Iterative Learning Control for Unknown Nonlinear Systems Based on Data-Driven Model-Free Feedback Linearization

Boyu Wen
China University of Geosciences
Wuhan, China
wenboyu1999@cug.edu.cn

Xin Chen*
China University of Geosciences
Wuhan, China
chenxin@cug.edu.cn

Wojciech Paszke
University of Zielona Gora
Zielona Gora, Poland
w.paszke@iee.uz.zgora.pl

***Abstract*—This paper proposes a novel iterative learning control (ILC) scheme for unknown affine nonlinear systems by integrating model-free feedback linearization with two-dimensional (2D) structure of the controlled dynamics. The approach eliminates the requirement for prior model knowledge by employing model reference adaptive control (MRAC) and *Q*-learning to achieve feedback linearization of unknown nonlinear systems. A computationally efficient method is developed to estimate feedback linearization parameters using historical data from previous trials. Upon obtaining the linearized system representation, the control design is performed within the 2D system setting, resulting in a set of linear matrix inequality (LMI) constraints that leads to the ILC law. The efficacy of the proposed approach is validated through numerical experiments on an inverted pendulum system, demonstrating high-precision trajectory tracking across iterative executions.**

***Index Terms*—Iterative learning control, feedback linearization, nonlinear systems, model reference adaptive control, *Q*-learning, adaptive dynamic programming**

I. Introduction

Iterative learning control has emerged as a powerful methodology for improving the performance of systems that execute repetitive tasks over a fixed time interval. By leveraging data from previous iterations, ILC systematically adjusts control inputs to enhance tracking accuracy from cycle to cycle [1], [2]. This approach has found significant applications in various industrial processes including robotic manipulation [3], chemical batch processing [4], and precision manufacturing [5], where operations are inherently repetitive in nature.

Traditional ILC schemes predominantly focus on linear systems or nonlinear systems with repetitive uncertainties. However, many practical systems exhibit substantial nonlinearities and non-repetitive uncertainties that challenge conventional ILC approaches. When applied to nonlinear systems with unknown dynamics or time-varying disturbances, the performance of standard ILC methods may deteriorate significantly due to their inherent linear assumptions [6].

Feedback linearization represents a fundamental nonlinear control technique that transforms nonlinear system dynamics into equivalent linear forms through appropriate coordinate transformations and state feedback. This transformation enables the application of well-established linear control theories to complex nonlinear systems [7]. The conventional implementation of feedback linearization, however, heavily relies on precise knowledge of system dynamics [8], which is often unavailable in practical applications due to unmodeled dynamics, parameter variations, or completely unknown system structures.

Recent research efforts have explored the integration of ILC with linearization to address nonlinear systems. Several adaptive ILC schemes have been developed that incorporate dynamic linearization techniques to handle system nonlinearities [9]. Other approaches have integrated neural networks as compensators to mitigate the effects of non-repetitive uncertainties [10]. While these methods demonstrate improved performance, they typically require partial model knowledge or must satisfy specific conditions such as Lipschitz continuity, potentially limiting their practical applicability.

The advancement of machine learning, particularly reinforcement learning (RL), has opened new avenues for data-driven control of complex systems without explicit model knowledge. Recent works have investigated the combination of RL with ILC to compensate for uncertainties and enhance learning performance. For instance, deep reinforcement learning has been employed as a compensator within 2D ILC frameworks to achieve nearly perfect tracking [11]. Meanwhile, Koopman operator theory has been introduced to lift nonlinear systems into high-dimensional linear spaces, facilitating linear controller design [12]. Nevertheless, these approaches often involve high-dimensional models and require substantial data for training, which may be impractical for real-time applications.

Despite these advancements, most existing methods still depend on nominal models or extensive prior knowledge. To the

This research was funded in part by the Technical Innovation Major Project of Hubei Province, China under Grant No. 2020AEA010, in part by National Science Centre in Poland, grant No. 2023/48/Q/ST7/00205, in part by the National Natural Science Foundation of China under Grant No. 61873248, in part by the Natural Science Foundation of Hubei Province, China under Grant No. 2020CFA031, in part by the CSC scholarship No. 202506410001, and in part by the 111 Project under Grant No. B17040.

best of our knowledge, the problem of designing a completely model-free ILC scheme for unknown affine nonlinear systems remains largely unexplored. This paper aims to bridge this gap by proposing a novel ILC strategy that integrates model-free feedback linearization with an ILC framework. The main contributions are summarized as follows:

- A model-free feedback linearization framework is developed using MRAC and Q-learning, which transforms unknown nonlinear systems into linear equivalents without requiring any prior model knowledge.
- ILC law design is performed with 2D system model and results in LMI-based conditions for controller design. And the complete control scheme is validated through numerical simulations on a discretized inverted pendulum system, demonstrating effective tracking performance under repetitive operation scenarios.

