Received August 27, 2019, accepted September 9, 2019, date of publication September 13, 2019, date of current version September 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2941226

Linear Time-Varying Data Model-Based Iterative Learning Recursive Least Squares Identifications for Repetitive Systems

NA LIN[1], RONGHU CHI[1], AND BIAO HUANG[2], (Fellow, IEEE)
[1]School of Automation and Electronics Engineering, Qingdao University of Science and Technology, Qingdao 266061, China
[2]Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB T6G 2G6, Canada

Corresponding authors: Ronghu Chi (ronghu_chi@hotmail.com) and Biao Huang (bhuang@ualberta.ca)

This work was supported in part by the National Science Foundation of China under Grant 61374102 and Grant 61873139, in part by the Taishan Scholar Program of the Shandong Province of China, in part by the Key Research and Development Program of Shandong Province under Grant 2018GGX101047, in part by the Natural Science Foundation of Shandong Province of China under Grant ZR2019MF036 and Grant ZR2017MEE071, and in part by the Natural Science and Engineering Research Council of Canada.

ABSTRACT In this paper, an iterative learning recursive least squares (ILRLS) identification method is developed by considering a class of repetitive systems. First, considering a repetitive discrete-time system corrupted by white noise, we present a linear time-varying data model to describe the input-output dynamic behavior of the system in iteration domain. On this basis, two ILRLS methods are proposed taking both white noises and colored noises into consideration. With an extensive analysis, the two proposed methods are shown applicable to repetitive nonlinear discrete-time systems owing to their data-driven nature by which no explicit models are required. The proposed ILRLS methods are executed pointwisely along the iteration direction, and they can also deal with time-varying uncertainties. The results are proved and verified by mathematical analysis along with simulations.

INDEX TERMS System identification, iterative learning recursive least squares, linear time-varying data model, repetitive discrete-time systems, data-driven approach.

I. INTRODUCTION

System identification is of great importance in many engineering fields such as in chemical process [1], [2], power system [3], [4], biomedical system [5], [6], and so on. Many identification methods [7]–[9] have been proposed over the last several decades to identify time-invariant parameters of the system. However, time-varying parameters are actually more common in practice. Examples include the armature resistance and flux linkage of DC motor [10], stiffness of rotational mechanical system [11], etc. Therefore, it is imperative to investigate the identification problem for time-varying parametric systems.

Time-varying system identification has also been extensively studied. In [12], an improved least squares algorithm is proposed for linear time-varying (LTV) systems by adding additional terms to the parameter estimation law and the covariance matrix update law. In [13], a recursive least squares (RLS) method is proposed for the permanent magnet synchronous motor by coupling identification with bias compensation. A robust forgetting factor based RLS algorithm [14] is proposed for time-varying disturbances. Some other identification methods for time-varying parametric systems include robust adaptive method [15], matrix gradient algorithm [16], variable structure systems theory and sliding-mode based identification method [17], etc. However, the identification performance of these methods may deteriorate significantly if a large change or a fast-rate variation of parameters occurs.

Many systems [18]–[20] perform repetitive operations over fixed time intervals. For such a repetitive system, the time-varying parameters can also vary along the iteration direction. From this perspective, several identification algorithms [21]–[27] have been proposed for repetitive systems by extending the traditional identification algorithms from the time domain to the iteration domain. The time-varying parameters are estimated in batch from iteration to iteration rather than from time to time. To be specific, Reference [21] proposes an iterative identification approach for linear continuous-time systems. An iterative

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

learning RLS algorithm has been developed in [22], [23] for time-varying autoregressive moving average exogenous models. Later, a two-dimensional RLS algorithm is developed in [24], [25] for time-varying autoregressive exogenous (ARX) model identification. In addition to the above parametric identification methods [21]–[25] for linear repetitive systems, Ref. [26] proposes an iterative learning identification projection algorithm for nonlinear systems. Further, Ref. [27] extends the results in [26] and proposes a modified iterative RLS algorithm for nonlinear high-speed train systems where the linear parametric structures are known a priori.

It is worth pointing out that either a mechanistic model structure or an identification model structure must be required in most of the existing system identification methods [12]–[17], [21]–[27], which causes difficulties in some practical applications where these model structures are not easily available. In addition, it is often costly to obtain a mathematical model by using physical and/or chemical principles because of the complicated dynamic behaviors of a practical large-scale plant. On the other hand, simpler linear identification methods may loss some information when linearizing a nonlinear system where unmodeled dynamics and other model-unmatching factors can occur. To deal with the above mentioned problem in system identification, a nonlinear system identification method is proposed in [28] by using a dynamic neural network. Further, a parallel and a series-parallel neural network identification models are proposed in [29] with a dynamic BP algorithm for network training. Reference [30] uses fuzzy neural modeling to identify the nonlinear time-varying plant. However, all of these nonlinear system identification methods [28]–[30] are only conducted in time domain and little result has been reported for nonlinear iterative identification of repetitive systems. Another limitation of the fuzzy neural network based nonlinear identification methods [28]–[30] is the need of proper choice of fuzzy rules and neural networks, which might be difficult especially when there is no available priori model knowledge about the nonlinear systems.

Recently, the data-driven approach [31]–[36], whether for control design or system identification, has become increasingly popular because it is independent from model information except for the use of I/O data. A data-driven hybrid ARX and markov-chain modeling method [31] is proposed to identify time-varying time delays. Reference [32] proposes a variational Bayesian approach to identify the ARX model. According to the I/O data statistical analysis and the prespecified state dependent parameter model, a data-based mechanistic modeling method [33] is proposed to determine the model parameters. Huang and Kadali [34] propose a data-driven approach to predictive control without requiring an explicit model. However, either an ARX model structure in [31]–[32] or the state dependent parameter model in [33], [34] is required although they are also called data-driven methods. Model structure is the prerequisite of conducting system identification, while actual systems are difficult to be captured completely by any specific model structure.

More recently, Hou and Jin [35] propose linear data models for nonlinear systems to reformulate the input-output mapping relationship, without needing complex mechanistic analysis, which only describes the dynamic behavior of the I/O data. Further, Chi *et al.* [36] propose iterative dynamic linearization (IDL) for nonlinear repetitive plants mainly for the design and analysis of control applications. However, the application of IDL for nonlinear system identification has not been reported up to now.

In view of the above considerations, we propose an iterative learning RLS (ILRLS) identification method for repetitive systems under a data-driven framework in this work. First, an iterative linear time-varying data model is derived for linear repetitive systems subject to white noise, and then the linear time-varying data model based ILRLS identifications algorithm is developed. Subsequently, its convergence is analyzed rigorously. Further, the algorithm is extended to the case with colored noise, and the corresponding parameter estimation convergence analysis is also provided. Moreover, by introducing an IDL technique of Ref. [36], an iterative linear time-varying data model of nonlinear systems is derived, and the corresponding ILRLS identifications algorithm for nonlinear systems is designed, which follows a similar design procedure of the ILRLS for linear systems. The main innovations of our work are summarized below.

(i) The presented ILRLS method does not use any explicit model knowledge but is data-driven using only I/O data, such that it is applicable to both linear and nonlinear systems. (ii) The proposed ILRLS is conducted along the iterative direction instead of the time direction, and thus it can address arbitrarily fast time-varying uncertainties. (iii) The iterative linear data model used in this work virtually exists in the computer and does not need to have a physical interpretation such as mechanistic models.

To clearly demonstrate the basic idea, the proposed method is analyzed by taking linear time-varying system as an example, and then is extended to nonlinear systems. Mathematical analysis and simulations verify that the presented ILRLS is efficient in applications with ability to attenuate influences of the disturbances.

This paper is structured as follows. Section II formulates the problems. Section III proposes the ILRLS method. The analysis is shown in Section IV. Section V extends the results to a repetitive nonlinear nonaffine system. Section VI gives simulations results, and Section VII concludes the paper.

II. PROBLEM FORMULATION

Consider a discrete-time LTV system,

$$\begin{cases} \boldsymbol{x}_k(t+1) = \boldsymbol{A}(t)\boldsymbol{x}_k(t) + \boldsymbol{B}(t)u_k(t) \\ y_k(t) = \boldsymbol{C}(t)\boldsymbol{x}_k(t) + v_k(t), \end{cases} \tag{1}$$

where $y_k(t) \in R$, $u_k(t) \in R$, $v_k(t) \in R$ and $\boldsymbol{x}_k(t) \in R^n$ represent system outputs, inputs, white noise and state, respectively; $t \in \{0, 1, \cdots, N\}$ is discrete-time sampling instants, where $N \in Z^{+}$ is finite; k is the iteration index; $\boldsymbol{A}(t) \in R^{n\times n}$, $\boldsymbol{B}(t) \in R^{n\times 1}$ and $\boldsymbol{C}(t) \in R^{1\times n}$ are system matrixes.

