Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

On Higher-Order Iterative Learning Control Algorithm in Presence of Measurement Noise

Samer S. Saab, Senior Member, *IEEE*

***Abstract*– Higher-Order Iterative Learning Control (HO-ILC) algorithms use past system control information from more than one past iterative cycle. This class of ILC algorithms have been proposed aiming at improving the learning efficiency and performance. This paper addresses the optimality of HO-ILC in the sense of minimizing the control error covariance matrix in the presence of measurement noise. It is shown that the optimal weighting matrices corresponding to the control information associated with more than one cycle preceding the current cycle are zero. Consequently, an optimal HO-ILC is automatically reduced to an optimal first-order ILC. The system under consideration is a linear discrete-time varying systems with different relative degree between the input and each output. Furthermore, a suboptimal second-order ILC is proposed for a class of nonlinear systems. Based on a numerical example, it is shown that a compatible suboptimal first-order ILC yields better performance than the proposed suboptimal second-order ILC algorithm.**

I. INTRODUCTION

In the past two decades Iterative Learning Control (ILC) has attracted considerable attention in many areas and applications. ILC is an approach aimed to improve the desired dynamic behaviors of systems that operate repetitively over a fixed time interval. It is useful for problems whose system must follow different types of inputs in the face of modeling uncertainty. The high accuracy output tracking potential of ILC algorithms makes them attractive despite the run-to-run implementations problems inherited to this approach. ILC incorporates past control information such as tracking errors and their corresponding control input signals into the construction of the present control action.

ILC algorithms may be categorized in terms of their order. Traditional ILC algorithms (e.g., [1]) are classified as first-order ILC scheme, which only employs control information of the previous cycle. The n^{th}-order updating law uses the data of the n previous cycles including control inputs and their corresponding output errors. The class of ILC corresponding to $n > 1$ are known as higher-order ILC (HO-ILC) algorithms. Many researchers have considered HO-ILC algorithms; e.g., [2]-[10]. Most of the proposed HO-ILC algorithms are shown to be robust and, in absence of measurement errors, drive the output error to zero for different classes of systems. It is also interesting to note that at the 15^{th} IFAC World Congress on Automation and Control a special session was devoted to HO-ILC; e.g., [7]-[9]. The basic incentive for using HO-ILC is to improve the control performance by using more of the past control information. However, there are $2 \times n$ gain or weighting matrices associated with a traditional n^{th}-order P-type ILC. Half of these gains multiply the previous control inputs and the other half multiplies the corresponding output errors. Appropriate selection of these gain matrices can be troublesome. In [6] lower order ILC algorithms are shown to outperform (speed of convergence) higher-order ILC in the sense of time weighted norm. However, the results in [6] are only applicable to SISO LTI systems in absence of measurement errors. Based on numerical examples, many have demonstrated that the speed of tracking convergence of a second-order ILC outperforms a first order ILC; e.g., [2], [3], [5]. In [10], it is shown that HO-ILC could be used to reduce the variance of the effect of measurement noise.

The problem addressed in this paper is based on the following question: consider a first-order ILC and assume that the present control action exploits "optimally" the past control information in every iterative cycle (in presence of measurement errors) could a HO-ILC algorithm further contribute to the optimality measure under consideration? The optimality measure considered in this paper is in the sense of minimizing the control error covariance matrix. The system under consideration is a MIMO linear discrete-time varying systems in presence of measurement noise with different relative degree between the input and each output. The problem is tackled by first considering a second-order ILC whereby the gains associated with the update law are established by minimizing the control error covariance matrix. It is shown that the gains corresponding to the control information associated with two cycles preceding the current cycle are zero. That is, the optimal second-order ILC is automatically reduced to an optimal first-order ILC. The generalization of HO-ILC is then readily deduced.

Since the optimal ILC algorithm requires knowledge of system dynamics and disturbance statistics, a suboptimal second-order ILC is proposed for a class of nonlinear systems. Based on a numerical example, it is shown that a compatible suboptimal first-order ILC yields better performance than the proposed suboptimal second-order ILC.

This paper is organized as follows. Section II formulates the problem addressed in this paper. The optimality of HO-ILC is presented in Section III. The proposed suboptimal second-order ILC is presented in Section IV. A numerical example, illustrating the performance of the proposed algorithm, is included in Section V.

II. PROBLEM FORMULATION

The system considered (similar to [11]-[14]) is a discrete-time-varying linear system described by the following dif-

This work was supported by the Lebanese American University.

Samer S. Saab is with the Faculty of Electrical and Computer Engineering, Lebanese American University, P.O. Box 36, Byblos, Lebanon ssaab@lau.edu.lb

ference equation

$$\begin{aligned} x(t+1,k) &= A(t)x(t,k)+B(t)u(t,k)+w(t,k) \\ y(t,k) &= C(t)x(t,k)+v(t,k) \end{aligned} \quad (1)$$

where $t \in [0,n_t]$, $x(t,k) \in \Re^n, u(t,k) \in \Re^r, w(t,k) \in \Re^n, y(t,k) \in \Re^m$, and $v(t,k) \in \Re^m$.

(A1) Relative Degree Assumptions: Assume that for all $t \in [0,n_t]$, and $\mu_q > 1$

$$C_q(t+\mu_q)\left[\prod_{j=1}^{\mu_q-1-i} A(t+\mu_q-j)\right] B(t+i)=0, \quad (2)$$

with $1 \leq i \leq \mu_q-1$, and $1 \leq q \leq m$, and where $\prod_{j=1}^{0}(.) = I$, and the $m \times r$ matrix

$$G(t)=\left[\begin{array}{c} C_1(t+\mu_1)\left[\prod_{i=1}^{\mu_1-1} A(t+\mu_1-i)\right] B(t) \\ \vdots \\ C_m(t+\mu_m)\left[\prod_{i=1}^{\mu_m-1} A(t+\mu_m-i)\right] B(t) \end{array}\right]$$

is either of full column rank (requires $m \geq r$) or of full row rank (requires $r \geq m$). Note that if $\mu_q = 1$, then $C_q(t+1)B(t) \neq 0$.