The remainder of this paper is organized as follows. Section II. formulates the problem and reviews preliminaries on feedback linearization and ILC. Section III. details the proposed model-free feedback linearization and ILC design over 2D system setting. Section IV. presents simulation results and analysis, and Section V. concludes the paper with future research directions.

II. Background and Problem Formulation

A. Iterative Learning Control

Consider the following discrete-time nonlinear time-invariant system

$$\begin{aligned} x_k(t+1) &= f(x_k(t)) + g(x_k(t))u_k(t), \quad x_k(0) = x_0, \\ y_k(t) &= h(x_k(t)), \end{aligned} \tag{1}$$

where k denotes the trial index and $t \in [0, N-1]$ denotes the time index along a trial of finite duration N. $x_k \in \mathbb{R}^m, u_k \in \mathbb{R}, y_k \in \mathbb{R}$ denote the state, control input and system output on trial k, respectively. It is assumed that at the end of each trial the state vector resets to the same vector denoted by $x_k(0)$. $f(\cdot)$, $g(\cdot)$ are unknown system functions.

For a discrete-time nonlinear system with unknown dynamics, the task in this work is to iteratively determine a control input sequence $u(t), t \in [0, N-1]$, such that the system output $y(t)$ accurately tracks a predefined reference trajectory $y_d(t)$ over the finite time horizon. Therefore, the goal is to construct a sequence of control inputs $\{u_k(t)\}$ ensuring convergence of the tracking error defined as

$$\bar{e}_k(t) = y_d(t) - y_k(t) \tag{2}$$

to 0 or else to within a specified tolerance after a finite number of trials. After each trial, the control input is updated using the information accumulated during the previous trial. The update mechanism follows the general ILC law:

$$u_k(t) = u_{k-1}(t) + \delta u_k(t), \tag{3}$$

where $\delta u_k(t)$ represents the corrective term, which will be specified subsequently.

While numerous ILC design strategies exist for linear time-invariant systems, extending these methods to unknown nonlinear discrete-time systems like (1) poses significant challenges. However, if the nonlinear input-output dynamics of system (1) can be transformed into a linear representation via a model-free feedback linearization approach, then established ILC design techniques, particularly those based on 2D system framework, can be directly applied. This will significantly reduce the design difficulty of ILC and provide a new and effective solution to the ILC design problem for unknown nonlinear systems.

B. Feedback Linearization

Assume the relative degree of system (1) is n. In order to obtain the explicit representation relationship between the system input and output, the difference Lie derivative of $y(t+n)$ is solved as follows:

$$y_k(t+n) = L_f^n h(x_k(t)) + L_g L_f^{n-1} h(x_k(t)) u_k(t), \tag{4}$$

where L is the symbol of Lie derivative. Define $v_k(t)$ as the new virtual control input. This can make the system output and the virtual control input satisfy $y_k(t+n) = v_k(t)$ by designing the following feedback linearization controller

$$u_k(t) = -\frac{L_f^n h(x_k(t))}{L_g L_f^{n-1} h(x_k(t))} + \frac{v_k(t)}{L_g L_f^{n-1} h(x(t))}. \tag{5}$$

Define $a(x_k(t)) = L_g L_f^{n-1} h(x_k(t))$ and $b(x_k(t)) = -L_f^n h(x_k(t))$, (5) can be rewritten as

$$u_k(t) = \frac{1}{a(x_k(t))} \left[b(x_k(t)) + v_k(t) \right]. \tag{6}$$

By applying the feedback linearization controller to nonlinear system (1), the nonlinear input-output relationship of (1) can be transformed into a linear relationship as follows

$$\begin{pmatrix} y_k(t) \\ y_k(t+1) \\ \vdots \\ y_k(t+n-1) \\ y_k(t+n) \end{pmatrix} = \begin{pmatrix} l_{1_k}(t) \\ l_{2_k}(t) \\ \vdots \\ l_{n-1_k}(t) \\ l_{n_k}(t) \end{pmatrix} = \begin{pmatrix} h(x_k(t)) \\ L_f h(x_k(t)) \\ \vdots \\ L_f^{n-1} h(x_k(t)) \\ L_f^n h(x_k(t)) + L_g L_f^{n-1} h(x_k(t)) u_k(t) \end{pmatrix}, \tag{7}$$

