1, pp. 260–273, Jan. 2016.

[5] S. Preitl, R.-E. Precup, Z. Preitl, S. Vaivoda, S. Kilyeni, and J. K. Tar, "Iterative feedback and learning control. Servo systems applications," *IFAC Proc. Volumes*, vol. 40, no. 8, pp. 16–27, 2007.

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

[7] K. K. Tan, S. Zhao, and J.-X. Xu, "Online automatic tuning of a proportional integral derivative controller based on an iterative learning control approach," *IET Control Theory Appl.*, vol. 1, no. 1, pp. 90–96, Jan. 2007.

[8] Z. Hou, R. Chi, and H. Gao, "An overview of dynamic-linearization-based data-driven control and applications," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4076–4090, May 2017.

[9] 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 Syst.*, vol. 26, no. 3, pp. 96–114, Jun. 2006.

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

[11] J.-X. Xu, "A survey on iterative learning control for nonlinear systems," *Int. J. Control*, vol. 84, no. 7, pp. 1275–1294, Jul. 2011.

[12] W. Xiong, D. W. C. Ho, and X. Yu, "Saturated finite interval iterative learning for tracking of dynamic systems with HNN-structural output," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 7, pp. 1578–1584, Jul. 2016.

[13] W. He, T. Meng, D. Huang, and X. Li, "Adaptive boundary iterative learning control for an Euler–Bernoulli beam system with input constraint," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1539–1549, May 2018.

[14] D. Shen, "Data-driven learning control for stochastic nonlinear systems: Multiple communication constraints and limited storage," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2429–2440, Jun. 2018.

[15] M. Sun, T. Wu, L. Chen, and G. Zhang, "Neural AILC for error tracking against arbitrary initial shifts," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 7, pp. 2705–2716, Jul. 2018.

[16] D. Meng and J. Zhang, "Convergence analysis of robust iterative learning control against nonrepetitive uncertainties: System equivalence transformation," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 25, 2020, doi: 10.1109/TNNLS.2020.3016057.

[17] Q.-Y. Xu and X.-D. Li, "HONN-based adaptive ILC for pure-feedback nonaffine discrete-time systems with unknown control directions," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 1, pp. 212–224, Jan. 2020.

[18] D. Meng and K. L. Moore, "Contraction mapping-based robust convergence of iterative learning control with uncertain, locally Lipschitz nonlinearity," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 2, pp. 442–454, Feb. 2020.

[19] Q. Yu, Z. Hou, X. Bu, and Q. Yu, "RBFNN-based data-driven predictive iterative learning control for nonaffine nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 4, pp. 1170–1182, Apr. 2020.

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

[21] Y. Chen and K. L. Moore, "An optimal design of PD-type iterative learning control with monotonic convergence," in *Proc. IEEE Int. Symp. Intell. Control*, Vancouver, BC, Canada, Oct. 2002, pp. 55–60.

[22] J.-X. Xu and Y. Tan, "Robust optimal design and convergence properties analysis of iterative learning control approaches," *Automatica*, vol. 38, no. 11, pp. 1867–1880, Nov. 2002.

[23] S. Yang, J.-X. Xu, D. Huang, and Y. Tan, "Optimal iterative learning control design for multi-agent systems consensus tracking," *Syst. Control Lett.*, vol. 69, pp. 80–89, Jul. 2014.

[24] 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 Syst. Technol.*, vol. 19, no. 1, pp. 166–180, Jan. 2011.

[25] D. H. Owens, C. T. Freeman, and B. Chu, "An inverse-model approach to multivariable norm optimal iterative learning control with auxiliary optimisation," *Int. J. Control*, vol. 87, no. 8, pp. 1646–1671, Aug. 2014.

[26] M. Volckaert, M. Diehl, and J. Swevers, "Generalization of norm optimal ILC for nonlinear systems with constraints," *Mech. Syst. Signal Process.*, vol. 39, nos. 1–2, pp. 280–296, Aug. 2013.

