2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Composite Learning Exponential Parameter Estimation for Discrete-Time Nonlinear Systems

Qian Wang[1], Tian Shi[2], Vladimir Nikiforov[3], and Yongping Pan[1]

Abstract— **As a significant part of system modeling and control, parameter identification of continuous and discrete-time systems has been extensively studied in the past decades. However, most existing parameter identifiers cannot guarantee exponential parameter convergence without a strict condition termed persistent excitation (PE). This paper presents a composite learning-based parameter estimator for discrete-time nonlinear systems with linear-in-the-parameters uncertainties. A generalized prediction error based on regressor extension with online data memory is incorporated into the normal prediction error to accelerate parameter estimation. The storage and forgetting of online data are determined by only active regressor channels, which removes the restriction that all regressor channels need to be activated simultaneously for parameter estimation. Exponential parameter convergence under the proposed estimator is achieved under an interval excitation (IE) or even partial IE condition that is strictly weaker than the PE condition. Simulation results have verified the effectiveness and superiority of the proposed estimator compared with state-of-the-art estimators.**

I. Introduction

Parameter identification refers to determining the parameters in the mathematical model of a system by observing input and output signals [1]. As a significant part of system modeling and control, parameter identification of continuous and discrete-time systems has been widely studied in the past decades, with recent survey papers in [2]–[8]. Despite the rich experience and progress gained in this field, several open issues still remain, especially those related to improving parameter convergence performance. In general, parameter convergence cannot be ensured without a stringent condition termed persistent excitation (PE), which implies that input signals must contain sufficiently rich spectral information to excite system models relevant to identified parameters [9]. In fact, the input signals may lead to undesired phenomena, such as exciting unmodeled dynamics that can destroy system performance or even stability [10].

Modern control systems are usually implemented through digital computing units, such that system inputs and outputs are discretely processed. Estimation and control based on discrete-time frameworks are beneficial for understanding the design and implementation of digital control systems, thereby improving system performance, reliability, and efficiency [11].

*This work was supported in part by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University, China, under Grant 23lgzy004
(*Corresponding author: Yongping Pan*).

[1]Qian Wang and Yongping Pan are with the School of Advanced Manufacturing, Sun Yat-sen University, Shenzhen 518100, China wangq666@mail2.sysu.edu.cn; panyongp@mail.sysu.edu.cn

[2]Tian Shi is with School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China shit23@mail2.sysu.edu.cn

[3]Vladimir Nikiforov is with the ITMO University, Saint Petersburg 197101, Russia nikiforov_vo@itmo.ru

Adaptive parameter estimation for discrete-time linear time-invariant systems under relaxed excitation conditions has been considered primarily based on dynamic regressor extension and mixing (DREM) [12]–[16] and memory regressor extension (MRE) [17]. In [12]–[14], some DREM parameter estimators were presented to ensure asymptotic parameter convergence under a non-square integrability condition. In [15], a linear regression equation with a regressor under PE was constructed based on DREM, while the original regressor only needs to satisfy a weaker condition named interval excitation (IE) to obtain exponential parameter convergence. Parameter convergence in finite time based on DREM under IE was achieved in [16], but it is valid only when specific initial conditions are imposed on the estimator. In [17], a parameter estimator based on MRE was introduced to achieve asymptotic parameter convergence under a weaker non-Lebesgue integrable condition. Nonlinear systems can more accurately describe real-world plants in which the changes of unknown parameters affect the dynamic behavior in a nonlinear manner. There are only a few results on parameter estimation for discrete-time nonlinear systems without the PE condition [18], [19]. In [18], a DREM-based estimator for nonlinearly parameterized systems was developed to achieve asymptotic parameter convergence under the non-square integrability condition. In [19], a DREM-based least-squares (LS) parameter estimator for the same system as in [18] was designed to guarantee exponential parameter convergence under the IE condition, where a stringent global Lipschitz condition is imposed on nonlinear terms.

Composite learning is an emerging methodology for adaptive estimation and control to ensure exponential parameter convergence without the PE condition [20]–[22] and has been applied to many real-world robotic systems [8]. In composite learning, a generalized prediction error is constructed utilizing regressor extension with online data memory and is incorporated as extra feedback for parameter estimation such that exponential parameter convergence is guaranteed under IE [23]. A major limitation of existing composite learning methods is that the storing and forgetting of online data are determined by the minimum singular value of an excitation matrix, which means that all channels of the regressor need to be activated simultaneously at a certain moment. A natural idea is whether it is possible to forget the information of inactive channels while ensuring partial parameter convergence. In addition, composite learning is expected to be established in a discrete-time framework for more convenient implementation.

Motivated by the above discussions, we propose a composite learning parameter estimator for discrete-time nonlinear systems with linear-in-the-parameters (LIP) uncertainties to achieve exponential parameter convergence without the PE condition. Compared with existing parameter estimators for discrete-time systems, the proposed estimator includes several advantages: 1) The normal prediction error is incorporated to accelerate parameter estimation; 2) the storage and forgetting of online data are determined by only active regressor channels, removing the restriction that all regressor channels must be activated simultaneously for parameter estimation; 3) exponential parameter convergence is achieved under IE or even partial IE. It is worth noting that this article provides the first composite learning scheme in the discrete-time framework, which does not simply discretize its continuous version.

Notation: $\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{R}^n$, and $\mathbb{R}^{N\times n}$ denote the spaces of real numbers, positive real numbers, real n-dimensional vectors, and real $N\times n$-dimensional matrices, respectively, $\mathbb{N}$ denotes the set of positive integers, $\|\boldsymbol{x}\|$ is the Euclidean norm of $\boldsymbol{x}$, I is an identity matrix, $\Omega_{\mathrm{r}} := \{\boldsymbol{x} | \|\boldsymbol{x}\| \leq r\}$ denotes a ball of radius $r\in\mathbb{R}^+$, L_∞ denotes the space of bounded signals, $\sigma_{\min}(A)$ is the minimum singular value of A, and $\arg\max_{k\in S} f(k) := \{k\in S | f(i) \leq f(k), \forall i \in S\}$ with $f:\mathbb{N}\mapsto\mathbb{R}$ and $S\subset\mathbb{N}$, where $N, n\in\mathbb{N}$, $A\in\mathbb{R}^{n\times n}$, and $\boldsymbol{x}\in\mathbb{R}^n$.