where $l_{j_k}(t), j = 1 \dots n$ represents the virtual linearization state. Furthermore, the state-space equation of the virtual linearized state can be written in the following standard form

$$\begin{aligned} l_k(t+1) &= A l_k(t) + B v_k(t) \\ y_k(t) &= C l_k(t), \end{aligned} \tag{8}$$

where

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}.$$

Based on the above derivation, if the feedback linearization controller (6) can be obtained, then the nonlinear system (1) can be corrected to the standard linear form as shown in (8).

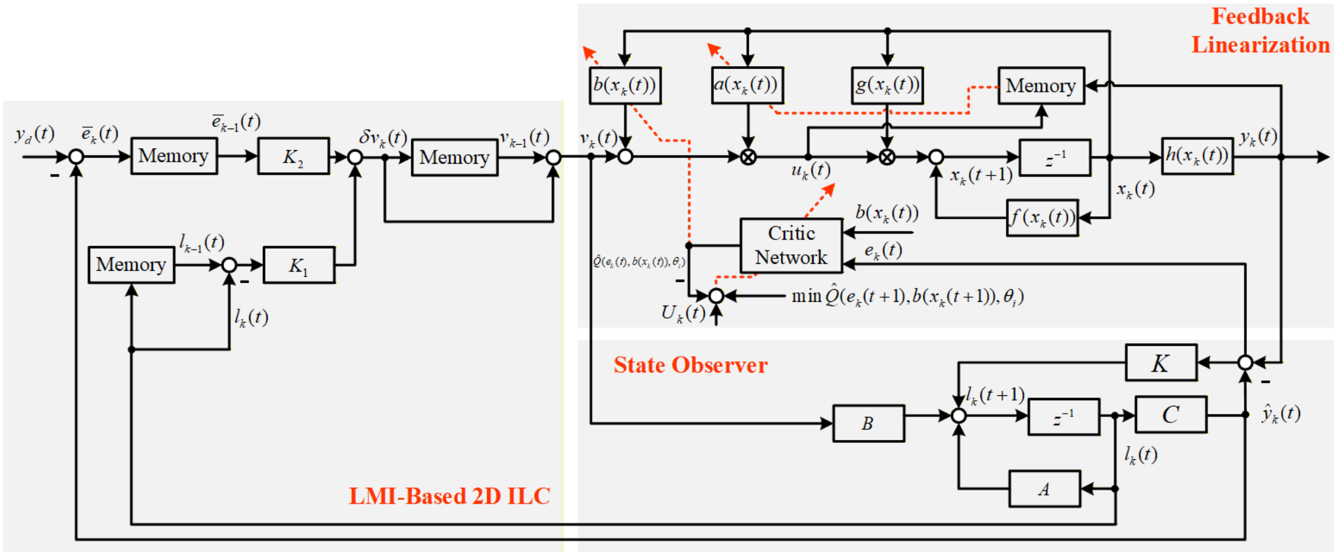

Fig. 1: Model-free feedback linearization based ILC framework diagram

However, solving the feedback linearization controller relies on the precise dynamics of the nonlinear system. For the unknown nonlinear system (1), directly solving for $a(x(t))$ and $b(x(t))$ is extremely difficult. To address the aforementioned issues, in the next section, we will construct an MRAC-based model-free feedback linearization framework and utilise data-driven reinforcement learning algorithms to solve for the feedback linearization controller.

III. Model-Free Feedback Linearization Based ILC Structure

In this section, a model-free feedback linearization framework based on MRAC and Q-learning is established to correct an unknown nonlinear system into a standard linear form. Then an LMI-based ILC design method can be used to achieve precise tracking of the reference signal by the nonlinear system.

The diagram of the model-free feedback linearization based ILC framework is shown in Fig. 1. On the right-hand side, a model-free feedback linearization framework based on MRAC is established. The feedback linearization controller is solved by Q-learning algorithm, thereby correcting the input-output dynamics of the unknown nonlinear system to a standard linear form. On the left-hand side, an ILC controller is designed based on the standard linearised system model and LMI, ultimately achieving precise tracking of the unknown nonlinear system's output $y_k(t)$ with respect to the reference signal $y_d(t)$. In the following we will describe in detail the process of solving for the feedback linearization controller and designing the ILC controller.