(A2) Realizable Trajectory Assumption: It is assumed that $y_d(t) = \left[\begin{array}{ccc} y_{1,d}(t) & \cdots & y_{m,d}(t) \end{array}\right]^T$ is a realizable desired output trajectory. That is, for any realizable output trajectory and an appropriate initial condition $x_d(0)$, there exists a unique control input $u_d(t) \in \Re^r$ generating the trajectory for the nominal plant. That is, the following difference equation is satisfied

$$\begin{aligned} x_d(t+1) &= A(t)x_d(t)+B(t)u_d(t) \\ y_d(t) &= C(t)x_d(t) \end{aligned} \quad (3)$$

Define the state and the input error vectors as $\delta x(t,k) \triangleq x_d(t)-x(t,k)$, and $\delta u(t,k) \triangleq u_d(t)-u(t,k)$, respectively. Denote E, and E_k to be the expectation operators with respect to time domain, and iteration domain, respectively.

(A3) Statistical Assumptions: It is assumed that the state disturbance $w(t,k)$, and the measurement noise $v(t,k)$ are modeled as zero-mean white noise and statistically independent. Furthermore, $E(w(t,k)w(t,k)^T) = Q_t$ is positive semi-definite matrix, $E(v(t+\mu,k)v(t+\mu,k)^T) = R_{t+\mu}$ is positive definite matrix for all k, The initial state error $\delta x(0,k)$ is a zero-mean statistically independent random variable, and without loss of generality, the initial input $u(t,0) = 0$, $t \in [0,n_t]$.

The learning update law under consideration is given by

$$\begin{aligned} u(t,k+1) = \; & L^{[0]}(t,k)u(t,k)+L^{[1]}(t,k)u(t,k-1) \\ & +K^{[0]}(t,k)e(t+\mu,k) \\ & +K^{[1]}(t,k)e(t+\mu,k-1) \end{aligned} \quad (4)$$

where $e(t+\mu,k) \triangleq [e_1(t+\mu_1,k),..., \; e_m(t+\mu_m,k)]^T$, $L^{[0,1]}(t,k)$ is the $(r\times r)$ input gain matrix, $K^{[0,1]}(t,k)$ is the $(r\times m)$ learning control gain matrix, with $L^{[0]}(t,0) = I$, $L^{[1]}(t,0) = 0$, and $K^{[1]}(t,0) = 0$, $t \in [0,n_t]$. Furthermore, $e_q(t,k) \triangleq y_{q,d}(t)-y_q(t,k)$, for $1 \leq q \leq m$, is q^{th} the output *measurement* error due to the control action $u(t,k)$; that is, $e_q(t,k) = y_{q,d}(t)-y_q(t,k)$.

Problem Statement: Let system (1) satisfy Assumptions (A2)-(A3), and updating law (4) be applied. The problem addressed in this paper consists of finding the optimal values of $L^{[0]}(t,k)$, $L^{[1]}(t,k)$, $K^{[0]}(t,k)$, and $K^{[1]}(t,k)$ such that the trace of the input error covariance matrix is minimized. The results are presented in Section III. Furthermore, another issue addressed in this paper is the development of an algorithm for finding the "suboptimal" values of $L^{[0]}(t,k)$, $L^{[1]}(t,k)$, $K^{[0]}(t,k)$, and $K^{[1]}(t,k)$ for a class of nonlinear systems (21). The suboptimal algorithm is presented in Section IV.

In what follows, we denote the matrix norm $\|M\| = \|M\|_2$ (l-2 norm).

III. Optimal Second-Order ILC

In this section, we develop the optimal values of $L^{[0]}(t,k)$, $L^{[1]}(t,k)$, $K^{[0]}(t,k)$, and $K^{[1]}(t,k)$, which minimize the trace of the input error covariance matrix.

In what follows, we derive an expression for the input error covariance matrix. Iterating in the time domain the state variable of (1) from t to $t+\mu_c$, with μ_c being ant positive integer, we get

$$x(t+\mu_c,k) = \widehat{A}_{\mu_c}^{\mu_c} x(t,k) + \sum_{i=0}^{\mu_c-1} \widehat{A}_{\mu_c}^{\mu_c-1-i}\left[B(t+i)\,u(t+i,k)+w(t+i,k)\right]$$

where $\widehat{A}_f^g \triangleq \prod_{j=1}^{g} A(t+f-j)$ with $\prod_{j=k}^{k-1}(.) \triangleq I$. It follows that the q^{th} output for $1 \leq q \leq m$ can be expressed as

$$y_q(t+\mu_q,k) = C_q\widehat{A}_{\mu_q}^{\mu_q} x(t,k) + C_q\widehat{A}_{\mu_q}^{\mu_q-1} B(t)\,u(t,k) + C_q\sum_{i=0}^{\mu_q-1}\widehat{A}_{\mu_q}^{\mu_q-1-i} w(t+i,k) + v(t+\mu_c,k)$$

where $C_q \triangleq C_q(t+\mu_q)$. Similarly, an expression for $y_{q,d}(t+\mu_q)$ can be derived. Consequently, the q^{th} output measurement error is given by

$$e_q = C_q\widehat{A}_{\mu_q}^{\mu_q}\delta x(t,k) - v_q(t+\mu_q,k) + C_q\widehat{A}_{\mu_q}^{\mu_q-1} B(t)\,\delta u(t,k) - C_q\sum_{i=0}^{\mu_q-1}\widehat{A}_{\mu_q}^{\mu_q-1-i} w(t+i,k)$$

