Received April 25, 2019, accepted May 13, 2019, date of publication May 20, 2019, date of current version June 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2917607

Optimized Iterative Learning Control for Linear Discrete-Time-Invariant Systems

YAN LIU, XIAOE RUAN, AND XIAOHUI LI
School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710001, China

Corresponding author: Xiaoe Ruan (wruanxe@mail.xjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61273135.

ABSTRACT In this paper, an optimized first-order iterative learning control (OILC) scheme is explored for a class of linear discrete-time-invariant systems with Markov parameters available and the system relative degree being unity. For the OILC scheme, the iteration-time-variable derivative learning-gain vector is argued by sequentially minimizing the sum of the tracking error energy and the learning effort intensity amplified by an iteration-wise tuning factor. In virtue of the optimization criterion, the existence and the uniqueness of the iteration-time-variable learning-gain vector is achieved. Then, by making use of the elementary transformations which exchange the rows and columns of a matrix and by taking advantage of the positivity relationship of the eigenvalues with the matrix-weighing quadratic function, the strictly monotone convergence of the OILC scheme is derived, which conveys that the strict monotonicity is guaranteed without any requirement to the system Markov parameters and the convergence rate is adjustable by scaling the tuning factor. Furthermore, an optimized higher-order iterative learning control mechanism is developed for the system relative degree is larger than unity, for which the existence and the uniqueness of the optimized higher-order iteration-time-variable learning-gain vector are discussed and the strictly non-conditional monotone convergence is analyzed. The numerical simulations demonstrate the validity and effectiveness.

INDEX TERMS Iterative learning control, iteration-time-variable learning-gain vector, elementary transformation, tuning factor, strictly monotone convergence.

I. INTRODUCTION

For trajectory tracking issue, the iterative learning control (ILC) has been acknowledged as an efficacious technique that fits for repetitive systems to operate over a fixed finite time interval [1]–[4]. The fundamental mechanism is to generate a sequence of upgraded control inputs in recursive mode by modifying the control input of the current operation with current or historical tracking errors integrated by a linear operator. The conventional ILC is of proportional-derivative-integral-type (PID-type) profile. One hand, owing to the simple algorithmic structure, less requirement prior to system knowledge and significant performance, the ILC has been attracted much attention to several applications as addressed in references [5]–[10]. On the other hand, it is noticed that, in the conventional PID-type ILC schemes, the learning mode is passive as the learning gains are fixed without any adaption to the systems parameters or the tracking error information.

The associate editor coordinating the review of this manuscript and approving it for publication was Dong Shen.

As such, the learning performance with less utility of the system knowledge may render poor improvement if the learning gains are selected improperly.

Ideally, the learning performance might be the expected if the system Markov parameters are available and utilizable. For the regard, the earlier devotions have proposed optimal learning control schemes by numerical techniques such as steepest descending method, Newton-Raphson method and Gauss-Newton method while minimizing a quadratic criterion of the time-wise tracking error for discrete-time or continuous-time systems [11], [12]. After that, by formulating a finite-length discrete-time system in a lifted vector-matrix form, the concept of the norm-optimal iterative learning control (NOILC) has been innovated, where the performance index is the sum of norms in terms of the tracking error and the increment of two adjacent control inputs and the argument is the control input for the next operation [13]. Following up the NOILC profile, numerous contributions have been done for the NOILC developments to kinds of systems [14]–[21]. Moreover, the non-lifted data-driven approach has been adopted for the NOILC schemes to reduce the computational complexity [18], [22]. For those NOILCs, in authors' opinion, the profile is originated from the concept of the optimal control for reference trajectory tracking. Thus, the so-called NOILC schemes have inherited the characteristics of the optimal control mechanism that the learning gain matrix of the NOILC scheme is irrelevant to the reference trajectory, namely, independent upon the tracking error. But it requires inversion computation of matrix. In the circumstance, the strictly monotonous convergence is conditional, such as, the tracking error is guaranteed if and only if when the Markov parameters matrix is nonsingular as mentioned in reference [13].

For the sake of deducing the computation complexity, Owens' group has proposed the parameter-optimal iterative learning control (POILC) scheme for discrete-time systems [23], where the POILC updating law is D-type that the iteration-variable-time-invariant learning gain is argued. For this, the performance index is defined as a sum of the tracking error measured in 2-norm and the learning effort intensity amplified by a fixed tuning factor. Here the learning effort intensity is defined as the square of the learning gain. The further POILC works have been involving the optimization techniques, improvement of convergence rate, as well as the robustness [24]–[33]. Perhaps because the iteration-wise argument of the learning gain of the POILCs is single-dimensional, its solution confines the learning performance. For example, the monotone convergence is only guaranteed under the assumption that the Markov parameter matrix is positive definite [23]. The condition is critical in deed. Progressively, a full-dimensional parameter-optimal iterative learning control scheme has been exploited in article [34], where the iteration-time-variable learning-gain vector is adapted. Here the notion ''full-dimensional'' means that the dimension of the argument in terms of the learning-gain vector is equal to the sampling number. However, due to the similar analytical technique that adopted for NOILCs and POILCs, the strictly monotone convergence of the tracking error is incomplete. Such as, in reference [34], the strict monotonicity in Proposition 7 is beyond comprehension, especially, for the case when part components of the current tracking error are zero. Besides, the effect of the tuning factor on the convergence rate has not been discussed. Motivated by the significance of the full-dimensional parameter-optimal mode and the incompleteness of the existing results, this paper develops the optimized first-order and higher-order iterative learning control (OILC) schemes for the first-order and higher-order linear time-invariant systems, respectively. For the OILC schemes, the iteration-time-variable learning-gain vector is argued by minimizing a sum of the tracking error energy and the learning effort intensity tuned by an iteration-variable factor. The strictly monotonous convergence of the tracking error energy is rigorously derived and the comparable results are remarked in detail.

The remainder of this paper is arranged as follows. In section II, an optimized first-order iterative learning control scheme is formulated and some necessary lemmas are presented. Section III focuses on the strictly monotone convergence analysis of the first-order optimized learning scheme and Section IV involves of the optimized higher-order iterative learning control strategy and the convergence. Section V exhibits the numerical simulations and the last section concludes the paper.

II. OPTIMIZED FIRST-ORDER ITERATIVE LEARNING CONTROL SCHEME AND SOME LEMMAS

Consider a single-input-single-output (SISO) linear discrete-time-invariant (LDTI) system whose dynamics is described as follows.

$$\begin{cases} \boldsymbol{x}(n+1) = \mathbf{A}x(n) + \mathbf{B}u(n), \\ y(n+1) = \mathbf{C}x(n+1), \\ \boldsymbol{x}(0) = 0, \quad n = 0, 1, \cdots, N-1, \end{cases} \tag{1}$$

where N denotes the total sampling number, $\boldsymbol{x}(n)$, $u(n)$ and $y(n)$ are p-dimensional state vector, scalar input and scalar output, respectively. $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices with appropriate dimensions, respectively. Assume that $\mathbf{CB} \neq 0$, which means that the relative degree of system (1) is unity.

Let $\mathbf{u}^* = \boldsymbol{\varepsilon}_1 = [1|0|\cdots|0]^{\mathrm{T}}$ denote the N-dimensional impulse signal sequence and stimulate system (1). Then the output takes a form of

$$\begin{aligned} \mathbf{y}^* = \mathbf{g}_1 &= [g_1(1)|g_1(2)|\cdots|g_1(r-1)|g_1(r)|\cdots|g_1(N)]^{\mathrm{T}} \\ &= \left[\mathbf{CB}|\mathbf{CAB}|\mathbf{CA}^2\mathbf{B}|\cdots|\mathbf{CA}^{N-1}\mathbf{B}\right]^{\mathrm{T}}. \end{aligned}$$

Here, the vector $\boldsymbol{g}_1$ is assigned as the impulse response sequence or Markov parameters of system (1).

Thus, for any input sequence $\mathbf{u} = [u(0)|u(1)|\cdots|u(N-1)]^{\mathrm{T}}$, the output sequence of system (1) is expressed as

$$y(n+1) = \sum_{l=0}^{n} g_1(n+1-l)u(l), \quad n = 0, 1, \cdots, N-1. \tag{2}$$

Let

$$\mathbf{H} = \begin{bmatrix} g_1(1) & 0 & 0 & \cdots & 0 \\ g_1(2) & g_1 & 0 & \cdots & 0 \\ g_1(3) & g_1(2) & g_1(1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_1(N) & g_1(N-1) & g_1(N-2) & \cdots & g_1(1) \end{bmatrix}.$$

Here, the matrix $\mathbf{H}$ is termed as Markov parameters matrix of system (1).

Denote

$$\mathbf{y} = [y(1)|y(2)|\cdots|y(N)]^{\mathrm{T}}.$$

Then system (1) or (2)becomes

$$\mathbf{y} = \mathbf{Hu}. \tag{3}$$

Suppose that the operation of system (3) is repetitive and attempts to track a predetermined desired trajectory

$\mathbf{y}_d = [y_d(1)|y_d(2)|\cdots|y_d(N)]^{\mathrm{T}}$. Let $\mathbf{u}_1 = [u_1(0)|u_1(1)|\cdots|u_1(N-1)]^{\mathrm{T}}$ be an arbitrarily initial-iteration input and $\mathbf{y}_1 = [y_1(1)|y_1(2)|\cdots|y_1(N)]^{\mathrm{T}}$ be its output, respectively.

Denote $\mathbf{e}_1 = \mathbf{y}_d - \mathbf{y}_1$ as the tracking error between the desired trajectory $\mathbf{y}_d$ and the output $\mathbf{y}_1$. Then, by compensating for $u_1(n)$ with its output error $e_1(n+1)$, the input $u_2(n)$ for the second iteration is generated. Recursively, a D-type iterative learning control updating law is generated as

$u_1(n)$: given arbitrarily;

$$\begin{aligned} u_{k+1}(n) &= u_k(n) + \Gamma_k(n+1)e_k(n+1), \\ n &= 0, 1, \cdots, \quad N-1, \ k = 1, 2, \cdots, \end{aligned} \tag{4}$$

where the subscript k stands for the iteration or operation index and $\Gamma_k(n+1)$ is assigned as the iteration-time-variable learning gain, respectively.