A. The MRAC-Based Model-Free Feedback Linearization Framework

Firstly, a model-free feedback linearization control framework based on MRAC is constructed. We select the feedback linearized standard linear system (8) as the reference model. The adjustable system consists of feedback linearization controller (6) and unknown nonlinear system (1).

In the MRAC-based feedback linearization framework, the virtual linearised states of the unknown nonlinear system cannot be directly measured. Considering that the expected dynamics of virtual states is the same as that of the standard linear system (8), therefore a state observer is constructed as follows to reconstruct the virtual linearised states

$$\begin{aligned} \hat{l}_k(t+1) &= A\hat{l}_k(t) + Bv_k(t) + K[y_k(t) - \hat{y}_k(t)] \\ \hat{y}_k(t) &= C\hat{l}_k(t), \end{aligned} \tag{9}$$

where K is the gain of the state observer.

We take the state observer as the reference model and provide the same input signal $v_k(t)$ to the reference model and the adjustable system, before the feedback linearization of system (1) is completed, there will be an output error vector as follows

$$e_k(t) = l_k(t) - \hat{l}_k(t). \tag{10}$$

If and only if $a(x_k(t)) = L_g L_f^{n-1} h(x_k(t))$ and $b(x_k(t)) = -L_f^n h(x_k(t))$ are satisfied, at this point the feedback linearization controller (6) will precisely correct the system (1) to the standard linear system as shown in (8), and then we can design the observer gain K to make the tracking errors asymptotically converge.

Next, we will solve $a(x_k(t))$ and $b(x_k(t))$ based on model-free RL algorithm.

B. Q-Learning Based Feedback Linearization Controller Design

1) First we consider $a(x_k(t))$. It is not difficult to obtain

$$a(x_k(t)) = \frac{\partial y_k(t+n)}{\partial u_k(t)}. \tag{11}$$

It should be noted that we cannot use the input-output data from future moments to calculate $a_k(x(t)$ at the current moment, but in ILC, we can use the data from future moments in the previous trial to compute $a(x_k(t))$ as follows

$$\begin{aligned}&a(x_k(t))\\&=\frac{y_{k-1}(t+n)-y_{k-1}(t)-[y_{k-1}(t+n-1)-y_{k-1}(t-1)]}{u_{k-1}(t+n-1)-u_{k-1}(t+n)}.\end{aligned} \tag{12}$$

2) Next, a value iteration based model-free *Q*-learning algorithm is applied to solve $b(x_k(t))$. The performance index function is defined as follows

$$\begin{aligned}J(e_k(t))&=\sum_{\beta=0}^{N-t}\gamma^{\beta}U_k(t+\beta)\\&=\sum_{\beta=0}^{N-t}\gamma^{\beta}\|e_k(t+\beta)\|_Q^2+\|b(x_k(t+\beta))\|_R^2,\end{aligned} \tag{13}$$

where $U_k(t)\in\mathbb{R}^+$ represents the utility function, $Q\in\mathbb{R}^{n\times n}$ and $R\in\mathbb{R}$ are positive definite constant matrices. Clearly, our goal is to find a function $b(x_k(t))$ that minimizes the performance index function $J(e_k(t)))$.

As a model-free RL algorithm, *Q*-learning can iteratively solve the optimal performance index function and the optimal control strategy. Here, we regard $b(x_k(t))$ as the control strategy, and based on the Bellman optimality principle, the optimal *Q* function satisfies the following equation

$$Q^*(e_k(t)b(x_k(t)))=U_k(t)+\gamma\min_{b(x_k(t+1))}Q^*(e_k(t+1),b(x_k(t+1))), \tag{14}$$

where $0<\gamma<1$ is the discount factor. Then the corresponding optimal performance index function and the optimal control strategy can be given as follows

$$J^*(e_k(t))=\min_{b(x_k(t))}Q^*(e_k(t),b(x_k(t))), \tag{15}$$

$$b^*(x_k(t))=\arg\min_{b(x_k(t))}Q^*(e_k(t),b(x_k(t))). \tag{16}$$

In the continuous state-action space, it is usually not feasible to precisely calculate the iterative *Q* function $Q_i((e_k(t),b(x_k(t))))$ and its associated policy $b_i(x_k(t))$. Typically, function approximators such as neural networks are employed to approximate $Q_i(\cdot)$ and $b_i(\cdot)$, which are respectively referred to as critic and actor networks. Therefore, we use an approximator $\hat{Q}(e_k(t),b(x_k(t)),\theta_i)$ to approximate the iterative *Q* function, and the update rule for the approximate *Q* function is given in the following.

