IEEE CSS

Received 24 January 2022; revised 2 May 2022; accepted 1 June 2022. Date of publication 21 June 2022; date of current version 2 August 2022. Recommended by Senior Editor Sonia Martinez.

Digital Object Identifier 10.1109/OJCSYS.2022.3185002

A Discrete Fractional Order Adaptive Law for Parameter Estimation and Adaptive Control

MOHAMED ABURAKHIS, RAÚL ORDÓÑEZ (Member, IEEE), AND OUBOTI DJANEYE-BOUNDJOU

Department of Electrical and Computer Engineering, University of Dayton, Dayton, OH 45469 USA

CORRESPONDING AUTHOR: Raúl Ordóñez (e-mail: ordonez@ieee.org)

ABSTRACT In this article, a discrete fractional order adaptive law (DFOAL) is designed based on the Caputo fractional difference to perform parameter estimation of structured uncertainties. The paper provides a rigorous stability analysis of the DFOAL parameter estimation method. The DFOAL is then modified in order to improve parameter estimator performance to show that, under certain conditions, it provides asymptotic convergence to the true parameter values even when the regressor is not persistently exciting. A method to allow for practical implementation of the DFOAL and the modified DFOAL is developed. Finally, the modified DFOAL is used to identify the plant parameters in an indirect adaptive control law for a class of nonlinear discrete-time systems with structured uncertainty.

INDEX TERMS Adaptive control, discrete fractional calculus, discrete-time control system, parameter estimation.

I. INTRODUCTION

Continuous time (CT) control systems have been extensively studied in the literature [1]–[3]. At the same time, there are many systems are best approached from a discrete time (DT) framework. DT control systems may be achieved by discretizing the CT system, or the systems are naturally DT. The integer index k will be used to denote these discrete points in time, and the sampling time will be normalized to be 1 for convenience. For different reasons, uncertainties naturally occur in control systems, where the robustness properties of adaptive control have been extensively investigated [1], [4]–[7].

Parameter estimation can be shown to be convergent using a stable adaptive control technique. A new adaptive controller is proposed in [8] to treat the uncertainties in the plant model and external disturbances. In their work, a finite-time model of learning is incorporated into the $\mathcal{L}_1$-based verified safe control to improve Simplex performance in unexpected environments. A unique and robust direct adaptive control mechanism has been designed and synthesized by the author in [9]. The study investigates how chaos is suppressed and synchronized in the chaotic nuclear spin generator system. They demonstrate that the proposed technique performs well when there are unknown model uncertainties and external disturbances.

In essence, the fundamental problem, therefore, is to design control laws to handle system uncertainties. Adaptive control techniques tune the controller's parameters on-line to achieve and improve closed-loop stability and robustness. In this paper, parameter estimation (for purposes of function approximation) and parametric adaptive control problems will be handled using discrete fractional calculus (DFC).

Fractional order differential equations and, consequently, difference equations can be used to model certain phenomena. Some studies take into account systems that are modeled using fractional order equations. The authors of [10] established a conformable derivative-based first-state estimation scheme for fractional-order systems. In this study, a variation of Barbalat's lemma is used to verify the convergence of estimate errors. The stability of the equilibrium point of the fractional order differential equation based on the Caputo–Katugampola definition is studied in [11]. The complexity of modern control systems has motivated researchers to apply fractional techniques for modeling and controlling these systems. The authors in [12] have studied and proven the stability of the delta fractional order Caputo difference equation using a discrete fractional Lyapunov direct method. In [13], the authors proposed a Lyapunov direct method to study the stability of nonlinear discrete fractional systems implemented based on backward difference. They provided conditions for asymptotic stability. The stability proof is obtained in [12], [13] by applying the fractional difference to a Lyapunov candidate. In our paper, the stability analysis is addressed based on the classical integer order analysis. The author in [14] proposes a new algorithm to identify SISO discrete fractional order linear dynamic systems with error-in-variable. The author in [15] investigated the design of fractional order discrete time controllers, where the performance of the controller was studied based on the Grünwald-Letnikov definition. Since the Grünwald-Letnikov and Riemann-Liouville definitions are equivalent [16], one can be substituted for the other [17].

However, to the best of our knowledge, and unlike the research presented in this paper, there are no studies that use discrete fractional order techniques for controlling classical discrete systems that have already been modeled by integer order difference equations. Here, fractional order calculus is used to increase the degrees of freedom of control and thus provide extra flexibility to the control designer and improve performance. It is worth emphasizing that even though we deal with the DT case, a DT FO derivative retains some of the desirable properties of the CT FO derivatives with the advantage that the DT FO derivatives are easier to implement.

In [18], the author designed a new algorithm for a least-squares parameter estimator, where the learning rate is updated every time step. However, the update law is in the regression form, while our technique, as we will see, is free from this restriction. The proposed adaptive law saves past information in order to yield more design flexibility and improved parameter convergence.

This paper deals with structured uncertainty, unlike unstructured uncertainty, where other mechanics can be used, such as neural networks, fuzzy methods, or further, as the authors in [19] model the lumped uncertainties with Bernstein-type operators.

This paper contributes to the control system community by employing DFC for both stand-alone parameter estimation (for function approximation) and indirect adaptive control (IAC) for a class of nonlinear systems with structured uncertainties. The use of Fractional Order (FO) integrals in adaptation laws brings with it a memory of sorts to the adaptive mechanism. The analysis in this paper is new and useful because it shows that this memory provides interesting stability and convergence properties to the adaptive element, properties that are otherwise not found in traditional integer order adaptation. The fractional order integrator has value within the context of adaptive control, and this paper is a novel attempt at demonstrating this value. The new discrete fractional order adaptive law (DFOAL) is designed to estimate the parameters of a structured uncertainty presented in an integer order DT system. In the stability analysis, we use the integer order backward difference equations to provide stability conditions, which is an additional contribution to the area because we do not require unrealistic fractional derivatives of the plant itself. The final contribution to this paper is the additional modification of the adaptive law to improve the parameter estimator's performance. Under some conditions, this modified DFOAL will be able to provide asymptotic convergence to the true value of the uncertain parameters without the need for persistency of excitation. It only requires sufficient excitation.

The flow of the paper is divided into two parts: first, we setup the problem of stand-alone function approximation; and second, we extend this concept to adaptive control by estimating the parameters of the controller on-line. The parameter estimation is performed using the gradient descent technique. We use the DFC technique to generalize the integer-order classical gradient descent law to a fractional order gradient descent law. The paper is further subdivided into eight sections that are organized as follows: After we present an overview of DFC in Section II, we introduce the problem statement and the classical Integer Order (IO) gradient descent adaptive law in Section III. Section IV gives an overview of the IO gradient update of parameters. Section V introduces the DFOAL, which is modified further in Section VI. Sections V and VI also present a numerical example for each case. Section VII extends the results to indirect adaptive control for a class of DT systems with structured uncertainties. Finally, Section VIII concludes the paper.

II. PRELIMINARIES

In this section, we introduce definitions and concepts of DFC that are relevant to this paper.

Definition 1. [20]–[22]: For a function $y(k)$, the forward and backward difference operators Delta (Δ) and Nabla (∇), respectively, are defined by

$$\Delta^m y(k) = \Delta(\Delta^{m-1} y(k)), \tag{1}$$

$$\nabla^m y(k) = \nabla(\nabla^{m-1} y(k)), \tag{2}$$

where $\Delta y(k) = y(k+1) - y(k)$, $\nabla y(k) = y(k) - y(k-1)$, $m = 1, 2, 3, \ldots$, and $k \in \mathbb{N}_a = \{a, a+1, \ldots\}$.

For convenience, here we choose the initial index to be $a = 0$. Consider the rising factorial $t^{\overline{m}}$, with $t \in \mathbb{R}$ and $m \in \mathbb{N}$, defined by

$$t^{\overline{m}} = t(t+1)(t+2)\ldots(t+m-1).$$

It is also known as t to m rising and can be generalized to a non-integer power.

Definition 2. [20], [21]: The generalization of the raising factorial for non-integer real power ν is

$$t^{\overline{\nu}} = \frac{\Gamma(t+\nu)}{\Gamma(t)}, \tag{3}$$

where $\Gamma(\cdot)$ is the Gamma function, $t \in \mathbb{R} - \{\ldots, -2, -1, 0\}$, and $\nu \in \mathbb{R}$.

The fractional order sum introduced next is a generalization of the integer order one.

Definition 3. [13], [20], [21]: For $f : \mathbb{N}_a \to \mathbb{R}$, the backward fractional sum ($\nabla^{-\nu}$) of order ν is defined by

$$\nabla_a^{-\nu} f(k) = \frac{1}{\Gamma(\nu)} \sum_{s=a}^{k} (k-s+1)^{\overline{\nu-1}} f(s), \tag{4}$$

where $\nu > 0$, a is the initial index and $k \in \mathbb{N}_a$.

In contrast to the operator $\Delta^{-\nu}$, which maps functions from $\mathbb{N}_a$ to $\mathbb{N}_{a+\nu}$, the $\nabla_a^{-\nu}$ operator maps functions from $\mathbb{N}_a$ to $\mathbb{N}_a$ [13], [20], [21]. This is the main reason behind choosing the backward operator to represent the DFOAL in this paper. The transformation between the nabla and the delta operators with an integer order power ($m = 1, 2, \dots$) can be done simply by using the shift operator. However, for a non-integer order, the transformation will also involve a change in the domain. More details about this can be found in [23].

