ISA Transactions 69 (2017) 31–42

Contents lists available at ScienceDirect

ISA Transactions

journal homepage: www.elsevier.com/locate/isatrans

Research article

Convergence of fractional adaptive systems using gradient approach

Javier A. Gallegos [a,b,*], Manuel A. Duarte-Mermoud [a,b]

[a] Department of Electrical Engineering, University of Chile, Av. Tupper 2007, Santiago, Chile
[b] Advanced Mining Technology Center, University of Chile, Av. Tupper 2007, Santiago, Chile

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form
23 February 2017
Accepted 23 April 2017
Available online 5 May 2017

Keywords:
Adaptive systems
Fractional systems
Gradient method

ABSTRACT

Conditions for boundedness and convergence of the output error and the parameter error for various Caputo's fractional order adaptive schemes based on the steepest descent method are derived in this paper. To this aim, the concept of sufficiently exciting signals is introduced, characterized and related to the concept of persistently exciting signals used in the integer order case. An application is designed in adaptive indirect control of integer order systems using fractional equations to adjust parameters. This application is illustrated for a pole placement adaptive problem. Advantages of using fractional adjustment in control adaptive schemes are experimentally obtained.

1. Introduction

It is well known the relevance of adaptive systems in modern engineering together with the fact that most of them are based on gradient decent algorithms mainly because of its simplicity, wide range of applications and effectiveness. For instances model reference adaptive control and adaptive observers, which have been used to resolve many engineering problems [1,2]. Although the gradient approach gives clues to design adaptive systems, analytic tools coming from system theory are additionally required to prove stability and convergence of its relevant variables.

Equations defined by fractional operators have a non-local property (entailed in the definition of fractional derivatives as integrals), which could potentially improve the performance of adaptive systems under parameter variations and external perturbations when the adaptive laws are defined using fractional derivatives. Roughly speaking, the non-locality would counterbalance the past data with the present one obtaining smoother solutions. This has already motivated the introduction of fractional calculus in some proposals of adaptive schemes. The pioneering design is found in [3], which showed through simulations some advantages in speed of convergence and stability, in comparison with integer adjustment by the adequate choice of the derivation order. As a consequence, the derivation order of adaptive laws is a relevant optimization variable. However, a theoretical foundation for the error convergence analysis has not been reported in the revised literature. This is the main motivation for this work.

Our contribution is to give explicit conditions for convergence of the error to its minimum for the simplest fractional scheme of the steepest descent approach. This problem remains unsolved in the literature as stated in the recent work [4]. An explicit application of the analytic results obtained in this paper to indirect adaptive control of integer systems is presented. This problem also has not an analytic solution in the revised literature, see for instance the recent paper [5] where a time discretization is required to adjust the parameters and no convergence condition is provided. The application itself is a contribution since in [4] no specific application was proposed. Looking relax the convergence condition, we will also study more complicated schemes to adjust parameters. These schemes show additional features as robustness under external perturbations and arbitrary speed of convergence.

Our approach is based on the continuous adjustment and can be applied in extreme seeking problems. In [6] a discrete approach is considered for an specific problem, though none condition on the convergence is obtained. A different approach using sliding mode is studied in [7], though no convergence to the minimum is achieved and the error remains oscillating around zero with small amplitude. Our approach uses the well studied Caputo derivative, however others authors have studied a distributed model approach to the fractional derivative concept applying it to error models (see [8]), though only stability is proved and many issues of the Lyapunov approach used remain unclear (e.g. $V(0)$ could be infinite, the stability is proved for internal variables but not for the relevant ones). Though a theoretical argumentation of practical advantages of using the fractional approach in adaptive schemes is not expressly undertaken here, some advantages in transient and robustness behavior as compared to integer parameter adjustments are obtained by simulations.

The paper is organized as follows. In Section 2 necessary background concepts and propositions are stated. Section 3 presents

* Corresponding author at: Department of Electrical Engineering, University of Chile, Av. Tupper 2007, Santiago, Chile.
E-mail addresses: jgallego@ing.uchile.cl (J.A. Gallegos), mduartem@ing.uchile.cl (M.A. Duarte-Mermoud).

http://dx.doi.org/10.1016/j.isatra.2017.04.021

convergence results for a simple fractional adaptive scheme, and the concept of sufficiently exciting signals is introduced. Section 4 posits convergence results for various fractional adaptive schemes based on the concept of sufficiently exciting signals. In Section 5, an example of application is developed. Finally, in Section 6 some conclusions and open problems on this subject are mentioned.

2. Preliminaries

Some definitions and properties used throughout the paper are presented in this section. These are mainly taken from [9] except where indicated.

Consider $f: [0, T] \to \mathbb{R}$ a function.

Definition 1. *The fractional integral of order* $\alpha \geq 0$ *of* $f \in \mathcal{L}^1([0, T])$, i.e. $\int_0^t |f(\tau)| d\tau < \infty$, *is defined as*

$$I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau) d\tau.$$

for $t \in (0, T]$.

Among the many fractional derivative definitions, the following assures implementable adaptive laws (as we will see in Section 3.1).

Definition 2. *The Caputo derivative of order* $\alpha > 0$ *of function* $f \in C^n([0, T])$, i.e. *f having continuous first n derivatives, is defined as*

$$D^\alpha f(t) = I^{n-\alpha} D^n f(t),$$

where $n = \lceil \alpha \rceil$.

An analogue to the fundamental theorem of integer calculus is stated in the next two properties for Caputo fractional derivative.

Property 1. *If f belongs to* $C^n[a, b]$, *then for all* $t \in [a, b]$

$$I^\alpha D^\alpha f(t) = f(t) - \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!} t^k. \tag{1}$$

Property 2. *If f belongs to* $\mathcal{L}^\infty[(a, b)]$, *the Lebesgue space of bounded functions on the interval* (*a*,*b*), *then for all* $t \in [a, b]$

$$D^\alpha I^\alpha f(t) = f(t). \tag{2}$$

The next results will be regularly cited along this work. It is assumed that $0 < \alpha \leq 1$. The first is an important property of Caputo Derivative (also proved for differentiable functions in [10]).