As usual, the following assumption is made.

Assumption 1: System (1) has identical initial condition, that is, $\boldsymbol{x}_k(0) = \boldsymbol{c}$, $\forall k$, where $\boldsymbol{c}$ represents a constant vector.

According to (1), we have

$$\begin{aligned} y_k(t+1) = &\boldsymbol{C}(t+1)\prod_{i=0}^{t}\boldsymbol{A}(i)\boldsymbol{x}_k(0) \\ &+\boldsymbol{C}(t+1)\sum_{j=1}^{t+1}\prod_{i=j}^{t}\boldsymbol{A}(i)\boldsymbol{B}(j-1) \\ &\times u_k(j-1) + v_k(t) \end{aligned} \tag{2}$$

Then, using Assumption 1, we can get

$$y_k(t+1) - y_{k-1}(t+1) = \Delta\boldsymbol{u}_k^T(t)\boldsymbol{\varphi}(t) + \Delta v_k(t), \tag{3}$$

where $\boldsymbol{u}_k(t) = [u_k(0), \cdots, u_k(t)]^T$, Δ is a difference operator along the direction of iteration, e.g., $\Delta\boldsymbol{u}_k(t) = \boldsymbol{u}_k(t) - \boldsymbol{u}_{k-1}(t)$, $\boldsymbol{\varphi}(t) = [\phi_0(t), \cdots, \phi_t(t)]^T \in R^{(t+1)}$, where $\phi_i(t) = \boldsymbol{C}(t+1)\prod_{j=i+1}^{t}\boldsymbol{A}(j)\boldsymbol{B}(i)$, $i \in \{0, \cdots, t\}$, and $\prod_{i=t+1}^{t}\boldsymbol{A}(i) = 1$.

In this work, we define $\hat{\boldsymbol{\varphi}}_k(t)$ and $\hat{y}_k(t+1)$ to represent the estimate of parameter $\boldsymbol{\varphi}(t)$ and output $y_k(t+1)$ respectively. We will design an estimator to estimate time-varying parameter $\boldsymbol{\varphi}(t)$, so that the predicted output $\hat{y}_k(t+1)$ can estimate the real output $y_k(t+1)$. And the predicted output can be calculated using the following equation:

$$\hat{y}_k(t+1) = \hat{y}_{k-1}(t+1) + \Delta\boldsymbol{u}_k^T(t)\hat{\boldsymbol{\varphi}}_k(t) \tag{4}$$

III. ALGORITHM DESIGN

From (3) and (4), we have the following equation:

$$\boldsymbol{Y}_k(t+1) = \boldsymbol{Y}_{k-1}(t+1) + \Delta\boldsymbol{U}_k(t)\boldsymbol{\varphi}(t) + \Delta\bar{\boldsymbol{V}}_k(t) \tag{5}$$

where

$$\begin{aligned} \boldsymbol{Y}_k(t+1) &= col\{y_1(t+1), \cdots, y_k(t+1)\} \in R^{k\times 1}, \\ \boldsymbol{Y}_{k-1}(t+1) &= col\{y_0(t+1), \cdots, y_{k-1}(t+1)\} \in R^{k\times 1}, \\ \boldsymbol{U}_k(t) &= [\boldsymbol{u}_1(t), \cdots, \boldsymbol{u}_k(t)]^T \in R^{k\times(t+1)}, \\ \bar{\boldsymbol{V}}_k(t) &= col\{v_1(t), \cdots, v_k(t)\} \in R^{k\times 1}, \end{aligned}$$

and $col\{\cdot\}$ is a notation for the column vector. Define an objective function,

$$J(\boldsymbol{\varphi}(t)) = \Delta\bar{\boldsymbol{V}}_k^T(t)\Delta\bar{\boldsymbol{V}}_k(t) \tag{6}$$

Solving (6), the least square estimation of the parameter vector $\boldsymbol{\varphi}(t)$ can be obtained

$$\hat{\boldsymbol{\varphi}}_k(t) = \left(\Delta\boldsymbol{U}_k^T(t)\Delta\boldsymbol{U}_k(t)\right)^{-1}\Delta\boldsymbol{U}_k^T(t)\boldsymbol{T}_k(t+1), \tag{7}$$

where $\boldsymbol{T}_k(t+1) = \boldsymbol{Y}_k(t+1) - \hat{\boldsymbol{Y}}_{k-1}(t+1)$ and $\hat{\boldsymbol{Y}}_{k-1}(t+1)$ is the estimate of $\boldsymbol{Y}_{k-1}(t+1)$.

Denote

$$\boldsymbol{P}_k^{-1}(t) = \sum_{j=1}^{k}\Delta\boldsymbol{u}_j(t)\Delta\boldsymbol{u}_j^T(t), \tag{8}$$

$$\boldsymbol{\zeta}_k(t+1) = \Delta\boldsymbol{U}_k^T(t)\boldsymbol{T}_k(t+1). \tag{9}$$

Then, (7) becomes

$$\hat{\boldsymbol{\varphi}}_k(t) = \boldsymbol{P}_k(t)\boldsymbol{\zeta}_k(t+1) \tag{10}$$

Further, one has

$$\boldsymbol{P}_k^{-1}(t) = \boldsymbol{P}_{k-1}^{-1}(t) + \Delta\boldsymbol{u}_k(t)\Delta\boldsymbol{u}_k^T(t) \tag{11}$$

Using the Matrix Inversion Lemma [35] and (11), it results

$$\boldsymbol{P}_k(t) = \boldsymbol{P}_{k-1}(t) - \frac{\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)}{1 + \Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)} \tag{12}$$

Further, according to (11), one has

$$\boldsymbol{P}_k^{-1}(t) = \Delta\boldsymbol{U}_k^T(t)\Delta\boldsymbol{U}_k(t) \tag{13}$$

Then, consideration of (7), (8) and (13), yields

$$\hat{\boldsymbol{\varphi}}_k(t) = \boldsymbol{P}_k(t)\left[\Delta\boldsymbol{U}_{k-1}^T(t), \Delta\boldsymbol{u}_k(t)\right]\begin{bmatrix}\boldsymbol{T}_{k-1}(t+1) \\ \vartheta_k(t+1)\end{bmatrix} \tag{14}$$

where $\vartheta_k(t+1) = y_k(t+1) - \hat{y}_{k-1}(t+1)$.

According to (9) and (10), Eq. (14) is further written as

$$\hat{\boldsymbol{\varphi}}_k(t) = \boldsymbol{P}_k(t)\left(\boldsymbol{P}_{k-1}^{-1}(t)\hat{\boldsymbol{\varphi}}_{k-1}(t) + \Delta\boldsymbol{u}_k(t)\vartheta_k(t+1)\right) \tag{15}$$

Substituting (11) into (15), yields

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_k(t) = &\hat{\boldsymbol{\varphi}}_{k-1}(t) + \boldsymbol{P}_k(t)\Delta\boldsymbol{u}_k(t) \\ &\times\left(\vartheta_k(t+1) - \Delta\boldsymbol{u}_k^T(t)\hat{\boldsymbol{\varphi}}_{k-1}(t)\right) \end{aligned} \tag{16}$$

Denoting $\boldsymbol{H}_k(t) = \boldsymbol{P}_k(t)\Delta\boldsymbol{u}_k(t)$, according to (12), one has

$$\begin{aligned} \boldsymbol{H}_k(t) &= \boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)\left(1 - \frac{\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)}{1 + \Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)}\right) \\ &= \frac{\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)}{1 + \Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)} \end{aligned} \tag{17}$$

In summary, the ILRLS algorithm can be presented as follows:

$$\hat{\boldsymbol{\varphi}}_k(t) = \hat{\boldsymbol{\varphi}}_{k-1}(t) + \boldsymbol{H}_k(t)\left(\vartheta_k(t+1) - \Delta\boldsymbol{u}_k^T(t)\hat{\boldsymbol{\varphi}}_{k-1}(t)\right) \tag{18}$$

$$\boldsymbol{H}_k(t) = \frac{\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)}{1 + \Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)} \tag{19}$$

$$\boldsymbol{P}_k(t) = \boldsymbol{P}_{k-1}(t) - \boldsymbol{H}_k(t)\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t) \tag{20}$$

IV. CONVERGENCE ANALYSIS

Before we give the convergence theorem of the presented algorithm, the other necessary assumptions and a useful property are given.