II. Problem Formulation

Consider a class of discrete-time nonlinear systems with LIP uncertainties as follows [24][1]:

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k)) + \Phi^T(\boldsymbol{x}(k), \boldsymbol{u}(k))\boldsymbol{\theta} \quad (1)$$

where $\boldsymbol{x}(k)\in\mathbb{R}^n$ is a system state, $\boldsymbol{u}(k)\in\mathbb{R}^m$ is a control input, $\boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k)) : \mathbb{R}^n\times\mathbb{R}^m \mapsto \mathbb{R}^n$ is a known vector field, $\Phi(\boldsymbol{x}(k), \boldsymbol{u}(k)) : \mathbb{R}^n\times\mathbb{R}^m\mapsto\mathbb{R}^{N\times n}$ is a known nonlinear regressor, $\boldsymbol{\theta}\in\Omega_{c_\theta}\subset\mathbb{R}^N$ is an unknown parameter vector, $c_\theta \in\mathbb{R}^+$ is a certain constant, and $N\in\mathbb{N}$ denotes the number of unknown parameters. For convenience, $\boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k))$ and $\Phi(\boldsymbol{x}(k), \boldsymbol{u}(k))$ are abbreviated as $\boldsymbol{f}(k)$ and $\Phi(k)$, respectively. Two definitions are given for theoretical analysis [25].

Definition 1: A bounded signal $\Phi(k)\in\mathbb{R}^{N\times n}$ is of PE if $\exists k_{\mathrm{d}}, \sigma\in\mathbb{R}^+$ such that $\sum_{\tau=k-k_{\mathrm{d}}}^{k}\Phi(\tau)\Phi^T(\tau)\geq\sigma I, \forall k\geq 0$.

Definition 2: A bounded signal $\Phi(k)\in\mathbb{R}^{N\times n}$ is of IE if $\exists k_{\mathrm{d}}, \sigma, k_{\mathrm{a}}\in\mathbb{R}^+$ such that $\sum_{\tau=k_{\mathrm{a}}-k_{\mathrm{d}}}^{k_{\mathrm{a}}}\Phi(\tau)\Phi^T(\tau)\geq\sigma I$.

Definition 3: A bounded signal $\Phi(k)\in\mathbb{R}^{N\times n}$ is of partial IE if $\exists k_{\mathrm{d}}, \sigma, k_{\mathrm{b}}\in\mathbb{R}^+$ such that $\sum_{\tau=k_{\mathrm{b}}-k_{\mathrm{d}}}^{k_{\mathrm{b}}}\Phi_\nu(\tau)\Phi_\nu^T(\tau)\geq\sigma I$, where $\Phi_\nu\in\mathbb{R}^{q\times n}$ is a sub-regressor composed of some row vectors of $\Phi(k)$ with $1\leq q<N$.

A row vector $\boldsymbol{\phi}_j(k)\in\mathbb{R}^n (j=1,\cdots,N)$ of a regressor $\Phi(k)\in\mathbb{R}^{N\times n}$ is named an active channel if $\|\boldsymbol{\phi}_j(k)\|\neq 0$ for the current moment k; otherwise it is an inactive channel[2]. So, $\Phi_\nu(k)$ in Definition 3 is composed of all active channels, which means that $\Phi_\nu(k)$ is of IE and $\Phi(k)$ is of partial IE.

Let $\hat{\boldsymbol{\theta}}\in\mathbb{R}^N$ be an estimate of the unknown parameter vector $\boldsymbol{\theta}$, and $\tilde{\boldsymbol{\theta}}(k) = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(k)\in\mathbb{R}^N$ be a parameter estimation error. To avoid using $\boldsymbol{x}(k+1)$ in parameter estimation and enhancing robustness against noises, we apply a discrete-time stable filter $L(z) := \frac{\alpha_1}{1-\alpha_2 z^{-1}}$ with $\alpha_1, \alpha_2\in\mathbb{R}^+$ to (1), which results in a filtered linear parameterized model

$$\boldsymbol{\chi}_{\mathrm{f}}(k) = \Phi_{\mathrm{f}}^T(k)\boldsymbol{\theta} \quad (2)$$

with $\boldsymbol{\chi}_{\mathrm{f}}(k) := L(z)[\boldsymbol{x}(k)] - z^{-1}L(z)[\boldsymbol{f}(k)]$ and $\Phi_{\mathrm{f}}(k) := z^{-1}L(z)[\Phi(k)]$. A state prediction model is given by

$$\hat{\boldsymbol{\chi}}_{\mathrm{f}}(k) = \Phi_{\mathrm{f}}^T(k)\hat{\boldsymbol{\theta}}(k)$$

where $\hat{\boldsymbol{\chi}}_{\mathrm{f}}(k)\in\mathbb{R}^n$ denotes an estimate of $\boldsymbol{\chi}_{\mathrm{f}}(k)$. Define a filtered state prediction error

$$\boldsymbol{\varepsilon}_{\mathrm{f}}(k) := \boldsymbol{\chi}_{\mathrm{f}}(k) - \Phi_{\mathrm{f}}^T(k)\hat{\boldsymbol{\theta}}(k) \quad (3)$$

and design a standard gradient-based estimator

$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + \gamma M_{\mathrm{f}}^{-1}(k)\Phi_{\mathrm{f}}(k)\boldsymbol{\varepsilon}_{\mathrm{f}}(k)$$