$e_q \triangleq e_q(t+\mu_q,k)$. It follows that the input error associated with (4)

$$\begin{aligned} \delta u(t,k+1) = \; & L^{[0]}(t,k)\delta u(t,k) - L^{[0]}(t,k)u_d(t) \\ & -K^{[0]}(t,k)\left[G(t)\delta u(t,k)+V_k\right] + u_d(t) \\ & +L^{[1]}(t,k)\delta u(t,k-1) - L^{[1]}(t,k)u_d(t) \\ & -K^{[1]}(t,k)\left[G(t)\delta u(t,k-1)+V_{k-1}\right] \end{aligned}$$

$$V_k \triangleq \left[F(t)\delta x(t,k) + \sum_{i=0}^{\overline{\mu}-1} W_i(t)w(t+i,k) + v(t+\mu,k)\right]$$

and the rows of $W_i(t) \triangleq [\{C_j\widehat{A}_{\mu_j}^{\mu_j-1-i}\}$, $F(t) \triangleq \{C_j\widehat{A}_{\mu_j}^{\mu_j}\}$, and $G(t) = \{C_j\widehat{A}_{\mu_j}^{\mu_j-1}\}$, $1 \leq j \leq m$. $v(t+\mu,k) \triangleq [v_1(t+$

$\mu_1, k), ..., v_m(t + \mu_m, k)]^T$ and with the condition that if $\mu_q < \overline{\mu} \triangleq \max_{1\leq q\leq m}(\mu_q)$, then the q^{th} row of $W_i(t)$, for $\mu_q < i \leq \overline{\mu}$, is zero.

$$\begin{aligned}\delta u(t,k+1) = &\left[L^{[0]}(t,k) - K^{[0]}(t,k)G(t)\right]\delta u(t,k)\\ &+\left[L^{[1]}(t,k) - K^{[1]}(t,k)G(t)\right]\delta u(t,k-1)\\ &+\left[I - L^{[0]}(t,k) - L^{[1]}(t,k)\right]u_d(t)\\ &-K^{[0]}(t,k)V_k - K^{[1]}(t,k)V_{k-1}\end{aligned}$$

Let the input error and the state error covariance matrices be defined as $P_{u(t,k)} = E[\delta u(t,k)\delta u(t,k)^T]$, and $P_{x(t,k)} = E[\delta x(t,k)\delta x(t,k)^T]$, respectively.

Notations. For compactness, the argument t is dropped as follows: $L_k^{[0,1]} \triangleq L^{[0,1]}(t,k)$, $K_k^{[0.1]} \triangleq K^{[0,1]}(t,k)$, $F \triangleq F(t)$, $G \triangleq G(t)$, $W_i \triangleq W_i(t)$, $\Gamma_k^{[0,1]} \triangleq \left[L^{[0,1]}(t,k) - K^{[0,1]}(t,k)G(t)\right]$, $\Lambda_k \triangleq \left[I - L^{[0]}(t,k) - L^{[1]}(t,k)\right]$, $P_{u_d} \triangleq E[u_d(t)u_d(t)^T]$, $P_{k,d} \triangleq E[\delta u(t,k)u_d(t)^T]$, $P_{k,k-1} \triangleq E[\delta u(t,k)\delta u(t,k-1)^T]$, $S_{k,k-1} \triangleq E[V_kV_{k-1}^T] = FE[\delta x(t,k)\delta x(t,k-1)^T]F^T$, and $S_k \triangleq E[V_kV_k^T] = FE[\delta x(t,k)\delta x(t,k)^T]F^T + \sum_{i=0}^{\overline{\mu}-1} W_iQ_iW_i^T + R_{t+\mu}$.

Thus, $\delta u(t,k+1) = \left[L_k^{[0]} - K_k^{[0]}G\right]\delta u(t,k) + \left[L_k^{[1]} - K_k^{[1]}G\right]\delta u(t,k-1) + \left[I - L_k^{[0]} - L_k^{[1]}\right]u_d(t) - K_k^{[0]}V_k - K_k^{[1]}V_{k-1}$. It follows that

$$\begin{aligned}P_{k+1,d} = &\left[L_k^{[0]} - K_k^{[0]}G\right]P_{k,d} + \left[L_k^{[1]} - K_k^{[1]}G\right]P_{k-1,d}\\ &+\left[I - L_k^{[0]} - L_k^{[1]}\right]P_{u_d}\end{aligned} \tag{5}$$

$$\begin{aligned}P_{u(t,k+1)} \triangleq\ & E[\delta u(t,k+1)\delta u(t,k+1)^T]\\ =\ &\left(L_k^{[0]} - K_k^{[0]}G\right)P_{u(t,k)}\left(L_k^{[0]} - K_k^{[0]}G\right)^T\\ &+\left(L_k^{[1]} - K_k^{[1]}G\right)P_{u(t,k-1)}\left(L_k^{[1]} - K_k^{[1]}G\right)^T\\ &+\left[I - L_k^{[0]} - L_k^{[1]}\right]P_{u_d}\left[I - L_k^{[0]} - L_k^{[1]}\right]^T\\ &+K_k^{[0]}S_k\left(K_k^{[0]}\right)^T + K_k^{[1]}S_{k-1}\left(K_k^{[1]}\right)^T\\ &+\left(L_k^{[0]} - K_k^{[0]}G\right)P_{k,k-1}\left(L_k^{[1]} - K_k^{[1]}G\right)^T\\ &+\left(L_k^{[1]} - K_k^{[1]}G\right)P_{k,k-1}^T\left(L_k^{[0]} - K_k^{[0]}G\right)^T\\ &+\left(L_k^{[0]} - K_k^{[0]}G\right)P_{k,d}\left(I - L_k^{[0]} - L_k^{[1]}\right)^T\\ &+\left[I - L_k^{[0]} - L_k^{[1]}\right]P_{k,d}^T\left(L_k^{[0]} - K_k^{[0]}G\right)^T\\ &+\left(L_k^{[1]} - K_k^{[1]}G\right)P_{k-1,d}\left(I - L_k^{[0]} - L_k^{[1]}\right)^T\\ &+\left[I - L_k^{[0]} - L_k^{[1]}\right]P_{k-1,d}^T\left(L_k^{[1]} - K_k^{[1]}G\right)^T\\ &+K_k^{[0]}S_{k,k-1}\left(K_k^{[1]}\right)^T + K_k^{[1]}S_{k,k-1}^T\left(K_k^{[0]}\right)^T\end{aligned}$$