Assumption 2: The noise in the system is white and satisfies following conditions,

$$E\{v_k(t)|\mathscr{F}_{k-1}(t)\}=0 \quad \text{and } E\left\{v_k^2(t)|\mathscr{F}_{k-1}(t)\right\}\le\sigma^2(t),$$

where $E\{\cdot\}$ is the expectation operator; $\mathscr{F}_k(t)$ represents the σ-algebra [22] formed by input and output data obtained through k has repeated runs, and $\sigma(t)$ is uniformly bounded with respect to t.

Assumption 3 [23]: There exists a constant k_0, such that the following persistent excitation condition is established when $k\ge k_0$:

$$\alpha(t)\boldsymbol{I}\le\frac{1}{k}\sum_{i=1}^{k}\Delta\boldsymbol{u}_i(t)\Delta\boldsymbol{u}_i^T(t)\le\beta(t)\boldsymbol{I}, \quad a.s.$$

where $\alpha(t)>0$ and $\beta(t)>0$.

Property 1 [37]: For a matrix $\boldsymbol{Q}$ and a vector $\boldsymbol{x}$, the following inequation is satisfied:

$$\lambda_{\min}(\boldsymbol{Q})\|\boldsymbol{x}\|^2\le\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}\le\lambda_{\max}(\boldsymbol{Q})\|\boldsymbol{x}\|^2,$$

where $\lambda_{\min}(\boldsymbol{Q})$ and $\lambda_{\max}(\boldsymbol{Q})$ represent the minimum and maximum eigenvalue of the matrix $\boldsymbol{Q}$, respectively.

To describe the results, we introduce the notation O. When $k\to\infty$, the sequences $f_k\to 0$ and $g_k\to 0$; $f_k=O(g_k)$ means that there are a positive constant c and a positive integer k_0 such that when $k>k_0$, $|f_k/g_k|\le c$.

The convergence of the proposed ILRLS method is shown in the following theorem.

Theorem 1: Consider the linear system (1) under Assumptions 1 – 3. Applying the ILRLS method (18) – (20), one can guarantee that the parameter estimation error converges with the increasing number of iterations, i.e.,

$$\|\tilde{\boldsymbol{\varphi}}_k(t)\|=O\left(\sqrt{\sum_{j=0}^{t}L_k(j)\Big/\lambda_{\min}\left(\boldsymbol{P}_k^{-1}(t)\right)}\right), \quad a.s., \tag{21}$$

where $L_k(j)=\left(\ln\left|\boldsymbol{P}_k^{-1}(j)\right|\right)^{\varsigma(j)}$, $\varsigma(j)>1, j=0,1,\cdots,t$, and $\tilde{\boldsymbol{\varphi}}_k(t)=\hat{\boldsymbol{\varphi}}_k(t)-\boldsymbol{\varphi}(t)$.

Proof: Denote $\eta_k(t)=\vartheta_k(t+1)-\Delta\boldsymbol{u}_k^T(t)\hat{\boldsymbol{\varphi}}_k(t)$, and $\varepsilon_k(t)=\vartheta_k(t+1)-\Delta\boldsymbol{u}_k^T(t)\hat{\boldsymbol{\varphi}}_{k-1}(t)$. Multiplying $\Delta\boldsymbol{u}_k^T(t)$ from both sides of (18), and then subtracting $\vartheta_k(t+1)$ from the both sides, one obtains

$$\eta_k(t)=\varepsilon_k(t)-\Delta\boldsymbol{u}_k^T(t)\boldsymbol{H}_k(t)\varepsilon_k(t) \tag{22}$$

Substituting Eq. (19) into Eq. (22), results in

$$\eta_k(t)=\frac{1}{1+\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)}\varepsilon_k(t) \tag{23}$$

Subtracting $\boldsymbol{\varphi}(t)$ from both sides of (18), results in

$$\tilde{\boldsymbol{\varphi}}_k(t)=\tilde{\boldsymbol{\varphi}}_{k-1}(t)+\boldsymbol{H}_k(t)\varepsilon_k(t) \tag{24}$$

According to (19) and (23), Eq. (24) becomes

$$\tilde{\boldsymbol{\varphi}}_k(t)=\tilde{\boldsymbol{\varphi}}_{k-1}(t)+\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)\eta_k(t) \tag{25}$$

Using (25) and (11), one obtains

$$\begin{aligned}\tilde{\boldsymbol{\varphi}}_{k-1}^T(t)\boldsymbol{P}_{k-1}^{-1}(t)\tilde{\boldsymbol{\varphi}}_k(t)&=\tilde{\boldsymbol{\varphi}}_k^T(t)\boldsymbol{P}_k^{-1}(t)\tilde{\boldsymbol{\varphi}}_k(t)\\&\quad-\tilde{\boldsymbol{\varphi}}_k^T(t)\Delta\boldsymbol{u}_k(t)\Delta\boldsymbol{u}_k^T(t)\tilde{\boldsymbol{\varphi}}_k(t)\\&\quad-\Delta\boldsymbol{u}_k^T(t)\tilde{\boldsymbol{\varphi}}_k(t)\eta_k(t)\end{aligned} \tag{26}$$

Again, using (25), we have

$$\begin{aligned}&\tilde{\boldsymbol{\varphi}}_{k-1}^T(t)\Delta\boldsymbol{u}_k(t)\eta_k(t)\\&\quad=\tilde{\boldsymbol{\varphi}}_k^T(t)\Delta\boldsymbol{u}_k(t)\eta_k(t)\\&\qquad-\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)\eta_k^2(t)\end{aligned} \tag{27}$$

Denoting $\gamma_k(t)=-\Delta\boldsymbol{u}_k^T(t)\tilde{\boldsymbol{\varphi}}_k(t)$, then according to (26)-(27), we can get

$$\begin{aligned}&\tilde{\boldsymbol{\varphi}}_k^T(t)\boldsymbol{P}_k^{-1}(t)\tilde{\boldsymbol{\varphi}}_k(t)\\&\quad=\tilde{\boldsymbol{\varphi}}_{k-1}^T(t)\boldsymbol{P}_{k-1}^{-1}(t)\tilde{\boldsymbol{\varphi}}_{k-1}(t)+\gamma_k^2(t)\\&\qquad-2\gamma_k(t)\eta_k(t)-\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)\eta_k^2(t)\\&\quad\le\tilde{\boldsymbol{\varphi}}_{k-1}^T(t)\boldsymbol{P}_{k-1}^{-1}(t)\tilde{\boldsymbol{\varphi}}_{k-1}(t)+\gamma_k^2(t)-2\gamma_k(t)\eta_k(t)\end{aligned} \tag{28}$$

Defining $V_k(t)=\tilde{\boldsymbol{\varphi}}_k^T(t)\boldsymbol{P}_k^{-1}(t)\tilde{\boldsymbol{\varphi}}_k(t)$, then from (28), we get

$$V_k(t)\le V_{k-1}(t)+\beta_k(t)-2v_k(t)\gamma_k(t) \tag{29}$$

where $\beta_k(t)=\gamma_k^2(t)-2(\eta_k(t)-v_k(t))\gamma_k(t)$.

In addition, in view of (25) and the definition of $\gamma_k(t)$, we can derive

$$\begin{aligned}-\gamma_k(t)&=\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)(\eta_k(t)-v_k(t))\\&\quad+\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)v_k(t)\\&\quad+\tilde{\boldsymbol{\varphi}}_{k-1}^T(t)\Delta\boldsymbol{u}_k(t)\end{aligned} \tag{30}$$

Substitute (30) into (29), gives

$$\begin{aligned}V_k(t)&\le V_{k-1}(t)+\beta_k(t)+2G_k(t)v_k(t)\\&\quad+2\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)v_k^2(t)\end{aligned} \tag{31}$$

where $G_k(t)=\tilde{\boldsymbol{\varphi}}_{k-1}^T(t)\Delta\boldsymbol{u}_k(t)+\Delta\boldsymbol{u}_k^T(t)\boldsymbol{P}_{k-1}(t)\Delta\boldsymbol{u}_k(t)$ $(\eta_k(t)-v_k(t))$.