Similar to the two famous cases in continuous time fractional calculus, for $k \in \mathbb{N}$, $m-1 < \nu < m$, and $m = 1, 2, \dots$ the nabla Riemann-Liouville fractional difference can be obtained by taking the integer order difference of the fractional sum of a discrete function $f(k)$, i.e., $\nabla^\nu f(k) = \nabla^m (\nabla^{\nu-m} f)(k)$ while the opposite of it yields the Caputo fractional difference, i.e., ${}^C\nabla^\nu f(k) = \nabla^{\nu-m}\nabla^m f(k)$ [13], [20], [21], [24], [25].

Therefore, the backward Riemann-Liouville fractional difference for order ν, such that $m-1 < \nu < m$, is

$$\nabla_a^{\nu} f(k) = \frac{\nabla^m}{\Gamma(m-\nu)} \sum_{s=a}^{k} (k-s+1)^{\overline{m-\nu-1}} f(s), \tag{5}$$

and the backward Caputo fractional difference for order ν is

$${}^C\nabla_a^{\nu} f(k) = \frac{1}{\Gamma(m-\nu)} \sum_{s=a}^{k} (k-s+1)^{\overline{m-\nu-1}} \nabla^m f(s). \tag{6}$$

The relation between the fractional order difference in the sense of Riemann-Liouville and Caputo is given in the following lemma.

Lemma 1: [26] For $\nu > 0$, the following equality holds:

$$\nabla^{-\nu}\nabla f(k) = \nabla\nabla^{-\nu} f(k) - \frac{(k-a)^{\overline{\nu}}}{\Gamma(1-\nu)} f(a), \quad k \in \mathbb{N}_a. \tag{7}$$

III. PROBLEM STATEMENT

The problems of parameter estimation and IAC when dealing with structured uncertainties will be tackled by using DT fractional order techniques. In the following section, we introduce the general background of the stand-alone parameter estimation to get a full picture of the procedure. As mentioned, in this study, structured uncertainty will be considered where unknown ideal parameters are multiplied by a known accessible regressor to form the uncertainty.

A. STAND-ALONE PARAMETER ESTIMATION

For stand-alone parameter estimation for function approximation, we will consider the integer order (IO) discrete-time system

$$F(k) = \theta^{*\top} \eta(x(k)), \tag{8}$$

which is a linear parametric model [4], [27] in the form of structured measurable uncertainty, and $x, F \in \mathbb{R}$. The unknown parameter vector $\theta^* \in \mathbb{R}^p$ will be estimated on-line, and its estimate is denoted $\hat{\theta}(k) \in \mathbb{R}^p$. Moreover, $\eta(x(k)) \in \mathbb{R}^p$ is known and accessible regressor. We define the computable approximation error

$$\begin{aligned} \psi(k) &= \hat{F}(k) - F(k), \\ &= \hat{\theta}^\top(k)\eta(x(k)) - \theta^{*\top}\eta(x(k)), \\ &= \tilde{\theta}^\top(k)\eta(x(k)), \end{aligned} \tag{9}$$

where $\tilde{\theta}(k) = \hat{\theta}(k) - \theta^*$ is the parameter error vector. In a limited sense, successful function approximation can be said to take place when the instantaneous approximation error $\psi(k)$ converges to zero. This instantaneous convergence, however, does not imply that $\tilde{\theta}(k)$ converges to zero, which is a broader definition of function approximation that implies convergence of $\psi(k)$ to zero. In this paper, we are interested in this broader definition.

B. INDIRECT ADAPTIVE CONTROL

For the IAC problem, we will study an IO dynamic nonlinear system transformed into canonical form. Consider the DT system in feedback linearizable canonical form

$$\begin{aligned} x_1(k+1) &= x_2(k), \\ &\vdots \\ x_{n-1}(k+1) &= x_n(k), \\ x_n(k+1) &= f(x(k)) + u(k), \\ y(k) &= x_1(k), \end{aligned} \tag{10}$$

where $f(x(k))$ are assumed Lipschitz continuous. The goal is to design a controller $u(k) \in \mathbb{R}$ such that the output $y(k) \in \mathbb{R}$ asymptotically tracks a known bounded reference sequence $r(k)$. We assume that the signals $r(k), \dots, r(k+n)$ are available (computable using a reference model or delay). For a bounded reference sequence, the system's states will be bounded if the error is forced to be bounded. The system error will be represented using an error manifold, defined as

$$\begin{aligned} e(k) &= k_1(x_1(k) - r(k)) + k_2(x_2(k) - r(k+1)) + \dots \\ &\quad + k_{n-1}(x_{n-1}(k) - r(k+n-2)) \\ &\quad + x_n(k) - r(k+n-1). \end{aligned} \tag{11}$$

This means that

$$e(k+1) = \chi(k) + f(x(k)) + u(k), \tag{12}$$

where

$$\begin{aligned} \chi(k) &= k_1(x_2(k) - r(k+1)) + k_2(x_3(k) - r(k+2)) + \dots \\ &\quad + k_{n-1}(x_n(k) - r(k+n-1)) - r(k+n). \end{aligned} \tag{13}$$

Recall, from (10), $y(k) = x_1(k)$. Taking the z-transform of both sides of (11), we have that

$$E(z) = k_1(Y(z) - R(z)) + k_2 z(Y(k) - R(z)) + \dots + k_{n-1}z^{n-1}(Y(k) - R(z)) + z^n(Y(k) - R(z)), \quad (14)$$

which means

$$\frac{(Y(z) - R(z))}{E(z)} = \frac{1}{z^{n-1} + k_{n-1}z^{n-2} + \dots + k_2 z + k_1}, \quad (15)$$

where, given the discrete-time z-transform operator $\mathcal{Z}$, $Y(z) = \mathcal{Z}\{y(k)\}$, $R(z) = \mathcal{Z}\{r(k)\}$, and $E(z) = \mathcal{Z}\{e(k)\}$. For stability purposes, the real coefficients k_i $(i = 1, \dots n-1)$ in (15) are chosen according to the Jury stability criterion for placing the poles of the transfer function (15) inside the unit circle. If the controller $u(k)$ is chosen as

$$u(k) = -\chi(k) - f(x(k)) + \kappa e(k), \quad (16)$$

then (12) will become

$$e(k+1) = \kappa e(k). \quad (17)$$

Hence, for $|\kappa| < 1$, the error sequence $e(k)$ will converge exponentially to 0. However, we assume that the plant is subject to uncertainties in $f(x(k))$, and therefore (16) cannot be implemented.

IV. INTEGER ORDER GRADIENT UPDATE LAW

In this paper, the normalized gradient descent adaptive law [4] will be applied to update the estimated parameter vector online. It is given by

$$\hat{\theta}(k) = \hat{\theta}(k-1) - \frac{\gamma\eta(x(k-1))\psi(k-1)}{\alpha + \|\eta(x(k-1))\|^2}, \quad (18)$$

or, using backward difference operator (∇),

$$\nabla\hat{\theta}(k) = \nabla^1\hat{\theta}(k) = -\frac{\gamma\eta(x(k-1))\psi(k-1)}{\alpha + \|\eta(x(k-1))\|^2}, \quad (19)$$

where $\alpha > 0$ and the learning rate $\gamma > 0$. Let $A_m(k) = \frac{\eta(x(k))\eta^\top(x(k))}{\alpha+\|\eta(x(k)\|^2}$. Notice that $A_m(k)$ is a p-by-p bounded symmetric positive semi-definite matrix, and $\lambda_{\max}(A_m(k)) < 1$ for all k, where. in this paper, $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ are respectively used to represent the maximum and minimum eigenvalues operator. Therefore, from (18), the parameter error vector dynamics can be written as

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) - \gamma A_m(k-1)\tilde{\theta}(k-1), \quad (20)$$

or

$$\tilde{\theta}(k+1) = \tilde{\theta}(k) - \gamma A_m(k)\tilde{\theta}(k), = [I_p - \gamma A_m(k)]\tilde{\theta}(k), \quad (21)$$

where I_p is the p-by-p identity matrix. The stability analysis can be done by using either Lyapunov direct method or by performing matrix convergence analysis. Due to space constraints, we only show the latter here.

Note that $\lambda_{\max}(A_m(k)) = \frac{\lambda_{\max}(\eta\eta^\top)}{\alpha+\|\eta\|^2}$, and the matrix $\eta\eta^\top$ has only one non-zero eigenvalue, which is shown in the next Lemma.

Lemma 2: [28], [29] The positive semi-definite matrix $A = \eta\eta^\top$, where η is a p-by-1 non-zero vector, has only one non-zero eigenvalue, given by

$$\lambda(A) = \|\eta\|^2.$$

Matrix $[I - \gamma A_m(k)]$, whose eigenvalues are $1 - \gamma\lambda(A_m(k))$, is a convergent matrix if $|1 - \gamma\lambda(A_m(k))| < 1$, or

$$-1 < 1 - \gamma\lambda(A_m(k)) < 1,$$
$$0 < \gamma\lambda(A_m(k)) < 2,$$
$$0 < \gamma < \frac{2}{\lambda_{\max}(A_m(k))}. \quad (22)$$

It can be verified that (22) would be the same stability condition acquired if using Lyapunov analysis. Now, from Lemma 2 it follows that $\lambda_{max}(A_m(k)) = \frac{\|\eta(x(k))\|^2}{\alpha+\|\eta(x(k))\|^2} < 1$ and subsequently, (22) can be (conservatively) simplified to $0 < \gamma < 2$. For a more robust (but expensive) result, the learning rate γ could be updated every iteration since we have access to the regressor $\eta(x(k))$ and can calculate $\lambda_{\max}(A_m(k))$. Thus, it is sufficient to make sure that, for each $s \in \mathbb{N}_a$,

$$0 < \gamma(s) < 2\left(\frac{\alpha + \|\eta(x(s))\|^2}{\|\eta(x(s))\|^2}\right) \quad (23)$$

is satisfied for $\gamma(s)$. Lastly, notice that (21) can be rewritten as

$$\tilde{\theta}(k) = \tilde{\theta}(0) - \sum_{s=0}^{k-1}\gamma(s)A_m(s)\tilde{\theta}(s). \quad (24)$$

After we present the integer order gradient update law, we transition to a more generalized adaptive law by exploiting fractional calculus techniques.