in which $M_{\mathrm{f}}(k) := I + \gamma\Phi_{\mathrm{f}}(k)\Phi_{\mathrm{f}}^T(k)\in\mathbb{R}^{N\times N}$ is a matrix normalization gain, and $\gamma\in\mathbb{R}^+$ is a learning rate. Then, the parameter error dynamics can be represented by

$$\tilde{\boldsymbol{\theta}}(k+1) = \tilde{\boldsymbol{\theta}}(k) - \gamma M_{\mathrm{f}}^{-1}(k)\Phi_{\mathrm{f}}(k)\Phi_{\mathrm{f}}^T(k)\tilde{\boldsymbol{\theta}}(k)$$

which implies that the normal prediction error $\boldsymbol{\varepsilon}_{\mathrm{f}}(k)$ in (3) converges to $\mathbf{0}$ asymptotically [25, Sec. 3.6.2]. However, the exponential convergence of $\tilde{\boldsymbol{\theta}}(k)$ depends on PE, which requires that the system state $\boldsymbol{x}(k)$ in (1) includes considerably rich spectral information all the time. This article aims to design a parameter estimation scheme to achieve the exponential convergence of $\tilde{\boldsymbol{\theta}}(k)$ under IE or even partial IE.

Remark 1: Three different excitation conditions are given in Definitions 1-3, respectively. The PE condition in Definition 1 is quite strict as it requires the excitation strength to be greater than σ all the time. The IE condition in Definition 2 weakens the PE condition as it only requires the excitation strength to satisfy the condition at a certain moment k_{a}, which may be satisfied during transient processes [20]–[22]. However, the PE or IE condition requires all channels $\boldsymbol{\phi}_j(k) (j=1,\cdots,N)$ to be activated simultaneously for all $k\geq 0$ and a certain time k_{a}, respectively, which is difficult to satisfy in most practical scenarios. The partial IE condition in Definition 3 ignores all inactive channels such that it is satisfied at the beginning and some certain moments later owing to $\boldsymbol{x}(0)\neq\mathbf{0}$ and $\boldsymbol{x}(k)\neq\mathbf{0}, \exists k>0$ in general, i.e., it is impossible for all channels $\boldsymbol{\phi}_j(k)\equiv 0\ (j=1,\cdots,N), \forall k\geq 0$.

III. Composite Learning Parameter Estimation

This section aims to design a composite learning parameter estimator for the system (1) to ensure exponential convergence of the parameter estimation error $\tilde{\boldsymbol{\theta}}(k)$ under IE or partial IE. First, define an excitation matrix

$$Q(k) := \sum_{\tau=k-k_{\mathrm{d}}}^{k}\Phi_{\mathrm{f}}(\tau)\Phi_{\mathrm{f}}^T(\tau). \quad (4)$$

Multiplying each side of (2) by $\Phi_{\mathrm{f}}(k)$ and applying (4), one gets a generalized regression model

$$\boldsymbol{\psi}(k) = Q(k)\boldsymbol{\theta}$$

[1]The LIP system (1) has nonlinearities in $\Phi(\boldsymbol{x}(k), \boldsymbol{u}(k))$ so that many results on linear systems cannot be directly applied to (1), and $\boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k))$ and $\Phi(\boldsymbol{x}(k), \boldsymbol{u}(k))$ are bounded and continuous in their arguments.

[2]All row vectors $\boldsymbol{\phi}_j(k)\in\mathbb{R}^n (j=1,\cdots,N)$ are linearly independent.

with $\boldsymbol{\psi}(k) := \sum_{\tau=k-k_{\mathrm{d}}}^{k} \Phi_{\mathrm{f}}(\tau)\boldsymbol{\chi}_{\mathrm{f}}(\tau)$. It is assumed that there exist $k_{\mathrm{a}}, \sigma \in \mathbb{R}^+$ to make the IE condition hold. Define a generalized prediction error

$$\boldsymbol{\xi}(k) := \begin{cases} \boldsymbol{\psi}(k_{\mathrm{e}\nu}) - Q(k_{\mathrm{e}\nu})\hat{\boldsymbol{\theta}}(k), \ k < k_{\mathrm{a}} \\ \boldsymbol{\psi}(k_{\mathrm{e}}) - Q(k_{\mathrm{e}})\hat{\boldsymbol{\theta}}(k), \ k \geq k_{\mathrm{a}} \end{cases} \tag{5}$$

with

$$k_{\mathrm{e}} := \arg\max_{\tau\in[k_{\mathrm{a}},k]} \sigma_{\min}(Q(\tau)),$$

$$k_{\mathrm{e}\nu} := \arg\max_{\tau\in[0,k]} \sigma_{\min}(Q_\nu(\tau))$$

in which $Q_\nu(k) := \sum_{\tau=k-k_{\mathrm{d}}}^{k} \Phi_{\mathrm{f}\nu}(\tau)\Phi_{\mathrm{f}\nu}^T(\tau)$ with $\Phi_{\mathrm{f}\nu}^T(k)$ being a sub-regressor composed of some columns $\boldsymbol{\phi}_{\mathrm{ft}_i}(k)$ of $\Phi_{\mathrm{f}}^T(k)$ that satisfies $\|\boldsymbol{\phi}_{\mathrm{ft}_i}(k)\| \neq 0$ for the current moment k, i.e., $\Phi_{\mathrm{f}\nu}^T(k) = [\boldsymbol{\phi}_{\mathrm{ft}_1}, \boldsymbol{\phi}_{\mathrm{ft}_2}, \cdots, \boldsymbol{\phi}_{\mathrm{ft}_{N_\nu}}]$ with $1 \leq t_i \leq N$ and $i = 1$ to $N_\nu < N$. The current maximal exciting strength is

$$\sigma_{\mathrm{c}}(k) := \begin{cases} \sigma_{\min}(Q_\nu(k_{\mathrm{e}\nu})), \ k < k_{\mathrm{a}} \\ \sigma_{\min}(Q(k_{\mathrm{e}})), \ k \geq k_{\mathrm{a}}. \end{cases} \tag{6}$$