Further, denoting $b_k(t)=\eta_k(t)-v_k(t)-\frac{1+q}{2}\gamma_k(t)$, where q is a constant. Then $\beta_k(t)$ becomes

$$\beta_k(t)=-2\gamma_k(t)b_k(t)-q\gamma_k^2(t) \tag{32}$$

Substituting (32) into (31), and then summing over both sides of it, one has

$$\begin{aligned}\sum_{j=0}^{t}V_k(j)&\le\sum_{j=0}^{t}V_{k-1}(j)-2\sum_{j=0}^{t}\gamma_k(j)b_k(j)\\&\quad-q\sum_{j=0}^{t}\gamma_k^2(j)+2\sum_{j=0}^{t}G_k(j)v_k(j)\\&\quad+2\sum_{j=0}^{t}\Delta\boldsymbol{u}_k^T(j)\boldsymbol{P}_{k-1}(j)\Delta\boldsymbol{u}_k(j)v_k^2(j)\end{aligned} \tag{33}$$

Denoting $S_k(t) = \sum_{l=k_0}^{k}\left\{2\sum_{j=0}^{t}\gamma_l(j)b_l(j)\right\}$, according to the positive real lemma, one has $S_k(t) \geq 0$.

Then, we can further obtain

$$S_k(t) = S_{k-1}(t) + 2\sum_{j=0}^{t}\gamma_k(j)b_k(j) \tag{34}$$

Denoting $M_k(t) = S_k(t) + \sum_{j=0}^{t}V_k(j)$, then from (33) and (34), one can derive

$$\begin{aligned} M_k(t) \leq\ & M_{k-1}(t) - q\sum_{j=0}^{t}\gamma_k^2(j) \\ & + 2\sum_{j=0}^{t}\Delta\boldsymbol{u}_k^T(j)\boldsymbol{P}_{k-1}(j)\Delta\boldsymbol{u}_k(j)v_k^2(j) \\ & + 2\sum_{j=0}^{t}G_k(j)v_k(j) \end{aligned} \tag{35}$$

Taking expectation on both sides of (35), one gets

$$\begin{aligned} & E\left\{M_k(t)\,|\mathscr{F}_{k-1}(t)\right\} \\ & \quad \leq M_{k-1}(t) - qE\left\{\sum_{j=0}^{t}\gamma_k^2(j)\,|F_{k-1}(t)\right\} \\ & \quad\quad + 2\sum_{j=0}^{t}\Delta\boldsymbol{u}_k^T(j)\boldsymbol{P}_{k-1}(j)\Delta\boldsymbol{u}_k(j)\sigma^2 \end{aligned} \tag{36}$$

Denote

$$D_k(t) = q\sum_{l=k_0}^{k}\sum_{j=0}^{t}\gamma_l^2(j) = D_{k-1}(t) + q\sum_{j=0}^{t}\gamma_k^2(j) \tag{37}$$

Let

$$\Xi_k(t) = \left(M_k(t) + D_k(t)\right)\Big/\sum_{j=0}^{t}L_k(j) \tag{38}$$

where $L_k(j) = \left(\ln\left|\boldsymbol{P}_k^{-1}(j)\right|\right)^{\varsigma(j)}$, $\varsigma(j) > 1$.

Taking expectation on both sides of (38), and using (36) and (37), we get

$$\begin{aligned} & E\left\{\Xi_k(t)\,|\mathscr{F}_{k-1}(t)\right\} \\ & \quad = \frac{1}{\sum_{j=0}^{t}L_k(j)}E\left\{M_k(t) + D_k(t)\,|\mathscr{F}_{k-1}(t)\right\} \\ & \quad \leq \Gamma_k(t)\left(M_{k-1}(t) + D_{k-1}(t) + \Theta_k(t)\right) \end{aligned} \tag{39}$$

where $\Gamma_k(t) = \frac{1}{\sum_{j=0}^{t}L_k(j)}$, and $\Theta_k(t) = 2\sum_{j=0}^{t}\Delta\boldsymbol{u}_k^T(j)\boldsymbol{P}_{k-1}(j)$ $\Delta\boldsymbol{u}_k(j)\sigma^2$.

Further, from (38) and (39), it is yielded,

$$\begin{aligned} & E\left\{\Xi_k(t)\,|F_{k-1}(t)\right\} \\ & \quad \leq \Xi_{k-1}(t) + \Gamma_k(t)\Theta_k(t) \\ & \quad\quad - \left(M_{k-1}(t) + D_{k-1}(t)\right)\left(\Gamma_{k-1}(t) - \Gamma_k(t)\right) \end{aligned} \tag{40}$$

Considering the martingale convergence theorem [38], we obtain that

$$\lim_{k\to\infty}M_k(t)\Gamma_k(t) < \infty \tag{41}$$

By virtue of the definition of $M_k(t)$, one obtains

$$\Gamma_k(t)\sum_{j=0}^{t}V_k(j) < \infty$$

That is,

$$V_k(t) = O\left(\sum_{j=0}^{t}L_k(j)\right) \tag{42}$$

Furthermore, according to Property 1, we can finally derive that

$$\left\|\tilde{\boldsymbol{\varphi}}_k(t)\right\| = O\left(\sqrt{\sum_{j=0}^{t}L_k(j)\Big/\lambda_{\min}\left(\boldsymbol{P}_k^{-1}(t)\right)}\right) \tag{43}$$

V. ALGORITHM EXTENDED

A. EXTENSION TO THE CASE WITH COLORED NOISE

Again consider system (1), and replace $v_k(t)$ by $D\left(z^{-1}\right)v_k(t)$, where $D\left(z^{-1}\right)$ represents the polynomial of the unit back-shift operator z^{-1}, *i.e.*,

$$D\left(z^{-1}\right) = 1 + d_1z^{-1} + \cdots + d_{n_d}z^{-n_d},$$

where $d_1, \cdots, d_{n_d}$ are some unknown constants, and n_d is the order of the polynomial. Thus, $D\left(z^{-1}\right)v_k(t)$ represents a colored noise.

Then, following the same steps from (2) - (3), we can get

$$y_k(t+1) = y_{k-1}(t+1) + \Delta\vec{\boldsymbol{u}}_k^T(t)\vec{\boldsymbol{\varphi}}(t) + \Delta v_k(t) \tag{44}$$

where $\vec{\boldsymbol{u}}_k(t) = \left[\boldsymbol{u}_k^T(t), v_k(t-1), \cdots, v_k\left(t-d_{n_d}\right)\right]^T$, and $\vec{\boldsymbol{\varphi}}(t) = \left[\boldsymbol{\varphi}^T(t), d_1, \cdots, d_{n_d}\right]^T$.

Further, following the same step as (5), we can get

$$\boldsymbol{Y}_k(t+1) = \boldsymbol{Y}_{k-1}(t+1) + \Delta\vec{\boldsymbol{U}}_k(t)\vec{\boldsymbol{\varphi}}(t) + \Delta\vec{\boldsymbol{V}}_k(t) \tag{45}$$

where $\vec{\boldsymbol{U}}_k(t) = \left[\vec{\boldsymbol{u}}_1^T(t), \cdots, \vec{\boldsymbol{u}}_k^T(t)\right]^T \in R^{k\times(t+1+n_d)}$.

Designing the same objective function as (6), and solving it, one has

$$\hat{\vec{\boldsymbol{\varphi}}}_k(t) = \left[\Delta\vec{\boldsymbol{U}}_k^T(t)\Delta\vec{\boldsymbol{U}}_k(t)\right]^{-1}\Delta\vec{\boldsymbol{U}}_k^T(t)\boldsymbol{T}_k(t+1) \tag{46}$$

Denote $\vec{\boldsymbol{P}}_k^{-1}(t) = \sum_{j=1}^{k}\Delta\vec{\boldsymbol{u}}_j(t)\Delta\vec{\boldsymbol{u}}_j^T(t) \in R^{(t+1+n_d)\times(t+1+n_d)}$, and following the similar derivation process as (9) - (17), we can obtain the ILRLS algorithm for linear systems with colored noise as follows,

$$\hat{\vec{\boldsymbol{\varphi}}}_k(t) = \hat{\vec{\boldsymbol{\varphi}}}_{k-1}(t) + \vec{\boldsymbol{H}}_k(t)\left(\vartheta_k(t+1) - \Delta\vec{\boldsymbol{u}}_k^T(t)\hat{\vec{\boldsymbol{\varphi}}}_{k-1}(t)\right) \tag{47}$$

$$\vec{\boldsymbol{H}}_k(t) = \frac{\vec{\boldsymbol{P}}_{k-1}(t)\Delta\vec{\boldsymbol{u}}_k(t)}{1 + \Delta\vec{\boldsymbol{u}}_k^T(t)\vec{\boldsymbol{P}}_{k-1}(t)\Delta\vec{\boldsymbol{u}}_k(t)} \tag{48}$$

$$\vec{\boldsymbol{P}}_k(t) = \vec{\boldsymbol{P}}_{k-1}(t) - \vec{\boldsymbol{H}}_k(t)\Delta\vec{\boldsymbol{u}}_k^T(t)\vec{\boldsymbol{P}}_{k-1}(t) \tag{49}$$

It can be found that the ILRLS with colored noises has a similar form as that of (18) - (20), but the dimensions of corresponding variables in the two methods are different.