V. STAND-ALONE PARAMETER ESTIMATION USING DFOAL

Now, we go back to the classical difference (19). This expression can be generalized to a fractional order difference equation of order $0 < \nu \le 1$. That is,

$$\nabla^\nu\hat{\theta}(k) = h(k), \quad (25)$$

where $h(k) = -\frac{\gamma(k)\eta(x(k-1))\psi(k-1)}{\alpha+\|\eta(x(k-1))\|^2}$. Next, we present a useful theorem on the nabla fractional order initial value problem based on the Caputo definition.

Theorem 1: [21] Let $h : \mathbb{N}_{a+1} \to \mathbb{R}$, $\nu > 0$, and $m = \lceil\nu\rceil$, where $\lceil\cdot\rceil$ is the ceiling function. Then, for $0 \le N \le m-1$, the Caputo-based nabla fractional order initial value problem

$${}^C\nabla_a^\nu z(k) = h(k), \; k \in \mathbb{N}_{a+1},$$
$$\nabla^N z(a) = c_N,$$

has the solution

$$z(k)=\sum_{N=0}^{m-1}\frac{(k-a)^{\overline{N}}}{(N+1)!}z(a)+\nabla_a^{-\nu}h(k). \tag{26}$$

Corollary 1: For $h:\mathbb{N}_1\rightarrow\mathbb{R}$ and $0<\nu\leq 1$, then, the Caputo-based nabla fractional order initial value problem

$$^{C}\nabla_0^{\nu}z(k)=h(k),\ k\in\mathbb{N}_1,$$
$$z(0)=z_0,$$

has the solution

$$z(k)=z_0+\nabla_0^{-\nu}h(k),\ k\in\mathbb{N}_1. \tag{27}$$

It can be noticed that the initial conditions take on the same form when dealing with an integer order difference equation, which is similar to its CT counterpart.

Now we are ready to introduce our first main result.

Theorem 2: Consider the integer order discrete system (8). The generalization of the integer order difference ($\nabla^{\nu=1}$) of the normalized gradient descent adaptive law (19) to non-integer ($0<\nu\leq 1$) order guarantees that the approximation error $\psi(k)$ tends to zero asymptotically and the parameter error vector $\tilde{\theta}(k)$ is bounded.

Proof: By applying Corollary 1 to (25), we have the DFOAL in explicit form,

$$\hat{\theta}(k)=\hat{\theta}(0)-\sum_{s=0}^{k-1}\gamma(s)\beta_{s_1}(k,s)\frac{\eta(x(s))}{\alpha+\|\eta(x(s))\|^2}\psi(s), \tag{28}$$

where

$$\beta_{s_1}(k,s)=\frac{1}{\Gamma(\nu)}\frac{\Gamma(k-(s+1)+\nu)}{\Gamma(k-(s+1)+1)}. \tag{29}$$

It is clear that, if $\nu=1$, the kernel $\beta_{s_1}(k,s)=1$. That makes the IO adaptive law a special case of FO one. The order ν is an additional design parameter, and therefore, design flexibility is increased, as will be later illustrated via examples. The fractional order parameter error vector is

$$\tilde{\theta}(k)=\tilde{\theta}(0)-\sum_{s=0}^{k-1}\gamma(s)\beta_{s_1}(k,s)A_m(s)\tilde{\theta}(s), \tag{30}$$

and since $0<\nu\leq 1$ and $s\leq k-1$, we have that $0<\beta_{s_1}(k,s)\leq 1$. By comparing (24) with (30), the boundedness condition for (30) can be deduced,

$$0<\gamma(s)<\frac{2}{\beta_{s_1}(k,s)\lambda_{\max}(A_m(s))}. \tag{31}$$

Since the kernel $\beta_{s_1}(k,s)$ is upper bounded by 1, notice that

$$0<\gamma(s)<\frac{2}{\lambda(A_m(s))}\leq\frac{2}{\beta_{s_1}(k,s)\lambda(A_m(s))}, \tag{32}$$

and thus the DFOAL may expand the range of γ with respect to the range of the IO case in (23). ■

Note that, similar to the IO update law (19), all we can guarantee at this point is boundedness and convergence of $\psi(k)$ to zero, and boundedness of $\tilde{\theta}(k)$, albeit with a greater degree of freedom, as will be illustrated in Section V-B. Parameter convergence can be obtained using the normalized discrete fractional gradient descent (28) if the regressor is persistently excited, as is well known in the literature [4]. However, as will be shown in Section VI, a stronger result can be obtained via a suitable modification of the DFOAL in (28).

Remark 1: The DFOAL in (28) keeps saving the past values of the regressor that is weighted with the kernel β_{s_1} as the time proceeds. The summation is carried out over time, unlike integer-order techniques that use the integer-order integral or sum, which merely update the past estimated values of the parameters. Moreover, the general structure of the DFOAL is similar to the solution of the LTV system, where it is divided into two terms: one of them depends on the initial condition, and the other depends on the input signal. However, the kernel in the DFOAL, i.e., $\beta_{s_1}(k,s)$, which is expressed in terms of the Gamma function, gives the advantage of keeping all the past values recorded in the memory. Notice from (29) that

$$\beta_{s_1}(k,s)=\beta_{s_1}(k-1,s)\frac{k-s+\nu-2}{k-s-1}. \tag{33}$$

Hence, computing the kernel $\beta_{s_1}(k,s)$ under the summation in (28) means that it is not sufficient to know $\beta_{s_1}(k-1,s)$, as (33) shows that the second term to the right of the equality depends on s. At each k step, the DFOAL will include all $k-1$ step in the memory in terms of s. The update approach in DFOAL could be problematic regarding the size of the memory. In the next section we will discuss and solve this problem.

A. IMPLEMENTATION OF DFOAL

An advantage of DFOAL (28) is its ability to save and use all the past values of the parameter vector $\hat{\theta}(k)$, therefore, taking advantage of memory. However, implementing (28) in practice may not be computationally possible for a large k. That is, as k grows larger, saving all previous values of $\hat{\theta}(k)$ becomes more computationally expensive. This problem can be overcome using the following two methods.

Method 1: A desired tolerance ϵ of the absolute approximation error $|\psi|$ can be chosen, i.e., $|\psi|\leq\epsilon$, for some positive scalar ϵ, where the estimated parameters $\hat{\theta}(k)$ reach a certain level of accuracy after which there is no need to save new values of $\hat{\theta}(k)$. Essentially, at the step time k when we stop adding new values because $|\psi|\leq\epsilon$, the kernel $\beta_{s_1}(k,k-1)=1$ and will be fixed there, so that the DFOAL (28) becomes the normalized gradient descent adaptive law (18). It should be noted that switching between two or more algorithms may necessitate reworking the overall stability analysis if the stability results when using either algorithms are different. Due to space constraints, this is not elaborated upon in this paper.

Method 2: This method consists in running the DFOAL (28) for a cycling M-length window, $M>0$, when $k>M$. That is, if $k\leq M$, the DFOAL is implemented as defined by (28).

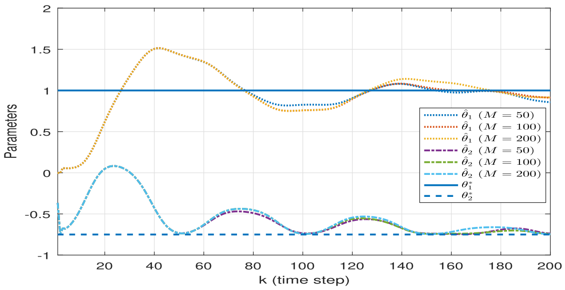


FIGURE 1. Parameter estimation using DFOAL with $\nu = 0.8$ for different choices for *M*.

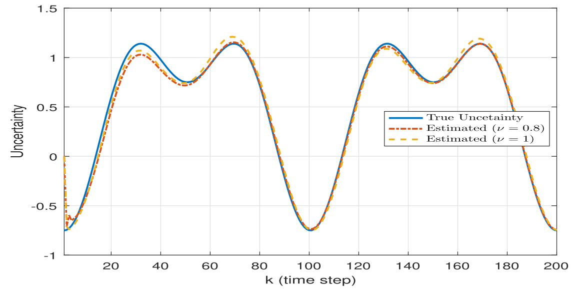


FIGURE 2. Uncertainty estimation using IO ($\nu = 1$) gradient descent adaptive law and DFOAL ($\nu = 0.8$).

However, for $k > M$, it is changed to

$$\hat{\theta}(k) = \hat{\theta}(k - M) - \sum_{s=k-M}^{k-1} \gamma(s)\beta_{s_1}(k, s)\frac{\eta(x(s))}{\alpha + \|\eta(x(s))\|^2}\psi(s). \tag{34}$$

Implementing (34) for $k > M$ means that we only have to save M instances (between $k - 1$ and $k - M$) of parameter vectors $\hat{\theta}(k)$, learning rate $\gamma(k)$, regressor $\eta(x(k))$, and approximation error $\psi(k)$, as opposed to k instance of the same entities if implementing (28) instead. Note that the size of M could be chosen such that it is large enough to conform with the available memory capability. In Lemma 3 below, we address the stability and performance consequences of using Method 2.