Expanding and rearranging the "square" terms, we obtain

$$\begin{aligned}P_{u(t,k+1)} = &P_{u_d} + L_k^{[0]}\Omega_k\left(L_k^{[0]}\right)^T + L_k^{[1]}\Omega_{k-1}\left(L_k^{[1]}\right)^T\\ &+K_k^{[0]}\left[GP_{u(t,k)}G^T + S_k\right]\left(K_k^{[0]}\right)^T\\ &+K_k^{[1]}\left[GP_{u(t,k-1)}G^T + S_{k-1}\right]\left(K_k^{[1]}\right)^T\\ &+\Upsilon_k + \Upsilon_k^T\end{aligned} \tag{6}$$

where $\Omega_k \triangleq P_{u(t,k)} + P_{u_d} - P_{k,d} - P_{k,d}^T$, and

$$\begin{aligned}\Upsilon_k \triangleq &-L_k^{[0]}P_{u(t,k)}(K_k^{[0]}G)^T - L_k^{[1]}P_{u(t,k-1)}(K_k^{[1]}G)^T\\ &-L_k^{[0]}P_{u_d} - L_k^{[1]}P_{u_d} + L_k^{[0]}P_{u_d}(L_k^{[1]})^T\\ &+L_k^{[0]}P_{k,k-1}(L_k^{[1]})^T - L_k^{[0]}P_{k,k-1}(K_k^{[1]}G)^T\\ &-L_k^{[1]}P_{k,k-1}^T(K_k^{[0]}G)^T + K_k^{[0]}GP_{k,k-1}(K_k^{[1]}G)^T\\ &+L_k^{[0]}P_{k,d} - L_k^{[0]}P_{k,d}(L_k^{[1]})^T + K_k^{[0]}GP_{k,d}(L_k^{[0]})^T\\ &+K_k^{[0]}GP_{k,d}(L_k^{[1]})^T - P_{k,d}^T(K_k^{[0]}G)^T + L_k^{[1]}P_{k-1,d}\\ &-L_k^{[1]}P_{k-1,d}(L_k^{[0]})^T + K_k^{[1]}GP_{k-1,d}(L_k^{[1]})^T\\ &+K_k^{[1]}GP_{k-1,d}(L_k^{[0]})^T - P_{k-1,d}^T(K_k^{[1]}G)^T\\ &+K_k^{[0]}S_{k,k-1}(K_k^{[1]})^T\end{aligned}$$

Theorem 1. Let system (1) satisfy Assumptions (A2) and (A3) and the updating law presented by (4) be applied. The optimal $L^{[0]}(t,k)$, $L^{[1]}(t,k)$, $K^{[0]}(t,k)$, and $K^{[1]}(t,k)$, which minimize the trace of the input error covariance matrix for all k and $t \in [0, n_t]$, are given by

$$\begin{aligned}L^{[1]}(t,k) &= 0,\ K^{[1]}(t,k) = 0,\ L^{[0]}(t,k) = I\\ K_k^{[0]} &= P_{u(t,k)}G^T\left[GP_{u(t,k)}G^T + S_k\right]^{-1}\end{aligned} \tag{7}$$

In addition, the recursive algorithm for the input error covariance matrix is given by

$$P_{u(t,k+1)} = (I - K_k^{[0]}G)P_{u(t,k)} \tag{8}$$

Proof. We differentiate the trace of $P_{u(t,k+1)}$, in (6), with respect to $L_k^{[0]}$, $L_k^{[1]}$, $K_k^{[0]}$ and $K_k^{[1]}$ and set each to zero.

$$\frac{\partial(\text{trace}(P_{u(t,k+1)}))}{\partial L_k^{[0]}} = 2L_k^{[0]}\Omega_k - 2H_k^{L^{[0]}} \equiv 0 \tag{9}$$

where $H_k^{Q^{[i]}} \triangleq \frac{1}{2}\frac{\partial(\text{trace}(\Upsilon_k + \Upsilon_k^T))}{\partial L_k^{[0]}}$, where $Q \in \{L, K\}$ and $i \in \{0, 1\}$.

$$\begin{aligned}H_k^{L^{[0]}} = &K_k^{[0]}GP_{u(t,k)} + P_{u_d} - L_k^{[1]}P_{u_d} - L_k^{[1]}P_{k,k-1}^T\\ &+K_k^{[1]}GP_{k,k-1}^T - P_{k,d}^T + L_k^{[1]}P_{k,d}^T - K_k^{[0]}GP_{k,d}\\ &+L_k^{[1]}P_{k-1,d} - K_k^{[1]}GP_{k-1,d}\end{aligned}$$

$$\frac{\partial(\text{trace}(P_{u(t,k+1)}))}{\partial L_k^{[1]}} = 2L_k^{[1]}\Omega_{k-1} - 2H_k^{L^{[1]}} \equiv 0 \tag{10}$$

$$\begin{aligned} H_k^{L^{[1]}} &= K_k^{[1]} G P_{u(t,k-1)} + P_{u_d} - L_k^{[0]} P_{u_d} - L_k^{[0]} P_{k,k-1} \\ &+ K_k^{[0]} G P_{k,k-1} - P_{k-1,d}^T + L_k^{[0]} P_{k,d} - K_k^{[0]} G P_{k,d} \\ &+ L_k^{[0]} P_{k-1,d}^T - K_k^{[1]} G P_{k-1,d} \end{aligned}$$

$$\begin{aligned} \frac{\partial(\text{trace}(P_{u(t,k+1)}))}{\partial K_k^{[0]}} &= 2K_k^{[0]} \left[G P_{u(t,k)} G^T + S_k \right] \\ -2H_k^{K^{[0]}} &\equiv 0 \end{aligned} \quad (11)$$