B. CONVERGENCE ANALYSIS

The following is an additional assumption.

Assumption 4 [23]: $\frac{1}{D(z^{-1})} - \frac{1}{2}$ is strictly positive.

The convergence property of the proposed ILRLS (47) - (49) for the system with colored noises is given in Theorem 2.

Theorem 2: Consider the linear system (1) with colored noises under assumptions 1 – 4. Applying the proposed ILRLS (47) - (49), one can guarantee that the parameter estimation error converges with increasing number of iterations, i.e.,

$$\left\|\tilde{\vec{\varphi}}_k(t)\right\| = O\left(\sqrt{\sum_{j=0}^{t}\vec{L}_k(j)\Big/\lambda_{\min}\left(\boldsymbol{P}_k^{-1}(t)\right)}\right), \quad a.s.,$$

where $\vec{L}_k(j) = \left(\ln\left|\vec{\boldsymbol{P}}_k^{-1}(j)\right|\right)^{\varsigma(j)}$, $\varsigma(j) > 1, j = 0, 1, \cdots, t$, $\tilde{\vec{\varphi}}_k(t) = \hat{\vec{\varphi}}_k(t) - \vec{\boldsymbol{u}}(t)$.

Proof: Since the proof is similar to that of Theorem 1, we omit it wherein for simplicity.

C. EXTENSION TO NONLINEAR SYSTEMS

Consider a nonlinear discrete-time system:

$$\begin{cases} x_k(t+1) = f(x_k(t), \cdots, x_k(t-n_x), u_k(t), \cdots, \\ \qquad\qquad u_k(t-n_u)) \\ y_k(t+1) = x_k(t+1) + v_k(t) \end{cases} \tag{50}$$

where $f(\cdot)$ is a nonlinear function and is continuously differentiable; n_x and n_u are positive integers.

Besides Assumption 1, another assumption [36] is made on (50), as shown below.

Assumption 5: $f(\cdot)$ satisfies globally Lipschitz condition, i.e., $|f(\chi_1, \boldsymbol{u}_1) - f(\chi_2, \boldsymbol{u}_2)| \leq L_\chi |\chi_1 - \chi_2| + L_u \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|$, where $L_\chi < \infty$ and $L_u < \infty$.

According to [36], one has

$$x_k(t+1) = g\left(x_k(0), \boldsymbol{u}_k^T(t)\right) \tag{51}$$

where $g(\cdot)$ is a composite function of $f(\cdot)$ and thus has the same properties as $f(\cdot)$, *i.e.*, $g(\cdot)$ also satisfies assumptions 1 and 5.

Then, the iterative dynamic linearization of (50) can be obtained as,

$$x_k(t+1) = x_{k-1}(t+1) + \Delta\boldsymbol{u}_k^T(t)\boldsymbol{\varphi}_k(t) \tag{52}$$

where $\boldsymbol{\varphi}_k(t) = \frac{\partial g(\cdot)}{\partial \boldsymbol{u}_k(t)} = [\varphi_k(0), \cdots, \varphi_k(t)]^T$.

Note that $\boldsymbol{\varphi}_k(t)$ represents the gradients of function $g(\cdot)$ with respect to input $u_k(\cdot)$. If the nonlinear system (50) is strongly repetitive, $\boldsymbol{\varphi}_k(t)$ can be iteration-independent or becomes slowly iteration-varying at least. Under this consideration, $\boldsymbol{\varphi}_k(t)$ can be denoted as $\boldsymbol{\varphi}(t)$, and then (52) can be rewritten as Eq. (3).

Therefore, the subsequent design and analysis of the ILRLS algorithm for the repetitive nonlinear nonaffine system (50) becomes the same as that for the repetitive linear system (1). In other words, the proposed ILRLS methods (18) - (20) as well as its extension to colored noises (47) - (49) are both capable of identifying the repetitive nonlinear nonaffine systems.

VI. SIMULATION

Example 1: Consider the LTV system (1), where

$$\boldsymbol{A}(t) = \begin{bmatrix} 0.2\exp(-t/100) & -0.6 & 0 \\ 0 & 0.5 & \sin t \\ 0 & 0 & 0.7 \end{bmatrix},$$
$$\boldsymbol{B}(t) = \begin{bmatrix} 0 & 0.3\sin t & 1 \end{bmatrix}^T,$$
$$\boldsymbol{C}(t) = \begin{bmatrix} 0 & 0.1 & 1 + 0.1\cos t \end{bmatrix},$$

$t \in \{0, 1, \cdots, 200\}$, and $v_k(t)$ is a random output noise varying with both the iteration and time. In the simulation, $v_k(t) = 0.1randn$ is shown in Fig. 1. The input $u_k(t)$ is a random sequence with mean 0 and variance 1.

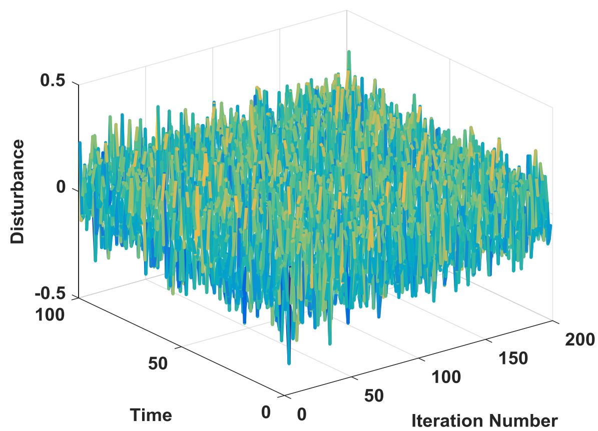

FIGURE 1. White noise sequence in Example 1.

Applying the proposed ILRLS method (18) - (20) and setting $\hat{\boldsymbol{\varphi}}_0(t) = 0.8\boldsymbol{I}$, and $\boldsymbol{P}_0(t) = 0.1\boldsymbol{I}$, where $\boldsymbol{I}$ represents an identity matrix, the relative parameter estimation error, defined as $\varepsilon_k(t) = \left\|\hat{\boldsymbol{\varphi}}_k(t) - \boldsymbol{\varphi}(t)\right\| / \|\boldsymbol{\varphi}(t)\|$, is shown in Fig. 2. From Fig. 2, it is clear that along the direction of increasing iteration number, parameter estimation error decreases. In addition, the output estimate performance of the 200th iteration is shown in Fig. 3, which illustrates that the estimated output approaches closely to the observed one and the presented ILRLS method is effective.

Further, define the mean square sum (MSS) of output estimation error as $\delta_k = \sum_{t=0}^{100}\left(\hat{y}_k(t) - y_k(t)\right)^2 \Big/ 100$ which is shown in Fig. 4. It is seen that the output estimation error decreases gradually along iteration direction, which indicates the effectiveness of ILRLS method.

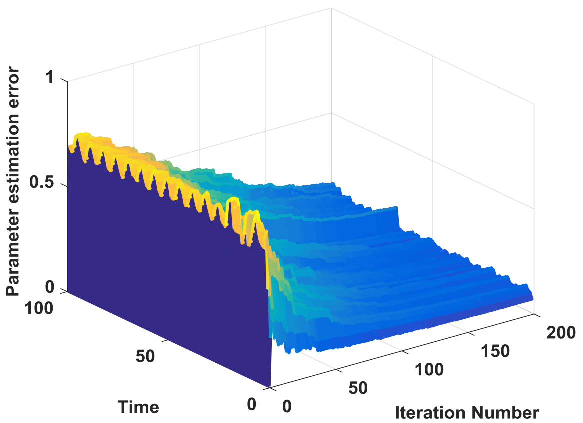

FIGURE 2. Relative parameter estimation error.