Property 3. [[11], *Lemma* 1]*Let* $x(\cdot)$ *be an absolutely continuous function, then for all* $t \geq 0$ *it holds that*

$$D^\alpha x^2(t) \leq 2x(t) D^\alpha x(t). \tag{3}$$

The next lemma is refereed as the Comparison Principle.

Lemma 1. *[[12], Lemma 6.1] If* $x(0) = y(0)$ *and* $D^\alpha x(t) \geq D^\alpha y(t)$ *for all* $t \geq 0$, *then* $x(t) \geq y(t)$ *for all* $t \geq 0$.

The following lemma is referred as Babalat Lemma.

Lemma 2. *[[13], Lemma 9] If* $f: [0, \infty) \to \mathbb{R}$ *is a uniformly continuous function such that* $\lim_{t\to\infty} \int_0^\infty |f(\tau)| d\tau$ *exists and is finite, then* $\lim_{t\to\infty} f(t) = 0$.

The following theorem plays a key part to establish convergence results.

Theorem 1. *[[14], Theorem 4] Consider the following Caputo system,*

$$D^\alpha x(t) = A(t) x(t) \tag{4}$$

where $x: [0, \infty) \to \mathbb{R}^n$ *and* $A: [0, \infty) \to \mathbb{R}^{n\times n}$. *Let* $f: [0, \infty) \to \mathbb{R}$ *be an non negative differentiable function such that* $I^\alpha[f] \to \infty$ *as* $t \to +\infty$. *Let* $x(0) \in \mathbb{R}^n$ *be any initial condition. If* $A(t) \leq -f(t)I$ *holds for all* $t \geq 0$, *where I is the identity matrix, and the components of matrix A are of class* $C^1(\mathbb{R}_+)$ *then x converges to zero.*

3. Fractional adaptation using gradient method

Based on the fact that the (integer) gradient of a function $J(\cdot) \in \mathbb{R}$ is orthogonal to its level curve, this method proposes to minimize the function $J: \mathbb{R}^n \to \mathbb{R}$ by means of a sequence given by

$$\phi_{n+1} = \phi_n - \gamma_n \nabla_\phi J(\phi_n), \quad n \geq 0, \tag{5}$$

where γ_n are suitably chosen real numbers.

A classical result states that if J is a convex differentiable function and ∇J is a Lipschitz continuous function, the sequence (5) converges to $\arg\min_\phi J(\phi)$.

We will look for a continuous generalization of the gradient method, using fractional operators. Although the level set is given by $J(\phi(\tau)) = C$ and $D^\alpha[J(\phi(\tau))] = 0$, we cannot conclude that $<\nabla_\phi J, d\phi/d\tau> = 0$, because the chain rule from integer calculus has not general validity in fractional calculus. Therefore, there are no reasons to replace in (5) the integer order gradient by the fractional one. Instead, for an objective function of the form $J = J(t, \phi)$, a continuous adjustment given by

$$D^\alpha \phi = -\gamma(t) \nabla_\phi J(t, \phi), \tag{6}$$

where the integer gradient of function $J = J(t, \phi)$ is still used.

To motivate this generalization, we note first that α allows us to have an extra degree of freedom whereby a most general optimization process is obtained. Second, we are imposing the minimizing direction $-\nabla_\phi J(t, \phi)$ to the adjustment. Third, for $\alpha = 1$, by infinitesimal integration, $\phi(t + dt) - \phi(t) = -\gamma(t) \nabla_\phi J(t, \phi)$ which gives the idea of a direct generalization of sequence (5) to the Eq. (6) and though for $\alpha < 1$ this fact does not hold since fractional derivative is non-local, we can find a similarity with the non-local equation $\phi(t) - \phi(0) = -I^\alpha[\gamma(t) \nabla_\phi J(t, \phi)]$, since for $\alpha = 1$ it gives $\phi(t) - \phi(0) = -\int_0^t \gamma(t) \nabla_\phi J(t, \phi) dt$.

When $J = J(\phi)$, $\alpha = 1$, and ϕ is on a Hilbert space, there exist results of weak convergence of ϕ to $\arg\min_\phi J(\phi)$, under smoothness and convexity hypotheses. Also, convergence of $|\nabla_\phi J|$ to zero is guaranteed (see [15]).

3.1. Convergence of fractional type I error model

We will study the case of a quadratic objective function $J(t, \phi) = e^2(t)$ with

$$e(t) = \phi(t)^T w(t) \tag{7}$$

where $w: [0, \infty) \to \mathbb{R}^n$ is a vector function of time, $\phi: [0, \infty) \to \mathbb{R}^n$ is called the parameter error and e is the error function. Using Caputo fractional derivative adaptive laws of the form (6) for $\gamma(t) = 1$ we get

$$D^\alpha \phi(t) = -e(t) w(t) = -w(t) w^T(t) \phi(t). \tag{8}$$

Eqs. (7) and (8) with $\alpha = 1$ are known as Error Model of Type I [1]. Eq. (7) is found in adaptive identification problems and adaptive control problems of integer order linear systems (Chapters 2 and 3 of [2], respectively). For instance, consider the problem of estimating the true constant parameter vector $\theta \in \mathbb{R}^n$ of the linear system expressed in non-minimal realization (linear regression) by

$$y(t) = \theta^T w(t) \tag{9}$$

where $y(t) \in \mathbb{R}$ is the output of the system and $w(t) \in \mathbb{R}^n$ is the state vector (in non-minimal realization) or the information vector (in linear regression), both being known signals. The estimated system output at the instant of time t is denoted as $\hat{y}(t) \in \mathbb{R}$ and it will be given by

$$\hat{y}(t) = \hat{\theta}^T(t) w(t) \tag{10}$$

where $\hat{\theta}(t) \in \mathbb{R}^n$ is the estimate of the true parameter θ at instant of time t. By defining $e(t) := \hat{y}(t) - y(t)$ and $\phi(t) := \hat{\theta}(t) - \theta$, Eqs. (7) and (8) are obtained.

By the election of Caputo derivative as the fractional derivative, Eq. (8) is implementable since it is equivalent to

$$D^\alpha \phi(t) = D^\alpha(\hat{\theta}(\cdot) - \theta) = D^\alpha \hat{\theta}(t) = -e(t)w(t). \tag{11}$$

since Caputo derivative of a constant is zero.