$$\begin{aligned} H_k^{K^{[0]}} &= L_k^{[0]} P_{u(t,k)} G^T + L_k^{[1]} P_{k,k-1}^T G^T - K_k^{[1]} G P_{k,k-1}^T G^T \\ &- L_k^{[0]} P_{k,d}^T G^T - L_k^{[1]} P_{k,d}^T G^T + P_{k,d}^T G^T - K_k^{[1]} S_{k,k-1}^T \end{aligned}$$

$$\begin{aligned} \frac{\partial(\text{trace}(P_{u(t,k+1)}))}{\partial K_k^{[1]}} &= 2K_k^{[1]} \left[G P_{u(t,k-1)} G^T + S_{k-1} \right] \\ -2H_k^{K^{[1]}} &\equiv 0 \end{aligned} \quad (12)$$

$$\begin{aligned} H_k^{K^{[1]}} &= L_k^{[1]} P_{u(t,k-1)} G^T + L_k^{[0]} P_{k,k-1} G^T \\ &- K_k^{[0]} G P_{k,k-1} G^T - L_k^{[1]} P_{k-1,d}^T G^T \\ &- L_k^{[0]} P_{k-1,d}^T G^T + P_{k-1,d}^T G^T - K_k^{[0]} S_{k,k-1} \end{aligned}$$

In the following, Equations (9)-(12) are substituted in (6).

$$\begin{aligned} P_{u(t,k+1)} &= P_{u_d} + H_k^{L^{[0]}} (L_k^{[0]})^T + H_k^{L^{[1]}} (L_k^{[1]})^T \\ &+ H_k^{K^{[0]}} (K_k^{[0]})^T + H_k^{K^{[1]}} (K_k^{[1]})^T + \Upsilon_k + \Upsilon_k^T \end{aligned}$$

After cancelling all zero terms, the above equation becomes

$$\begin{aligned} P_{u(t,k+1)} &= \left(I - L_k^{[0]} - L_k^{[1]} \right) P_{u_d} + \left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} \\ &+ \left(L_k^{[1]} - K_k^{[1]} G \right) P_{k-1,d} \end{aligned} \quad (13)$$

Comparing (13) with (5), we find $P_{k,d} = P_{u(t,k)}$. Thus, Equations (9)-(12) become

$$\begin{aligned} 0 &= L_k^{[0]} \left[P_{u_d} - P_{k,d} \right] - \left[(I - L_k^{[1]}) (P_{u_d} - P_{k,d}) \right. \\ &+ \left. \left(L_k^{[1]} - K_k^{[1]} G \right) \left(P_{k-1,d} - P_{k,k-1}^T \right) \right] \end{aligned} \quad (14)$$

$$\begin{aligned} 0 &= L_k^{[1]} \left[P_{u_d} - P_{k-1,d} \right] - \left[(I - L_k^{[0]}) (P_{u_d} - P_{k-1,d}) \right. \\ &+ \left. \left(L_k^{[0]} - K_k^{[0]} G \right) (P_{k,d} - P_{k,k-1}) \right] \end{aligned} \quad (15)$$

$$\begin{aligned} 0 &= K_k^{[0]} \left[G P_{k,d} G^T + S_k \right] - \left[(I - L_k^{[1]}) P_{k,d} G^T \right. \\ &+ \left. \left(L_k^{[1]} - K_k^{[1]} G \right) P_{k,k-1}^T G^T - K_k^{[1]} S_{k,k-1}^T \right] \end{aligned} \quad (16)$$

$$\begin{aligned} 0 &= K_k^{[1]} \left[G P_{k-1,d} G^T + S_{k-1} \right] - \left[\left(I - L_k^{[0]} \right) P_{k-1,d} G^T \right. \\ &+ \left. \left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,k-1} G^T - K_k^{[0]} S_{k,k-1} \right] \end{aligned} \quad (17)$$

Next, we show that

$$L_k^{[0]} + L_k^{[1]} = I \text{ and } L_k^{[1]} = K_k^{[1]} G \quad (18)$$

and $P_{k,d} = P_{k,k-1}$. Since $u(t,0) = 0$, then $P_{0,d} = P_{u_d}$. For $k = 0$, $L_0^{[1]} = 0$, $K_0^{[1]} = 0$, or $L_0^{[1]} = K_0^{[1]} G$ and $L_0^{[0]} = I$, $P_{1,0} = \left(L_0^{[0]} - K_0^{[0]} G \right) P_{u_d} = P_{1,d}$. By an induction argument, we assume that $L_k^{[0]} + L_k^{[1]} = I$, $L_k^{[1]} = K_k^{[1]} G$ and $P_{k,d} = P_{k,k-1}$ and show these are true for $k+1$. Since $P_{k,d} = P_{u(t,k)}$ and $P_{u(t,k)}$ is symmetric then $P_{k,k-1} = P_{k,k-1}^T$.

$$\begin{aligned} P_{k+1,k} &= \left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} + \left(L_k^{[1]} - K_k^{[1]} G \right) P_{k,k-1}^T \\ &+ \left(I - L_k^{[0]} - L_k^{[1]} \right) P_{k,d} = \left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} \end{aligned}$$

Note that (13) implies

$$P_{k+1,d} = \left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} \quad (19)$$