Suppose that $Q_i(e_k(t),b(x_k(t)))$ is able to be approximated by a continuous function as

$$\hat{Q}(e_k(t),b(x_k(t)),\theta_i)=\sum_{d=1}^{H}\theta_{i,d}w_d(e_k(t),b(x_k(t))), \tag{17}$$

where $\theta_i\in\mathbb{R}^H$ represents the model parameter vector, $w_d(e_k(t),b(x_k(t)))\in\mathbb{R}$ is the homeomorphism mapping of virtual state error and control.

For $i=0,1,2,\ldots$, the model parameter vector θ_i is updated as

$$\begin{aligned}&\theta_{i+1}=\theta_i+\alpha\eta_i c^T\left(e_k(t),b(x_k(t)),\theta_i\right)\Big[U_k(t)\\&+\gamma\min_{b_k(t+1)}\hat{Q}(e_k(t+1),b_k(t+1),\theta_i)-\hat{Q}(e_k(t),b(x_k(t)),\theta_i)\Big],\end{aligned} \tag{18}$$

where $\alpha\in\mathbb{R}^+$ represents the learning rate,

$$c\left(e_k(t),b(x_k(t)),\theta_i\right)=\frac{\partial\hat{Q}(e_k(t),b(x_k(t)),\theta_i)}{\partial\theta_i}\in\mathbb{R}^H,$$

$$\eta_i=\left[I-\gamma\,\mathrm{diag}\left\{\frac{c\left(e_k(t+1),b^*(x_k(t+1))|_{\theta_i},\theta_i\right)}{c\left(e_k(t),b(x_k(t)),\theta_i\right)}\right\}\right]^T\in\mathbb{R}^{H\times H},$$

$I\in\mathbb{R}^{H\times H}$ is an identity matrix, $b^*(x_k(t+1))|_{\theta_i}=\arg\min_{b(x_k(t+1))}\hat{Q}(e_k(t+1),b(x_k(t+1)),\theta_i)$.

When θ_i converges to the optimal model parameter vector θ^*, then the corresponding optimal control strategy $b^*(x_k(t))$ can be derived from the optimal approximate *Q* function as follows

$$b^*(x_k(t))=\arg\min_{b(x_k(t))}\hat{Q}(e_k(t),b(x_k(t)),\theta^*). \tag{19}$$

The executable algorithm for the value iteration based *Q*-learning algorithm is presented in Algorithm 1.

Algorithm 1 Value Iteration Based *Q*-Learning Algorithm.

Initialization

Give $0<\epsilon<1$ as the parameter of greedy policy for control selection
Let iteration index $i=0$
Start with an initial state $x_k(t)$ and an initial virtual state error $e_k(t), k=1, t=0$
while not terminated **do**
  if $t<N$ **then**
    Choose a control $b(x_k(t))$ by using greedy policy with parameter ϵ
    Compute $U_k(t)+\gamma\min_{b_k(t+1)}\hat{Q}(e_k(t+1),b_k(t+1),\theta_i)-\hat{Q}(e_k(t),b(x_k(t)),\theta_i)$
    Compute θ_{i+1} by (18)
    $i\leftarrow i+1$, $t\leftarrow t+1$
  else
    $k\leftarrow k+1$, $t=0$
  end if
  if $\|\theta_i-\theta_{i-1}\|\leq\zeta$, where ζ is a small positive constant **then**
    $\theta^*\leftarrow\theta_i$
    break
  end if
end while
Calculate the optimal performance index function as $J^*(e_k(t))=\min_{b(x_k(t))}\hat{Q}(e_k(t),b(x_k(t)),\theta^*)$
Calculate the optimal control policy as $b^*(x_k(t))=\arg\min_{b(x_k(t))}\hat{Q}(e_k(t),b(x_k(t)),\theta^*)$

C. LMI-Based Controller Design for ILC scheme

After convergence of the value iteration based Q-learning algorithm, the feedback linearization controller can correct the nonlinear system (1) into the linear form (9). By designing an appropriate observer gain K, the dynamics of system (9) can be corrected to that of the standard linear system (8). Subsequently, an LMI-based ILC input $v_k(t)$ is designed based on the standard linearization form (8). Since the input-output relationship of the nonlinear system (1) has been corrected to match that of system (8), thus the ILC law enables the output of the unknown nonlinear system to precisely track the reference signal.