The composite learning estimator that integrates the normal prediction error $\boldsymbol{\varepsilon}_{\mathrm{f}}(k)$ in (3) and the generalized prediction error $\boldsymbol{\xi}(k)$ in (5) is designed as follows:

$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + \frac{\gamma}{2}\left[M_{\mathrm{f}}^{-1}(k)\Phi_{\mathrm{f}}(k)\boldsymbol{\varepsilon}_{\mathrm{f}}(k) + M^{-1}(k)\boldsymbol{\xi}(k)\right] \tag{7}$$

in which

$$M(k) := \begin{cases} I + \gamma Q(k_{\mathrm{e}\nu}), \ k < k_{\mathrm{a}} \\ I + \gamma Q(k_{\mathrm{e}}), \ k \geq k_{\mathrm{a}} \end{cases}$$

is a matrix normalization gain. Let $\hat{\boldsymbol{\theta}}_\nu(k) \in \mathbb{R}^{N_\nu}$ denote a part of $\hat{\boldsymbol{\theta}}(k)$ regarding the sub-regressor $\Phi_{\mathrm{f}\nu}(k)$, and $\tilde{\boldsymbol{\theta}}_\nu(k) \in \mathbb{R}^{N_\nu}$ denote the corresponding parameter estimation error. The following theorem shows our main results.

Theorem 1: Let $[0, k_{\mathrm{f}})$ with $k_{\mathrm{f}} \in \mathbb{R}^+$ denote the maximal iteration set of the existence of solutions of the system (1). For any given $\hat{\boldsymbol{\theta}}(0) \in \Omega_{\mathrm{c}_\theta}$ and $\gamma \in \mathbb{R}^+$, the composite learning update law of $\hat{\boldsymbol{\theta}}(k)$ in (7) guarantees that:

1) $\hat{\boldsymbol{\theta}}(k)$ and $\boldsymbol{\varepsilon}_{\mathrm{f}}(k)$ are of L_∞, $\forall k \leq k_{\mathrm{f}}$;
2) $\tilde{\boldsymbol{\theta}}(k) \to \mathbf{0}$ exponentially at $k \geq k_{\mathrm{a}}$ as $k \to \infty$, if the IE condition $Q(k_{\mathrm{a}}) \geq \sigma I$ in Definition 2 is satisfied for certain constants $k_{\mathrm{a}}, \sigma \in \mathbb{R}^+$;
3) $\tilde{\boldsymbol{\theta}}_\nu(k) \to \mathbf{0}$ exponentially at $k \geq k_{\mathrm{b}}$ as $k \to \infty$, if the partial IE condition $Q_\nu(k_{\mathrm{b}}) \geq \sigma I$ in Definition 3 holds for certain constants $k_{\mathrm{b}}, \sigma \in \mathbb{R}^+$.

Proof: The proof is given in the Appendix. ■

Remark 2: It follows from Theorem 1 that the proposed parameter estimator (7) ensures exponential parameter convergence under the strictly weaker IE or partial IE condition, which endows robustness against external disturbances and measurement noise, where the rigorous proof can be referred to [9]. Increasing the learning rate γ in (7) can accelerate parameter convergence but may also potentially increase the sampling frequency required and noise sensitivity.

Remark 3: Compared with the MRE estimator in [17], the proposed parameter estimator (7) has the following characteristics: 1) The normal prediction error $\boldsymbol{\varepsilon}_{\mathrm{f}}(k)$ in (3) is involved additionally to enhance parameter convergence, while the MRE estimator only uses a generalized prediction error to estimate unknown parameters; 2) the exciting strength $\sigma_{\mathrm{c}}(k)$ in (6) is monotonically non-decreasing, attributing to the storage and forgetting of online data, which is beneficial to improve the rate of parameter convergence; 3) exponential parameter convergence can be achieved under IE or partial IE instead of PE for the MRE estimator.

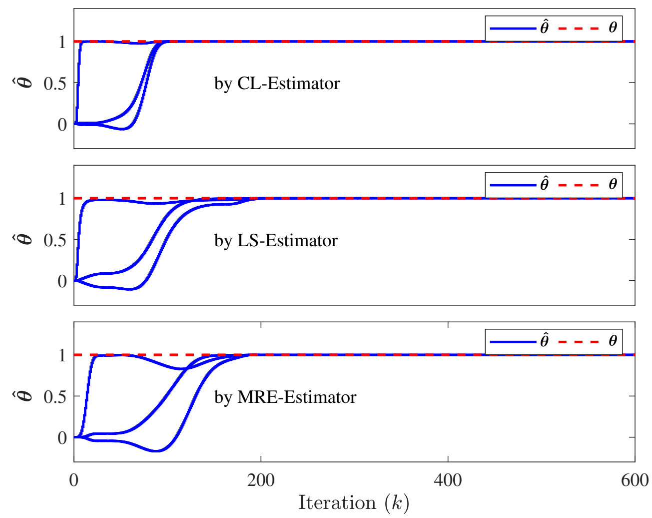

Fig. 1. Parameter estimates $\hat{\boldsymbol{\theta}}$ by three estimators under the PE condition.