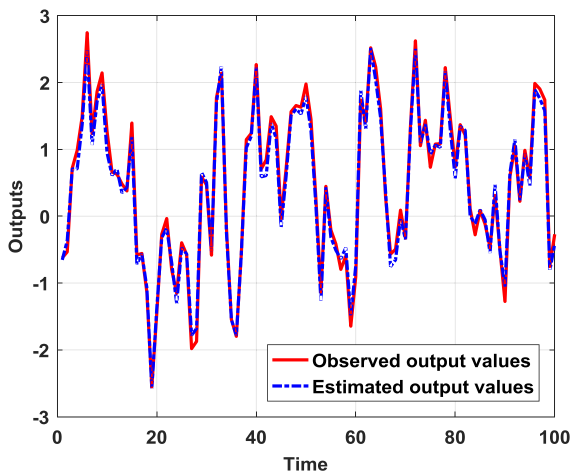

FIGURE 3. The identification performance of the 200th iteration in Example 1.

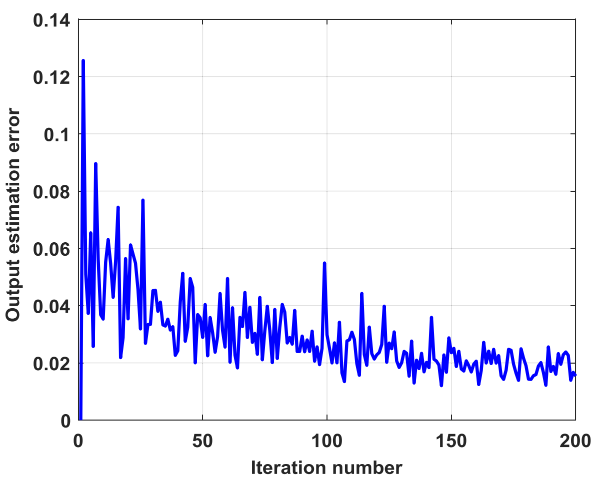

FIGURE 4. The MSS of output estimation error in Example 1.

Example 2: Consider a nonlinear nonaffine system:

$$y_k(t+1) = \frac{-0.9y_k(t) + (1+a(t))\, u_k(t)}{1+y_k(t)^2} + v_k(t)$$

where $a(t) = 4 \times round(t/100) + \sin(t/100)$, and $t \in \{0, 1, \cdots, 500\}$.

In the simulation, $v_k(t)$ is taken as white noise, as shown in Fig. 5. The input signal $u_k(t)$ is generated by $u_k(t) = 0.1\sin(t\pi/100) + 0.1randn$.

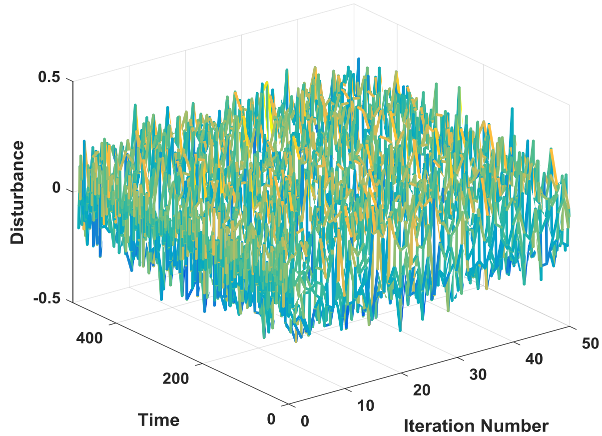

FIGURE 5. White noise sequence in Example 2.

Applying the proposed ILRLS algorithm (18) - (20) and selecting $\hat{\boldsymbol{\varphi}}_0(t) = 40\boldsymbol{I}$, and $\boldsymbol{P}_0(t) = 5\boldsymbol{I}$, Fig. 6 shows the observed and estimated output profiles at the 50th iteration, which indicates that the developed method is applicable for nonlinear systems and achieves a satisfactory identification performance. Accordingly, the MSS of output estimation error δ_{50} is shown in Fig. 7, and we can see from Fig. 7 that the MSS of output estimation error decreases with the increase of iteration number.

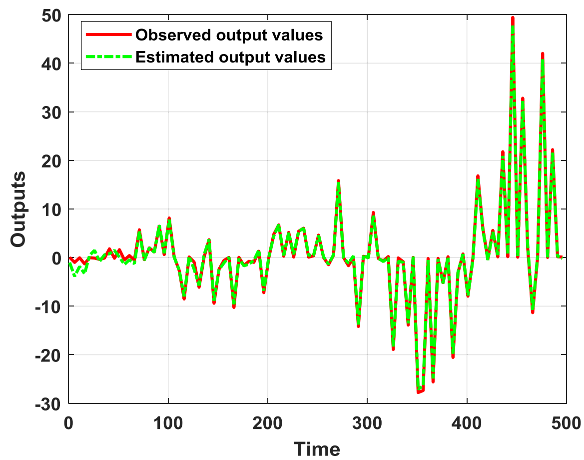

FIGURE 6. The identification performance of the 50th iteration in Example 2.

Further, consider the nonlinear case with colored noise, i.e.,

$$y_k(t+1) = \frac{-0.9y_k(t) + (1+a(t))\, u_k(t)}{1+y_k(t)^2} + D\left(z^{-1}\right) v_k(t)$$

where $D\left(z^{-1}\right) = 1 + 2z^{-1} - 0.8z^{-2}$.

Applying the proposed ILRLS algorithm (47) - (49) and choosing $\hat{\vec{\boldsymbol{\varphi}}}_0(t) = 45\boldsymbol{I}$, $\vec{\boldsymbol{P}}_0(t) = 2\boldsymbol{I}$, the results are given

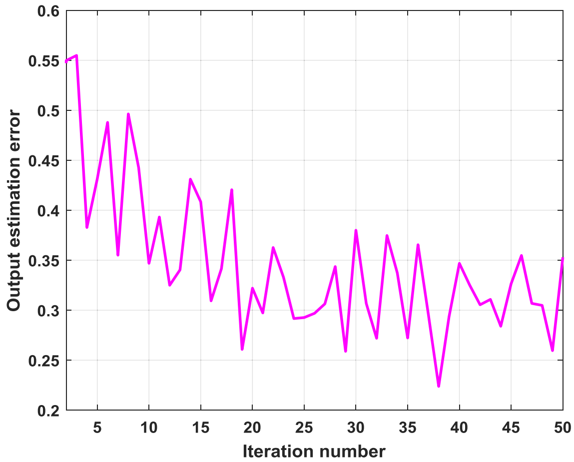

FIGURE 7. The MSS of output estimation error in Example 2.

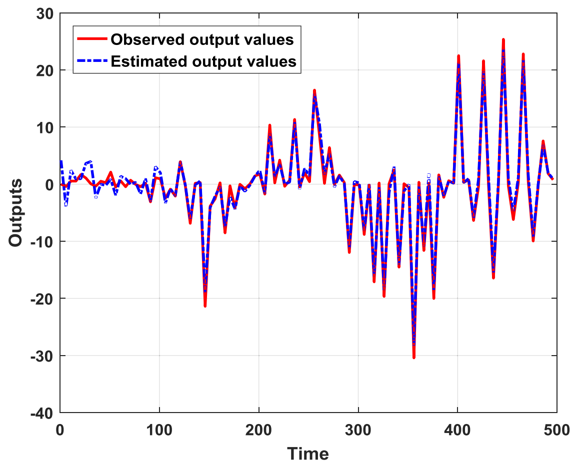

FIGURE 8. The identification performance at the 50th iteration in case of colored noises in Example 2.

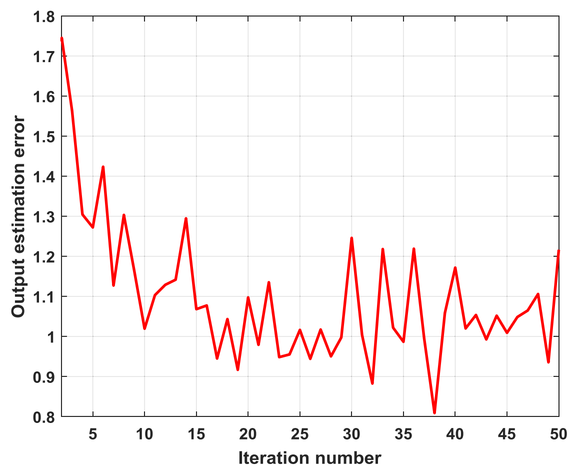

FIGURE 9. The MSS of output estimation error in case of colored noises in Example 2.

in Figs. 8 and 9. According to figures 8 - 9, one can see that the proposed identification algorithm is also applicable to nonlinear systems with colored noise and achieves a good identification performance.