Let us define the iterative control law as

$$v_k(t) = v_{k-1}(t)+K_1(l_k(t)-l_{k-1}(t))+K_2\bar{e}_{k-1}(t+1). \quad (20)$$

Note that the above formula is a version of (3) where the corrective term is computed as a combination of state feedback and the previous error feedforward. Clearly the goal is design K_1 and K_2 for both trial-to-trial error convergence and stability along the trial. To achieve this we use a 2D systems setting for ILC analysis. Firstly, it follows from (2) and (8) that

$$\bar{e}_k(t)-\bar{e}_{k-1}(t) = -CA\delta l_k(t) - CB\delta v_k(t-1), \quad (21)$$

where

$$\delta l_k(t)=l_k(t-1)-l_{k-1}(t-1),\ \delta v_k(t)=v_k(t)-v_{k-1}(t). \quad (22)$$

Secondly, one can find that

$$\delta l_k(t+1)= A\delta l_k(t) + B\delta v_k(t-1) \quad (23)$$

and hence the controlled dynamics can be written as

$$\begin{aligned}\delta l_k(t+1) &=(A+BK_1)\delta l_k(t)+BK_2\bar{e}_{k-1}(t)\\ \bar{e}_k(t) &=(-CA-CBK_1)\delta l_k(t)+(I-CBK_2)\bar{e}_{k-1}(t).\end{aligned} \quad (24)$$

The above set of equations has the form of a 2D Roesser model where $\delta l_k(t)$ plays the role of horizontal state vector and $\bar{e}_k(t)$ plays the role of the vertical state vector. Hence 2D Roesser model stability results can be applied for computing K_1 and K_2 which ensure both trial-to-trial error convergence and along the trial performance. Note that the system (8) has a relative degree ρ equal to the number of virtual states. Consequently, $CA^{\rho-1}B$ is nonsingular, while CB through $CA^{\rho-2}B$ are zero matrices. To compensate for the effect of ρ, a anticipatory gain operator can be incorporated, and ILC law (20) can be rewritten as

$$v_k(t)=v_{k-1}(t) + K_1\delta l_k(t + 1) + K_2\bar{e}_{k-1}(t + \rho). \quad (25)$$

and hence (24) is transformed to (see [2] for details)

$$\begin{aligned}\delta l_k(t + 1) =&(A + BK_1)\delta l_k(t) + BK_2\bar{e}_{k-1}(t + p - 1),\\ \bar{e}_k(t + p - 1) =&(-CA^\rho - CA^{\rho-1}BK_1)\delta l_k(t)\\ &+ (I - CA^{\rho-1}BK_2)\bar{e}_{k-1}(t + p - 1).\end{aligned} \quad (26)$$

Following result characterizes asymptotic stability of 2D model (26) and hence simultaneously guarantees the trial-to-trial convergence and transient response goals.

Lemma 1: [2] An ILC scheme described as Roesser model of the form (26) is asymptotically stable and hence trial-to-trial error convergence occurs if there exist $X_1 \succ 0$, $X_2 \succ 0$, M_1 and M_2 such that the following LMI is feasible

$$\begin{bmatrix} -X_1 & 0 & (\star) & (\star)\\ 0 & -X_2 & (\star) & (\star)\\ AX_1{+}BM_1 & BM_2 & -X_1 & 0\\ -CA^\rho X_1{-}CA^{\rho-1}BM_1 & X_2{-}CA^{\rho-1}BM_2 & 0 & -X_2 \end{bmatrix} \prec 0, \quad (27)$$

where $(\star)$ denotes transposed elements. Moreover, the required K_1 and K_2 are computed as

$$K_1 = M_1X_1^{-1},\ K_2 = M_2X_2^{-1}$$

IV. Simulation based Case Study

In order to verify the performance of the proposed model-free feedback linearization based ILC approach, numerical simulations are conducted on a discretized inverted pendulum system with the following dynamics

$$\begin{aligned}\begin{bmatrix}x_{1k}(t{+}1)\\ x_{2k}(t{+}1)\end{bmatrix} &= \begin{bmatrix} x_{1k}(t) + 0.1x_{2k}(t)\\ 0.1\frac{g}{\ell}\sin(x_{1k}(t)){+}(1{-}0.1\kappa\ell)x_{2k}(t)\end{bmatrix} + \begin{bmatrix}0\\ \frac{0.1}{m\ell^2}\end{bmatrix}u_k(t)\\ y_k(t) &= x_{1k}(t),\end{aligned} \quad (28)$$

where $m = 1/2$ kg and $\ell = 1/3$ m are the mass and length of the pendulum bar, respectively. Let $\kappa = 0.2$ and $g = 9.8$ m/s^2 be the frictional factor and the gravitational acceleration, respectively. The above parameters and model structure are unknown to the learning process.