Denote lifted vectors as

$$\begin{aligned} \mathbf{u}_k &= [u_k(0)|u_k(1)|\cdots|u_k(N-1)]^{\mathrm{T}}, \\ \mathbf{y}_k &= [y_k(1)|y_k(2)|\cdots|y_k(N)]^{\mathrm{T}}, \\ \mathbf{e}_k &= y_d - y_k = [e_k(1)|e_k(2)|\cdots|e_k(N)]^{\mathrm{T}}, \\ \mathbf{\Gamma}_k &= [\Gamma_k(1)|\Gamma_k(2)|\cdots|\Gamma_k(N)]^{\mathrm{T}}. \end{aligned}$$

Let

$$\mathbf{E}_k = \mathrm{diag}(e_k(1), e_k(2), \cdots, e_k(N)).$$

Then equation (3) is rewritten as

$$\mathbf{y}_k = \mathbf{H}\mathbf{u}_k.$$

Analogously, the D-type ILC (4) becomes

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{E}_k\mathbf{\Gamma}_k. \tag{5}$$

For the algorithm (5), $\mathbf{\Gamma}_k$ is termed as the iteration-time-variable derivative learning-gain vector, which is optimized by solving the following minimization problem.

Mind that the purpose of generating the ILC algorithm (5) is to drive system (3) to follow the predetermined desired trajectory $\mathbf{y}_d$ as closely as possible, namely, to decrease the tracking error as much as possible with some learning effort.

For this, define an iteration-wise performance index as

$$\arg\min_{\Gamma_k} J(\Gamma_k) = \|\mathbf{e}_{k+1}\|^2 + w_k \cdot \|\mathbf{\Gamma}_k\|^2, \tag{6}$$

where $w_k > 0$ is an appropriate iteration-wise tuning factor.

As the time index of the tracking error $e_k(n+1)$ in the algorithm (4) is one sampling lagged than that of $u_k(n)$, thus the ILC scheme (4) or (5) is termed as the optimized first-order iterative learning control (OILC) scheme.

Remark 1: It should point out that the iteration-time-variable learning-gain vector-based ILC (5) has been firstly innovated in reference [34], where the optimal learning gain vector (OLGV) $\mathbf{\Gamma}_k$ is a solution of the optimization problem as

$$\min_{\Gamma_k} J(\mathbf{\Gamma}_k) = \|\mathbf{e}_{k+1}\|^2 + w \cdot \|\mathbf{\Gamma}_k\|^2, \tag{7}$$

where the parameter w is a fixed tuning factor.

In the investigation of [34], Proposition 5 has conducted the monotone convergence of the algorithm (5) in virtue of the relationship as

$$\|\mathbf{e}_{k+1}\|^2 \le J\left(\mathbf{\Gamma}_k^*\right) = \|\mathbf{e}_{k+1}\|^2 + w\left\|\mathbf{\Gamma}_k^*\right\|^2 \le J(\mathbf{0}) \le \|\mathbf{e}_k\|^2, \tag{8}$$

where $\mathbf{\Gamma}_k^*$ is the optimal solution of the optimization problem (7).

In authors' opinion, the last right inequality $J(\mathbf{0}) \le \|\mathbf{e}_k\|^2$ of (8) is defaulted but not proved in a rigorous manner. In fact, it is evident to yield that $J(\mathbf{0}) = \|\mathbf{e}_{k+1}\|^2 = \|\mathbf{e}_k\|^2$ because $\mathbf{\Gamma}_k = \mathbf{0}$ implies $\mathbf{u}_{k+1} = \mathbf{u}_k$ in algorithm (5). Therefore, the inequality (8) conveys that

$$\|\mathbf{e}_{k+1}\|^2 \le J\left(\mathbf{\Gamma}_k^*\right) = \|\mathbf{e}_{k+1}\|^2 + w\left\|\mathbf{\Gamma}_k^*\right\|^2 \le J(\mathbf{0}) = \|\mathbf{e}_{k+1}\|^2. \tag{9}$$

This implies that Proposition 5 in reference [34] says nothing. For the circumstance, the so-called monotonicity based results need to be clarified in rigorous manner.

Nevertheless, reference [34] has raised the iteration-time-variable learning-gain vector-argued optimal iterative learning control profile and has provided algebraic approach. By devotion, the authors have observed that the strictly monotone convergence is ensured definitely. The rigorous derivation depends upon the elementary transformation of a matrix as addressed in Section III.

Before going to the convergence, some lemmas are presented as follows.

Lemma 1 [35]: For a square matrix $\mathbf{M}$, denoting $\lambda(\mathbf{M})$ as its eigenvalue and $\rho(\mathbf{M})$ as its spectral radius. Then

$$(\mathbf{I} - \mathbf{M})^{-1} = \sum_{k=0}^{+\infty} \mathbf{M}^k,$$

if and only if $\rho(\mathbf{M}) < 1$.

Lemma 2: For given constants $0 < \delta$ and $0 < a_1 \le a_2$ define functions

$f_1(w) = \frac{2}{w+a_2} - \frac{\delta}{(w+a_1)^2}$ and $f_2(w) = \frac{2}{w+a_1} - \frac{\delta}{(w+a_2)^2}$, $w \in (0, +\infty)$.

Then there exists at least a point $w \in (0, +\infty)$ so that $0 < f_1(w) \le f_2(w) < 1$ for $0 < w^* < w < +\infty$.

Proof: By considering the assumptions $0 < \delta$ and $0 < a_1 \le a_2$ into account, the relation $f_1(w) \le f_2(w)$ is obvious. What follows is to prove the fact that $0 < f_1(w)$ and $f_2(w) < 1$, for $0 < w^* < w < +\infty$.

Notice that

$$\begin{aligned} f_1(w) &= \frac{2}{w+a_2} - \frac{\delta}{(w+a_1)^2} \\ &= \frac{\left(w + 2a_1 - \frac{\delta}{2}\right)^2 + \left(w^2 + 2a_1^2 - \delta a_2 - \left(2a_1 - \frac{\delta}{2}\right)^2\right)}{(w+a_2)(w+a_1)^2} \end{aligned}$$

Therefore, there exists a sufficient larger $\overline{w} > 0$ so that $0 < f_1(w)$, for $0 < \overline{w} < w < +\infty$.

Besides

$$f_2(w) = \frac{2}{w+a_1} - \frac{\delta}{(w+a_2)^2} = \frac{\left(w+2a_2-\frac{\delta}{2}\right)^2 + \left(w^2+2a_2^2-\delta a_1-\left(2a_2-\frac{\delta}{2}\right)^2\right)}{(w+a_1)(w+a_2)^2}$$

Then, we have

$$\lim_{w\to+\infty} f_2(w) = \lim_{w\to+\infty} \frac{2(w+a_2)^2-\delta(w+a_1)}{(w+a_1)(w+a_2)^2} = 0.$$

This implies that there exists such a constant $\tilde{w}$ $(\overline{w} < \tilde{w})$ that

$$f_2(\tilde{w}) < 1. \tag{10}$$

Further

$$\frac{\mathrm{d}f_2(w)}{\mathrm{d}w} = \frac{-2}{(w+a_1)^2} + \frac{2\delta}{(w+a_2)^3} = \frac{-2(w+a_2)^3+2\delta(w+a_1)^2}{(w+a_1)^2(w+a_2)^3}.$$

Therefore, there exists a sufficiently larger $0 < \hat{w}$ so that $\frac{\mathrm{d}f_2(w)}{\mathrm{d}w} < 0$, for $0 < \hat{w} < w < +\infty$, which means that $f_2(w)$ is descending, for $0 < \hat{w} < w < +\infty$.

Together with the inequality (10), there exists an appropriate w^* $(w^* > \max\{\bar{w}, \tilde{w}, \hat{w}\})$ so that the following inequality holds.

$$0 < f_1(w) < f_2(w) < 1, \quad \text{for } 0 < w^* < w < +\infty.$$

III. STRICTLY MONOTONE CONVERGENCE AND CONVERGENCE RATE

Theorem 1: Assume that $w_k > \lambda\left(\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)$ *. Then there exists a unique OLGV* $\mathbf{\Gamma}_k$ *for the minimization problem (6) so that the OILC scheme (5) drives the system (3) satisfying*

$$\|\mathbf{e}_{k+1}\|^2 \le \sigma_k \|\mathbf{e}_k\|^2 \quad \textit{with } 0 < \sigma_k < 1.$$

Proof: Substituting (5) into system (3) gives rise to

$$\mathbf{e}_{k+1} = \mathbf{e}_k - \mathbf{H}(\mathbf{u}_{k+1} - \mathbf{u}_k) = \mathbf{e}_k - \mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k. \tag{11}$$

Thus

$$\begin{aligned}\|\mathbf{e}_{k+1}\|^2 &= \mathbf{e}_{k+1}^{\mathrm{T}}\mathbf{e}_{k+1} = (\mathbf{e}_k - \mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k)^{\mathrm{T}}(\mathbf{e}_k - \mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k) \\ &= \|\mathbf{e}_k\|^2 - \mathbf{e}_k^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k - \mathbf{\Gamma}_k^{\mathrm{T}}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{e}_k \\ &\quad + \mathbf{\Gamma}_k^{\mathrm{T}}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k.\end{aligned} \tag{12}$$

From (6) and (12), we have

$$J(\mathbf{\Gamma}_k) = \mathbf{\Gamma}_k^{\mathrm{T}}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k - 2\mathbf{e}_k^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k + \mathbf{e}_k^{\mathrm{T}}\mathbf{e}_k + w_k\mathbf{\Gamma}_k^{\mathrm{T}}\mathbf{\Gamma}_k.$$