Example 3: To further verify the wide applicability of the proposed method, we apply the proposed ILRLS method to a linear motor drive system. The dynamic model of a linear motor system [27], [39] is given as follows:

$$M\dot{v} = u - F_{fric}(v) + d \tag{53}$$

$$F_{fric}(v) = Bv + \Lambda \arctan(9000v) \tag{54}$$

where $v(m/s)$, $u(V)$, and $M\left(V/m/s^2\right)$ denote the speed, the control input voltage and the nominal mass of the linear motor, respectively. $d(V)$ denotes the time-varying external disturbance. $B = 0.2$ is an equivalent viscous friction coefficient. Λ is a time-varying amplitude.

The running time of the linear motor system is 2 seconds. In this simulation, the sampling time is $h = 0.004\,(s)$. $\Lambda(t) = 0.2\cos(50t/N)$ with $N = 500$. The nominal mass $M = 0.00088$. The disturbance $d(t) = 0.5randn(V)$. Applying the proposed method (18) - (20), the parameters are set as $\hat{\boldsymbol{\varphi}}_0(t) = 7\boldsymbol{I}$ and $\boldsymbol{P}_0(t) = 20\boldsymbol{I}$. The MSS of output estimation error δ_{50} is shown as the red dotted line in Fig. 10, and the Fig. 11 shows the identification performance at the 50th iteration. It is seen from Figs. 10 – 11 that the proposed method can achieve a good output identification performance for the practical linear motor system.

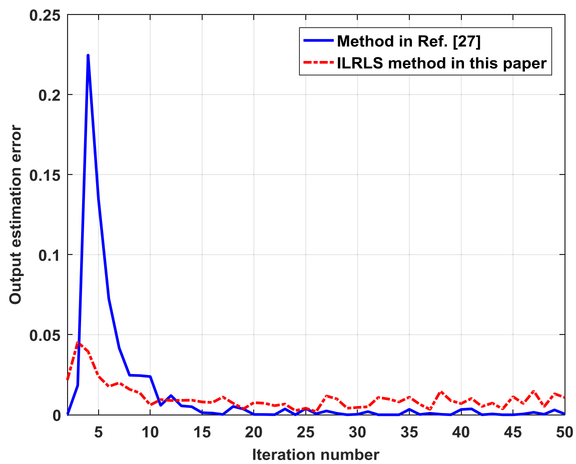

FIGURE 10. The MSS of output estimation error in Example 3.

For the sake of comparison, the iterative identification method in [27] is also applied for the identification of linear motor system (53) - (54). In [27], the linear motor system is transformed into a discrete-time parametric one shown as follows,

$$v_k(t+1) = \boldsymbol{\Phi}_k^T(t)\boldsymbol{\theta}(t) + h/M\left(u_k(t) + M/hv_k(t)\right) \tag{55}$$

where $\boldsymbol{\Phi}_k(t) = [-(h/M)\,v_k(t), -(h/M)\arctan(9000v_k(t)), h/M]$ and the parameters to be identified are $\boldsymbol{\theta}(t) = [B, \Lambda(t), d(t)]^T$. The iterative RLS method proposed in [27] is given as follows:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + \boldsymbol{P}_k(t)\boldsymbol{\Phi}_k(t)\,(y_k(t+1) \\ - \boldsymbol{\Phi}_k^T(t)\hat{\boldsymbol{\theta}}_{k-1}(t) - h/Mu_k(t)\Big)\end{aligned} \tag{56}$$

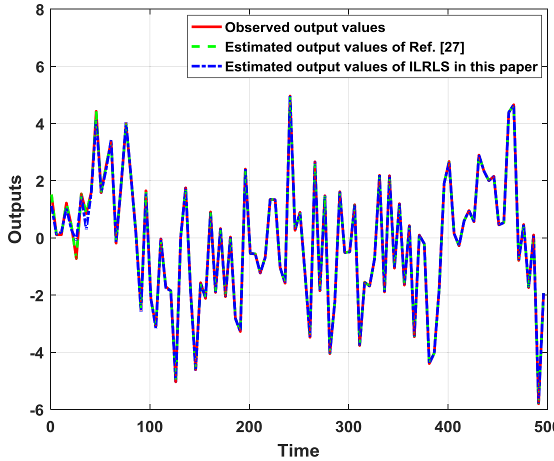

FIGURE 11. The identification performance at the 50th iteration in Example 3.

$$+\boldsymbol{P}_k(t)\boldsymbol{P}_0^{-1}(t)\Delta\hat{\boldsymbol{\theta}}_{k-1}(t)$$

$$\boldsymbol{P}_k(t) = \boldsymbol{P}_{k-1}(t) - \frac{\boldsymbol{P}_{k-1}(t)\boldsymbol{\Phi}_k(t)\boldsymbol{\Phi}_k^T(t)\boldsymbol{P}_{k-1}(t)}{1+\boldsymbol{\Phi}_k^T(t)\boldsymbol{P}_{k-1}(t)\boldsymbol{\Phi}_k(t)} \tag{57}$$

Applying the identification method (56) – (57) with $\boldsymbol{P}_0(t) = diag(10, 0.02, 0.1)$ and $\hat{\boldsymbol{\theta}}_0(t) = [0.5, 0, 0.5]^T$, the simulation results are also shown in figures 10 – 11.

One can see that the traditional iterative identification method [27] achieves a slightly better performance than the proposed one because it utilizes the known model structure information of the linear motor system, i.e., the linear parametric structure. However, the traditional iterative identification method [27] is hard to be directly applied to a nonlinear plant since it is difficult to transfer the plant to a linear parametric model.

VII. CONCLUSION

Two linear time-varying data model based ILRLS identification methods are proposed for repetitive discrete-time systems with white noises and colored noises, respectively. The linear time-varying data model is purely data based without any physical meaning and is used for the identification algorithm design and analysis only. Both the two proposed methods are data-driven, which can also be applied to nonlinear repetitive systems. The arbitrarily fast time-varying uncertainties can well be addressed by using the proposed ILRLS since the identification is done along iterations where the time-varying uncertainties become invariant along with the iteration direction. Mathematical analysis and simulations both verify the theoretical results. However, the nonlinear systems considered in this work are required to be global Lipschitz continuous, and the ILRLS design and analysis for local Lipschitz nonlinear systems are still an open problem, which will be considered in our future work.

REFERENCES

[1] L. Chen, A. Tulsyan, B. Huang, and F. Liu, "Multiple model approach to nonlinear system identification with an uncertain scheduling variable using EM algorithm," *J. Process Control*, vol. 23, no. 10, pp. 1480–1496, Nov. 2013.

[2] P. Gil, F. Santos, L. Palma, and A. Cardoso, "Recursive subspace system identification for parametric fault detection in nonlinear systems," *Appl. Soft Comput.*, vol. 37, pp. 444–455, Dec. 2015.

[3] A. E. David and G. Sansavini, "Identification of critical states in power systems by limit state surface reconstruction," *Int. J. Elect. Power Energy Syst.*, vol. 101, pp. 162–175, Oct. 2018.

[4] J. Wang, J. Su, Y. Zhao, X. Pang, J. Li, and Z. Bi, "Performance assessment of primary frequency control responses for thermal power generation units using system identification techniques," *Int. J. Elect. Power Energy Syst.*, vol. 100, pp. 81–90, Sep. 2018.

[5] T. Tokuda, J. Yoshimoto, Y. Shimizu, G. Okada, M. Takamura, Y. Okamoto, S. Yamawaki, and K. Doya, "Identification of depression subtypes and relevant brain regions using a data-driven approach," *Sci. Rep.*, vol. 8, Sep. 2018, Art. no. 14082. doi: 10.1038/s41598-018-32521-z.

[6] Y. Yang, A. T. Connolly, and M. M. Shanechi, "A control-theoretic system identification framework and a real-time closed-loop clinical simulation testbed for electrical brain stimulation," *J. Neural Eng.*, vol. 15, pp. 1–24, Dec. 2018.

[7] B. Zhang and Z. Mao, "Bias compensation principle based recursive least squares identification method for Hammerstein nonlinear systems," *J. Franklin Inst.*, vol. 354, pp. 1340–1355, Feb. 2017.

[8] M. Kazemi and M. M. Arefi, "A fast iterative recursive least squares algorithm for Wiener model identification of highly nonlinear systems," *ISA Trans.*, vol. 67, pp. 382–388, Mar. 2017.

[9] Y. Zhang, Z. Zhao, and G. Cui, "Auxiliary model method for transfer function estimation from noisy input and output data," *Appl. Math. Model.*, vol. 39, no. 15, pp. 4257–4265, 2015.