The following elementary result establishes the first difference between fractional and integer order adjustment,

Proposition 1. *For system (7) and (8) assume that w is a bounded, continuously differentiable and uniformly continuous function and $0 < \alpha \le 1$. Then,*

(i) e and ϕ are bounded uniformly continuous functions, $I^\alpha e^2 < \infty$ and $\liminf_{t\to\infty} e^2(t) = 0$. In addition, $\phi = 0$ is an stable global point of the system.

(ii) For $\alpha = 1$, e converges to zero.

(iii) For $0 < \alpha < 1$ there exists a uniformly continuous function w such that e does not converge to zero. But for every $\epsilon > 0$, $e^2(t) > \epsilon$ only can occur at time intervals of finite length with such intervals necessarily finite or occur with an unbounded increasing separation of time. Moreover, the root mean square value (RMS) of e converges to zero.

Proof. (i) For $0 < \alpha \le 1$, by choosing $2V = \phi^T\phi$ and assuming that w is continuous, it follows that $D^\alpha V \le -e^2 \le 0$ (Property 3 can be applied since ϕ is continuously differentiable for $t > 0$ by Property 12 in [13]). By $\alpha-$ integrating the last inequality, $V(t) - V(0) \le -I^\alpha e^2$. Hence $V(\phi(t)) \le V(\phi(0))$ and thus ϕ and e are bounded. Hence, $I^\alpha e^2$ is also bounded. Then, by Property 15 in [13], $\liminf_{t\to\infty} e^2 = 0$. Moreover, e^2 is uniformly continuous because ϕ is uniformly continuous by Proposition 1 in [13] using that ϕ has bounded $\alpha-$ derivative and using the boundedness of ϕ and e. Finally, since $\|\phi(t)\| \le \|\phi(0)\|$ for every $t \ge 0$ and any $\phi(0) \in \mathbb{R}^n$, we conclude global stability of the origin, using the stability concept suited to fractional systems ([16], Section 3.1).

(ii) For $\alpha = 1$, $e^2 \in \mathcal{L}^1(\mathbb{R})$. By Lemma 2, e will converge to zero and, in particular, ϕ will be asymptotically orthogonal to w.

(iii) For $0 < \alpha < 1$, let us consider first the scalar case. Without loss of generality, we suppose that $\phi(0) > 0$. Let $w^2 = f$ be a function that does not converge to zero with $I^\alpha f(t) < 1/2$ for all $t \ge 0$. This function exists because of Proposition 14 in [13] where it was proved that there exists a uniformly continuous function g not converging to zero such that $I^\alpha g < C$. Thus, function f is obtained by choosing $f = g/(2C)$, which is also uniformly continuous.

By noting that $\phi(t) = \phi(0) - I^\alpha[\phi w^2](t)$ is a solution of Eq. (8) and $0 < \phi(t) < \phi(0)$ (Property 9 in [14]), it follows that $I^\alpha[\phi w^2](t) < \phi(0) I^\alpha w^2(t) < 1/2\phi(0)$, whereby

$\phi(t) \ge \phi(0)(1 - I^\alpha w^2(t))$.

Therefore $\phi(t) \ge 1/2\phi(0)$, and since w does not converge to zero, e does not converge to zero either. The same arguments go through for the case when it is assumed that $\phi(0) < 0$.

For the vector case, it is enough take $w^T = (f, 0, 0, \ldots, 0)$ since in that case $e(t) = \phi_1(t) f(t)$ and the analysis for ϕ_1 is therefore the same as the scalar case.

For $\alpha < 1$ and for every $\epsilon > 0$, $e^2(t) > \epsilon$ only occurs at time intervals of finite length, from the uniform continuity of e^2. Such intervals necessarily occur with an increasing separation of time, otherwise $I^\alpha e^2$ diverges, since there would exist a finite time T large enough where always occurs one of those intervals in the intervals $[iT, (i+1)T]$ for any $i \in \mathbb{N}$ whereby, using Example 5 of [14], $I^\alpha e^2$ diverges. That the RMS value of e converges to zero, follows (by continuity of the root function) from the fact that

$$t^{-1}\int_0^t e^2(\tau)d\tau = t^{-1} I^{1-\alpha}[I^\alpha e^2] \le C t^{-1} t^{1-\alpha} \to 0 \tag{12}$$

when $t \to \infty$, where C is the bound of $I^\alpha e^2$.

Considering Proposition 1 there are two ways of solving the optimization problem; to make ϕ asymptotically orthogonal to w or to constraint the set of functions w to those that converge to zero. In the following, we study the first alternative by making ϕ to tend to zero. To simplify the notation, we will introduce the following definitions.

Definition 3. Let I_n be the $n \times n$ identity matrix and $\mathbb{R}^+$ the set of all real numbers greater or equal to zero. Define in the space of bounded, continuously differentiable functions the following subset:

$$SE(n, \alpha) := \Big\{ w: [0, \infty) \to \mathbb{R}^n | (\exists f_w: [0, \infty) \to [0, \infty))\ st. \lim_{t\to\infty} I^\alpha f_w(t) = \infty \wedge (\forall t > 0) w(t)w^T(t) \ge f_w(t) I_n \Big\}, \tag{13}$$

$$PE(n, 1) := \Big\{ w: [0, \infty) \to \mathbb{R}^n | (\exists \epsilon, T_0 > 0)\ st. (\forall t > 0) \int_t^{t+T_0} w(\tau)w^T(\tau)d\tau \ge \epsilon I_n \Big\}. \tag{14}$$

We will say that $w \in SE(n, \alpha)$ is sufficiently exciting and $w \in PE(n, 1)$ is persistently exciting.

Remark 1. (i) There is a qualitative difference between both definitions. The condition for $PE(n, 1)$ is established for any finite interval, whereas the one for $SE(n, \alpha)$ involves the limit of whole real positive line. Essentially, this corresponds to the local character of the integer order Eq. (8) which allows integrating locally to get an evaluation from t to $t + T_0$ and therefore, a local condition. On the other hand, being fractional derivative a non-local operator, one gets instead an evaluation between t=0 and t.