Letting the gradient $\nabla J(\mathbf{\Gamma}_k)$ be zero yields

$$\left(w_k\mathbf{I} + \mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)\mathbf{\Gamma}_k = \mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{e}_k. \tag{13}$$

As the matrix $\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k$ is nonnegative, the matrix $\left(w_k\mathbf{I} + \mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)$ is positive definite and thus nonsingular for $w_k > 0$. Therefore, the solution $\mathbf{\Gamma}_k$ to the equation (13) is existent and unique. It is evident that

$$\mathbf{\Gamma}_k = \left(w_k\mathbf{I} + \mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{e}_k. \tag{14}$$

It should point that the solution (14) is existent only if $w_k > 0$. But the convergence requires the assumption that $w_k > \lambda\left(\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)$ to ensure Theorem 1 holds, that is,

$$\lambda\left(w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right) = w_k^{-1}\lambda\left(\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right) < 1.$$

From Lemma 1, we have

$$\begin{aligned}\mathbf{e}_{k+1} &= \mathbf{e}_k - \mathbf{H}\mathbf{E}_k\mathbf{\Gamma}_k \\ &= \mathbf{e}_k - \mathbf{H}\mathbf{E}_k\left(w_k\mathbf{I} + \mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{e}_k \\ &= \mathbf{e}_k - w_k^{-1}\mathbf{H}\mathbf{E}_k\left(\mathbf{I} + w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{e}_k \\ &= \mathbf{e}_k - w_k^{-1}\mathbf{H}\mathbf{E}_k\Big(\mathbf{I} - w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k \\ &\quad + (-w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k)^2 + \cdots\Big)\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{e}_k \\ &= \left(\mathbf{I} - w_k^{-1}\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}} + (-w_k^{-1}\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}})^2 + \cdots\right)\mathbf{e}_k \\ &= \left(\mathbf{I} + w_k^{-1}\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\right)^{-1}\mathbf{e}_k.\end{aligned}$$

Then

$$\begin{aligned}\|\mathbf{e}_{k+1}\|^2 &= \mathbf{e}_k^{\mathrm{T}}\left(\mathbf{I} + w_k^{-1}\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\right)^{-2}\mathbf{e}_k \\ &\le \frac{1}{\left(1 + w_k^{-1}\lambda_{\min}\left(\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\right)\right)^2}\|\mathbf{e}_k\|^2.\end{aligned} \tag{15}$$

What follows is to discuss the convergence factor $q_k = \frac{1}{\left(1+w_k^{-1}\lambda_{\min}\left(\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\right)\right)^2}$ for the cases when $\mathbf{E}_k$ is nonsingular and singular, respectively.

Case1: Matrix $\mathbf{E}_k$ is nonsingular.

Taking the assumption $\mathbf{CB} \neq 0$ that means the Markov parameters matrix $\mathbf{H}$ is nonsingular into account, it is confirmative that the matrix $\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}$ is positive definite. Therefore, $\lambda_{\min}\left(\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\right) > 0$. This means that

$$0 < q_k = \frac{1}{\left(1 + w_k^{-1}\lambda_{\min}\left(\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\right)\right)^2} < 1.$$

Thus inequality (15) becomes

$$\|\mathbf{e}_{k+1}\|^2 \le q_k\|\mathbf{e}_k\|^2. \tag{16}$$

Case2: Matrix $\mathbf{E}_k$ is singular.

Suppose that $rank(\mathbf{E}_k) = s_k < N$ with $e_k(i_m) \neq 0$, for $m = 1, 2, \cdots, s_k$, and $e_k(l_p) = 0$, for $p = 1, 2, \cdots, N - s_k$, respectively.

Denote an index set as

$$\Omega = \{1, 2, \cdots, N\} \setminus \left\{l_1, l_2, \cdots, l_{N-s_k}\right\} = \left\{i_1, i_2, \cdots, i_{s_k}\right\}.$$

Here

$$\begin{aligned}&i_1 < i_2 < \cdots < i_{s_k}, \\ &l_1 < l_2 < \cdots < l_{N-s_k}.\end{aligned}$$

Let

$$\begin{aligned}\tilde{e}_k &= \left[e_k(i_1)|e_k(i_2)|\cdots|e_k(i_{s_k})\right]^{\mathrm{T}},\\ \hat{e}_k &= \left[e_k(l_1)|e_k(l_2)|\cdots|e_k(l_{N-s_k})\right]^{\mathrm{T}},\\ \tilde{\mathbf{E}}_k &= \mathrm{diag}\left(e_k(i_1), e_k(i_2), \cdots, e_k(i_{s_k})\right),\\ \hat{\mathbf{E}}_k &= \mathrm{diag}\left(e_k(l_1), e_k(l_2), \cdots, e_k(l_{N-s_k})\right),\\ \tilde{\boldsymbol{\Upsilon}}_k &= \left[\Upsilon_k(i_1)|\Upsilon_k(i_2)|\cdots|\Upsilon_k(i_{s_k})\right]^{\mathrm{T}},\\ \hat{\boldsymbol{\Upsilon}}_k &= \left[\Upsilon_k(l_1)|\Upsilon_k(l_2)|\cdots|\Upsilon_k(l_{N-s_k})\right]^{\mathrm{T}}.\end{aligned}$$

Then, $\hat{\mathbf{e}}_k = \mathbf{0}$, $\hat{\mathbf{E}}_k = \mathbf{0}$ and $\tilde{\mathbf{E}}_k$ is invertible.

Denote

$$\mathbf{P}_k = \left[\boldsymbol{\varepsilon}_{i_1}|\boldsymbol{\varepsilon}_{i_2}|\cdots|\boldsymbol{\varepsilon}_{i_{s_k}}|\boldsymbol{\varepsilon}_{l_1}|\boldsymbol{\varepsilon}_{l_2}|\cdots|\boldsymbol{\varepsilon}_{l_{N-s_k}}\right]^{\mathrm{T}},$$

where ε_j stands for the unity vector whose the j-th component is unity and the others are zero. It is obvious that the matrix $\mathbf{P}_k$ is orthogonal. Thus $\mathbf{P}_k$ is invertible and $\mathbf{P}_k^{-1} = \mathbf{P}_k^{\mathrm{T}}$.

Because the rows of the matrix $\mathbf{P}_k$ are the i_1−th, $\cdots$, i_s−th, l_1−th, $\cdots$, l_{N-s}−th rows of the identity matrix $\mathbf{I}$, it is learnt that $\mathbf{P}_k$ is an elementary matrix which may exchanges the rows positions of a matrix while pre-multiplied by $\mathbf{P}_k$. Thus its linear transformation to a matrix is termed as an elementary transformation of a matrix. According to the denotation of $\mathbf{P}_k$, we have

$$\mathbf{P}_k\mathbf{H}\mathbf{P}_k^{\mathrm{T}} = \begin{bmatrix} h_1 & 0 & \cdots & 0 & *\cdots* \\ h_{i_2-i_1+1} & h_1 & \cdots & 0 & *\cdots* \\ \vdots & \vdots & \ddots & \vdots & \vdots\cdots\vdots \\ h_{i_{s_k}-i_1+1} & h_{i_{s_k}-i_2+1} & \cdots & h_1 & *\cdots* \\ * & * & \cdots & * & *\cdots* \\ \vdots & \vdots & \ddots & \vdots & \vdots\cdots\vdots \\ * & \cdots & * & \cdots & **\cdots* \end{bmatrix},$$

where $h_j = g_1(j)$ and "$*$" remarks some elements of matrix $\mathbf{H}$.

Denote

$$\mathbf{P}_k\mathbf{H}\mathbf{P}_k^{\mathrm{T}} = \begin{bmatrix} \tilde{\mathbf{H}}_k & \check{\mathbf{H}}_k \\ \hat{\mathbf{H}}_k & \widehat{\mathbf{H}}_k \end{bmatrix}.$$

Here

$$\tilde{\mathbf{H}}_k = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ h_{i_2-i_1+1} & h_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{i_{s_k}-i_1+1} & h_{i_{s_k}-i_2+1} & \cdots & h_1 \end{bmatrix}$$

is an $s_k \times s_k$ -dimensional square matrix an $\hat{\mathbf{H}}_k$, $\check{\mathbf{H}}_k$ and $\widehat{\mathbf{H}}_k$ are matrices with appropriate dimensions, respectively. It is obvious that $\tilde{\mathbf{H}}_k$ is invertible as $h_1 = g_1(1) \neq 0$.