Consequently, $P_{k+1,k} = P_{k+1,d}$. Note that

$$\begin{aligned} 0 &= L_{k+1}^{[1]} \left[P_{u_d} - P_{k,d} \right] - \left[(I - L_{k+1}^{[0]}) (P_{u_d} - P_{k,d}) \right. \\ &+ \left. \left(L_{k+1}^{[0]} - K_{k+1}^{[0]} G \right) (P_{k+1,d} - P_{k+1,k}) \right] \end{aligned}$$

or $(I - L_{k+1}^{[0]} - L_{k+1}^{[1]}) (P_{u_d} - P_{k,d}) = 0$. Note $P_{u_d} - P_{k,d} = -E\left[u(t,k) u_d^T \right]$ is nonsingular for $k \geq 1$. Consequently, $L_{k+1}^{[0]} + L_{k+1}^{[1]} = I$. Next consider

$$\begin{aligned} 0 &= L_{k+1}^{[0]} \left[P_{u_d} - P_{k+1,d} \right] - \left[(I - L_{k+1}^{[1]}) (P_{u_d} - P_{k+1,d}) \right. \\ &+ \left. \left(L_{k+1}^{[1]} - K_{k+1}^{[1]} G \right) \left(P_{k,d} - P_{k+1,k}^T \right) \right] \end{aligned}$$

or $\left(L_{k+1}^{[1]} - K_{k+1}^{[1]} G \right) \left(P_{k,d} - P_{k+1,k}^T \right) = 0$. Similarly, since $P_{k,d} - P_{k+1,k}^T = E\left[(\delta u(t,k) u(t,k+1)^T \right]$ is nonsingular, then $L_{k+1}^{[1]} = K_{k+1}^{[1]} G$. Next, we derive a useful expression relating $K_k^{[0]}$ and $K_k^{[1]}$. Equation (16) reduces to $-\left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} G^T = -K_k^{[1]} S_{k,k-1}^T - K_k^{[0]} S_k$. Equation (17)

$$\begin{aligned} &- \left(L_k^{[1]} - K_k^{[1]} G \right) P_{k-1,d} G^T + K_k^{[1]} S_{k-1} \\ &- \left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} G^T + K_k^{[0]} S_{k,k-1} = 0 \end{aligned}$$

Making use of $-\left(L_k^{[0]} - K_k^{[0]} G \right) P_{k,d} G^T = -K_k^{[1]} S_{k,k-1}^T - K_k^{[0]} S_k$, then $-K_k^{[1]} S_{k,k-1}^T - K_k^{[0]} S_k + K_k^{[1]} S_{k-1} + K_k^{[0]} S_{k,k-1} = 0$, or

$$K_k^{[1]} \left(S_{k-1} - S_{k,k-1}^T \right) = K_k^{[0]} \left(S_k - S_{k,k-1} \right) \quad (20)$$

Finally, we show that the gain values presented in (7), that is, $L_k^{[1]} = 0$, $K_k^{[1]} = 0$, $L_k^{[0]} = I$, and $K_k^{[0]} = P_{u(t,k)} G^T \left[G P_{u(t,k)} G^T + S_k \right]^{-1}$, along with the fact that (just shown) $P_{k,d} = P_{u(t,k)} = P_{k,k-1}$, are consistent with the condition given in (18) and solve all the optimal conditions presented in Equations (14)-(17). We only collaborate the condition of (17). For $K_k^{[1]} = 0$, (20) implies that $K_k^{[0]} S_k = K_k^{[0]} S_{k,k-1}$. Consequently, (17) reduces to $\left(I - K_k^{[0]} G \right) P_{k,k-1} G^T - K_k^{[0]} S_k = 0$. Since $P_{u(t,k)} = P_{k,k-1}$, then the above equation or (17) can be written as $K_k^{[0]} \left(G P_{u(t,k)} G^T + S_k \right) = P_{u(t,k)} G^T$. The rest, that is, condition given in (18) and the optimal conditions presented in Equations (14)-(16), are readily satisfied.■

Remark 1. The optimality presented in Theorem 1 indicates that the second-order ILC depends only on the control information generated from the previous iteration. Since the information from two previous iterations is not needed, then the control information generated from all previous m iterations, with $m \geq 2$, are also not needed. In general, an optimal higher-order ILC algorithm is nothing but an optimal first-order ILC algorithm.

Theorem 2. Let system (1) satisfy Assumptions (A1)-(A3), and the updating law presented by (4), (7) and (8) be applied. Then the boundedness of all trajectories is guaranteed for all k and $t \in [0, n_t]$. Furthermore, If G is a full column rank matrix for $t \in [0, n_t]$, then there exists a positive constant $c_{\delta u}$ such that $\|\delta u(t,k)\|^2 \leq \frac{c_{\delta u}}{k}$ and $\lim_{k\to\infty} \delta u(t,k) = 0$. In addition, if $\|\delta x(0,k)\|^2 \leq \frac{c_{\delta x0}}{k}$ for all k and $w(t,k) = 0$, then there exists a positive constant c_{δ_x} such that

$$\|\delta x(t,k)\|^2 \leq \frac{c_{\delta_x}}{k} \text{ and } \lim_{k\to\infty} \delta x(t,k) = 0$$

and the output error converges to zero at a rate inversely proportional to $\sqrt{k}$.

On the other hand, if G is a full row rank matrix for all k and $t \in [0, n_t]$, $w(t,k) = 0$, and there exists a positive constant c_{x0} such that $\|x(0,k+1) - x(0,k)\| \leq \frac{\overline{c}_{x0}}{k}$, then there exist positive constants c_ψ and c_η, such that

$$\|x(t,k+1) - x(t,k)\| \leq \frac{c_\eta}{k}$$

$$\|\delta\psi(t+\mu,k)\|^2 \leq \frac{c_\psi}{k} \text{ and } \lim_{k\to\infty} \psi(t+\mu,k) = 0$$

where where the output error $\delta\psi_q(\cdot,k) \triangleq C_q(\cdot)[x_d(\cdot) - x(\cdot,k)]$ with $1 \leq q \leq m$. ■

Proof. The proof follows similar arguments that are presented in [14], thus, omitted.

Remark 2. A suboptimal recursive algorithm, presented in [14], is based on unknown system dynamics and unknown disturbance statistics. The robustness and convergence is shown to possess similar characteristics to the ones of the optimal recursive algorithm summarized in Theorem 2 of this manuscript.