B. ILLUSTRATIVE EXAMPLE

Here, we test the performance of DFOAL parameter estimation and compare the choice of fractional order ($\nu = 0.8$) with integer order ($\nu = 1$) for the update law (28). When implementing the DFOAL, we use Method 2 described above.

Consider the DT system described in (8), where the true unknown parameter is $\theta^* = \left[1, -\frac{3}{4}\right]^\top$, and the regressor is $\eta(x(k)) = [\sin(2x(k)), \cos(2x(k))]^\top$, where $x(k) = \pi k/100$ spans the interval $[0, 2\pi]$ and $k = 0, 1, \ldots 200$. It can be noticed that it is persistently exciting since

$$\sum_{x=1}^{2} \eta(x)\eta^\top(x) = \begin{bmatrix} 1.0117 & -0.7185 \\ -0.7185 & 0.6004 \end{bmatrix} \geq 0.0588I.$$

Fig. 1 shows the behavior of the estimated parameters using DFOAL with order $\nu = 0.8$ for different choices for M $(50,\ 100,$ and $200)$.

Fig. 2 plots the true function and its estimate for a fixed learning rate $\gamma = 1.9$ when using DFOAL with $\nu = 0.8$ and $M = 200$ and the IO adaptive law. For the same orders, Fig. 3 shows the behavior of the parameter estimates. The plot shows that the use of the FO adaptive law improves the convergence behavior. The most notable feature is that DFOAL keeps past values of the parameters. On the other hand, the classical integer order adaptive law does not save past values, and instead it performs an instantaneous update. Noticeably, the use of FO adaptation law to estimate an IO system parameter exhibits good behavior and faster parameter convergence.

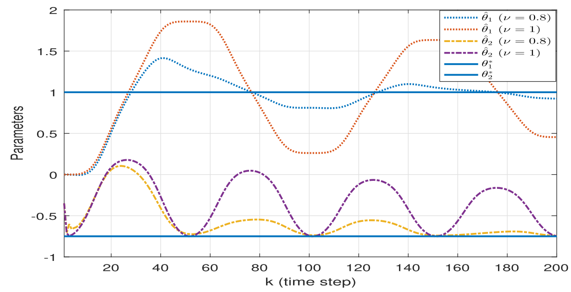


FIGURE 3. Parameter estimation using IO ($\nu = 1$) gradient descent adaptive law and DFOAL ($\nu = 0.8$).

However, the convergence to the true value is still restricted by the need for persistence of excitation of the regressor. In the next section, we will modify the FO adaptive law (28) to obtain a new adaptive law that can help drive the estimated parameters progressively closer to their true value even without persistence of excitation, provided sufficient excitation is present. This section presents the modified adaptive law, the stability analysis, and a practical way to implement it, which is the main intent of this paper.

VI. MODIFIED DFOAL

The DFOAL (28) at time-step $k + 1$ is

$$\begin{aligned}
\hat{\theta}(k+1) &= \hat{\theta}(0) - \sum_{s=0}^{k} \gamma(s)\beta_{s_1}(k+1, s)\frac{\eta(x(s))\psi(s)}{\alpha + \|\eta(x(s))\|^2}, \\
&= \hat{\theta}(0) - \sum_{s=0}^{k-1} \gamma(s)\beta_{s_1}(k+1, s)\frac{\eta(x(s))\psi(s)}{\alpha + \|\eta(x(s))\|^2} \\
&\quad - \gamma(k)\beta_{s_1}(k+1, k)\frac{\eta(x(k))}{\alpha + \|\eta(x(k))\|^2}\psi(k), \\
&= \hat{\theta}(0) - \sum_{s=0}^{k-1} \gamma(s)\beta_{s_1}(k+1, s)\frac{\eta(x(s))\psi(s)}{\alpha + \|\eta(x(s))\|^2} \\
&\quad - \gamma(k)\frac{\eta(x(k))\psi(k)}{\alpha + \|\eta(x(k))\|^2}.
\end{aligned} \tag{35}$$

From the gamma function's property $\Gamma(\nu+1)=\nu\Gamma(\nu)$, we have

$$\begin{aligned}\beta_{s_1}(k+1,s) &= \frac{\Gamma((k+1)-(s+1)+\nu)}{\Gamma(\nu)\Gamma((k+1)-(s+1)+1)},\\ &= \frac{\Gamma(k-(s+1)+\nu)}{\Gamma(k-(s+1)+1)}\\ &\quad+(\nu-1)\frac{\Gamma(k-(s+1)+\nu)}{\Gamma(\nu)\Gamma(k-(s+1)+2)},\\ &= \beta_{s_1}(k,s)-\nu_c\beta_{s_2}(k,s),\end{aligned}$$

with $\nu_c = 1-\nu$ and $\beta_{s_2} = \frac{\Gamma(k-(s+1)+\nu)}{\Gamma(\nu)\Gamma(k-(s+1)+2)}$. Thus,

$$\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(0)-\sum_{s=0}^{k-1}\gamma(s)\left[\beta_{s_1}(k,s)-\nu_c\beta_{s_2}(k,s)\right]\\ &\quad\times\frac{\eta(x(s))}{\alpha+\|\eta(x(s))\psi(s)\|^2}-\gamma(k)\frac{\eta(x(k))\psi(k)}{\alpha+\|\eta(x(k))\|^2},\\ &= \hat{\theta}(0)-\sum_{s=0}^{k-1}\gamma(s)\beta_{s_1}(k,s)\frac{\eta(x(s))\psi(s)}{\alpha+\|\eta(x(s))\|^2}\\ &\quad+\sum_{s=0}^{k-1}\gamma(s)\nu_c\beta_{s_2}(k,s)\frac{\eta(x(s))\psi(s)}{\alpha+\|\eta(x(s))\|^2}\\ &\quad-\gamma(k)\frac{\eta(x(k))\psi(k)}{\alpha+\|\eta(x(k))\|^2},\end{aligned} \tag{36}$$

$$\begin{aligned}&= \hat{\theta}(k)-\gamma(k)\frac{\eta(x(k))\psi(k)}{\alpha+\|\eta(x(k))\|^2}\\ &\quad+\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)\frac{\eta(x(s))\psi(s)}{\alpha+\|\eta(x(s))\|^2}.\end{aligned} \tag{37}$$

From (37), the parameter error vector can be written as

$$\begin{aligned}\tilde{\theta}(k+1) &= \tilde{\theta}(k)-\gamma(k)A_m(k)\tilde{\theta}(k)\\ &\quad+\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)A_m(s)\tilde{\theta}(s),\\ &= \tilde{\theta}(k)-\gamma(k)A_m(k)\tilde{\theta}(k)\\ &\quad+\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)A_m(s)\left[\tilde{\theta}(k)-\tilde{\theta}(k)+\tilde{\theta}(s)\right],\\ &= \tilde{\theta}(k)-\gamma(k)A_m(k)\tilde{\theta}(k)\\ &\quad-\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)A_m(s)\tilde{\theta}(k)\\ &\quad+\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)A_m(s)\left[\tilde{\theta}(k)+\tilde{\theta}(s)\right].\end{aligned} \tag{38}$$

By modifying the DFOAL to be

$$\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(k)-\gamma(k)\frac{\eta(x(k))\psi(k)}{m(k)^2}\\ &\quad-\gamma(k)\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)\frac{\eta(x(s))\psi(s)}{m(s)^2}\\ &\quad-\gamma(k)\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)A_m(s)[\hat{\theta}(k)-\hat{\theta}(s)],\end{aligned} \tag{39}$$

where $m^2(\cdot)=\alpha+\|\eta(x(\cdot))\|^2$, the parameter error vector becomes

$$\begin{aligned}\tilde{\theta}(k+1) &= \tilde{\theta}(k)-\gamma(k)A_m(k)\tilde{\theta}(k)\\ &\quad-\gamma(k)\nu_c\sum_{s=0}^{k-1}\gamma(s)\beta_{s_2}(k,s)\frac{A(s)}{m_s^2}\tilde{\theta}(k).\end{aligned} \tag{40}$$

By letting $L^2(s,k)=\gamma(s)\beta_{s_2}(s,k)$, we have

$$\begin{aligned}\tilde{\theta}(k+1) &= \tilde{\theta}(k)-\gamma(k)\frac{A(k)}{m^2(k)}\tilde{\theta}(k)\\ &\quad-\gamma(k)\nu_c\sum_{s=0}^{k-1}\frac{L(s,k)\eta(x(s))L(s,k)\eta^\top(x(s))}{m^2(s)}\tilde{\theta}(k).\end{aligned} \tag{41}$$

Hence,

$$\begin{aligned}\tilde{\theta}(k+1) &= \tilde{\theta}(k)-\gamma(k)\frac{A(k)}{m^2(k)}\tilde{\theta}(k)-\gamma(k)\nu_c\Omega(k)\tilde{\theta}(k),\\ &= \left(I-\gamma(k)\frac{A(k)}{m^2(k)}-\gamma(k)\nu_c\Omega(k)\right)\tilde{\theta}(k),\end{aligned} \tag{42}$$

where

$$\Omega(k)=\sum_{s=0}^{k-1}\frac{L(s,k)\eta(x(s))L(s,k)\eta^\top(x(s))}{m^2(s)}=WW^\top, \tag{43}$$

and $W\in\mathbb{R}^{p\times k}$ is such that

$$\begin{aligned}W(k) = \Bigg[&\frac{L(0,k)\eta(x(0))}{m^2(0)},\frac{L(1,k)\eta(x(1))}{m^2(1)},\\ &\ldots,\frac{L(k-1,k)\eta(x(k-1))}{m^2(k-1)}\Bigg].\end{aligned} \tag{44}$$

Notice that $\Omega(k)$ is at least positive semi-definite, i.e., $\Omega(k)\geq 0$. We designate matrix Ω as the Information Accumulation Matrix (IAM). Since $\eta\in\mathbb{R}^p$, $\Omega\in\mathbb{R}^{p\times p}$. At the iteration $k=1$, the rank of the IAM is 1 since it results from the multiplication of two vectors [29], and if the regressor η is exciting enough, the IAM can become full rank at iteration $k=k_p$, for some $k_p\geq p$. Therefore, after the k_p^{th} iteration, if the IAM matrix is full rank, then $\lambda_{\min}(\Omega)>0$, which we call full rank condition. Note that this condition is different from the stronger requirement of persistent excitation. Instead, here we only require η to be sufficiently exciting, until some index k_p at which Ω becomes full rank.