(ii) Note that it is required for w and f_w to be bounded and continuously differentiable functions.

We prove the following general theorem based on the definition of these sets.

Theorem 2. *Let us consider system of Eqs. (7) and (8). If $w \in PE(n, 1)$ and $\alpha = 1$, then (ϕ, e) converges uniformly to zero. If $w \in SE(n, \alpha)$ and $0 < \alpha \le 1$ then (ϕ, e) converges asymptotically to zero.*

Proof. When $\alpha = 1$ the equation for ϕ is just $D\phi(t) = -w(t)w^T(t)\phi(t)$ and if $w \in PE(n, 1)$ then ϕ converges uniformly to zero [17]. Since w is bounded, from (7) e converges uniformly to zero.

Now consider $w \in SE(n, \alpha)$ and $0 < \alpha \le 1$. The equation for ϕ has the form $D^\alpha \phi(t) = -w(t)w^T(t)\phi(t) = A(t)\phi(t)$. The condition of Theorem 1 takes in this case the form $w(t)w^T(t) \ge f_w(t) I_n$, guaranteeing the convergence of ϕ to zero, provided that f_w is absolutely continuous, which holds because $w \in SE(n, \alpha)$ and therefore f_w is continuously differentiable. Hence, ϕ converges to zero and from (7) e converges to zero.

Remark 2. If $w \in SE(n, \alpha)$, for any constant not null vector $u \in \mathbb{R}^n$ we have $(w^T(t)u)^2 \geq f(t)\|u\|$. For example, if $w^T = (1\ 1)$ and $u^T = (-1\ 1)$ the condition does not hold since $(w^T u) = 0$ and $f(t) = 0$, which does not have a divergent fractional integral. By Cauchy-Schwarz inequality, $\|w(t)\|^2\|u\|^2 \geq (w^T(t)u)^2 \geq f(t)\|u\|$. Therefore, a necessary condition for $w \in SE(n, \alpha)$ is $\|w(t)\|^2 \geq f(t)$. For instance, if $\| w \|^2$ has bounded fractional integral, then $w \notin SE(n, \alpha)$.

Remark 3. In [17] it is proved that any signal in $PE(n, 1)$ guarantees uniform asymptotic stability of the equilibrium point $\phi \equiv 0$ for Eq. (8) when $\alpha = 1$ (the uniformity can be inferred from the invariance under initial time translations of its definition). The set $PE(n, 1)$ can be equivalently characterized as

$$PE(n, 1) = \Big\{ w: [0, \infty) \to \mathbb{R}^n | (\forall x \neq 0 \in \mathbb{R}^n)(\exists (\epsilon, \eta))(\forall t>0) \times \int_{t_0}^{t} x^T w(\tau) w^T(\tau) x d\tau \geq \epsilon(t - t_0) + \eta \Big\}.$$

Therefore, as a natural generalization to the fractional case it would be possible to define

$$PE(n, \alpha) := \Big\{ w: \mathbb{R}^+ \to \mathbb{R}^n | (\forall x \neq 0 \in \mathbb{R}^n)(\exists (\epsilon, \eta))(\forall t > 0) \times I^\alpha [x^T w w^T x](t) \geq \epsilon(t - t_0)^\alpha + \eta \Big\} \tag{15}$$

However, the concept of uniformity (relative to the initial condition) becomes ill-defined for fractional systems since the fractional derivative, as it has been used up to now, has a fixed initial time t_0 and hence, fractional equations has a fixed time for initial condition. Thus, the uniformity concept cannot unambiguously be applied, and therefore, $PE(n, \alpha)$ has no special meaning in this sense. A way to preserve this property could be changing together with the initial time for the initial conditions, the initial time of the integral defining the fractional derivative (indeed, this is equivalent to consider the uniformity of the associated integral equation).

Even if ϕ and w do not converge to zero, e can converge to zero as it is shown in the following example.

Example 1. If w is a not null constant vector, by Remark 3, ϕ does not necessarily converge to zero. But, multiplying (8) by w^T we have $D^\alpha[w^T\phi](t) = -(w^T w) w^T \phi(t)$ i.e. $D^\alpha e(t) = -(w^T w) e(t)$, whereby e converges to zero. If w can be written as $w = (\xi^T\ 0^T)^T$ where $\xi \in SE(s, \alpha)$ with $s < n$, then ϕ has its first s components converging to zero, and the rest $n - s$ components are constants, whereby $e = \phi^T w$ converges to zero.

3.2. Characterization of $SE(n, \alpha)$

The set $PE(n, 1)$ has been characterized in the literature [1,18]. For example, quasi periodic functions belong to this set. The following Lemma connects this known set with $SE(n, \alpha)$.

Lemma 3. *Let $w \in \mathbb{R}^n$ be a uniformly continuous function. If $w \in PE(n, 1)$ then $w \in SE(n, \alpha)$ for every $0 < \alpha < 1$.*

Proof. By hypothesis, for all $t > 0$ there exists ϵ, T_0 such that $\varepsilon I_n \leq \int_t^{t+T_0} w(\tau) w^T(\tau) d\tau$, which is equivalent to $\varepsilon\|u\| \leq \int_t^{t+T_0} (w^T(\tau)u)^2 d\tau$, where $u \in \mathbb{R}^n - \{0\}$ is any constant vector. Applying mean value Theorem and since w is a continuous function, it follows that $(\exists\ \xi)$ with $t < \xi < t + T_0$ such that $\varepsilon\|u\| \leq T_0(w^T(\xi)u)^2$. Also, by continuity, there exists a finite interval $\mathbb{I} \subseteq (t, t + T_0)$ where $\forall\ \tau \in \mathbb{I}$ it is satisfied $(2T_0)^{-1}\varepsilon\|u\| \leq (w^T(\tau)u)^2$ or equivalently

$(2T_0)^{-1}\varepsilon I_n \leq w(\tau)w^T(\tau)$.