[27] P. Janssens, G. Pipeleers, and J. Swevers, "A data-driven constrained norm-optimal iterative learning control framework for LTI systems," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 2, pp. 546–551, Mar. 2013.

[28] T. D. Son, G. Pipeleers, and J. Swevers, "Robust monotonic convergent iterative learning control," *IEEE Trans. Autom. Control*, vol. 61, no. 4, pp. 1063–1068, Apr. 2016.

[29] X. Ge, J. L. Stein, and T. Ersal, "Frequency-domain analysis of robust monotonic convergence of norm-optimal iterative learning control," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 2, pp. 637–651, Mar. 2018.

[30] A. Haber, R. Fraanje, and M. Verhaegen, "Linear computational complexity robust ILC for lifted systems," *Automatica*, vol. 48, no. 6, pp. 1102–1110, Jun. 2012.

[31] H. Sun and A. G. Alleyne, "A computationally efficient norm optimal iterative learning control approach for LTV systems," *Automatica*, vol. 50, no. 1, pp. 141–148, Jan. 2014.

[32] P. Axelsson, R. Karlsson, and M. Norrlöf, "Estimation-based norm-optimal iterative learning control," *Syst. Control Lett.*, vol. 73, pp. 76–80, Nov. 2014.

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

[34] J. K. Rice and M. Verhaegen, "A structured matrix approach to efficient calculation of LQG repetitive learning controllers in the lifted setting," *Int. J. Control*, vol. 83, no. 6, pp. 1265–1276, Jun. 2010.

[35] 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.

[36] Z. Hou and S. Jin, *Model Free Adaptive Control: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2013.

[37] R. Chi, Z. Hou, B. Huang, and S. Jin, "A unified data-driven design framework of optimality-based generalized iterative learning control," *Comput. Chem. Eng.*, vol. 77, pp. 10–23, Jun. 2015.

[38] R. Chi, Z. Hou, S. Jin, and B. Huang, "Computationally efficient data-driven higher order optimal iterative learning control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 5971–5980, Dec. 2018.

[39] R. Chi, Z. Hou, S. Jin, D. Wang, and C.-J. Chien, "Enhanced data-driven optimal terminal ILC using current iteration control knowledge," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 11, pp. 2939–2948, Nov. 2015.

[40] 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. Syst., Man, Cybern. Syst.*, vol. 49, no. 4, pp. 677–686, Apr. 2019.

[41] 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. Syst., Man, Cybern., Syst.*, vol. 51, no. 3, pp. 1881–1891, Mar. 2021.

[42] M.-B. Radac, R.-E. Precup, and E. M. Petriu, "Model-free primitive-based iterative learning control approach to trajectory tracking of MIMO systems with experimental validation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 11, pp. 2925–2938, Nov. 2015.

[43] W. J. Rugh, *Linear System Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[44] 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 Appl.*, vol. 9, no. 8, pp. 1221–1228, May 2015.

[45] H. K. Khalil, *Nonlinear Systems*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[46] D. Meng and K. L. Moore, "Robust iterative learning control for nonrepetitive uncertain systems," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 907–913, Feb. 2017.

[47] D. Meng and J. Zhang, "Robust tracking of nonrepetitive learning control systems with iteration-dependent references," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 2, pp. 842–852, Feb. 2021.

[48] D. Meng and K. L. Moore, "Convergence of iterative learning control for SISO nonrepetitive systems subject to iteration-dependent uncertainties," *Automatica*, vol. 79, pp. 167–177, May 2017.

[49] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

Deyuan Meng (Senior Member, IEEE) received the B.S. degree in mathematics and applied mathematics from the Ocean University of China (OUC), Qingdao, China, in June 2005, and the Ph.D. degree in control theory and control engineering from Beihang University (BUAA), Beijing, China, in July 2010.

From November 2012 to November 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). His current research interests include iterative learning control, data-driven control, and multiagent 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, where he is currently pursuing the Ph.D. degree with the School of Automation Science and Electrical Engineering.

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

Mr. Zhang was a co-recipient of the "Best Paper Award" from the IEEE Seventh Data Driven Control and Learning Systems Conference in 2018.