In specific,

$$\mathbf{P}_k\mathbf{e}_k = \begin{bmatrix}\tilde{\mathbf{e}}_k\\ \hat{\mathbf{e}}_k\end{bmatrix} = \begin{bmatrix}\tilde{\mathbf{e}}_k\\ \mathbf{0}\end{bmatrix},$$

$$\mathbf{P}_k\boldsymbol{\Gamma}_k = \begin{bmatrix}\tilde{\boldsymbol{\Gamma}}_k\\ \hat{\boldsymbol{\Gamma}}_k\end{bmatrix},$$

$$\mathbf{P}_k\mathbf{E}_k\mathbf{P}_k^{\mathrm{T}} = \begin{bmatrix}\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & \hat{\mathbf{E}}_k\end{bmatrix} = \begin{bmatrix}\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}.$$

Pre-multiplying $\mathbf{P}_k$ to equation (13) and considering the property $\mathbf{P}_k^{\mathrm{T}} = \mathbf{P}_k^{-1}$ makes

$$\begin{aligned} w_k\mathbf{P}_k\boldsymbol{\Gamma}_k + (\mathbf{P}_k\mathbf{E}_k\mathbf{P}_k^{\mathrm{T}})(\mathbf{P}_k\mathbf{H}\mathbf{P}_k^{\mathrm{T}})^{\mathrm{T}}(\mathbf{P}_k\mathbf{H}\mathbf{P}_k^{\mathrm{T}})(\mathbf{P}_k\mathbf{E}_k\mathbf{P}_k^{\mathrm{T}})(\mathbf{P}_k\boldsymbol{\Gamma}_k)\\ = (\mathbf{P}_k\mathbf{E}_k\mathbf{P}_k^{\mathrm{T}})(\mathbf{P}_k\mathbf{H}\mathbf{P}_k^{\mathrm{T}})^{\mathrm{T}}(\mathbf{P}_k\mathbf{e}_k).\end{aligned}$$

That is

$$\begin{aligned} & w_k\begin{bmatrix}\tilde{\boldsymbol{\Gamma}}_k\\ \hat{\boldsymbol{\Gamma}}_k\end{bmatrix}\\ & + \begin{bmatrix}\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{H}}_k^{\mathrm{T}} & \hat{\mathbf{H}}_k^{\mathrm{T}}\\ \check{\mathbf{H}}_k^{\mathrm{T}} & \widehat{\mathbf{H}}_k\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{H}}_k & \check{\mathbf{H}}_k\\ \hat{\mathbf{H}}_k & \widehat{\mathbf{H}}_k\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{\Gamma}}_k\\ \hat{\boldsymbol{\Gamma}}_k\end{bmatrix}\\ & = \begin{bmatrix}\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{H}}_k^{\mathrm{T}} & \hat{\mathbf{H}}_k^{\mathrm{T}}\\ \check{\mathbf{H}}_k^{\mathrm{T}} & \widehat{\mathbf{H}}_k^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{e}}_k\\ \mathbf{0}\end{bmatrix}.\end{aligned}$$

Equivalently, equation (13) becomes

$$\begin{bmatrix} w_k\mathbf{I}_{s_k} + \tilde{\mathbf{E}}_k\left(\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k^{\mathrm{T}}\hat{\mathbf{H}}_k\right)\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & w_k\mathbf{I}_{N-s_k}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{\Gamma}}_k\\ \hat{\boldsymbol{\Gamma}}_k\end{bmatrix} = \begin{bmatrix}\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}} & \tilde{\mathbf{E}}_k\hat{\mathbf{H}}_k^{\mathrm{T}}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{e}}_k\\ \mathbf{0}\end{bmatrix}.$$

Therefore

$$\left(w_k\mathbf{I}_{s_k} + \tilde{\mathbf{E}}_k\left(\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k^{\mathrm{T}}\hat{\mathbf{H}}_k\right)\tilde{\mathbf{E}}_k\right)\tilde{\boldsymbol{\Gamma}}_k = \tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k, \tag{17}$$

$$w_k\hat{\boldsymbol{\Gamma}}_k = \mathbf{0}. \tag{18}$$

As $w_k \neq 0$ equation (18) makes

$$\hat{\boldsymbol{\Gamma}}_k = \mathbf{0}.$$

Notice that matrix $w_k\mathbf{I}_{s_k} + \tilde{\mathbf{E}}_k\left(\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k^{\mathrm{T}}\hat{\mathbf{H}}_k\right)\tilde{\mathbf{E}}_k$ is positive definite and thus invertible. Equation (17) therefore there exists a unique solution

$$\tilde{\boldsymbol{\Gamma}}_k = \left(w_k\mathbf{I}_{s_k} + \tilde{\mathbf{E}}_k\left(\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k^{\mathrm{T}}\hat{\mathbf{H}}_k\right)\tilde{\mathbf{E}}_k\right)^{-1}\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k.$$

Equivalently

$$\tilde{\mathbf{E}}_k\tilde{\boldsymbol{\Gamma}}_k = \tilde{\mathbf{E}}_k\left(w_k\mathbf{I}_{s_k} + \tilde{\mathbf{E}}_k\left(\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k^{\mathrm{T}}\hat{\mathbf{H}}_k\right)\tilde{\mathbf{E}}_k\right)^{-1}\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k. \tag{19}$$

Denote

$$\mathbf{M}_k = \tilde{\mathbf{E}}_k\left(\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k^{\mathrm{T}}\hat{\mathbf{H}}_k\right)\tilde{\mathbf{E}}_k, \tag{20}$$

$$\mathbf{Q}_k = \left(w_k\mathbf{I}_{s_k} + \mathbf{M}_k\right)^{-1}. \tag{21}$$

Then (19) is rewritten as

$$\tilde{\mathbf{E}}_k\tilde{\boldsymbol{\Gamma}}_k = \tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k. \tag{22}$$

Pre-multiplying equation (11) by $\mathbf{P}_k$ makes

$$\begin{aligned}\mathbf{P}_k\mathbf{e}_{k+1} &= \begin{bmatrix}\tilde{\mathbf{e}}_{k+1}\\ \hat{\mathbf{e}}_{k+1}\end{bmatrix}\\ &= \mathbf{P}_k\mathbf{e}_k - \left(\mathbf{P}_k\mathbf{H}\mathbf{P}_k^{\mathrm{T}}\right)\left(\mathbf{P}_k\mathbf{E}_k\mathbf{P}_k^{\mathrm{T}}\right)\left(\mathbf{P}_k\mathbf{\Gamma}_k\right)\\ &= \begin{bmatrix}\tilde{\mathbf{e}}_k\\ \mathbf{0}\end{bmatrix} - \begin{bmatrix}\tilde{\mathbf{H}}_k & \breve{\mathbf{H}}_k\\ \hat{\mathbf{H}}_k & \widehat{\mathbf{H}}_k\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{E}}_k & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{\Gamma}}_k\\ \hat{\mathbf{\Gamma}}_k\end{bmatrix}\\ &= \begin{bmatrix}\tilde{\mathbf{e}}_k - \tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{\Gamma}}_k\\ -\hat{\mathbf{H}}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{\Gamma}}_k\end{bmatrix}\\ &= \begin{bmatrix}\tilde{\mathbf{e}}_k - \tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k\\ -\hat{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k\end{bmatrix}. \end{aligned} \tag{23}$$

Calculating 2-norm to the above equation (23) gives rise to

$$\begin{aligned}&\|\mathbf{P}_k\mathbf{e}_{k+1}\|^2\\ &\quad= \|\mathbf{e}_{k+1}\|^2 = \tilde{\mathbf{e}}_{k+1}^{\mathrm{T}}\tilde{\mathbf{e}}_{k+1} + \hat{\mathbf{e}}_{k+1}^{\mathrm{T}}\hat{\mathbf{e}}_{k+1}\\ &\quad= \tilde{\mathbf{e}}_k^{\mathrm{T}}\left(\mathbf{I} - \tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\right)^{\mathrm{T}}\left(\mathbf{I} - \tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\right)\tilde{\mathbf{e}}_k\\ &\qquad+ \tilde{\mathbf{e}}_k^{\mathrm{T}}(\hat{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}})^{\mathrm{T}}(\hat{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}})\tilde{\mathbf{e}}_k\\ &\quad= \tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k - \tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\left(2\mathbf{Q}_k - \mathbf{Q}_k\mathbf{M}_k\mathbf{Q}_k\right)\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k. \end{aligned} \tag{24}$$

Let

$$\boldsymbol{\eta}_k = \tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k.$$

Mind that the matrix $\tilde{\mathbf{H}}_k$ is invertible and all of components of the vector $\tilde{\mathbf{e}}_k$ are nonzero. Then

$$\|\boldsymbol{\eta}_k\|^2 = \left\|\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k\right\|^2 = \mu_k\|\tilde{\mathbf{e}}_k\|^2 \neq 0.$$

In addition

$$\begin{aligned}\lambda_{\min}(\mathbf{Q}_k)\|\boldsymbol{\eta}_k\|^2 &\le \tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\mathbf{Q}_k\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k\\ &= \boldsymbol{\eta}_k^{\mathrm{T}}\mathbf{Q}_k\boldsymbol{\eta}_k \le \lambda_{\max}(\mathbf{Q}_k)\|\boldsymbol{\eta}_k\|^2. \end{aligned} \tag{25}$$

And

$$\tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\left(\mathbf{Q}_k\mathbf{M}_k\mathbf{Q}_k\right)\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k = \left(\mathbf{Q}_k\boldsymbol{\eta}_k\right)^{\mathrm{T}}\left(\mathbf{M}_k\right)\mathbf{Q}_k\boldsymbol{\eta}_k. \tag{26}$$

From the denotations of $\mathbf{M}_k$ in (20) and $\mathbf{Q}_k$ in (21), it is easy to assert that the matrices $\mathbf{M}_k$ and $\mathbf{Q}_k$ are symmetric and positive definite. Therefore, equation (26) leads to

$$\begin{aligned}\tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\left(\mathbf{Q}_k\mathbf{M}_k\mathbf{Q}_k\right)\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k &= \left(\mathbf{Q}_k\boldsymbol{\eta}_k\right)^{\mathrm{T}}\left(\mathbf{M}_k\right)\mathbf{Q}_k\boldsymbol{\eta}_k\\ &= \delta_k\cdot\boldsymbol{\eta}_k^{\mathrm{T}}\mathbf{Q}_k\mathbf{Q}_k\boldsymbol{\eta}_k, \end{aligned} \tag{27}$$

where δ_k is an appropriate constant satisfying $0 < \lambda_{\min}(\mathbf{M}_k) \le \delta_k \le \lambda_{\max}(\mathbf{M}_k)$ with $\lambda_{\min}(\mathbf{M}_k)$ and $\lambda_{\max}(\mathbf{M}_k)$ respectively present the minimum and the maximum of the eigenvalues of the matrix $\mathbf{M}_k$.