It is worth noting that if $\nu = 1$ in the modified DFOAL (39), which implies that $\nu_c = 0$, then it will be reduced to the classical normalized gradient descent adaptive law (18).

Assumption 1: The regressor $\eta(x(k))$ is sufficiently exciting so that there exists an index $k_p \geq p$ at which the matrix $\Omega(k)$ is full rank.

Theorem 3: The use of modified DFOAL (39) for estimating the parameters of the integer order discrete system (8) guarantees that both the approximator and parameter errors $(\psi(k),\ \tilde{\theta}(k))$ converge to zero if the Assumption 1 is satisfied.

Proof: The proof can be given in two ways: the Lyapunov direct method or matrix convergence analysis. Here, we present both analyses to highlight the fact that the matrix convergence method yields a less conservative result.

First, by designing a positive definite, decrescent, radially unbouded function

$$V_{\tilde{\theta}}(\tilde{\theta}(k)) = \tilde{\theta}(k)^\top \tilde{\theta}(k) = \|\tilde{\theta}(k)\|^2, \tag{45}$$

the classical difference operator of (45) is

$$\begin{aligned}\Delta V_{\tilde{\theta}}(\tilde{\theta}(k)) &= \tilde{\theta}(k+1)^\top \tilde{\theta}(k+1) - \tilde{\theta}(k)^\top \tilde{\theta}(k),\\ &= \tilde{\theta}^\top(k)[I - \gamma(k)A_m(k) - \gamma(k)\nu_c\Omega]^2\tilde{\theta}(k)\\ &\quad - \tilde{\theta}(k)^\top\tilde{\theta}(k),\\ &= \tilde{\theta}^\top(k)[-2\gamma A_m(k) + \gamma^2 A_m^2(k)\\ &\quad + 2\gamma^2\nu_c A_m(k)\Omega(k) - 2\gamma\nu_c\Omega(k)\\ &\quad + \gamma^2\nu_c^2\Omega(k)^2]\tilde{\theta}(k).\end{aligned} \tag{46}$$

In this case, we split the matrix

$$\begin{aligned}&-2\gamma A_m(k) + \gamma^2 A_m^2(k) + 2\gamma^2\nu_c A_m(k)\Omega(k)\\ &-2\gamma\nu_c\Omega(k) + \gamma^2\nu_c^2\Omega(k)^2\end{aligned}$$

into two matrices,

$$-2\gamma A_m(k) + \gamma^2 A_m^2(k)$$

and

$$-2\gamma\nu_c\Omega(k) + 2\gamma^2\nu_c A_m(k)\Omega(k) + \gamma^2\nu_c^2\Omega(k)^2.$$

The first matrix would be positive semi-definite if we pick the learning rate γ according to (22).

The second matrix,

$$\begin{aligned}&\left[-2\gamma\nu_c\Omega(k) + 2\gamma^2\nu_c A_m(k)\Omega(k) + \gamma^2\nu_c^2\Omega(k)^2\right]\\ &\quad = -2\gamma\nu_c\Omega(k)\left[I - \gamma A_m(k) - \frac{1}{2}\gamma\nu_c\Omega(k)\right],\end{aligned}$$

is strictly negative if

$$\gamma A_m(k) + \gamma\frac{\nu_c}{2}\Omega(k) < I,$$

or

$$\begin{aligned}&\gamma A_m(k) + \gamma\frac{\nu_c}{2}\Omega(k)\\ &\quad < [\gamma\lambda_{\max}(A_m(k)) + \gamma\frac{\nu_c}{2}\lambda_{\max}(\Omega(k))]I < I.\end{aligned}$$

Thus, we arrive at the condition

$$\gamma\lambda_{\max}(A_m(k)) + \gamma\frac{\nu_c}{2}\lambda_{\max}(\Omega(k)) < 1.$$

Hence, if we pick

$$\begin{aligned}0 < \gamma &< \frac{1}{\lambda_{\max}(A_m(k)) + \frac{\nu_c}{2}\lambda_{\max}(\Omega(k))},\\ &= \frac{2}{2\lambda_{\max}(A_m(k)) + \nu_c\lambda_{\max}(\Omega(k))},\end{aligned} \tag{47}$$

then $\Delta V_{\tilde{\theta}}(\tilde{\theta}(k)) < 0$. Given that $\Delta V_{\tilde{\theta}} = V_{\tilde{\theta}}(k+1) - V_{\tilde{\theta}}(k) = \|\tilde{\theta}(k+1)\|^2 - \|\tilde{\theta}(k)\|^2$ and $\Delta V_{\tilde{\theta}}(\tilde{\theta}(k)) < 0$ also means

$$\|\tilde{\theta}(k+1)\| < \|\tilde{\theta}(k)\|, \tag{48}$$

or, equivalently, $\lim_{k\to\infty}\|\tilde{\theta}(k)\| = 0$.

Second, the stability analysis may also be performed by studying the matrix converge properties. Equation (42) can be written as

$$\tilde{\theta}(k+1) = [I - \gamma(k)A_m(k) - \gamma(k)\nu_c\Omega]\tilde{\theta}(k). \tag{49}$$

Now, recalling Theorem (4.3.1) (Weyl) in [29],

$$\begin{aligned}&\lambda_{\max}(\gamma(k)A_m(k) + \gamma(k)\nu_c\Omega)\\ &\quad \leq \gamma(k)\lambda_{\max}(A_m(k)) + \gamma(k)\nu_c\lambda_{\max}(\Omega).\end{aligned} \tag{50}$$

The eigenvalues of the matrix $[I - \gamma(k)A_m(k) - \gamma(k)\nu_c\Omega]$ are given by $1 - \lambda(\gamma(k)A_m(k) + \gamma(k)\nu_c\Omega)$. Thus, we pick the learning rate $\gamma(k)$ such that the eigenvalues are located within the unit circle, that is,

$$\left|1 - \lambda(\gamma(k)A_m(k) + \gamma(k)\nu_c\Omega)\right| < 1,$$

or

$$0 < \lambda(\gamma(k)A_m(k) + \gamma(k)\nu_c\Omega) < 2. \tag{51}$$

From (50), (51) results in

$$0 < \gamma(k)\left(\lambda_{\max}(A_m(k)) + \nu_c\lambda_{\max}(\Omega)\right) < 2.$$

Hence,

$$0 < \gamma(k) < \frac{2}{\lambda_{\max}(A_m(k)) + \nu_c\lambda_{\max}(\Omega)}. \tag{52}$$

Notice that this condition is less conservative than (47). In the end, from (49), if γ is chosen according to (52), then (51) is verified and, consequently, (48) is also verified.

Therefore, picking γ such that (47) or (52) is satisfied, when using the modified DFOAL, we have proven (48), which means that $\|\tilde{\theta}(k)\|$ is bounded and it asymptotically converges to zero. ■

The ultimate result of [18] is asymptotic stability of the approximation error and boundedness of the parameter error. Here, in the ideal case where all the past information can be saved, the asymptotic stability of the origin $e = 0$, $\tilde{\theta} = 0$ follows. However, computational limit of memory in practice imposes a restriction on the ideal results. In the next remark, this problem is addressed and a practical solution is presented, similar to Section V-A, that recovers the ideal results and yields practical stability of $[e, \tilde{\theta}^\top]^\top$.

Remark 2: To remedy the computational limit problem, in practice, the implementation of the modified DFOAL can be done as described by Method 2 in Section V-A. That is, the modified DFOAL is implemented for a cycling M-length window. More specifically, if $k \leq M$, the modified DFOAL, as given by (39), should be used as is. However, for $k > M$, it is then changed to

$$\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(k) - \gamma(k)\frac{\eta(x(k))\psi(k)}{m(k)^2}\\ &\quad - \gamma(k)\nu_c \sum_{s=k-M}^{k-1} \gamma(s)\beta_{s_2}(k,s)\frac{\eta(x(s))\psi(s)}{m(s)^2}\\ &\quad - \gamma(k)\nu_c \sum_{s=k-M}^{k-1} \gamma(s)\beta_{s_2}(k,s)A_m(s)[\hat{\theta}(k) - \hat{\theta}(s)].\end{aligned} \tag{53}$$

It is worth mentioning that the practical expression of the modified DFOAL, i.e. (53), is obtained by using (34) and following the same steps from (35) to (39).