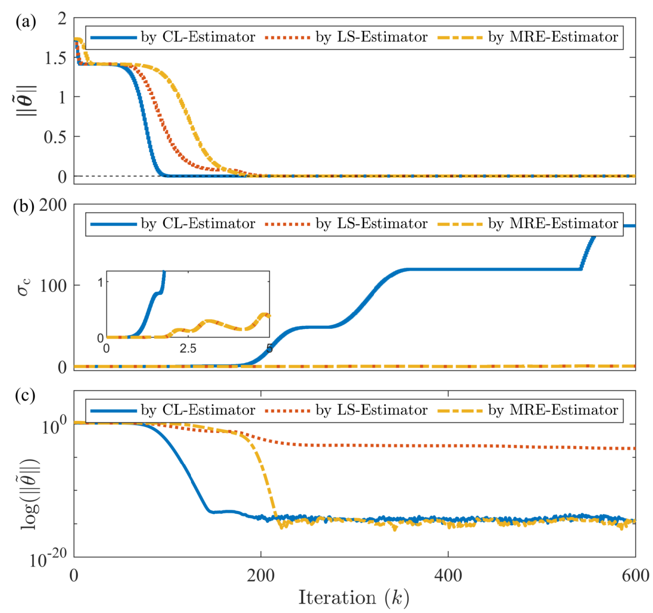

Fig. 2. Performance comparisons of three estimators under the PE condition. (a) The estimation error norms $\|\tilde{\boldsymbol{\theta}}\|$. (b) The exciting strengths σ_{c}. (c) The estimation error norms $\|\tilde{\boldsymbol{\theta}}\|$ in logarithm.

IV. Simulation Validation

Consider a Van der Pol oscillator modeled in the discrete-time form as follows [26]:

$$\begin{cases} x_1(k+1) = x_1(k) + T_{\mathrm{s}}x_2(k), \\ x_2(k+1) = x_2(k) + T_{\mathrm{s}}\boldsymbol{\varphi}^T(k)\boldsymbol{\theta} \end{cases}$$

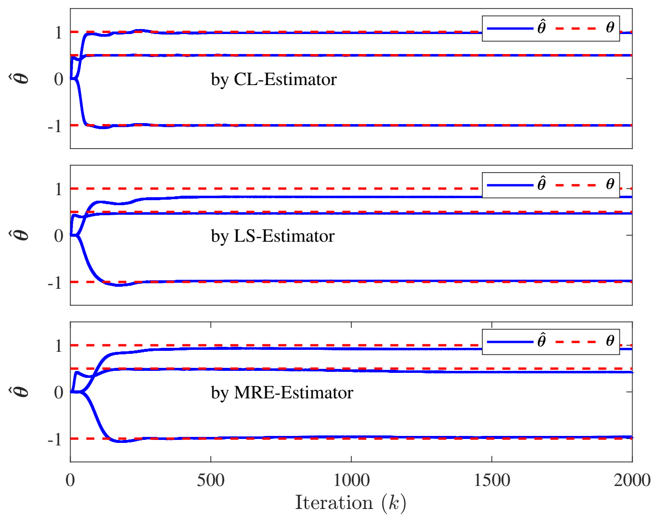

Fig. 3. Parameter estimates $\hat{\boldsymbol{\theta}}$ by three estimators under the IE condition.

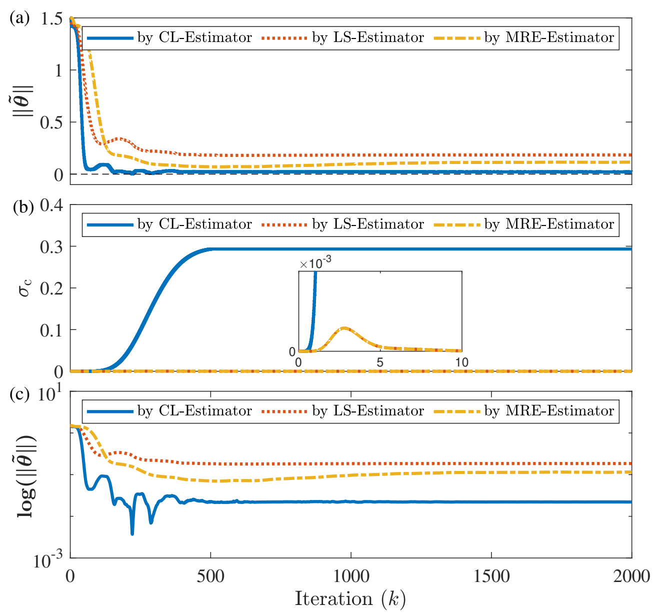

Fig. 4. Performance comparisons of three estimators under the IE condition. (a) The estimation error norms $\|\tilde{\boldsymbol{\theta}}\|$. (b) The exciting strengths σ_{c}. (c) The estimation error norms $\|\tilde{\boldsymbol{\theta}}\|$ in logarithm.

with $\boldsymbol{\varphi}(k) = [-x_1(k), x_2(k), -x_1^2(k)x_2(k)]^T$ and $T_{\mathrm{s}} \in \mathbb{R}^+$ being a sampling time. We get $\boldsymbol{x}(k) = [x_1(k), x_2(k)]^T$, $\boldsymbol{f}(k) = [x_1(k) + T_{\mathrm{s}}x_2(k), x_2(k)]^T$, and $\Phi(k) = [\mathbf{0}, T_{\mathrm{s}}\boldsymbol{\varphi}(k)]$. For simulations, set $\hat{\boldsymbol{\theta}}(0) = \mathbf{0}$, $T_{\mathrm{s}} = 0.01$, $\alpha_1 = 5$, and $\alpha_2 = e^{-5T_{\mathrm{s}}}$. Gaussian white noise with 0 mean and 0.0002 standard deviation is added to the measurement of $\boldsymbol{x}(k)$. The classical LS estimator in [27] and the MRE estimator in [17] are selected as baselines of the proposed composite learning (CL) estimator, where the shared parameters of the three estimators are set to be the same values for fair comparisons.