IV. SUBOPTIMAL SECOND-ORDER ILC

The proposed optimal ILC algorithm requires knowledge of system dynamics and disturbance statistics. In the case of unknown system dynamics and disturbance statistics, the optimality theory cannot be achieved. In this section, a suboptimal second-order ILC is proposed for a class of nonlinear systems. We consider the class of discrete-time affine nonlinear systems described by the following difference equation

$$\begin{aligned} x(t+1,k) &= f(x(t,k)) + B(x(t,k))u(t,k) \\ y(t,k) &= g(x(t,k)) + v(t,k) \end{aligned} \tag{21}$$

The descriptions of the above system is similar to the ones considered in [5] except for $f(.)$, $B(.)$ and $g(.)$ can grow as fast as any polynomial with arbitrary order (Lipschitz condition is not required), and the $m \times r$ matrix

$$G(x) = \begin{bmatrix} \frac{\partial}{\partial u} g_1 \circ f^{\mu_1 - 1}[f(x) + B(x)u] \\ \vdots \\ \frac{\partial}{\partial u} g_m \circ f^{\mu_m - 1}[f(x) + B(x)u] \end{bmatrix}$$

is either of full column rank (requires $m \geq r$) or of full row rank (requires $m \leq r$).

The proposed suboptimal algorithm is partially motivated by the optimal theory presented in the previous section, in particular, Equations (14), (18), (19) and (20). Although, we drop the argument t for compact presentation, in what follows, the proposed gains are generally time varying. In the following a recursive algorithm is presented to update the gain matrices $\widetilde{L}_k^{[0]}$, $\widetilde{L}_k^{[1]}$, $\widetilde{K}_k^{[0]}$, and $\widetilde{K}_k^{[1]}$, in (4), which correspond to $L^{[0]}(t,k)$, $L^{[1]}(t,k)$, $K^{[0]}(t,k)$, and $K^{[1]}(t,k)$, respectively.

$$\widetilde{K}_k^{[0]} = \widetilde{L}_k^{[0]} \widetilde{P}_k (G_k \widetilde{P}_k G_k^T + \widetilde{S}_k)^{-1} \tag{22}$$

$$\widetilde{P}_{k+1} = \left(\widetilde{L}_k^{[0]} - \widetilde{K}_k^{[0]} G_k\right) \widetilde{P}_k, \; \widetilde{K}_{k+1}^{[1]} = \widetilde{K}_{k+1}^{[0]} \tag{23}$$

$$\widetilde{L}_{k+1}^{[1]} = \widetilde{K}_{k+1}^{[1]} G_k \text{ and } \widetilde{L}_{k+1}^{[0]} = I - \widetilde{L}_{k+1}^{[1]} \tag{24}$$

where $G_k \triangleq G(x(t,k))$, $\widetilde{L}_0^{[1]} = 0$, $\widetilde{K}_0^{[1]} = 0$, $\widetilde{L}_0^{[0]} = I$. Furthermore, $\widetilde{P}_0$ and $\widetilde{S}_k$ are symmetric and positive-definite matrices. The convergence characteristics of the proposed algorithm are only illustrated in the following example and compared with other ILC algorithms.

V. NUMERICAL EXAMPLE

In this example, three different ILC algorithms are applied to the same example considered in [5], where performance of the algorithms are compared. The ILC algorithms under examination are:

1) The first-order suboptimal ILC algorithm presented in [14], denoted by ILC1, where the recursive update of the learning gain is given by

$$\begin{aligned} \widetilde{K}_k &= \widetilde{P}_k \widetilde{G}_k^T \left[\widetilde{G}_k \widetilde{P}_k \widetilde{G}_k^T + \widetilde{S}_k\right]^{-1} & (25) \\ \widetilde{P}_{k+1} &= (I - \widetilde{K}_k \widetilde{G}_k) \widetilde{P}_k & (26) \end{aligned}$$

2) The proposed second-order ILC algorithm given by (4), and (22)-(24), denoted by ILC2.
3) The second-order ILC algorithm presented in [5], denoted by ILC3.

The same nonlinear discrete-time system, with relative degree four, presented in [5], is considered. The system is described by the following difference equation

$$\begin{aligned} x_1(t+1,k) &= 0.5\sin(x_2(t,k)) \\ &\quad + (1 + 0.5\cos(x_1(t,k))u(t,k) \\ x_{j+1}(t+1,k) &= x_j(t,k), \; j \in \{1,2,3\} \\ y(t,k) &= x_4(t,k) + v(t,k) \end{aligned}$$

The desired output is given by $y_d(t) = t(1 - t/96)/96$, for $0 \leq t \leq 100$. The measurement error $v(t,k)$ is zero-mean white Gaussian noise with standard deviation $\sigma_v = 0.05$. The initial state variables are given by $x_j(0,k) = y_d(4-j) + \delta x_0/k$, where δx_0 is a zero-mean white Gaussian noise with standard deviation $= 0.1$.

The learning gains, as employed in [5], are $\Phi_k(t) = 0.8 \times 0.95\left[1 + 0.5\cos(x_1(t,k))\right]^{-1}$ and $\Phi_{k-1}(t) = 0.2 \times 0.95\left[1 + 0.5\cos(x_1(t,k-1))\right]^{-1}$ and the gains that multiplies $u(t,k)$ and $u(t,k-1)$ are 0.8 and 0.2, respectively. For consistency, the parameters used for the used in (25) and (26), and for the proposed algorithm presented in (22)-(24) are: $\widetilde{G}_k(t) = \left[1 + 0.5\cos(x_1(t,k))\right]/0.95$, $\widetilde{P}_0 = 1$ for all t, and $\widetilde{S}_k = \sigma_v^2 + 1/k$. The time average of the matrices, $\widetilde{L}_k^{[0]}$, $\widetilde{L}_k^{[1]}$, $\widetilde{K}_k^{[0]}$ $(= \widetilde{K}_k^{[1]})$, and $\widetilde{P}_k$, employed in the proposed algorithm (ILC2) are shown in Figure 1. As expected $\lim_{k\to\infty}\left\|\widetilde{L}_k^{[0]}\right\| = 1$ where as $\lim_{k\to\infty}\|M_k\| = 0$, $M \in \left\{\widetilde{L}_k^{[1]}, \widetilde{K}_k^{[0]}, \widetilde{P}_k\right\}$. The performance index used in [5] is given by $J_k = \max_{4\leq t\leq 100}|y_d(t) - x_4(t,k)|$. The performance of the three algorithms is summarized in Figure 2. It can be noted that ILC1 and ILC2 outperforms ILC3. Furthermore, unlike ILC3, ILC1 and ILC2 drive the output error to zero. The maximum absolute output error, J_k, decrease two orders of magnitude in three decades of learning iterations. It is also worth noting that J_k (corresponding to ILC1 and ILC2) becomes less than the output measurement errors standard deviation for all $k \geq 10$. Although the overall performance of ILC1 and ILC2 are compatible, it can be noted that ILC1 results in smaller tracking errors for $k \leq 12$.