The desired trajectory for $y_k(t)$ is $y_d(t) = \sin(0.1\pi t)$. Subsequently, the MRAC-based model-free feedback linearization framework needs to be established. The system matrix of the reference model (9) are designed as follows

$$A = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \quad B = \begin{pmatrix}0\\ 1\end{pmatrix}, \quad C = \begin{pmatrix}1 & 0\end{pmatrix}.$$

The weight matrices Q and R in the utility function are designed as $[10, 0; 0, 1]$ and 1. The gain of observer $K = [10, 0; 0, 20]$. Denote $e_{1k}(t), e_{2k}(t)$ and $b(x_k(t))$ as e_1, e_2 and b respectively, the structure of the approximate Q function is designed as follows

$$\hat{Q}(e, b, \theta_i)=\theta_{i,1}e_1^2+\theta_{i,2}e_2^2+\theta_{i,3}b^2+\theta_{i,4}e_1e_2+\theta_{i,5}e_1b+\theta_{i,6}e_2b,$$

where $\theta_{i,1} \sim \theta_{i,6}$ represent the model parameters. Set the discount factor and learning rate to $\gamma = 0.99$ and $\alpha = 0.05$ respectively.

Next, an ILC law (20) for the standard linear system is designed based on solution to LMI (27) and the subsequent controller gain matrices are $K_1 = [0.2, -0.1]$, $K_2 = 0.03$.

Before the convergence of the value iteration based Q-learning algorithm, the output trajectories of the nonlinear system (28) and the desired trajectory are shown in Fig. 2. Due to the presence of the observer gain K, the system is able to maintain bounded stability under the control signal generated with ILC law. However, due to the structural mismatch between the nonlinear system and the reference model, the desired trajectory cannot be precisely tracked.

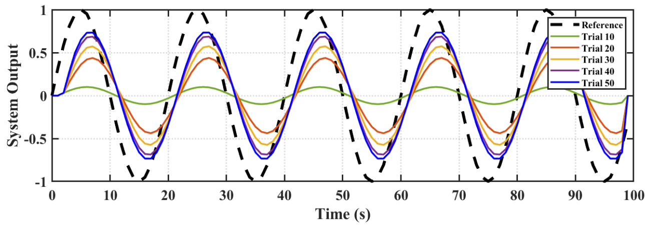

Fig. 2: Output trajectories of nonlinear system before feedback linearization

After 1000 iterations of Q-learning, the output trajectories of the nonlinear system (28) under the proposed control scheme are presented in Fig. 3, the corresponding tracking errors are illustrated in Fig. 4, the input trajectories are presented in Fig. 5. It can be observed from Fig. 3 that, as the iteration index increases, the system output progressively converges to the desired trajectory. The maximum steady-state tracking error and tracking root mean square error (RMSE) for the 10th and 50th trials are 0.7262, 9.382e-4 (steady-state error) and 0.5147, 0.0806 (RMSE), respectively. By the 50th trial, the output almost perfectly matches the reference signal across the entire time horizon.

These results validate that, once the model-free feedback linearization is accomplished via Q-learning, the LMI-based ILC controller (designed for the reference model system) effectively enables the original unknown nonlinear system to achieve high-precision tracking of the repetitive reference trajectory.

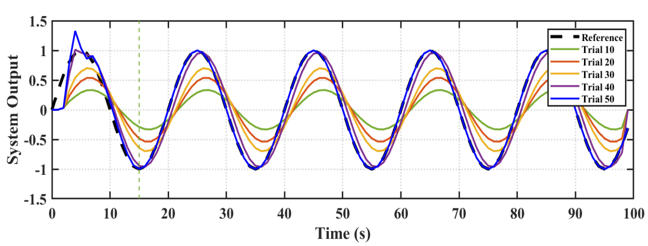

Fig. 3: Output trajectories of nonlinear system after feedback linearization

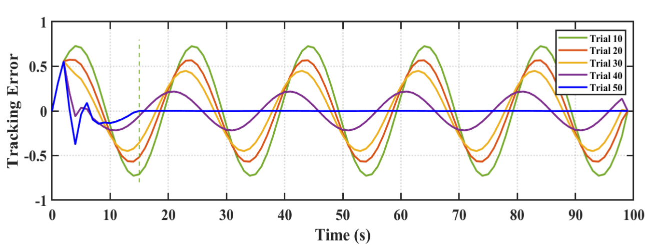

Fig. 4: Tracking error trajectories of nonlinear system after feedback linearization