Case 1: Under PE. Consider the case of PE with $\boldsymbol{x}(0) = [1, 0]^T$, $\boldsymbol{\theta} = [1, 1, 1]^T$, and $\gamma = 20$. Performance comparisons of the three estimators are shown in Figs. 1-2. The estimation error $\tilde{\boldsymbol{\theta}}$ by the CL estimator converges to $\mathbf{0}$ after running about

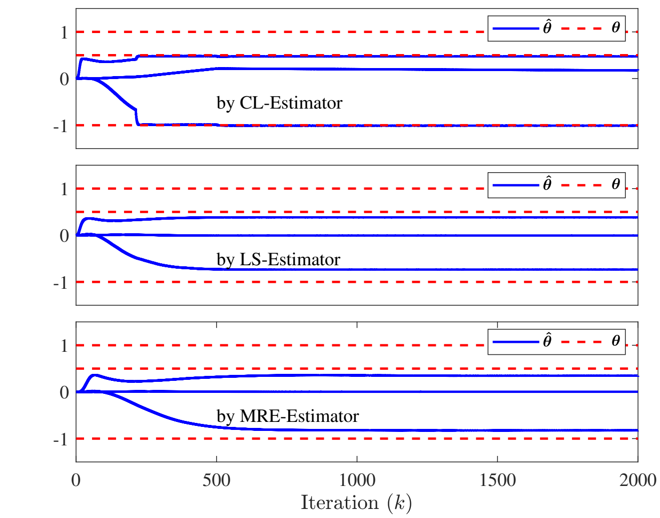

Fig. 5. Parameter estimates $\hat{\boldsymbol{\theta}}$ by three estimators under partial IE.

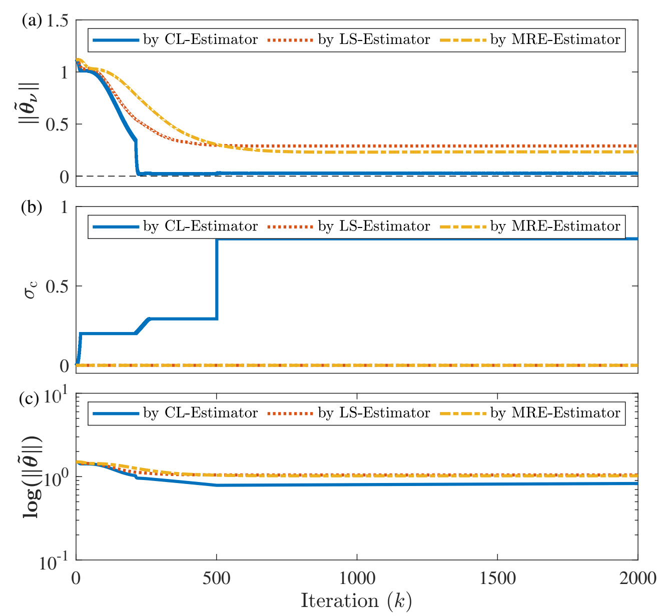

Fig. 6. Performance comparisons of three estimators under partial IE. (a) The norms of the partial estimation error $\tilde{\boldsymbol{\theta}}_\nu$. (b) The exciting strengths σ_{c}. (c) The estimation error norms $\|\tilde{\boldsymbol{\theta}}\|$ in logarithm.

100 iterations [see Fig. 2(a)], and the exciting strength σ_{c} in (6) keeps at a high level throughout [see Fig. 2(b)]. The LS and MRE estimators converge slower than the CL estimator [see Figs. 2(a) and (c)], although they also achieve exponential convergence due to the presence of PE [see Fig. 1].

Case 2: Under IE. Consider the case of IE with $\boldsymbol{x}(0) = [1, 0]^T$, $\boldsymbol{\theta} = [0.5, -1, 1]^T$, and $\gamma = 20$. Performance comparisons of the three estimators are shown in Figs. 3-4. The LS and MRE estimators perform much worse under IE [see Fig. 4(a)], as their exciting strengths σ_{c} decrease to 0 after running around 500 iterations [see Fig. 4(b)]. In contrast, the CL estimator achieves the convergence of $\tilde{\boldsymbol{\theta}}$ to $\mathbf{0}$ rapidly after about 200 iterations [see Fig. 4(a)], where σ_{c} in (6) is monotonically non-decreasing and keeps at a high level after about 250 iterations [see Fig. 4(b)] due to the storage of online data. The convergence rate of the CL estimator is higher than those of the other two estimators [see Fig. 4(c)].

Case 3: Under partial IE. Consider the case of partial IE with $\boldsymbol{x}(0) = [0.3, 0.001]^T$, $\boldsymbol{\theta} = [0.5, -1, 1]^T$, and $\gamma = 5$. Performance comparisons of the three estimators are shown in Figs. 5-6, where $\tilde{\boldsymbol{\theta}}_\nu := [\tilde{\theta}_1, \tilde{\theta}_2]^T$ is a partial estimation error on active channels, and an estimated value always 0 corresponds to an inactive channel. The LS and MRE estimators perform much worse under partial IE [see Fig. 6(a)], as their exciting strengths σ_c are 0 all the time [see Fig. 6(b)]. In contrast, the CL estimator exhibits the convergence of $\tilde{\boldsymbol{\theta}}_\nu$ to $\mathbf{0}$ after about 250 iterations [see Fig. 6(a)], where σ_c in (6) is monotonically non-decreasing and keeps at a high level after about 500 iterations [see Fig. 6(b)]. The CL estimator converges faster than the other two estimators [see Fig. 6(c)].

V. Conclusions

This paper has developed a CL parameter estimator for discrete-time nonlinear systems with LIP uncertainties, where exponential parameter convergence can be achieved under IE or even partial IE, which is strictly weaker than PE. The storage and forgetting of online data are determined by only active regressor channels, which gets rid of the restriction that all regressor channels must be activated simultaneously for parameter estimation. The effectiveness of the proposed method has been validated on a discrete-time model of the Van der Pol oscillator, where simulation results illustrate the superiority of the proposed estimator compared to the LS-based and MRE-based estimators. Future studies would apply the proposed estimator to more complicated systems.