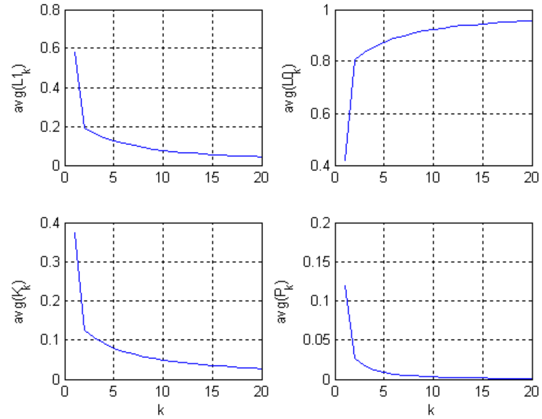

Fig. 1. Time average of the gain matrices.

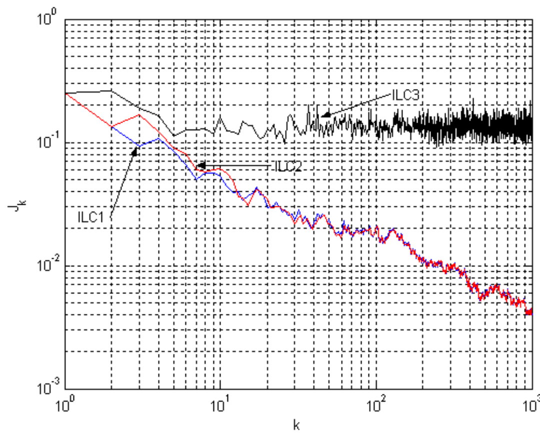

Fig. 2. Tracking performance of the three ILC algorithms in presence of random disturbances.

VI. Conclusion

This paper addressed the optimality of HO-ILC in the sense of minimizing the control error covariance matrix in the presence of measurement noise. It was shown that if the present control action exploits optimally or suboptimally the past control information in every iterative cycle, then HO-ILC algorithms are not needed for linear discrete-time varying systems with different relative degree. Furthermore, a suboptimal second-order ILC was proposed for a class of nonlinear systems. Based on a numerical example, it was shown that a compatible suboptimal first-order ILC resulted in better performance than the proposed suboptimal second-order ILC algorithm.

References

[1] Arimoto, S., Kawamura, S., and Miyazaki, F., "Bettering operation of robots by learning," *Journal of Robotics Systems*, vol. 1, 1984, pp. 123-140.

[2] Z. Bien, and K. M. Huh, "Higher-order iterative learning control algorithm," *IEE Proc. Part-D*, vol. 136, no. 3, 1989, pp. 105-112.

[3] Y. Chen, M. Sun, B. Huang, and H. Dou, "Robust higher order repetitive learning control. Algorithm for tracking control of delayed repetitive systems," *Proc. of the 31st IEEE Conference on Decision and Control*, December 1992, Tucson, AZ, pp. 2504-2510.

[4] Y. Q. Chen, C. Wen and M. Sun, "A robust higher-order P-type iterative learning control using current iteration tracking error," *Int. J. of Control*, vol. 68, no. 2, 1997. pp. 331-342.

[5] M. Sun and D. Wang, "Analysis of Nonlinear Discrete-Time Systems with Higher-Order Iterative Learning Control," *Dynamics and Control*, vol. 11, 2001, pp. 81-96.

[6] J.-X. Xu and Y. Tan, "On the convergence speed of a class of higher-order ILC," *Proc. of the 40th IEEE Conference on Decision and Control*, December 2001, Orlando, FA, pp. 4932-4937.

[7] K. L. Moore and Y. Chen, "On the monotonic convergence of higher order iterative learning update laws," *IFAC Proc. Triennial World Congress*, July 2002, Barcelona, Spain.

[8] M. Q. Phan and R. W. Longman, "Higher-order iterative learning control by pole placement and noise filtering, *IFAC Proc. Triennial World Congress*, July 2002, Barcelona, Spain.

[9] T. S. Al-Towaim, P. Lewin, and E. Rogers, "Higher order ILC versus alternative applied to chain conveyor systems," *IFAC Proc. Triennial World Congress*, July 2002, Barcelona, Spain.

[10] M. Norrlöf and S. Gunnarsson, "Disturbance aspects of high order ILC algorithm," *in the 15th IFAC Proc. Triennial World Congress*, July 2002, Barcelona, Spain.

[11] S. S. Saab, "A discrete-time stochastic learning control algorithm," *IEEE Trans. on Auto. Contr.*, vol. 46, no. 6, June, 2001, pp. 877-887.

[12] S. S. Saab, "On a discrete-time stochastic learning control algorithm," *IEEE Trans. on Auto. Contr.*, vol. 46, no. 8, August 2001, pp. 1333-1336.

[13] S. S. Saab, "Stochastic P-type/D-type iterative learning control algorithms," *Int. J. of Contr.*, vol. 76, no. 2, 2003, pp. 139-148.

[14] S. S. Saab, "A stochastic iterative learning control algorithm with application to an induction motor," *Int. J. of Contr.*, vol. 77, no. 2, 2004, pp. 144-163.