Since this is valid for any $t > 0$ and the continuity is uniform, there exists a divergent sequence $(t_i)_{i\in\mathbb{N}}$ where $t_{i+1} = t_i + T_0$, which defines intervals $\mathbb{I}_i$ of length independent of i, and a C^1 function $f(t)$ null everywhere except at intervals $\mathbb{I}_i$ where it takes values smaller than $(2\ T)^{-1}\varepsilon$. Therefore, $f(t)I_n \leq ww^T(t)$. By Example 5 in [14], f has divergent fractional integral for every $0 < \alpha \leq 1$, and therefore $w \in SE(n, \alpha)$.

In the following, further properties of the set $SE(n, \alpha)$ are established.

Property 4. *If $0 < \alpha \leq \beta$ then $SE(n, \alpha) \subseteq SE(n, \beta)$.*

Proof. Let us consider $w \in SE(n, \alpha)$, then we have $w(t)w^T(t) \geq f(t)I_n$, where $I^\alpha f \to \infty$. Thereby, there exists $T_C > 0$ such that $I^\alpha f(t) \geq C$ for any $t > T_C$. Since $I^\beta f = I^{\beta-\alpha} I^\alpha f$, we note that the integrand has divergent integral, since for $t < T_C$ the integral is bounded. Consequently $I^\beta f \to \infty$, and therefore $w \in SE(n, \beta)$.

By Lemma 3, the sets $SE(n, \alpha)$ are not empty, and $\bigcap_{\alpha\in[\beta,1]} SE(n, \alpha) = SE(n, \beta) \neq \phi$ for any $\beta > 0$.
The next property shows that the sets are not equal.

Property 5. *If $0 < \alpha < 1$ then $SE(1, \alpha) \subset SE(1, 1)$.*

Proof. By Property 4, it is enough to consider the pulse train function p with fixed pulse width defined such that the separation between its not null values, $1's$ tends to infinity (as times goes to infinity). The associated solution for ϕ of (8) with $\alpha = 1$ and using $w(t) = p(t)$, tends to zero since the pulse has a divergent integer integral, whereby it belongs to $SE(1, 1)$ but does not belong to $SE(1, \alpha)$.

The following property shows that the sets $SE(n, \alpha)$ are invariant under linear transformations.

Property 6. *Let $M \in \mathbb{R}^{n\times n}$ be a real constant matrix of full rank. If $w \in SE(n, \alpha)$ then $Mw \in SE(n, \alpha)$.*

Proof. Since $w \in SE(n, \alpha)$, $Mww^T M^T \geq fMM^T$. Since $MM^T > 0$ because M is a full rank matrix, and since is a constant matrix, there exists $\epsilon > 0$ (in fact, the smallest eigenvalue of MM^T) such that $MM^T > \epsilon I_n$, whereby $Mww^T M^T \geq f\epsilon I_n$.

From Property 6 spaces $SE(n, \alpha)$ and $SE(m, \alpha)$ can be related when $m < n$, through a matrix M of rank m. Next property is a simple but illustrative case.

Property 7. *Let $u \in \mathbb{R}^n$ be a constant non zero vector. If $w \in SE(n, \alpha)$ then u^{Tw} is in $SE(1, \alpha)$. In particular, if $w \in SE(n, \alpha)$ then each component is in $SE(1, \alpha)$, though the converse is not necessarily true (for instance, consider w a constant vector).*

Proof. Since $w \in SE(n, \alpha)$ this implies that $(u^T w(t)) \geq f(t)u^T u =: \hat{f}(t)$ and the first part follows. For the second part we take $u = e_i$ with $(e_i)_i$ for $i = 1, 2, \ldots n$ the canonical basis of $\mathbb{R}^n$.

As shown in the following property, if $w \in SE(n, \alpha)$ then w is a vector repeatedly moving along every direction in $\mathbb{R}^n$.

Property 8. *If w belongs to a proper subset of $\mathbb{R}^n$ then $w \notin SE(n, \alpha)$.*

Proof. Since w is in a proper subset, there exists a constant not null vector u orthogonal to that subspace, such that $w^T(t)u = 0$ for all $t \geq 0$, thereby $u^T w(t) w^T(t) u = 0$ for all $t \geq 0$.

The following property gives conditions for the invariance of $SE(n, \alpha)$ under $\mathcal{L}^1_\alpha$ (the set of functions such that its modulus has fractional α integral bounded) translations.

Property 9. *Let w_1, w_2 be two bounded functions such that $(w_1 - w_2) \in \mathcal{L}^1_\alpha$. Suppose further that there exists $f_i(t) > 0$ such that $w_i w_i^T > f_i I$, for $i=1,2$. Then $w_1 \in SE(n, \alpha)$ if and only if $w_2 \in SE(n, \alpha)$.*

Proof. We demonstrate first that if $w_1 \in SE(n, \alpha)$ then $w_2 \in SE(n, \alpha)$. Let C be defined as $C = \sup_t \{w_1(t), w_2(t)\}$. Let x be a constant vector. Then $x^T w_2 w_2^T x = x^T w_1 w_1^T x - x^T (w_1 - w_2) x^T (w_1 + w_2)$. By Cauchy-Schwarz inequality, $x^T w_2 w_2^T x \geq x^T w_1 w_1^T x - 2Cx^T x \, \|(w_1 - w_2)\|$. Therefore $x^T w_2 w_2^T x = f_2 x^T x \geq x^T w_1 w_1^T x - 2Cx^T x \, \|(w_1 - w_2)\|$. By α-integrating, $I^\alpha f_2 x^T x \geq I^\alpha (x^T w_1 w_1^T x) - 2Cx^T x I^\alpha \, \|(w_1 - w_2)\|$. Since $w_1 \in SE(n, \alpha)$ there exists f_1 such that $I^\alpha f_2 x^T x \geq I^\alpha f_1 x^T x - 2Cx^T x I^\alpha \, \|(w_1 - w_2)\|$ and using that $(w_1 - w_2) \in \mathcal{L}^1_\alpha$, we conclude that $[I^\alpha f_2] x^T x \geq [I^\alpha f_1] x^T x - \hat{C} \to \infty$ as t goes to infinity, whereby $w_2 \in SE(n, \alpha)$.

Since $(w_1 - w_2) \in \mathcal{L}^1_\alpha$ if and only if $(w_2 - w_1) \in \mathcal{L}^1_\alpha$, by interchanging the roles of w_1 and w_2, the reverse implication follows.