Appendix: The Proof of Theorem 1

Proof: The parameter error model derived by the parameter estimation law (7) is represented by

$$\tilde{\boldsymbol{\theta}}(k+1) = \Big[I - \frac{\gamma}{2}\Lambda(k)\Big]\tilde{\boldsymbol{\theta}}(k) \tag{8}$$

in which $\Lambda(k) \in \mathbb{R}^{N\times N}$ is given by

$$\Lambda(k) := \begin{cases} M_\mathrm{f}^{-1}(k)\Phi_\mathrm{f}\Phi_\mathrm{f}^T + M^{-1}(k)Q(k_{\mathrm{e}\nu}),\ k < k_\mathrm{a}, \\ M_\mathrm{f}^{-1}(k)\Phi_\mathrm{f}\Phi_\mathrm{f}^T + M^{-1}(k)Q(k_\mathrm{e}),\ k \geq k_\mathrm{a}. \end{cases}$$

Consider a Lyapunov function candidate

$$V(k) = \tilde{\boldsymbol{\theta}}^T(k)\tilde{\boldsymbol{\theta}}(k). \tag{9}$$

In view of (8) and (9), $V(k+1)$ is calculated as

$$\begin{aligned} V(k+1) = \tilde{\boldsymbol{\theta}}^T(k)\Big[&I - \frac{\gamma}{2}M_\mathrm{f}^{-1}(k)\Phi_\mathrm{f}(k)\Phi_\mathrm{f}^T(k) \\ &- \frac{\gamma}{2}M^{-1}(k)Q(k)\Big]^2\tilde{\boldsymbol{\theta}}(k). \end{aligned}$$

It follows from the following inequality

$$0 < I - \frac{\gamma}{2}M_\mathrm{f}^{-1}(k)\Phi_\mathrm{f}(k)\Phi_\mathrm{f}^T(k) - \frac{\gamma}{2}M^{-1}(k)Q(k) < I \tag{10}$$

to get $V(k+1) \leq V(k)$, which implies $\tilde{\boldsymbol{\theta}}(k), \hat{\boldsymbol{\theta}}(k) \in L_\infty$, and hence, $\boldsymbol{\varepsilon}_\mathrm{f}(k) \in L_\infty$ on $k \in [0, k_\mathrm{f})$.

Second, consider the convergence problem under the IE condition for $k \geq k_\mathrm{a}$, i.e., there exist $k_\mathrm{a}, \sigma \in \mathbb{R}^+$ such that $\sigma_\mathrm{c}(k_\mathrm{a}) \geq \sigma$. It follows from (10) that

$$\begin{aligned} &\Big[I - \frac{\gamma}{2}M_\mathrm{f}^{-1}(k)\Phi_\mathrm{f}(k)\Phi_\mathrm{f}^T(k) - \frac{\gamma}{2}M^{-1}(k_\mathrm{e})Q(k_\mathrm{e})\Big]^2 \\ \leq &\Big[I - \frac{\gamma}{2}M^{-1}(k_\mathrm{e})Q(k_\mathrm{e})\Big]^2 \\ = &\Big[I - \frac{\gamma}{2}\big(I + \gamma Q(k_\mathrm{e})\big)^{-1}Q(k_\mathrm{e})\Big]^2 \\ = &\Big[\big(I + \gamma Q(k_\mathrm{e})\big)^{-1}\big(I + \gamma Q(k_\mathrm{e}) - \frac{\gamma}{2}Q(k_\mathrm{e})\big)\Big]^2 \\ = &[I + \gamma Q(k_\mathrm{e})]^{-2}\Big[I + \frac{\gamma}{2}Q(k_\mathrm{e})\Big]^2 \\ = &[I + \gamma Q(k_\mathrm{e})]^{-2}\Big[\frac{1}{2}\big(I + \gamma Q(k_\mathrm{e})\big) + \frac{1}{2}I\Big]^2 \\ = &\frac{1}{4}\Big[I + \big(I + \gamma Q(k_\mathrm{e})\big)^{-1}\Big]^2. \end{aligned}$$

As $Q(k_\mathrm{e}) \geq \sigma_\mathrm{c}(k_\mathrm{a})I \geq \sigma I$ according to (4), one has

$$\begin{aligned} V(k+1) &\leq \frac{1}{4}\Big(1 + \frac{1}{1+\gamma\sigma}\Big)^2\tilde{\boldsymbol{\theta}}^T(k)\tilde{\boldsymbol{\theta}}(k) \\ &= \frac{1}{4}\Big(1 + \frac{1}{1+\gamma\sigma}\Big)^2 V(k) \end{aligned}$$

with $0 < \frac{1}{4}\big(1 + \frac{1}{1+\gamma\sigma}\big)^2 < 1$, implying that $\tilde{\boldsymbol{\theta}}(k)$ converges exponentially to $\mathbf{0}$ for $k \geq k_\mathrm{a}$.

Third, consider the parameter convergence under partial IE for $k \geq k_\mathrm{b}$. From Definition 3, there are $k_\mathrm{b}, \sigma \in \mathbb{R}^+$ such that $Q_\nu(k_\mathrm{b}) \geq \sigma I$. Choose a Lyapunov function candidate

$$V_\nu(k) = \tilde{\boldsymbol{\theta}}_\nu^T(k)\tilde{\boldsymbol{\theta}}_\nu(k). \tag{11}$$

The proof of parameter convergence under partial IE is similar to that under IE, so we follow the similar steps as above to obtain $V_\nu(k+1)$ based on (11) as follows:

$$V_\nu(k+1) \leq \tilde{\boldsymbol{\theta}}_\nu^T(k)\Big[\frac{1}{2}\Big(I + \big(I + \gamma Q_\nu(k_\mathrm{b})\big)^{-1}\Big)\Big]^2\tilde{\boldsymbol{\theta}}_\nu(k).$$

It follows from $Q_\nu(k_\mathrm{b}) \geq \sigma I$ and the above result that

$$\begin{aligned} V_\nu(k+1) &\leq \frac{1}{4}\Big(1 + \frac{1}{1+\gamma\sigma}\Big)^2\tilde{\boldsymbol{\theta}}_\nu^T(k)\tilde{\boldsymbol{\theta}}_\nu(k) \\ &= \frac{1}{4}\Big(1 + \frac{1}{1+\gamma\sigma}\Big)^2 V_\nu(k) \end{aligned}$$

with $0 < \frac{1}{4}\big(1 + \frac{1}{1+\gamma\sigma}\big)^2 < 1$, implying that $\tilde{\boldsymbol{\theta}}_\nu(k)$ exponentially converges to $\mathbf{0}$ for $k \geq k_\mathrm{b}$. ∎

References

[1] J. Voros, "Modeling and parameter identification of systems with multi-segment piecewise-linear characteristics," *IEEE Trans. Autom. Control*, vol. 47, no. 1, pp. 184–188, Jan. 2002.