Thus, equation (27) delivers

$$\begin{aligned}\delta_k\cdot\lambda_{\min}\left(\mathbf{Q}_k^2\right)\|\boldsymbol{\eta}_k\|^2 &\le \tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\left(\mathbf{Q}_k\mathbf{M}_k\mathbf{Q}_k\right)\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k\\ &\le \delta_k\cdot\lambda_{\max}\left(\mathbf{Q}_k^2\right)\|\boldsymbol{\eta}_k\|^2. \end{aligned} \tag{28}$$

Then equations (25) and (28) result in

$$\begin{aligned}&\left(2\lambda_{mim}(\mathbf{Q}_k) - \delta_k\lambda_{\max}\left(\mathbf{Q}_k^2\right)\right)\mu_k\|\mathbf{e}_k\|^2\\ &\qquad\le \tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\left(2\mathbf{Q}_k - \mathbf{Q}_k\mathbf{M}_k\mathbf{Q}_k\right)\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k, \end{aligned} \tag{29}$$

and

$$\begin{aligned}&\tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{H}}_k\tilde{\mathbf{E}}_k\left(2\mathbf{Q}_k - \mathbf{Q}_k\mathbf{M}_k\mathbf{Q}_k\right)\tilde{\mathbf{E}}_k\tilde{\mathbf{H}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k\\ &\qquad\le \left(2\lambda_{\max}(\mathbf{Q}_k) - \delta_k\lambda_{\min}\left(\mathbf{Q}_k^2\right)\right)\mu_k\|\mathbf{e}_k\|^2. \end{aligned} \tag{30}$$

From the expression (21), we have

$$\lambda(\mathbf{Q}_k) = \frac{1}{w_k + \lambda(\mathbf{M}_k)},\qquad \lambda\left(\mathbf{Q}_k^2\right) = \frac{1}{(w_k + \lambda(\mathbf{M}_k))^2}.$$

Denote

$$\begin{aligned}\xi_{k,\min}^* &= \left(2\lambda_{mim}(\mathbf{Q}_k) - \delta_k\lambda_{\max}\left(\mathbf{Q}_k^2\right)\right)\mu_k\\ &= \frac{2\mu_k}{w_k + \lambda_{\max}(\mathbf{M}_k)} - \frac{\delta_k\mu_k}{(w_k + \lambda_{\min}(\mathbf{M}_k))^2}, \end{aligned} \tag{31}$$

$$\begin{aligned}\xi_{k,\max}^* &= \left(2\lambda_{\max}(\mathbf{Q}_k) - \delta_k\lambda_{\min}\left(\mathbf{Q}_k^2\right)\right)\mu_k\\ &= \frac{2\mu_k}{w_k + \lambda_{\min}(\mathbf{M}_k)} - \frac{\delta_k\mu_k}{(w_k + \lambda_{\max}(\mathbf{M}_k))^2}. \end{aligned} \tag{32}$$

According to Lemma 2, there exists a constant $w_k > 0$ so that

$$0 < \xi_{k,\min}^* \le \xi_{k,\max}^* < 1. \tag{33}$$

Let

$$v_k = 1 - \xi_{k,\min}^*.$$

Then

$$0 < v_k = 1 - \xi_{k,\min}^* < 1.$$

Thus, equations (24) and (29) reduce

$$0 \le \|\mathbf{e}_{k+1}\|^2 \le v_k\tilde{\mathbf{e}}_k^{\mathrm{T}}\tilde{\mathbf{e}}_k = v_k\|\mathbf{e}_k\|^2. \tag{34}$$

Let $\sigma_k = \max\{q_k, v_k\}$.

Then inequalities (16) and (34) makes

$$\|\mathbf{e}_{k+1}\|^2 \le \sigma_k\|\mathbf{e}_k\|^2, \quad \text{with } 0 < \sigma_k < 1.$$

This completes the proof. ∎

Remark 2: It is no difficult to calculate that the matrix $\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}$ *in this paper turns to be the matrix* $\tilde{\mathbf{G}}_{e,k+1}\tilde{\mathbf{G}}_{e,k+1}^{\mathrm{T}}$ *in the reference [34]. It is recalled that Proposition 7 in reference [34] has exhibited that the monotone convergence is strictly ensured for the case when the smallest eigenvalue of the matrix* $\tilde{\mathbf{G}}_{e,k+1}\tilde{\mathbf{G}}_{e,k+1}^{\mathrm{T}}$ *or* $\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}$ *are nonzero. But the reference did not consider the insurance condition of the tuning factor w for the following formula.*

$$\begin{aligned}&\left(\mathbf{I} + w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right)^{-1}\\ &\qquad= \mathbf{I} - w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k + (-w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k)^2 + \cdots. \end{aligned} \tag{35}$$

In fact, from Taylor expansion, it is learnt that the equation (35) holds if and only if $\rho\left(w_k^{-1}\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right) < 1$.

Besides, for the case when some of the eigenvalues of matrix $\tilde{\mathbf{G}}_{e,k+1}\tilde{\mathbf{G}}_{e,k+1}^{\mathrm{T}}$ *or* $\mathbf{H}\mathbf{E}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}$ *are zero, Proposition 7 in reference [34] has presented that the projected tracking error is strictly monotonously convergent. Here, as the reference mentioned, the projection of the tracking error means to project the tracking error vector on the eigenvalues of* $\tilde{\mathbf{G}}_{e,k+1}\tilde{\mathbf{G}}_{e,k+1}^{\mathrm{T}}$ *that correspond to non-zero eigenvalues of* $\tilde{\mathbf{G}}_{e,k+1}\tilde{\mathbf{G}}_{e,k+1}^{\mathrm{T}}$ *. In authors' opinion, the projection concept here is beyond comprehension because the referencedid not give an explicit formulation or explanation to express the projection. Thus it is obscure that the projection of the tracking error vector isexactly to present the projection on the nonzero components of the tracking error vector. Meanwhile, as the projection operator is not an invertible transformation, thus the monotonicity of the projected tracking error does not definitely mean that the tracking error itself is strictly monotonously convergent.*

Remark 3: From the derivation of the case 2 which is equivalent to that part components of the tracking error vector $\mathbf{e}_k$ *are zero, it is noticed that the strictly monotone convergence is guaranteed only if the tuning factor* w_k *is chosen to ensure the inequality (33) holds, though the derivation is quite distinct from the case 1. From the expression of* $\xi^*_{k,\min}$ *in (31), the range of* w_k *is relevant to the eigenvalues of matrix* $\mathbf{M}_k$ *that is expressed by nonzero components of the tracking error vector* $\mathbf{e}_k$ *and the corresponding Markov parameters. From the proof of Lemma 2, it is worth y to emphasize that the larger of the tuning factor* w_k *guaranteeing (33) may yield the smaller of the values* $\xi^*_{k,\min}$ *. This is equivalent to the slower of the convergent rate* $\left(1-\xi^*_{k,\min}\right)$ *. This implies that we must choose the tuning factor* w_k *in an appropriate range neither too small nor too large to ensure an expected convergence rate* $\left(1-\xi^*_{k,\min}\right)$ *. For the circumstance, it is evident that the optimal learning gain vector* Γ_k *in (14) implies* $\lim_{k\to\infty}\Gamma_k = 0$.

IV. OPTIMIZED HIGHER-ORDER ITERATIVE LEARNING CONTROL SCHEME

Consider a single-input-single-output (SISO) linear discrete-time-invariant (LDTI) system whose dynamical description is as follows.

$$\begin{cases} x(n+1) = \mathbf{A}x(n) + \mathbf{B}u(n), \\ y(n+1) = \mathbf{C}x(n+1), \\ x(0) = 0, \quad n = 0, 1, \cdots, N-1, \end{cases} \tag{36}$$

where N denotes the total sampling number, $x(n)$, $u(n)$ and $y(n)$ are p-dimensional state vector, scalar input and scalar output, respectively. $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices with appropriate dimensions, respectively. Assume that the system (36) satisfies the following conditions

$$\begin{cases} \mathbf{C}\mathbf{A}^{i-1}\mathbf{B} = 0, & 1 \le i \le r-1 \\ \mathbf{C}\mathbf{A}^{i}\mathbf{B} \ne 0, & i = r \end{cases} \tag{37}$$

Here, r is an integer larger than unity and $A^i (i = 0, 1, \cdots, r-1)$ represents the i−th power of the matrix $\mathbf{A}$. In the circumstance, the relative degree of the system (36) is said to be r. The definition can refer to reference [36].

From the assumption of (37), it is easy to acquire that $g_1(i) = \mathbf{C}\mathbf{A}^{i-1}\mathbf{B} = 0$, for $1 \le i \le r-1$, and $g_1(r) \ne 0$. Then, for any input $\mathbf{u} = [u(0)|u(1)|\cdots|u(N-1)]^{\mathrm{T}}$, equation (2) makes

$$y(1) = y(2) = \cdots = y(r-1) = 0. \tag{38}$$

For a predetermined desired trajectory $\mathbf{y}_d = [0|\cdots|0|y_d(r)|y_d(r+1)|\cdots|y_d(N)]^{\mathrm{T}}$, formulate an r-th-order D-type ILC as

$u_1(n)$, arbitrary given;

$$u_{k+1}(n) = \begin{cases} u_k(n) + \Gamma_k(n+r)e_k(n+r), & n = 0, 1, \cdots, N-r; \\ u_k(n), & n = N-r+1, \cdots, N-1. \end{cases} \tag{39}$$

In algorithm (39), the subscript k, $k = 1, 2, \cdots$ denotes the iteration index and $\Gamma_k(n+r)$ is termed as the r-th-order iteration-time-variable derivative learning gain, respectively.