In [19] it is established a relationship among the functions w of the set $PE(n, 1)$, their auto-correlation functions R_w and their associated spectral densities S_w, which allows to formalize the idea of a 'sufficiently rich' signal containing enough frequencies for identification purposes. The next proposition extends the relationship for the set $SE(n, \alpha)$. We recall first the definition of the auto correlation function R_w for a stationary function w and its spectral measure S_w (Chapter 1 of [2]).

$$R_w(\tau) := \lim_{T \to +\infty} \frac{1}{T} \int_0^T w(t) w(t + \tau)^T \, dt. \tag{16}$$

$$S_w(d\nu) := \int_{-\infty}^{\infty} e^{-i\nu\tau} R_w(\tau) d\tau. \tag{17}$$

Property 10. *If w is a uniformly continuous function such that $R_w(0) > 0$ then $w \in SE(n, \alpha)$ for all $0 < \alpha \leq 1$. Conversely, if $w \in SE(n, \alpha)$ with its associated f_w such that $\lim_{T \to +\infty} \frac{1}{T} \int_s^{s+T} f_w \, dt > 0$ then $R_w(0) > 0$.*

Proof. The first implication follows from the fact that $R_w(0) > 0$ implies that $w \in PE(n, 1)$. By uniform continuity and Lemma 2, one concludes that $w \in SE(n, \alpha)$. The second implication is straightforward from the definition of $SE(n, \alpha)$.

The invariance analysis of set $PE(n, 1)$ under dynamical transformations can be obtained with the following property, which extends the result of [19] to fractional filters $H(s) = \frac{\sum_{i=1}^m b_i s^{\beta_i}}{\sum_{i=1}^n a_i s^{\alpha_i}}$ with $\alpha_1 < \ldots < \alpha_n$ and $\beta_1 < \ldots < \beta_m$ numbers in $\mathbb{R}$. The relative degree of H is $\alpha_n - \beta_m$.

Property 11. *Let u: $\mathbb{R}^+ \longrightarrow \mathbb{R}$ be a stationary function and y: $\mathbb{R}^+ \longrightarrow \mathbb{R}^n$ related to u through a vector filter H of relative degree >0 and BIBO stable in each component, written in Laplace domain as $y(s) = H(s) u(s)$. Suppose that for all $w_1, \ldots, w_n \in \mathbb{R}$, $H(jw_1), \ldots, H(jw_n)$ are linearly independent on $\mathbb{C}^n$. Then, $y \in PE(n, 1)$ if and only if the spectral measure of u is not concentrated on $k < n$ points.*

Proof. Since H is of relative degree >0 in each component, its impulse response is L^1 by Theorem 2 in [20]. Hence, by using the same arguments of the proof Proposition 1.6.2 in [2], y is stationary. The claim follows by using the same arguments of the proof Theorem 2.7.2 in [2].

Remark 4. There is a similar analysis for the spectral line at frequency ν of the signal u defined in [21] as

$$\hat{u}(\nu) := \lim_{T \to +\infty} \frac{1}{T} \int_s^{s+T} u(t) \exp(-j\nu t) \, dt.$$

In fact, if $y(t) \in \mathbb{R}^n$ is related with $u(t) \in \mathbb{R}^m$ by a (fractional or integer) linear time invariant filter $H(s)$ and u has spectral line at frequency ν, then y has spectral line at frequency ν too. Indeed, by convolution properties, $\hat{y}(\nu) = H(j\nu) \hat{u}(\nu)$.

4. Alternative schemes

We will now study extensions of the results presented in Section 3. Our intention is first to examine alternative schemes of fractional adjustment of parameters of Type I, determining in each case if the set of information signals w that assures convergence of the error e is enlarged or not. Note that in the fractional case, to assure convergence of the error, we require parametric convergence which is not necessary in the integer case as is showed by Proposition 1. These include schemes using different types of objective functional to be minimized as well as filtering of signals. Secondly, multiple error models and dynamic errors models are proposed. We will show that the newly introduced concept of sufficiently exciting signals plays a part to obtain sufficient conditions for convergence of the error even if the structure of fractional equations differs from which it was developed.

We will assume in what follows that the input signal w is bounded, uniformly continuous and continuously differentiable.

4.1. Non unity adaptive gains

In the previous analyses we assumed $\gamma = I_n$. It easy to see that the choice of $\gamma(\cdot)$ does not enlarge the set of functions w producing convergence of ϕ to zero. For instance, in the scalar case let us assume for practical reasons that $\gamma > 0$ is bounded (say by C). Then $\|I^\alpha[\gamma w^2]\| \leq C \|I^\alpha [w^2]\|$. If w does not make ϕ converge to zero, then $I^\alpha[w^2]$ is necessarily bounded. Then $I^\alpha[\gamma w^2]$ is bounded and therefore ϕ does not converges to zero (Theorem 3 in [13]).

However $\gamma(t)$ allows to handle the convergence rate. Take for instance the scalar case. Suppose that we want ϕ to converge as fast as it does with an input w_1, then we must choose $\gamma(t)$ such that $\gamma(t) w(t)^2 \geq w_1^2(t)$ (Proposition 3 in [14]).

Another use of $\gamma(t)$ allows to modify the convergence properties of standard gradient method with fractional derivatives or integrals. Let us take Eq. (7), the adaptive law (8) with $\alpha = 1$ and $\gamma = \exp\left(\frac{1}{1 + I^\alpha (e^2)}\right)$, to obtain

$$D\phi(t) = -w(t) w^T(t) \phi(t) \exp\left(\frac{1}{1 + I^\alpha[e^2](t)}\right). \tag{18}$$

By defining $2V = \phi^T \phi$, it follows that $DV(t) = -e^2(t) \exp\left(\frac{1}{1 + I^\alpha[\phi^2 w^2](t)}\right) \leq 0$. Whereby $e^2 \exp\left(\frac{1}{1 + I^\alpha[\phi^2 w^2](t)}\right)$ is in $\mathcal{L}^1$ and ϕ is bounded. Then $D\phi$ is bounded and ϕ is uniformly continuous, whereby e is also uniformly continuous. Since $1 < \exp\left(\frac{1}{1+I^{\alpha}[\phi^2 w^2](t)}\right)$ it follows that $e^2 \in \mathcal{L}^1$. By applying Lemma 1, the error converges to zero.