V. Conclusion

This paper has presented a novel model-free ILC framework for unknown affine nonlinear systems. By integrating model-free feedback linearization with ILC, the proposed approach successfully addresses the challenge of controlling nonlinear systems without prior model knowledge. The combination of MRAC and Q-learning enables effective feedback linearization, while the LMI-based design procedure produces controllers for stability and convergence. Simulation studies on an inverted pendulum system demonstrate that the method achieves precise trajectory tracking performance. Future research will focus on extending this approach to more general nonlinear systems and enhancing its robustness against measurement noise and non-repetitive disturbances.

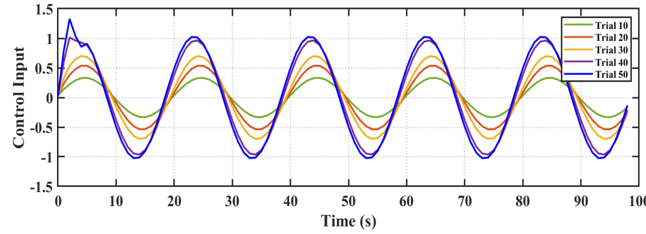

Fig. 5: Input trajectories of the feedback linearization control framework based on MRAC

References

[1] D. R. Smith and A. G. Alleyne, "Set-to-set iterative learning control," *Automatica*, vol. 179, p. 112422, 2025.

[2] E. Rogers, B. Chu, C. Freeman, P. Lewin, "Iterative Learning Control Algorithms and Experimental Benchmarking," *Wiley*, ISBN: 978-1-118-53537-0, 2023.

[3] C. Wang, M. Zheng, Z. Wang, C. Peng, and M. Tomizuka, "Robust iterative learning control for vibration suppression of industrial robot manipulators," *ASME J. Dyn. Sys., Meas., Control*, vol. 140, no. 1, p. 011003, 2018.

[4] R. Chi, Z. Hou, and J. Xu, "Adaptive ILC for a class of discrete-time systems with iteration-varying trajectory and random initial condition," *Automatica*, vol. 44, no. 8, pp. 2207–2213, 2008.

[5] I. Rotariu, M. Steinbuch, and R. Ellenbroek, "Adaptive iterative learning control for high precision motion systems," *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 5, pp. 1075–1082, 2008.

[6] X. Li and D. Huang, "Adaptive iterative learning control for nonsquare nonlinear systems with various nonrepetitive uncertainties: A unified approach," *IEEE Trans. Autom. Control*, vol. 69, no. 3, pp. 1736–1743, 2024.

[7] D. Gadginmath, V. Krishnan, and F. Pasqualetti, "Data-driven feedback linearization using the Koopman generator," *IEEE Trans. Autom. Control*, vol. 69, no. 12, pp. 8844–8851, 2024.

[8] Y. Sun, X. Chen, W. He, Z. Zhang, E. F. Fukushima, and J. She, "Q-learning based model-free input-output feedback linearization control method," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 9534–9539, 2023.

[9] N. Lin, R. Chi, B. Huang, and Z. Hou, "Multi-lagged-input iterative dynamic linearization based data-driven adaptive iterative learning control," *J. Franklin Inst.*, vol. 356, no. 1, pp. 457–473, 2019.

[10] Z. Chen, Y. Hou, R. Huang, and Q. Cheng, "Neural network compensator-based robust iterative learning control scheme for mobile robots nonlinear systems with disturbances and uncertain parameters," *Appl. Math. Comput.*, vol. 469, p. 128549, 2024.

[11] J. Liu, Z. Zhou, W. Hong, and J. Shi, "Two-dimensional iterative learning control with deep reinforcement learning compensation for the non-repetitive uncertain batch processes," *J. Process Control*, vol. 131, p. 103106, 2023.

[12] H. Tao, Y. Huang, T. Liu, and W. Paszke, "Reinforcement learning based iterative learning control for nonlinear batch process with non-repetitive uncertainty via Koopman operator," *J. Process Control*, vol. 148, p. 103402, 2025.