For statement simplicity, denote

$$\begin{aligned} \mathbf{y}_d^{N-r} &= [y_d(r)|y_d(r+1)|\cdots|y_d(N)]^{\mathrm{T}}, \\ \mathbf{u}_k^{N-r} &= [u_k(0)|u_k(1)|\cdots|u_k(N-r)]^{\mathrm{T}}, \\ \mathbf{u}_k^{r} &= [u_k(N-r+1)|u_k(N-r+2)|\cdots|u_k(N-1)]^{\mathrm{T}}, \\ \mathbf{y}_k^{N-r} &= [y_k(r)|y_k(r+1)|\cdots|y_k(N)]^{\mathrm{T}}, \\ \mathbf{y}_k^{r} &= [y_k(1)|y_k(2)|\cdots|y_k(r-1)]^{\mathrm{T}}. \end{aligned}$$

From (38), we have

$$\mathbf{y}_k^{r} = [y_k(1)|y_k(2)|\cdots|y_k(r-1)]^{\mathrm{T}} = \mathbf{0}.$$

Then

$$\begin{aligned} \mathbf{y}_d &= \left[\mathbf{0}^{\mathrm{T}}|(\mathbf{y}_d^{N-r})^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \mathbf{u}_k &= [u_k(0)|u_k(1)|\cdots|u_k(N-1)]^{\mathrm{T}} = \left[(\mathbf{u}_k^{N-r})^{\mathrm{T}}|(\mathbf{u}_k^{r})^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \mathbf{y}_k &= [y_k(1)|y_k(2)|\cdots|y_k(N)]^{\mathrm{T}} = \left[\mathbf{0}^{\mathrm{T}}|(\mathbf{y}_k^{N-r})^{\mathrm{T}}\right]^{\mathrm{T}}. \end{aligned}$$

Let

$$\mathbf{e}_k = \mathbf{y}_d - \mathbf{y}_k = \left[(\mathbf{e}_k^{r})^{\mathrm{T}}|(\mathbf{e}_k^{N-r})^{\mathrm{T}}\right]^{\mathrm{T}}.$$

Here

$$\begin{aligned} \mathbf{e}_k^{r} &= [e_k(1)|e_k(2)|\cdots|e_k(r-1)]^{\mathrm{T}} = \mathbf{0}, \\ \mathbf{e}_k^{N-r} &= [e_k(r)|e_k(r+1)|\cdots|e_k(N)]^{\mathrm{T}}. \end{aligned}$$

Then $\mathbf{e}_k = \left[\mathbf{0}^{\mathrm{T}}|(\mathbf{e}_k^{N-r})^{\mathrm{T}}\right]^{\mathrm{T}}$.

Meanwhile, equation (36) is reformed as

$$\mathbf{y} = \mathbf{H}\mathbf{u}, \tag{40}$$

where $\mathbf{H} = \begin{bmatrix} \mathbf{0}_{11} & \mathbf{0}_{12} \\ \mathbf{H}^{N-r} & \mathbf{0}_{22} \end{bmatrix}$ with $\mathbf{0}_{11}$, $\mathbf{0}_{12}$ and $\mathbf{0}_{22}$ being nullity block matrices with appropriate dimensions and

$$\mathbf{H}^{N-r} = \begin{bmatrix} g_1(r) & 0 & 0 & \cdots & 0 \\ g_1(r+1) & g_1(r) & 0 & \cdots & 0 \\ g_1(r+2) & g_1(r+1) & g_1(r) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_1(N) & g_1(N-1) & g_1(N-2) & \cdots & g_1(r) \end{bmatrix}.$$

Then $\mathbf{H}^{N-r}$ is invertible.

Denote

$$\begin{aligned} \mathbf{\Gamma}_k^{N-r} &= [\Gamma_k(r)|\Gamma_k(r+1)|\cdots|\Gamma_k(N)]^{\mathrm{T}}, \\ \mathbf{\Gamma}_k &= [\Gamma_k(1)|\Gamma_k(2)|\cdots|\Gamma_k(N)]^{\mathrm{T}} = \left[\mathbf{0}^{\mathrm{T}}|(\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \Phi_k^r &= diag(e_k(1), e_k(2), \cdots, e_k(r-1)) = \mathbf{0}, \\ \Phi_k^{N-r} &= diag(e_k(r), e_k(r+1), \cdots, e_k(N)), \\ \Phi_k &= diag(e_k(1), e_k(2), \cdots, e_k(N)) = diag(\mathbf{0}, \Phi_k^{N-r}). \end{aligned}$$

Then Φ_k^{N-r} is invertible.

The r-th-order D-type ILC (39) is thus rewritten as

$$\begin{aligned} \mathbf{u}_{k+1} &= \left[\left(\mathbf{u}_{k+1}^{N-r}\right)^{\mathrm{T}} | \left(\mathbf{u}_{k+1}^{r}\right)^{\mathrm{T}}\right]^{\mathrm{T}} \\ &= \left[(\mathbf{u}_k^{N-r} + \Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}|(\mathbf{u}_k^r)^{\mathrm{T}}\right]^{\mathrm{T}}. \end{aligned} \quad (41)$$

According to the denotations, we have

$$\begin{aligned} &\mathbf{e}_{k+1} \\ &= \begin{bmatrix} \mathbf{e}_{k+1}^r \\ \mathbf{e}_{k+1}^{N-r} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_k^r \\ \mathbf{e}_k^{N-r} \end{bmatrix} - \left(\begin{bmatrix} \mathbf{y}_{k+1}^r \\ \mathbf{y}_{k+1}^{N-r} \end{bmatrix} - \begin{bmatrix} \mathbf{y}_k^r \\ \mathbf{y}_k^{N-r} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{e}_k^r \\ \mathbf{e}_k^{N-r} \end{bmatrix} \\ &\quad - \begin{bmatrix} \mathbf{0}_{11} \mathbf{0}_{12} \\ \mathbf{H}^{N-r} \mathbf{0}_{22} \end{bmatrix} \left(\begin{bmatrix} \mathbf{u}_k^{N-r} + \Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r} \\ \mathbf{u}_k^r \end{bmatrix} - \begin{bmatrix} \mathbf{u}_k^{N-r} \\ \mathbf{u}_k^r \end{bmatrix}\right). \end{aligned} \quad (42)$$

Then

$$\mathbf{e}_{k+1} = \begin{bmatrix} \mathbf{e}_{k+1}^r \\ \mathbf{e}_{k+1}^{N-r} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{e}_k^{N-r} - \mathbf{H}^{N-r}\Phi_k^{N-r}\Gamma_k^{N-r} \end{bmatrix}. \quad (43)$$

Thus

$$\begin{aligned} \|\mathbf{e}_{k+1}\|^2 &= \left\|\mathbf{e}_{k+1}^{N-r}\right\|^2 = (\mathbf{e}_{k+1}^{N-r})^{\mathrm{T}}\mathbf{e}_{k+1}^{N-r} \\ &= \left(\mathbf{e}_k^{N-r} - \mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r}\right)^{\mathrm{T}} \\ &\quad \times \left(\mathbf{e}_k^{N-r} - \mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r}\right) \\ &= \left\|\mathbf{e}_k^{N-r}\right\|^2 - (\mathbf{e}_k^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r} \\ &\quad + (\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}\Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r} \\ &\quad - (\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}\Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{e}_k^{N-r}. \end{aligned} \quad (44)$$

According to the expressions of $\mathbf{e}_{k+1}$ and $\mathbf{\Gamma}_k$, the minimization problem (6) becomes

$$\begin{aligned} \underset{\Gamma_k}{\operatorname{argmin}} J(\mathbf{\Gamma}_k) &= \underset{\Gamma_k^{N-r}}{\operatorname{argmin}} J\left(\mathbf{\Gamma}_k^{N-r}\right) \\ &= \left\|\mathbf{e}_{k+1}^{N-r}\right\|^2 + \tilde{w}_k \cdot \left\|\mathbf{\Gamma}_k^{N-r}\right\|^2 \end{aligned} \quad (45)$$

Substituting expression (44) into (45), we have

$$\begin{aligned} J\left(\mathbf{\Gamma}_k^{N-r}\right) &= (\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}\Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r} \\ &\quad - (\mathbf{e}_k^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r} \\ &\quad - (\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}\Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{e}_k^{N-r} \\ &\quad + (\mathbf{e}_k^{N-r})^{\mathrm{T}}\mathbf{e}_k^{N-r} + \tilde{w}_k(\mathbf{\Gamma}_k^{N-r})^{\mathrm{T}}\mathbf{\Gamma}_k^{N-r}. \end{aligned} \quad (46)$$

The monotone convergence of the algorithm (41) is as follows.

Theorem 2: Assume that the relative degree of the system (36) is $r > 1$ and an appropriate tuning factor in (45) satisfies $\tilde{w}_k > \lambda\left(\Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\right)$. Then there exists a unique OLGV Γ_k^{N-r} for the minimization problem (45) so that the ILOC law (39) drives the system (36) satisfying

$$\|\mathbf{e}_{k+1}\|^2 \le \bar{\sigma}_k \|\mathbf{e}_k\|^2, \quad \text{with } 0 < \bar{\sigma}_k < 1$$

Proof: Calculating gradient of $J\left(\mathbf{\Gamma}_k^{N-r}\right)$ with respect to argument $\mathbf{\Gamma}_k^{N-r}$ in (46) reduces

$$\begin{aligned} &\nabla J(\mathbf{\Gamma}_k^{N-r}) \\ &= 2\Phi_k^r\left((\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\mathbf{\Gamma}_k^{N-r} - (\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{e}_k^{N-r}\right) \\ &\quad + 2\tilde{w}_k\mathbf{\Gamma}_k^{N-r}. \end{aligned}$$

Letting $\nabla J(\mathbf{\Gamma}_k^{N-r}) = 0$ results in

$$\begin{aligned} &(\tilde{w}_k\mathbf{I}^{N-r} + \Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r})\mathbf{\Gamma}_k^{N-r} \\ &\qquad = \Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{e}_k^{N-r}. \end{aligned}$$

Then

$$\begin{aligned} \mathbf{\Gamma}_k^{N-r} &= \left(\tilde{w}_k\mathbf{I}^{N-r} + \Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{H}^{N-r}\Phi_k^{N-r}\right)^{-1} \\ &\qquad \cdot \Phi_k^{N-r}(\mathbf{H}^{N-r})^{\mathrm{T}}\mathbf{e}_k^{N-r}. \end{aligned} \quad (47)$$