Since $1 < \exp\left(\frac{1}{1+I^{\alpha}(e^2)}\right)$ it follows that $DV \le -e^2$. By using the Comparison Principle and the fact that $DV = -e^2$ is exponentially asymptotically stable if $w \in PE(n, 1)$ [22], we have the same set of convergence as for $\alpha = 1$.

One can also use in (18) any bounded function r other than the exponential provided that $r(t) \ge c > 0$ for all t.

4.2. Commuted scheme

A simple scheme that combines both integer method, with its known convergence properties, and the fractional adjustment is given by Eq. (7) together with the adaptive law

$$D^{\alpha}\phi(t) = -e(t)w(t), \text{ with } \begin{cases} \alpha = \alpha_0 < 1 & \text{if } e^2 > e_0 \text{ and } \alpha \ne 1 \\ \alpha = 1 & \text{else} \end{cases} \tag{19}$$

where $0 < \alpha_0 < 1$ is a fixed number and $e_0 > 0$ is an arbitrarily chosen design parameter.

This adjustment assures asymptotic convergence of the error to zero for w bounded and uniformly continuous. Indeed, note that the condition $e^2 > e_0$and$\alpha \ne 1$ implies that once $\alpha = 1$ it remains in $\alpha = 1$. Hence, if $e(0)^2 < e_0$ then $\alpha = 1$ for all t and we apply Proposition 1. Else, by Proposition 2 there exists an instant of time T where $e^2(T) = e_0$ and with the integer law ($\alpha = 1$) it is assured convergence of e considering as initial condition the instant of time T (to preserve continuity). Similarly, it follows that if $w \in PE(n, 1)$ then ϕ converges to zero.

4.3. Error functional

Two type of objective functions can be considered in the gradient approach; algebraic and dynamic functional.

i) Algebraic functional

Le us consider system (7) together with an objective function of the type $J = F(e^2)$ instead of $J = e^2$ as previously, where $F(x) > 0$ for all $x > 0$. Then, the adaptive law based on the gradient approach (6) takes the form

$$D^{\alpha}\phi(t) = -w(t)w^T(t)\phi(t)F'(e^2(t)), \tag{20}$$

where $F' = dF/dx$. If $F'(x) > 0$ for all $x > 0$ then Eq. (20) has bounded solutions. Furthermore, if $\alpha = 1$ and F' is continuous then e and F converge to zero. Indeed, by defining $2V = \phi^T\phi$, we have $D^{\alpha}V \le -e^2F'(e^2) \le 0$, whereby the first claim follows. For $\alpha = 1$, $e^2F'(e^2)$ is in $\mathcal{L}^1$. Since e^2 is bounded and $F'(\cdot)$ is continuous, $F'(e^2)$ turns out to be bounded and uniformly continuous. By applying Lemma 1, and since e^2 is uniformly continuous (because ϕ has bounded derivative), it follows that e^2 converge to zero and therefore F tends to its minimum $F(0)$. Then, for $\alpha < 1$, $\liminf_{t\to\infty} F = F(0)$, by the continuity of F since it is differentiable.

Examples of objective functions satisfying the above conditions are: $\tanh(\cdot)$, class K functions, $\exp(\cdot)$, $F(e) = e^{2n}$ with $n \in \mathbb{N}$, among others.

ii) Dynamic functional

Let us consider the system defined by Eqs. (9), (10) and (7). Let L be any linear operator. Define the functional $J = L[e^2] = L[(w^T\hat{\theta} - w^T\theta)^2]$. Then, the following adaptive law based on the gradient approach (6) can be postulated with $\gamma > 0$

$$D^{\alpha}\hat{\theta} = -\gamma L[w\, w^T]\hat{\theta} + L[w\, y]. \tag{21}$$

Using $\phi = \hat{\theta} - \theta$, the previous equation can be expressed as

$$D^{\alpha}\phi(t) = -\gamma L[w(t)w^T(t)]\phi(t). \tag{22}$$

Example 2. Let us consider the linear operator $L[\cdot] = I^{\beta}[\cdot]$ with $\beta > 0$, Eq. (22) takes the form $D^{\alpha}\phi(t) = -I^{\beta}[w\, w^T](t)\phi(t)$ for $\gamma = 1$.

Note that for x a constant vector, $x^T(I^{\beta}[w\, w^T])x$ $= I^{\beta}[x^T\, w\, w^T\, x] = I^{\beta}[(x^T\, w)^2] \ge 0$. Hence, $I^{\alpha}[w\, w^T]$ is positive semi-definite and therefore the trajectories ϕ given by (22) are bounded. The reason is that by calling $A(t) = I^{\beta}[w\, w^T](t)$, we have $D^{\alpha}\phi(t) = -A(t)\phi$ and then, by Property 3, $D^{\alpha}[\phi^T\phi](t)$ $\le 2\phi^T D^{\alpha}\phi(t) = -2\phi^T A(t)\phi \le 0$ and hence $\|\phi(t)\| \le \|\phi(0)\|$ for all $t > 0$.

For asymptotic convergence, the condition is now imposed on $I^{\beta}[w\, w^T]$ instead $w\, w^T$. Therefore the set of possible signals guaranteeing convergence of ϕ is enlarged as compared with $SE(n, \beta)$, because if $w\, w^T \ge fI$ then $I^{\beta}[w\, w^T] \ge I^{\beta}[f]I$, and since f diverges then $I^{\beta}[f]$ also diverges. Indeed, using Lemma 9 in [13], if $\beta > 2$ and $ww^T \ge fI$ for f non negative, then $I^{\beta}f$ diverges.