[2] M. Gevers, "A personal view of the development of system identification: A 30-year journey through an exciting field," *IEEE Control Syst. Mag.*, vol. 26, no. 6, pp. 93–105, Dec. 2006.

[3] L. Ljung, "Perspectives on system identification," *Annu. Rev. Control*, vol. 34, no. 1, pp. 1–12, Apr. 2010.

[4] O. Sigaud, C. Salaün, and V. Padois, "On-line regression algorithms for learning mechanical models of robots: A survey," *Robot. Auton. Syst.*, vol. 59, no. 12, pp. 1115–1129, Dec. 2011.

[5] M. Schoukens and K. Tiels, "Identification of block-oriented nonlinear systems starting from linear approximations: A survey," *Automatica*, vol. 85, pp. 272–292, Nov. 2017.

[6] R. Ortega, V. Nikiforov, and D. Gerasimov, "On modified parameter estimators for identification and adaptive control. A unified framework and some new schemes," *Annu. Rev. Control*, vol. 50, pp. 278–293, Sep. 2020.

[7] D. N. Gerasimov and V. O. Nikiforov, "On key properties of the Lion's and Kreisselmeier's adaptation algorithms," *Int. J. Adapt. Control Signal Process.*, vol. 36, no. 6, pp. 1285–1304, Jul. 2022.

[8] K. Guo and Y. Pan, "Composite adaptation and learning for robot control: A survey," *Annu. Rev. Control*, vol. 55, pp. 279–290, May 2023.

[9] S. Sastry, M. Bodson, and J. F. Bartram, *Adaptive Control: Stability, Convergence, and Robustness*. Englewood Cliffs, NJ, USA: Prentice Hall, 1989.

[10] N. T. Nguyen, *Model-Reference Adaptive Control: A Primer*. Cham, Switzerland: Springer, 2018.

[11] K. Ogata, *Discrete-Time Control Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1995.

[12] A. A. Belov, R. Ortega, and A. A. Bobtsov, "Parameter identification of linear discrete-time systems with guaranteed transient performance," *IFAC-PapersOnLine*, vol. 51, no. 15, pp. 1038–1043, Jul. 2018.

[13] R. Ortega, S. Aranovskiy, A. A. Pyrkin, A. Astolfi, and A. A. Bobtsov, "New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases," *IEEE Trans. Autom. Control*, vol. 66, no. 5, pp. 2265–2272, Jun. 2020.

[14] A. Pyrkin, R. Ortega, V. Gromov, A. Bobtsov, and A. Vedyakov, "A globally convergent direct adaptive pole-placement controller for non-minimum phase systems with relaxed excitation assumptions," *Int. J. Adapt. Control Signal Process.*, vol. 33, no. 10, pp. 1491–1505, Aug. 2019.

[15] M. Korotina, J. G. Romero, S. Aranovskiy, A. Bobtsov, and R. Ortega, "A new on-line exponential parameter estimator without persistent excitation," *Syst. Control Lett.*, vol. 159, Art. No. 105079, Jan. 2022.

[16] R. Ortega, A. Bobtsov, and N. Nikolaev, "Parameter identification with finite-convergence time alertness preservation," *IEEE Control Syst. Lett.*, vol. 6, pp. 205–210, Feb. 2021.

[17] D. N. Gerasimov, M. E. Belyaev, and V. O. Nikiforov, "Improvement of transient performance in MRAC by memory regressor extension," *Eur. J. Control*, vol. 59, pp. 264–273, May 2021.

[18] R. Ortega, V. Gromov, E. Nuño, A. Pyrkin, and J. G. Romero, "Parameter estimation of nonlinearly parameterized regressions without overparameterization: Application to adaptive control," *Automatica*, vol. 127, Art. No. 109544, May 2021.

[19] R. Ortega, J. G. Romero, and S. Aranovskiy, "A new least squares parameter estimator for nonlinear regression equations with relaxed excitation conditions and forgetting factor," *Syst. Control Lett.*, vol. 169, Art. No. 105377, Nov. 2022.

[20] Y. Pan and H. Yu, "Composite learning from adaptive dynamic surface control," *IEEE Trans. Autom. Control*, vol. 61, no. 9, pp. 2603–2609, Sep. 2016.

[21] Y. Pan, J. Zhang, and H. Yu, "Model reference composite learning control without persistency of excitation," *IET Control. Theory Appl.*, vol. 10, no. 16, pp. 1963–1971, Sep. 2016.

[22] Y. Pan and H. Yu, "Composite learning robot control with guaranteed parameter convergence," *Automatica*, vol. 89, pp. 398–406, Mar. 2018.

[23] Y. Pan and T. Shi, "Adaptive estimation and control with online data memory: A historical perspective," *IEEE Control Syst. Lett.*, vol. 8, pp. 267–278, Feb. 2024.

[24] V. Adetola and M. Guay, "Performance improvement in adaptive control of linearly parameterized nonlinear systems," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2182–2186, Jun. 2010.

[25] G. Tao, *Adaptive Control Design and Analysis*. Hoboken, NJ, USA: Wiley, 2003.

[26] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.

[27] J.-J. E. Slotine and W. Li, *Applied Nonlinear Control*. Englewood Cliffs, NJ, USA: Prentice Hall, 1991.