Notice that the form of equation (47) is no other than the form of equation (14) if neglecting the dimension difference and the superscripts. Analogous to the derivation of Theorem 1, the conclusion of Theorem 2 is true. ■

Remark 4: As mentioned in Remark 1 that the matrix $\mathbf{HE}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}$ in this paper turns to be the matrix $\tilde{\mathbf{G}}_{e,k+1}\tilde{\mathbf{G}}_{e,k+1}^{\mathrm{T}}$ in the reference [34]. From the derivations of Theorems 1 and 2 of this paper, it is observed that the nullity of eigenvalues of the matrix $\mathbf{HE}_k\mathbf{E}_k\mathbf{H}^{\mathrm{T}}$ depends upon not only the singularity of the system Markov parameters matrix $\mathbf{H}$ but also the singularity of the iteration-wise tracking error-relevant diagonal matrix $\mathbf{E}_k$. In particular, for the higher-order relative degree of system (36), which implies the system Markov parameters matrix $\mathbf{H}$ is singular, the order of the D-type ILC (39) must match the relative degree of the system. For the circumstance, *in virtue of rigorous derivation of Theorems 1 and 2, it is significant that the strictly monotone convergence are guaranteed without any requirement to the system Markov parameters matrix* $\mathbf{H}$*. Therefore, Theorems 1 and 2 in this paper are regarded as non-conditional convergence conclusions. This is the true essence what the ILC pursues as the learning-gain vector is iteration-time-variable and full-dimensionally adapted. Whilst the reference [34] has presented incomplete monotone convergence, as remarked in Remark 2, under the assumption that* $\mathbf{H}+\mathbf{H}^{\mathrm{T}}$ *is positive definite.*

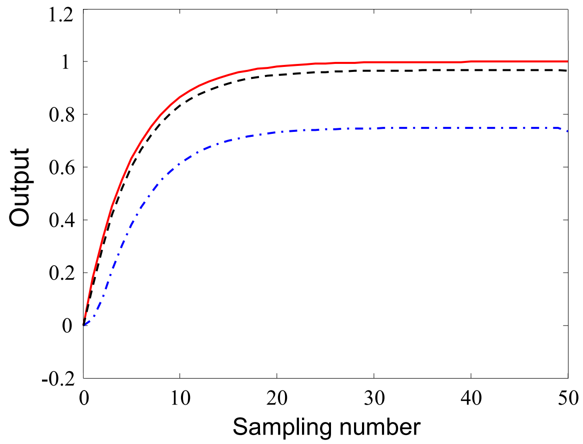

FIGURE 1. Tracking behavior.

V. NUMERICAL SIMULATIONS

Example 1: Consider an SISO linear LDTI system as follows.

$$\begin{cases} x_{k+1}(n+1) = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ -0.04 & -0.2 & -1.8 \end{bmatrix} x_{k+1}(n) \\ \qquad + \begin{bmatrix} -2 \\ 0 \\ 0.4 \end{bmatrix} u_{k+1}(n), \\ y_{k+1}(n+1) = \begin{bmatrix} 0.04 & 0.2 & 0 \end{bmatrix} x_{k+1}(n+1); \quad n \in S, \end{cases} \tag{48}$$

where the operation time samplings are set as $S = \{0, 1, \cdots, 49\}$, the initial state is set as $[x_1(0), x_2(0), x_3(0)]^{\mathrm{T}} = \mathbf{0}$. Obviously, the eigenvalues of the state matrix 0.1, 0.1 and -1.8, respectively. This means that system (48) is unstable as the modulus of eigenvalue -1.8 is greater than unity. The desired trajectory is defined as $y_d(n+1) = 1 - e^{-0.2n}$, $n \in S$. It is testified $\mathbf{CB} = -0.08 \neq 0$, which implies that the relative degree of the system (48) is unity. By imposing the OILC (5) with the learning gain vector expressed by (14) on system (48) and choosing the tuning factor as $w_k = \lambda_{\max}\left(\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right) + 10^{-6}$, the tracking behavior is exhibited in Fig.1, where the solid curve stands for the desired trajectory, the dash-dotted and dash ones are the outputs at the third and sixth iterations, respectively. It is seen from Fig.1 that the output tracks the desired trajectory very well.

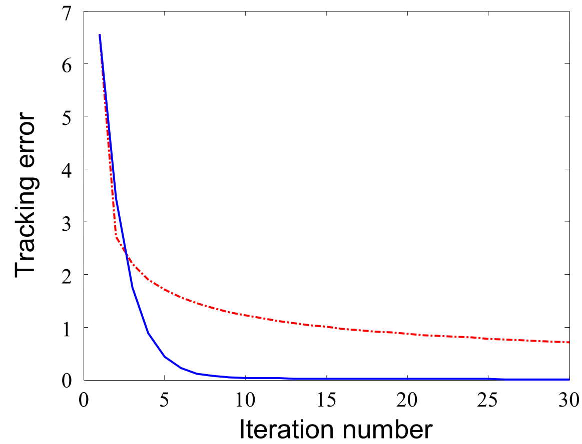

FIGURE 2. Comparison of tracking error tendency.

Fig.2 displays a comparison of the tracking errors tendencies of the proposed OILC (5) with the iteration-wise tuning factor $w_k = \lambda_{\max}\left(\mathbf{E}_k\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{E}_k\right) + 10^{-6}$ to that of with the fixed tuning factor $w_k = 0.005$. It is testified that the value $w_k = 0.005$ does not satisfy convergence condition of Theorem 1 for some iterations but satisfies convergence condition in reference [34]. In Fig.2, the solid curve is the tracking error for the former case whilst the dash-dotted one is for the latter case, respectively. The tracking error is measured in the form of the 2-norm as $\|\mathbf{e}_k\| = \sqrt{\sum_{n=1}^{50}(e_k(n))^2}$. From Fig.2, it is noticed that the tracking error of the proposed OILC (5) with the iteration-wise tuning factor converges much faster than that of the fixed tuning factor.

Example 2: Consider an SISO LDTI system as follows.

$$\begin{cases} x_{k+1}(n+1) = \begin{bmatrix} -0.1 & -0.1 \\ 0.1 & 0.78 \end{bmatrix} x_{k+1}(n) \\ \qquad + \begin{bmatrix} 0 \\ 0.8 \end{bmatrix} u_{k+1}(n), \\ y_{k+1}(n+1) = \begin{bmatrix} 0.5 & 0 \end{bmatrix} x_{k+1}(n+1); \quad n \in S, \end{cases} \tag{49}$$

where the operation time interval is set as $S = \{0, 1, \cdots, 99\}$, the initial state is set as $[x_1(0), x_2(0)]^{\mathrm{T}} = \mathbf{0}$. It is testified that $\mathbf{CB} = 0$ and $\mathbf{CAB} = -0.4 \neq 0$, which means that the relative degree of system (48) is $r = 2$. For this, the desired trajectory is chosen as $y_d(1) = 0$ and $y_d(n+1) = 0.5e^{(\frac{n}{100})}\sin(\frac{6n}{50}), n \in S\backslash\{0\}$, respectively. It is testified that the symmetric matrix $\frac{1}{2}\left(\mathbf{H}+\mathbf{H}^{\mathrm{T}}\right)$ exists negative eigenvalues, which implies that the system (49) does not satisfy the convergence assumption addressed in reference [34].

For comparison, the optimized second-order ILC (41) is adopted. The iteration-variable tuning factor is selected as $\tilde{w}_k = \lambda_{\max}(\Phi_k^{N-r}(\mathbf{H}^{N-r})^T\mathbf{H}^{N-r}\Phi_k^{N-r}) + 10^{-6}$ and the iteration-invariant tuning factor is fixed at $\tilde{w}_k = 0.01$ that fits to the circumstance $w > 0$ discussed in reference [34] but does not satisfy convergence condition of Theorem 2 for some iterations. The tracking behavior is depicted in Fig.3, where the solid curve plots the desired trajectory, the dash-dotted

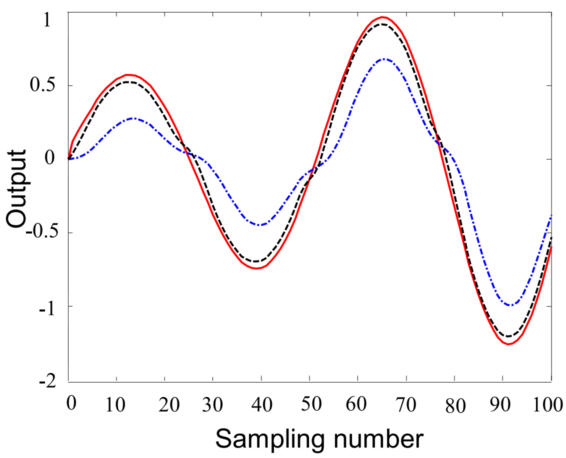

FIGURE 3. Tracking behavior.

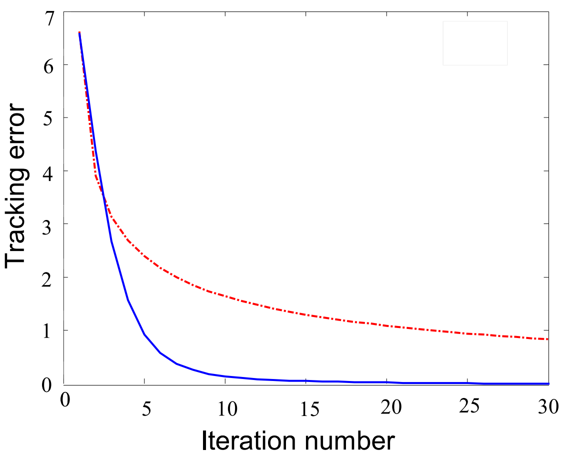

FIGURE 4. Comparison of tracking error tendency.

and the dash ones denote the outputs at the third and sixth iterations, respectively. It is found that the OILC (41) possesses perfect learning performance.