[10] H. Ríos, D. Efimov, J. A. Moreno, W. Perruquetti, and J. G. Rueda-Escobedo, "Time-varying parameter identification algorithms: Finite and fixed-time convergence," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3671–3678, Jul. 2017.

[11] F. Martel, C. Chochol, D. Rancourt, S. Chesne, and D. Rèmond, "Identification of time-varying parameters using the derivative formulation of Chebyshev polynomials," in *Proc. 16th IFAC Symp. Syst. Identif.*, Brussels, Belgium, 2012, pp. 1288–1293.

[12] R. Lozano, D. Dimogianopoulos, and R. Mahony, "Identification of linear time-varying systems using a modified least-squares algorithm," *Automatica*, vol. 36, no. 7, pp. 1009–1015, Jul. 2000.

[13] Z. Shi, Y. Wang, and Z. Ji, "Bias compensation based partially coupled recursive least squares identification algorithm with forgetting factors for MIMO systems: Application to PMSMs," *J. Franklin Inst.*, vol. 353, pp. 3057–3077, Sep. 2016.

[14] T. Pu and J. Bai, "Robust identification of discrete-time linear systems with unknown time-varying disturbance," *Digit. Signal Process.*, vol. 83, pp. 271–279, Dec. 2018.

[15] J. Na, J. Yang, X. Ren, and Y. Guo, "Robust adaptive estimation of nonlinear system with time-varying parameters," *Int. J. Adapt Control Signal Process.*, vol. 29, no. 8, pp. 1055–1072, Aug. 2015.

[16] M.-S. Chen, J.-S. Li, and P.-C. Wu, "A matrix gradient algorithm for identification of parameterized time-varying parameters," *Asian J. Control*, vol. 11, no. 1, pp. 59–65, Jan. 2009.

[17] J.-X. Xu, Y.-J. Pan, and T.-H. Lee, "A VSS identification scheme for time-varying parameters," *Automatica*, vol. 39, no. 4, pp. 727–734, Apr. 2003.

[18] O. Sörnmo, B. Bernhardsson, O. Kröling, P. Gunnarsson, and R. Tenghamn, "Frequency-domain iterative learning control of a marine vibrator," *Control Eng. Pract.*, vol. 47, pp. 70–80, Feb. 2016.

[19] N. Ajjanaromvat and M. Parnichkun, "Trajectory tracking using online learning LQR with adaptive learning control of a leg-exoskeleton for disorder gait rehabilitation," *Mechatronics*, vol. 51, pp. 85–96, May 2018.

[20] R. H. Chi, Z. S. Hou, and B. Huang, "Optimal iterative learning control of batch processes: From model-based to data-driven," *Acta Autom. Sinica*, vol. 43, no. 6, pp. 917–934, Jun. 2017.

[21] M. C. Campi, T. Sugie, and F. Sakai, "An iterative identification method for linear continuous-time systems," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1661–1669, Aug. 2008.

[22] M. X. Sun and H. B. Bi, "Learning identification: Least squares algorithms and their repetitive consistency," *Acta Autom. Sinica*, vol. 38, no. 5, pp. 698–706, May 2012.

[23] M.-X. Sun, H.-B. Bi, B.-X. Chen, and H.-G. He, "Learning identification of a class of stochastic time-varying systems with colored noise," *Control Theory Appl.*, vol. 29, no. 8, pp. 974–984, Aug. 2012.

[24] Z. Cao, Y. Yang, J. Lu, and F. Gao, "Two dimensional recursive least squares for batch processes system identification," in *Proc. 10th IFAC Int. Symp. Dyn. Control Process Syst.*, Mumbai, India, 2013, pp. 780–784.

[25] Z. Cao, Y. Yang, J. Lu, and F. Gao, "Constrained two dimensional recursive least squares model identification for batch processes," *J. Process Control*, vol. 24, no. 6, pp. 871–879, Jun. 2014.

[26] Z. Hou, Q. Yu, J.-X. Xu, and C. Yin, "A simultaneous iterative learning control and dynamic modeling approach for a class of nonlinear systems," in *Proc. SYSID*, Beijing, China, 2015, pp. 368–373.

[27] Q. Yu, Z. Hou, and J.-X. Xu, "D-type ILC based dynamic modeling and norm optimal ILC for high-speed trains," *IEEE Trans. Control Syst. Technol.*, vol. 2, no. 2, pp. 652–663, Mar. 2018.

[28] W. Yu and X. Li, "Some new results on system identification with dynamic neural networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 412–417, Mar. 2001.

[29] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Trans. Neural Netw.*, vol. 1, no. 1, pp. 4–27, Mar. 1990.

[30] R. H. Abiyev and O. Kaynak, "Type 2 fuzzy neural structure for identification and control of time-varying plants," *IEEE Trans. Ind. Electron.*, vol. 57, no. 12, pp. 4147–4159, Dec. 2010.

[31] Y. Zhao, A. Fatehi, and B. Huang, "A data-driven hybrid ARX and Markov chain modeling approach to process identification with time-varying time delays," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4226–4236, May 2017.

[32] J. Chen, B. Huang, F. Ding, and Y. Gu, "Variational Bayesian approach for ARX systems with missing observations and varying time-delays," *Automatica*, vol. 94, pp. 194–204, Aug. 2018.

[33] G. M. Ochieng and F. A. O. Otieno, "Data-based mechanistic modelling of stochastic rainfall-flow processes by state dependent parameter estimation," *Environ. Modell. Softw.*, vol. 24, pp. 279–284, Feb. 2009.

[34] B. Huang and R. Kadali, *Dynamic Modeling, Predictive Control and Performance Monitoring: A Data-Driven Subspace Approach*. London, U.K.: Springer-Verlag, 2008, pp. 121–141.

[35] Z. Hou and S. Jin, *Model Free Adaptive Control: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2013, pp. 33–89.

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

[37] L. Huang, *Linear Algebra in System and Control Theory*. Beijing, China: Science Press, 1984, pp. 224–269.

[38] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1984, pp. 411–415.

[39] L. Xu and B. Yao, "Adaptive robust precision motion control of linear motors with negligible electrical dynamics: Theory and experiments," *IEEE/ASME Trans. Mechatronics*, vol. 6, no. 4, pp. 444–452, Dec. 2001.

NA LIN received the M.Sc. degree in automatic control from the Qingdao University of Science and Technology, Qingdao, China, in 2017, where she is currently pursuing the Ph.D. degree with the Institute of Artificial Intelligence and Control, School of Automation and Electronic Engineering. Her research interests include data-driven control and iterative learning control.

RONGHU CHI received the Ph.D. degree in systems engineering from Beijing Jiaotong University, Beijing, China, in 2007. He was a Visiting Scholar with Nanyang Technological University, Singapore, from 2011 to 2012, and a Visiting Professor with the University of Alberta, Edmonton, AB, Canada, from 2014 to 2015. In 2007, he joined the Qingdao University of Science and Technology, Qingdao, China, and is currently a Full Professor with the School of Automation and Electronic Engineering. He has published over 100 articles in important international journals and conference proceedings. His current research interests include iterative learning control, data-driven control, and intelligent transportation systems, and so on.

Dr. Chi was a recipient of the "Taishan Scholarship," in 2016. He served in various positions at international conferences and was an Invited Guest Editor of the *International Journal of Automation and Computing*. He has also served as a Council Member of the Shandong Institute of Automation and a Committee Member of the Data-Driven Control, Learning and Optimization Professional Committee.

BIAO HUANG (M'97–SM'11–F'18) received the B.Sc. and M.Sc. degrees in automatic control from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1983 and 1986, respectively, and the Ph.D. degree in process control from the University of Alberta, Edmonton, AB, Canada, in 1997. In 1997, he joined the Department of Chemical and Materials Engineering, University of Alberta, as an Assistant Professor, where he is currently a Professor. He is the Industrial Research Chair in control of oil sands processes with the Natural Sciences and Engineering Research Council of Canada, and was the Industry Chair of Process Control with Alberta Innovates. He has applied his expertise extensively in industrial practice. His current research interests include process control, data analytics, system identification, control performance assessment, Bayesian methods, and state estimation.

Dr. Huang is a Fellow of the Canadian Academy of Engineering and the Chemical Institute of Canada. He was a recipient of a number of awards, including the Alexander von Humboldt Research Fellowship from Germany, the Best Paper Award from the IFAC *Journal of Process Control*, the APEGA Summit Award in Research Excellence, and the Bantrel Award in Design and Industrial Practice.