The following lemma discusses the stability consequences of using the practical implementation of the modified DFOAL. It shows that the parameter error vector is bounded and, furthermore, decreases and tends toward zero with every iteration granted Ω is positive definite.

Lemma 3: Using the practical implementation of the modified DFOAL given by (53), $\|\tilde{\theta}(k)\|$ remains bounded. Furthermore, every time Ω is positive definite, the size of $\|\tilde{\theta}(k)\|$ further shrinks, and $\tilde{\theta}(k)$ asymptotically approaches zero until the next cycling window.

Proof: From (53), the parameter error can be expressed as (42), where, this time around,

$$\Omega(k) = \sum_{s=k-M}^{k-1} \frac{L(s,k)\eta(x(s))L(s,k)\eta^\top(x(s))}{m^2(s)} = W\,W^\top, \tag{54}$$

and $W \in \mathbb{R}^{p\times M}$ is such that

$$W(k) = \left[\frac{L(M-k,k)\eta(x(M-k))}{m^2(M-k)}, \ldots, \frac{L(k-1,k)\eta(x(k-1))}{m^2(k-1)}\right]. \tag{55}$$

The matrix $\Omega = WW^\top$ may become positive definite within the time frame $k \leq M$, as long as, first, $M \geq p$ and, second, p linearly independent regressors $\frac{L(s,k)\eta(x(s))}{m^2(s)}$, $s = k-M,\ k-M+1, \ldots, k-1$, can be found within this time frame. This is because, as mentioned above, having p linearly independent regressors in W makes W full row rank and, subsequently, Ω positive definite.

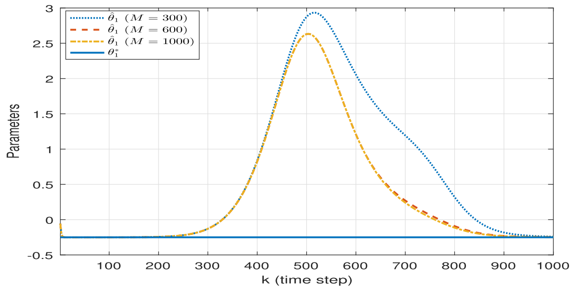
For $p \leq k \leq M$, we have two cases: either Ω is positive definite, i.e., $\Omega > 0$, or Ω is only positive semi-definite, i.e., $\Omega \geq 0$. If $\Omega > 0$, then from (48),

$$\|\tilde{\theta}(k)\| < \|\tilde{\theta}(k-1)\| < \|\tilde{\theta}(0)\|.$$

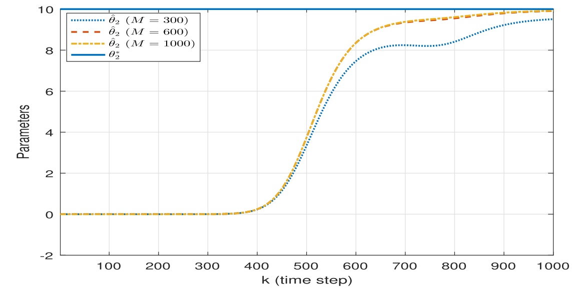
Otherwise, with $\Omega \geq 0$,

$$\|\tilde{\theta}(k)\| \leq \|\tilde{\theta}(k-1)\| \leq \|\tilde{\theta}(0)\|.$$

For $k > M$, we have the previous two cases again. If $\Omega > 0$, then again from (48), $\|\tilde{\theta}(k)\| < \|\tilde{\theta}(k-1)\|$. However, if $\Omega \geq 0$, then $\|\tilde{\theta}(k)\| \leq \|\tilde{\theta}(k-1)\|$. ■



(a) $\hat{\theta}_1$ behavior.



(b) $\hat{\theta}_2$ behavior.

FIGURE 4. Parameter estimation using modified DFOAL with $\nu = 0.8$ for different choices for *M*.

A. MODIFIED DFOAL ILLUSTRATIVE EXAMPLE

The modified adaptive law (39) will be applied to the example presented in [27],

$$F(k) = [\theta_1^*, \theta_2^*]\begin{bmatrix} 1 \\ \exp\left(-\left(\frac{x(k)-\frac{\pi}{2}}{2}\right)^2\right)\end{bmatrix}, \tag{56}$$

where $[\theta_1^*, \theta_2^*]^\top = [-0.25, 10]^\top$ is a true unknown parameter vector, $x(k) \in [-2\pi, 2\pi]$, where $k = 1, 2, \ldots, 1000$.

Fig. 4 shows the behavior of the estimated parameters $\hat{\theta}(k)$ using modified DFOAL with order $\nu = 0.8$ for different choices for M (200, 600), and the ideal case approximated by $M = 1000$. It can be observed that, the larger M is, the faster the convergence of the parameter estimates.

Fig. 5 shows the plot of the true function and its estimate using the modified DFOAL for $M = 1000$ and $\nu = 0.8$, $\nu = 0.9$, and $\nu = 1$. In this example, the regressor is not persistently exciting. However, it is exciting for some period of time, where the modified DFOAL is sufficiently stimulated and can

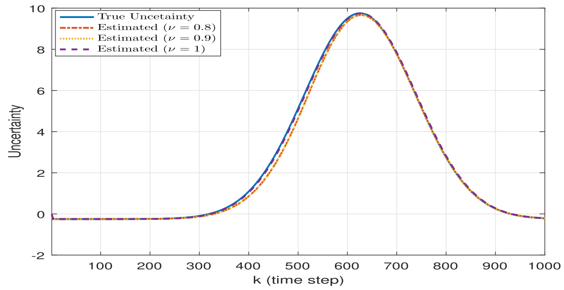
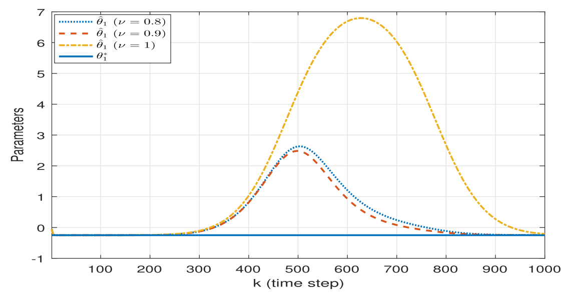
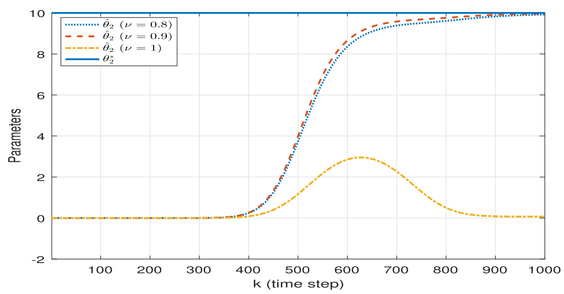


FIGURE 5. Uncertainty estimation using modified DFOAL ($\nu = 0.8, 0.9$) and IO ($\nu = 1$).



(a) $\hat{\theta}_1$ behavior.



(b) $\hat{\theta}_2$ behavior.

FIGURE 6. Parameter estimation using modified DFOAL ($\nu = 0.8, 0.9$) and IO ($\nu = 1$).

therefore identify the unknown parameters. Fig. 6 shows the behavior of the parameter estimates. The plot shows that the estimated parameters using (53) converge to their true values, in spite of the lack of persistency of excitation, whereas the parameters do not converge when using the IO update law (18).

The above analysis, results, and simulation for the function approximation will now be extended and applied to the DT control system. The modified DFOAL will be used to estimate the parameters of the indirect adaptive controller in the next section.

VII. INDIRECT ADAPTIVE CONTROL

In this section, we will approximate the unknown scalar plant dynamic $f(x(k))$ in (10) and use their approximations to construct the adaptive controller. Consider the following two assumptions, show in (57).

Assumption 2: The function $f(x(k))$ is in the form of structured uncertainty, i.e., they are formed by the multiplication of an unknown ideal parameter vector and a known accessible regressor vector.

According to Assumption 2, there exist ideal but unknown constant parameter vectors $\theta^* \in \mathbb{R}^p$ with $p \geq 1$ and corresponding measurable regressor vectors $\eta(x(k)) \in \mathbb{R}$ such that

$$f(x(k)) = \theta^{*\top}\eta(x(k)), \tag{57}$$

By letting $\hat{\theta}(k)$ be the approximations of θ^* then their corresponding parameter error vectors are $\tilde{\theta} = \hat{\theta} - \theta^*$.

Based on the *linear parametric models* of $f(x(k))$ in (57) its approximations is given as

$$\hat{f}(x(k)) = \hat{\theta}^\top(k)\eta(x(k)). \tag{58}$$

Referring back to (16), we define the feedback indirect adaptive controller as

$$\begin{aligned} u_I(k) &= -\chi(k) - \hat{f}(x(k)) + \kappa e(k), \\ &= -\chi(k) - \hat{\theta}^\top(k)\eta(x(k)) + \kappa e(k), \end{aligned} \tag{59}$$

where $|\kappa| < 1$. By replacing $u(k)$ with $u_I(k)$ in (12), the expression $e(k+1)$ becomes

$$e(k+1) = \kappa e(k) - \tilde{\theta}^\top(k)\eta(x(k)). \tag{60}$$

Now, let

$$\psi(k) = \tilde{\theta}^\top(k)\eta(x(k)). \tag{61}$$

From (60) and (61), notice that

$$\psi(k) = \kappa e(k) - e(k+1), \tag{62}$$

which means that

$$\psi(k-1) = \kappa e(k-1) - e(k). \tag{63}$$

It should be added that $\psi(k-1)$ is computable for all $k \geq 1$ because we always have access to the error $e(k)$.