Fig.4 presents comparative tracking error tendencies of the proposed OILC (41) with the selected iteration-variable and iteration-invariant tuning factors, respectively, where the solid curve remarks the tracking error for the former case whilst the dash-dotted one is for the latter case, respectively. It is observed that the tracking error produced by the OILC (41) with the iteration-variable tuning factor possesses faster convergence rate whist the tracking error made by the OILC (41) with the fixed tuning factor renders weaker convergence.

VI. CONCLUSION

For linear discrete-time-invariant systems operated over finite time length with relative degree being unity and larger, respectively, two optimization criterions are respectively investigated for the first-order and the higher-order D-type iterative learning controls whose iteration-time-variable learning-gain vectors are argued by minimizing the sum of the tracking error energy and the learning effort intensity proportioned by an iteration-variable tuning factor. By solving the minimization problems, it is noticed that the learning-gain matrices are strongly relevant to the tracking error. This means that the addressed ILCs are nonlinear. This is the key distinction from the existing NOILC and POILC. Aiming at the explicit expressions of the achieved iteration-time-variable learning-gain vectors, the strictly monotonous convergences of the OILCs are arrived for the cases when none of the components and part of the components of the current tracking error is null, respectively. In particular, the convergence derivations are quite complicated for the case when part of the components is null. It needs to utilize elementary transformations of a matrix to exploit the relationship between the norm of the tracking error and the learning gain matrices. From the derivations of Theorems 1 and 2, it is acquired that when the order of the optimized ILC is coincident to the relative degree of the system, the strictly monotonous convergence is ensured without any additional requirement to the system Markov parameter except the range of the tuning factor. Besides, the convergence rate is adjustable while scaling the tuning factor. Nevertheless, it should point out that the solutions of the iteration-time-variable learning-gain vectors and the monotonicity analysis are based on the precision of the system Markov parameters and the reset of the system initial state. For practical executions, the proposed OILCs must be robust to the parameter uncertainty and the measurement perturbation. Meanwhile, it is inspiring when the order does not match the relative degree. In addition, it would become complicated when the system dynamics is time-variable. The issues will be involved in future works.

REFERENCES

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

[2] C.-J. Chien and J.-S. Liu, "A P-type iterative learning controller for robust output tracking of nonlinear time-varying systems," *Int. J. Control*, vol. 64, no. 2, pp. 319–334, 1996.

[3] H.-S. Lee and Z. Bien, "Study on robustness of iterative learning control with non-zero initial error," *Int. J. Control*, vol. 64, no. 3, pp. 345–359, 1996.

[4] K. H. Park, "A study on the robustness of a PID-type iterative learning controller against initial state error," *Int. J. Syst. Sci.*, vol. 30, no. 1, pp. 49–59, Jan. 1999.

[5] S. S. Saab, "A stochastic iterative learning control algorithm with application to an induction motor," *Int. J. Control*, vol. 77, no. 2, pp. 144–163, Jan. 2004.

[6] Y. Q. Chen and K. L. Moore, "A practical iterative learning path-following control of an omni-directional vehicle," *Asian J. Control*, vol. 4, no. 1, pp. 90–98, Mar. 2002.

[7] C. Mi, H. Lin, and Y. Zhang, "Iterative learning control of antilock braking of electric and hybrid vehicles," *IEEE Trans. Veh. Technol.*, vol. 54, no. 2, pp. 486–494, Mar. 2005.

[8] Y. Wang, F. Gao, and F. J. Doyle, "Survey on iterative learning control, repetitive control, and run-to-run control," *J. Process Control*, vol. 19, no. 10, pp. 1589–1600, Dec. 2009.

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

[10] D. Shen and Y. Wang, "Survey on stochastic iterative learning control," *J. Process Control*, vol. 24, no. 12, pp. 64–77, Dec. 2014.

[11] M. Togai and O. Yamano, "Analysis and design of an optimal learning control scheme for industrial robots: A discrete system approach," in *Proc. 24th IEEE Conf. Decis. Control*, Fort Lauderdale, FL, USA, Dec. 1985, pp. 1399–1404.

[12] K. Furuta and M. Yamakita, "The design of a learning control system for multivariable systems," in *Proc. IEEE Int. Symp. Intell. Control*, Philadelphia, PA, USA, Jan. 1987, pp. 371–376.
[13] 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.
[14] J. H. Lee, K. S. Lee, and W. C. Kim, "Model-based iterative learning control with a quadratic criterion for time-varying linear systems," *Automatica*, vol. 36, no. 5, pp. 641–657, May 2000.
[15] 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.
[16] D. H. Owens, *Iterative Learning Control: An Optimization Paradigm*. London, U.K.: Springer-Verlag, 2016.
[17] N. Amann, D. H. Owens, and E. Rogers, "Predictive optimal iterative learning control," *Int. J. Control*, vol. 69, no. 2, pp. 203–226, 1998.
[18] H. Q. 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.
[19] B. Chu and D. H. Owens, "Accelerated norm-optimal iterative learning control algorithms using successive projection," *Int. J. Control*, vol. 82, no. 8, pp. 1469–1484, 2009.
[20] Y. Gua and S. Mishra, "Constrained optimal iterative learning control with mixed-norm cost functions," *Mechatronics*, vol. 43, no. 1, pp. 56–65, May 2017.
[21] 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, Feb. 2014.
[22] R. Chi, Z. Hou, S. Jin, and B. Huang, "Computationally-light non-lifted data-driven norm-optimal iterative learning control," *Asian J. Control*, vol. 20, no. 1, pp. 115–124, Jan. 2018.
[23] D. H. Owens and K. Feng, "Parameter optimization in iterative learning control," *Int. J. Control*, vol. 76, no. 11, pp. 1059–1069, Jul. 2003.
[24] S. Gunnarsson and M. Norrlöf, "On the design of ILC algorithms using optimization," *Automatica*, vol. 37, no. 12, pp. 2011–2016, Dec. 2001.
[25] D. H. Owens, B. Chu, and M. Songjun, "Parameter-optimal iterative learning control using polynomial representations of the inverse plant," *Int. J. Control*, vol. 85, no. 5, pp. 533–544, May 2012.
[26] X. Jin and J.-X. Xu, "A barrier composite energy function approach for robot manipulators under alignment condition with position constraints," *Int. J. Robust Nonlinear Control*, vol. 24, no. 17, pp. 2840–2851, Nov. 2014.
[27] R.-H. Chi, Z.-S. Hou, B. Huang, and S.-T. Jin, "A unified data-driven design framework of optimality-based generalized iterative learning control," *Comput. Chem. Eng.*, vol. 77, pp. 10–23, Jun. 2015.
[28] H.-J. Li, X.-H. Hao, and W.-T. Xu, "A fast parameter optimal iterative learning control algorithm," in *Proc. Int. Conf. Embedded Softw. Syst. Symp.*, Chengdu, SC, China, Jul. 2008, pp. 375–379.
[29] D. H. Owens, J. J. Hatonen, and S. Daley, "Robust monotone gradient-based discrete-time iterative learning control," *Int. J. Robust Nonlinear Control*, vol. 19, no. 6, pp. 634–661, Apr. 2009.
[30] J. J. Hätönen, K. Feng, and D. H. Owens, "New connections between positivity and parameter-optimal iterative learning control," in *Proc. IEEE Int. Symp. Intell. Control*, Houston, TX, USA, Oct. 2003, pp. 69–74.
[31] D. H. Owens, C. T. Freeman, and B. Chu, "Norm optimal iterative learning control with auxiliary optimization: A switching approach," in *Proc. 11th IFAC Int. Workshop Adaptation Learn. Control Signal Process.*, Caen, France, 2013, pp. 140–145.
[32] J. J. Hätönen, D. H. Owens, and K. Feng, "Basis functions and parameter optimisation in high-order iterative learning control," *Automatica*, vol. 42, no. 2, pp. 287–294, Feb. 2006.
[33] T. J. Harte, J. Hatonen, and D. H. Owens, "Discrete-time inverse model-based iterative learning control: Stability, monotonicity and robustness," *Int. J. Control*, vol. 78, no. 8, pp. 577–586, May 2005.
[34] J. Hätönen, D. H. Owens, and K. L. Moore, "An algebraic approach to iterative learning control," *Int. J. Control*, vol. 77, no. 1, pp. 45–54, Jan. 2004.
[35] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge Univ. Press, 2012.
[36] I. Alberto, *Nonlinear Control Systems*. Berlin, Germany: Springer-Verlag, 1989.

YAN LIU received the B.S. degree in mechanical engineering from Ningxia University, China, in 2000, and the M.S. degree in electronic engineering from Xi'an Jiaotong University, China, in 2012, where she is currently pursuing the Ph.D. degree. Her research interests include iterative learning control and optimization.

XIAOE RUAN received the B.S. and M.S. degrees in mathematics from Shaanxi Normal University, China, in 1988 and 1995, respectively, and the Ph.D. degree in control science and engineering from the Institute of Systems Science, Xi'an Jiaotong University, China, in 2002, where she has been with the Department of Applied Mathematics, School of Mathematics and Statistics, since 1995. From 2003 to 2004, she was a Postdoctoral Researcher with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology. From 2009 to 2010, she was a Visiting Scholar with the Department of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology, South Korea. Her current research interests include steady state hierarchical optimization of large-scale industrial processes, iterative learning control, and optimal control.

XIAOHUI LI received the B.S. degree in information and computing science from Northwest A&F University, China, in 2005, and the M.S. degree in probability and mathematical statistics from Shanghai Normal University, in 2011. He is currently pursuing the Ph.D. degree with Xi'an Jiaotong University, China. His research interests include iterative learning control and optimal control.