Theorem 4: Consider the discrete-time feedback linearized system (10) under Assumption 2. Under the full rank condition, the use of modified DFOAL (39) will guarantee that the error manifold of the system $e(k)$ (11) and the parameter error $\tilde{\theta}(k)$ converge to zero asymptotically.

Proof: The proof will be performed by using Lyapunov direct method. By referring to [1], [30], consider the Lyapunov function candidate

$$V(e, \tilde{\theta}) = c_1 V_e(e) + c_2 V_{\tilde{\theta}}(\tilde{\theta}), \tag{64}$$

where c_1, c_2 are constants to be picked, and

$$V_e(e) = e^2(k), \tag{65}$$

$$V_{\tilde{\theta}}(\tilde{\theta}) = \tilde{\theta}(k)^\top\tilde{\theta}(k). \tag{66}$$

Clearly, V is a positive definite, decrescent, and radially unbounded function. From (62) we have

$$e^2(k+1) - e^2(k) = \psi^2(k) - 2\kappa\psi(k)e(k) + (\kappa^2 - 1)e^2(k). \quad (67)$$

We know that for any $a,\ b \in \mathbb{R}$, $(a+b)^2 \geq 0$, which leads to $-2a^2 \pm 2ab \leq b^2 - a^2$, thus we can bound $e^2(k+1) - e^2(k)$ by

$$\begin{aligned}\Delta V_e(e) &= e^2(k+1) - e^2(k),\\ &\leq -\frac{1-\kappa^2}{2}e^2(k) + c_3\psi^2(k),\end{aligned} \quad (68)$$

where $c_3 = \frac{1+\kappa^2}{1-\kappa^2} > 1$. For IAC we will use the modified DFOAL for estimating the plant parameters. We repeat (45) through (47) to suit (68), thus

$$\begin{aligned}\Delta V_{\tilde{\theta}}(\tilde{\theta}(k)) &= \tilde{\theta}^\top(k)[-2\gamma A_m(k) + \gamma^2 A_m^2(k)]\tilde{\theta}(k)\\ &+ \tilde{\theta}^\top(k)\left[2\gamma^2\nu_c A_m(k)\Omega(k) - 2\gamma\nu_c\Omega(k)\right.\\ &\left.+\gamma^2\nu_c^2\Omega(k)^2\right]\tilde{\theta}(k),\\ &\leq -\left[2 - \frac{\eta^\top(x(k))\eta(x(k))}{m(k)^2}\right]\frac{\gamma\psi(k)^2}{m(k)^2}\\ &- 2\gamma\nu_c\lambda_{\min}(\Omega(k))\left[1 - \gamma\left(\lambda_{\max}(A_m(k)).\right.\right.\\ &\left.\left.+\frac{\nu_c}{2}\lambda_{\max}(\Omega(k))\right)\right]\tilde{\theta}^\top(k)\tilde{\theta}(k).\end{aligned} \quad (69)$$

Denote $c_4 = 2 - \frac{\gamma\|\eta(k)\|^2}{m(k)^2}$ and let

$$c_5 = \left[1 - \gamma\left(\lambda_{\max}(A_m(k)) + \frac{\nu_c}{2}\lambda_{\max}(\Omega(k))\right)\right]. \quad (70)$$

By picking

$$0 < \gamma < \frac{2m(k)^2}{\|\eta(k)\|^2}, \quad (71)$$

we obtain $0 < c_4 < 2$. Note that, by using Lemma 2, (71) can also be satisfied by using (22). The more restrictive condition of the learning rate γ presented in (47) would force c_5 to be $0 < c_5 < 1$.

From (64), (68), and (69),

$$\begin{aligned}\Delta V(e,\tilde{\theta}) &= c_1\Delta V_e(e) + c_2\Delta V_{\tilde{\theta}}(\tilde{\theta}),\\ &\leq -c_1\frac{1-\kappa^2}{2}e^2 - c_2c_5\gamma\nu_c\lambda_{\min}(\Omega(k))\tilde{\theta}^\top(k)\tilde{\theta}(k)\\ &-\left[\frac{c_2c_4\gamma}{m(k)^2} - c_1c_3\right]\psi^2(k).\end{aligned} \quad (72)$$

Let $c_6 = \frac{c_2c_4\gamma}{m(k)^2} - c_1c_3$. If we choose c_1, c_2 to satisfy $c_2 \geq \frac{c_1c_3}{\gamma c_4}m(k)^2$, then $c_6 > 0$. Hence,

$$\Delta V(e,\tilde{\theta}) \leq -c_7V(e,\tilde{\theta}), \quad (73)$$

where $c_7 = \max\{\frac{1-\kappa^2}{2}, c_5\gamma\nu_c\lambda_{\min}(\Omega(k))\}$. Hence, all signals are bounded, and the tracking error converges to zero as k tends to infinity. Moreover, if the IAM becomes full rank at some index k_p, then the error parameter vector will also converge to zero, and consequently $\hat{\theta}(k)$ will asymptotically converge to its true value θ^*. ■

Remark 3: The procedure from Remark 2 applies to the IAC case, as well as its corresponding parameter convergence conclusions.

A. ILLUSTRATIVE NUMERICAL EXAMPLES

In this section, we will apply the IAC technique to a DT system that is subjected to structured uncertainty.

Consider the DT system in canonical form,

$$\begin{aligned}x_1(k+1) &= x_2(k),\\ x_2(k+1) &= x_3(k),\\ x_3(k+1) &= \theta^{*\top}\eta(x(k)) + u(k),\\ y(k) &= x_1(k),\end{aligned} \quad (74)$$

where $\theta^* = [1, -1]^\top$ is considered as an unknown parameter vector and $\eta(x(k)) = [x_1(k)^2,\ x_3(k)]^\top$. The control law (59) will be applied to the DT system (74) to drive the output $y(x)$ to track a reference sequence defined by

$$r(k) = \exp\left(-\frac{1}{4}\frac{\left(k - \frac{k_f}{2}\right)^2}{k_f}\right), \quad (75)$$

where k_f is the final iteration. The modified DFOAL (39) is utilized to approximate the controller parameters for different orders, as is the classical integer order adaptive law, which, essentially, is the DFOAL for $\nu = 1$ for the sake of comparison. The error manifold, in this example, is defined by

$$e(k) = k_1(x_1(k) - r(k)) + k_2(x_2(k) - r(k+1)) + x_3 - r(k+2), \quad (76)$$

and consequently

$$\chi(k) = k_1(x_2(k) - r(k+1)) + k_2(x_3(k) - r(k+2)) - r(k+3). \quad (77)$$

We choose $\kappa = 0.3$, $k_1 = 0.8$ and $k_2 = -1.6$ so that the roots of the polynomial $z^2 + k_2z + k_1$ are $0.8 \pm i0.4$, which are located inside the unit circle. The initial vector state is $x(0) = [1, 1, 1]^\top$.

Fig. 7 shows the plot of the output $y(k)$ and the reference sequences $r(k)$ when using the modified DFOAL for orders $\nu = 0.6$, $\nu = 0.8$, and $\nu = 1$ where the structure is reduced to the classical integer order adaptive law. Fig. 8 shows the behavior of the parameter convergence for $\hat{\theta}_1$ and $\hat{\theta}_2$ for the same orders respectively. The plot shows that the estimated parameters converge to their true value when using the modified DFOAL, but they do not converge when using the classical IO adaptive law.

As expected from the stability proof, the classical integer order adaptive law updates the controller's parameters to bring the tracking error to zero and the parameters error remains bounded, while the modified DFOAL forces the tracking error to zero and additionally brings the estimated

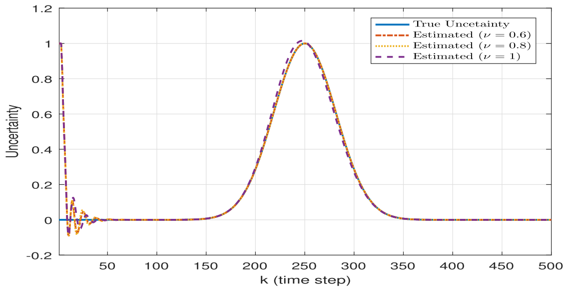
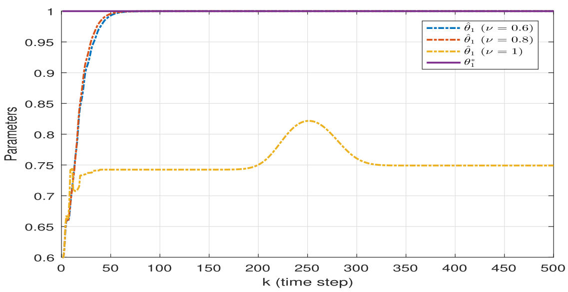
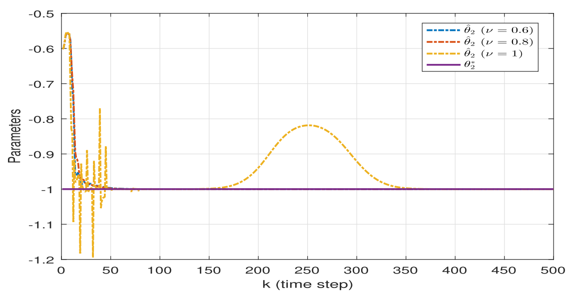


FIGURE 7. **Uncertainty estimation using modified DFOAL ($\nu = 0.6,\ 0.8$) and IO ($\nu = 1$).**



(a) $\hat{\theta}_1$ behavior.



(b) $\hat{\theta}_2$ behavior.

FIGURE 8. **Parameter estimation using modified DFOAL ($\nu = 0.6,\ 0.8$) and IO ($\nu = 1$).**

parameters to their true values. However, the parameter identification is still governed by the full rank condition of the IAM.

Remark 4: Compared to the concurrent learning technique presented in [27], the regressor η in this technique is naturally perturbed by the kernel β_{s_2}, which could help with the excitation level. Note that this technique does not select items that are stored in the IAM. However, the concurrent learning technique picks the information such that the minimum eigenvalue of the information matrix is maximized. Also, be aware that neither technique guarantees parameters convergence unless the full rank condition of the information matrix is satisfied.

VIII. CONCLUSION

We studied the stability inherent in using the classical gradient descent adaptive law for estimating DT systems that are subjected to structured uncertainty. We showed that the stability condition can also be used if we generalize the difference equation for the adaptive law to a fractional order difference equation. Because it uses the past values of the regressor, DFOAL provides a better parameter estimation performance than the classical gradient descent algorithm. Conversely, the classical gradient descent algorithm only performs an instantaneous update of the values of the parameters. This can clearly be noticed from (37) and its subsequent comments. We further modified the DFOAL to improve its performance, in order to provide asymptotic parametric convergence in the parameter estimation problem without the need for persistency of excitation. We presented a way to implement the DFOAL and the modified DFOAL to remedy the computational limit problem and make them practical, which yields asymptotic parametric convergence over a cycling M-window, as long as the regressor is sufficiently (not presciently) excited and the IAM is positive definite within some cycling windows. Furthermore, we used the modified DFOAL to identify the parameters of the control law for a class of nonlinear DT systems as a means of achieving closed-loop stability. The modified DFOAL saves past information in the IAM. The identification of the true values of the unknown parameters depends on the definiteness of the IAM, for which, convergence is restricted by the full rank condition of the IAM.

REFERENCES

[1] J. T. Spooner, M. Maggiore, R. Ordóñez, and K. M. Passino, *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximation Techniques*, vol. 15. New York, NY, USA: Wiley, 2002.
[2] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*. Hoboken, NJ, USA: Wiley, 1995.
[3] H. K. Khalil, *Noninear Systems*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, pp. 5–1, 1996.
[4] G. Tao, *Adaptive Control Design and Analysis*, vol. 37. Hoboken, NJ, USA: Wiley, 2003.
[5] B. Egardt, *Stability of Adaptive Controllers*, vol. 20. Berlin, Germany: Springer, 1979.
[6] B. Egardt, "Global stability analysis of adaptive control systems with disturbances," in *Proc. Joint Autom. Control Conf.*, 1980, no. 17, pp. 59–64.
[7] P. A. Ioannou and J. Sun, *Robust Adaptive Control*, vol. 1. Englewood Cliffs, NJ, USA: Prentice-Hall PTR, 1996.
[8] Y. Mao, Y. Gu, N. Hovakimyan, L. Sha, and P. Voulgaris, "Sl1-simplex: Safe velocity regulation of self-driving vehicles in dynamic and unforeseen environments," 2020, *arXiv:2008.01627*.
[9] I. Ahmad, "A Lyapunov-based direct adaptive controller for the suppression and synchronization of a perturbed nuclear spin generator chaotic system," *Appl. Math. Comput.*, vol. 395, 2021, Art. no. 125858.
[10] A. Jmal, M. Elloumi, O. Naifar, A. Ben Makhlouf, and M. A. Hammami, "State estimation for nonlinear conformable fractional-order systems: A healthy operating case and a faulty operating case," *Asian J. Control*, vol. 22, no. 5, pp. 1870–1879.
[11] A. Jmal, A. Ben Makhlouf, A. Nagy, and O. Naifar, "Finite-time stability for Caputo–Katugampola fractional-order time-delayed neural networks," *Neural Process. Lett.*, vol. 50, no. 1, pp. 607–621, 2019.
[12] D. Baleanu, G.-C. Wu, Y.-R. Bai, and F.-L. Chen, "Stability analysis of Caputo–like discrete fractional systems," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 48, pp. 520–530, 2017.
[13] G.-C. Wu, D. Baleanu, and W.-H. Luo, "Lyapunov functions for Riemann–Liouville-like fractional difference equations," *Appl. Math. Comput.*, vol. 314, pp. 228–236, 2017.
[14] D. Ivanov, "Identification discrete fractional order linear dynamic systems with errors-in-variables," in *Proc. IEEE East-West Des. Test Symp.*, 2013, pp. 1–4.
[15] J. Machado, "Discrete-time fractional-order controllers," *Fractional Calculus Appl. Anal.*, vol. 4, pp. 47–66, 2001.

[16] M. Aburakhis and A. Abdusamad, "Stand-alone function approximation using fractional order techniques," in *Proc. IEEE 1st Int. Maghreb Meeting Conf. Sci. Techn. Autom. Control Computer Eng. MI-STA*, 2021, pp. 98–103.
[17] S. Pooseh, R. Almeida, and D. F. Torres, "Discrete direct methods in the fractional calculus of variations," *Comput. Math. Appl.*, vol. 66, no. 5, pp. 668–676, 2013.
[18] M. Krstic, "On using least-squares updates without regressor filtering in identification and adaptive control of nonlinear systems," *Automatica*, vol. 45, no. 3, pp. 731–735, 2009.
[19] A. Izadbakhsh, H. Gholizade-Narm, and A. Deylami, "Observer-based adaptive controller design for chaos synchronization using Bernstein-type operators," *Int. J. Robust Nonlinear Control,* vol. 32, no. 7, pp. 4318–4335, May 2022.
[20] F. M. Atıcı and P. W. Eloe, "Discrete fractional calculus with the nabla operator," *Electron. J. Qualitative Theory Differ. Equ.*, vol. 2009, no. 3, pp. 1–12, 2009.
[21] C. Goodrich and A. C. Peterson, *Discrete Fractional Calculus*. Berlin, Germany: Springer, 2015.
[22] W. G. Kelley and A. C. Peterson, *Difference Equations: An Introduction With Applications*. Cambridge, MA, USA: Academic, 2001.
[23] T. Abdeljawad and F. M. Atici, "On the definitions of nabla fractional operators," in *Abstract Applied Analysis*, vol. 2012. London, U.K.: Hindawi, 2012.
[24] T. Abdeljawad, "On Riemann and Caputo fractional differences," *Comput. Math. Appl.*, vol. 62, no. 3, pp. 1602–1611, 2011.
[25] T. Abdeljawad, "On delta and nabla Caputo fractional differences and dual identities," *Discrete Dyn. Nature Soc.*, vol. 2013, pp. 1–12, 2013.
[26] T. Abdeljawad and D. Baleanu, "Fractional differences and integration by parts.," *J. Comput. Anal. Appl.*, vol. 13, no. 3, pp. 574–582, 2011.
[27] O. Djaneye-Boundjou and R. Ordóñez, "Parameter identification in structured discrete-time uncertainties without persistency of excitation," in *Proc. IEEE Eur. Control Conf.*, 2015, pp. 3149–3154.
[28] M. Aburakhis and Y. N. Raffoul, "Function approximation using a discrete fractional order gradient descent law," *Int. J. Difference Equ.*, vol. 15, no. 1, pp. 1–10, 2020.
[29] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[30] O. Djaneye-Boundjou and R. Ordonez, "Discrete-time indirect adaptive control of a class of single state systems using concurrent learning for parameter adaptation," in *Proc. IEEE Int. Symp. Intell. Control*, 2016, pp. 1–6.

MOHAMED ABURAKHIS received the B.S. degree from the University of Tripoli, Tripoli, Libya, in 2002, and the Ph.D. degree in electrical and computer engineering from the University of Dayton, Dayton, OH, USA, in 2019. In 2007, he continued to pursue the M.S. degree in control engineering from the University of Tripoli. His research interests include employing fractional calculus in adaptive control and parameters estimation.

RAÚL ORDÓÑEZ (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from Ohio State University, Columbus, OH, USA, in 1996 and 1999, respectively. He spent two years as an Assistant Professor with the Department of Electrical and Computer Engineering, Rowan University, Glassboro, NJ, USA, and then joined the ECE Department with the University of Dayton, Dayton, OH, where he has been since 2001 and is currently a full Professor. He has worked with the IEEE Control Systems Society as a member of the Conference Editorial Board of the IEEE Control Systems Society since 1999, Publicity Chair for the 2001 International Symposium on Intelligent Control, member of the Program Committee and Program Chair for the 2001 Conference on Decision and Control, and Publications Chair for the 2008 IEEE Multi-conference on Systems and Control. Dr. Ordóñez has served as an Associate Editor for the *International Control Journal Automatica* since 2006. He is a coauthor of the textbook *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques* (Wiley, 2002). He is also coauthor of the research monograph *Extremum Seeking Control and Applications - A Numerical Optimization Based Approach* (Springer, 2011).

OUBOTI DJANEYE-BOUNDJOU received the M.S. and Ph.D. degrees in electrical engineering, with an emphasis on control systems engineering, from the University of Dayton, Dayton, OH, USA, in 2013 and 2017, respectively. He spent two and a half years as a joint Adjunct Faculty and a Postdoctoral Researcher with the Department of Electrical Engineering, University of Dayton, and is currently working as a data scientist in the industry. His research interests include system identification, adaptive learning, adaptive control, artificial intelligence, and machine learning.