Example 3. Let us consider the linear operator $L(s) = \frac{1}{s^{\beta}+p}$ (written in the Laplace domain). By defining Γ and δ as

$$\Gamma^{(\beta)} = -p\Gamma + w\, w^T,$$

$$\delta^{(\beta)} = -p\delta + w\, y,$$

with $\Gamma(0) = 0$ and $\delta(0) = 0$, the adaptive law (21) can be written as

$$D^{\alpha}\hat{\theta} = -\gamma\Gamma\hat{\theta} + \delta y, \tag{23}$$

or equivalently, using (22),

$$D^{\alpha}\phi(t) = -\gamma\Gamma(t)\phi(t). \tag{24}$$

When $\alpha = \beta = 1$, this case corresponds to the design proposed by Kreisselmeier in [23], where the set $PE(n, 1)$ guarantees convergence of ϕ to zero. The advantage of this approach over schemes with simpler adaptive laws is that arbitrary speed rates can be achieved, by choosing γ such that Γ has an arbitrary spectrum. The counterpart is that a more intensive use of computational resources is needed in the implementation.

For $\beta \le 1$ and $\gamma = 1$, if x is a constant vector then

$$x^T\Gamma^{(\beta)}x = -px^T\Gamma x + x^T w\, w^T x.$$

By defining $z := x^T\Gamma x$ and $f(t) := x^T w(t)w^T(t)x \ge 0$, with $z(0) = 0$, it follows that

$$z^{(\beta)}(t) = -pz(t) + f(t). \tag{25}$$

By using the analytic solution of Eq. (25) as given in [9], it follows that $z = E_{\beta,\beta} * f$ and thus $z(t) \ge 0$ for all $t \ge 0$. Hence, Γ is positive semi-definite and, if there exists an instant of time such that $x^T w\, w^T x \ne 0$, then it is positive definite for $t > 0$. If $w \in PE(n, 1)$ or $w \in SE(n, \beta)$ such instant of time always exists by definition.

It follows from Eq. (24) that ϕ is bounded and the eigenvalue of Γ can be suitably modified through the adaptive gain γ, improving the convergence rate. For asymptotic convergence of Eq. (24) a sufficient condition is $\Gamma(t) \ge \epsilon I$ or that $\Gamma(t) \ge f(t)I$ with $I^{\alpha}f \to \infty$. For the former, by choosing $p > 0$, we have $E_{\beta,\beta}(t) > E_0 > 0$ for any $t > T$ and some $T > 0$, then $E_{\beta,\beta} * x^T w\, w^T x(t) > E_0 \int_T^t x^T w\, w^T x d\tau$ and thus $E_{\beta,\beta} * ww^T > E_0\epsilon I$ for any $t > T$ provided that $w \in PE(n, 1)$.

Remark 5. The Kreisselmeier law ($\alpha = \beta = 1$) shares the same philosophy as the adjustment with fractional derivative, in the sense that both make use of a memory effect, where all past information is (not uniformly) considered in the current updating. For the fractional case, this comes from the definition of fractional derivative. It can be shown [1,23] that the Kreisselmeier law optimizes the following functional

$$J(t) := \int_0^t \exp(-p(t-\tau))e^2(t,\tau)d\tau, \tag{26}$$

where

$$e(t,\tau) := \hat{\theta}^T(t)w(\tau) - y(\tau) = \hat{\theta}^T(t)w(\tau) - \theta^T w(\tau) = \phi(t)^T w(\tau). \tag{27}$$

Whereby, the adaptive law is given by ($\gamma = 1$)

$$\frac{d\phi}{dt}(t) = -\int_0^t e(t,\tau)w(\tau)d\tau, \tag{28}$$

which makes evident the use of the past data.

A simple fractional generalization of the last equation is

$$\frac{d\phi}{dt}(t) = -I^\alpha\ [e(t,\tau)w(\tau)], \tag{29}$$

where the Kernel is now $(t)^{-\alpha}$ instead of $\exp(-t)$. It must be noted that since $e(t,\tau)$ is defined in (27), this law is in principle realizable and the error defined by (27), is evaluated in the past but using the current estimate of θ at t and not the past estimates. It follows that using (27), the adaptive law (29) can be expressed as $\frac{d\phi}{dt}(t) = -[I^\alpha\ w^2](t)\phi(t)$. Thereby the equation is stable and the condition for convergence of ϕ to zero (and therefore, convergence of ϕw) is $I^\alpha\ w^2 \to \infty$. In the vector case, the condition turns out to be $I^\alpha\ [ww^T] \in PE(n,1)$.

We next give a robustness result for this law.

Proposition 2. *Consider the following equation*

$$D^\alpha\phi(t) = -\Gamma(t)\phi(t) + \nu(t) \tag{30}$$

where $\Gamma \geq \epsilon I$ is a continuously differentiable matrix function and ν is a bounded vector function converging to zero such that $I^\alpha \nu$ is bounded. Then ϕ converges to zero.

Proof. By Corollary 4 in [14], ϕ is a bounded vector function. Hence $\phi^T\nu(t) = f(t)$ is bounded and converges to zero.

By noting that $\phi^T(t)D^\alpha\phi(t) = -\phi^T(t)\Gamma(t)\phi(t) + f(t)$ and since $D^\alpha[\phi^T\phi](t) \leq 2\phi^T(t)D^\alpha\phi(t)$ by Property 3, it follows that $D^\alpha[\phi^T\phi](t) \leq -2\epsilon\phi^T(t)\phi(t) + f(t)$. By Theorem 5 in [14], it follows for equation $D^\alpha[\phi^T\phi](t) = -2\epsilon[\phi^T(t)\phi(t)] + f(t)$ that $[\phi^T\phi]$ converges to zero. The claims follows by Comparison Principle.

4.4. Multiple errors

We study here the case when the error is a vector function rather than scalar function.

i) Vector case

Let us consider again relationships (9) and (10). By filtering the input $w \in \mathbb{R}^n$ and the output $y \in \mathbb{R}$ by n linearly independent filters H_i (possibly fractional ones), we get $Y = W^T\theta$, where $Y = [y_1, y_2, \ldots, y_m]^T \in \mathbb{R}^m$ and $W = [w_1|w_2|\ldots|w_m] \in \mathbb{R}^{n\times m}$, with $y_i = H_i y$ and $w_i = H_i I_n w$, where I_n denotes the identity matrix of order n. Therefore, an error vector defined as $E(t) = [e_1(t), e_2(t), \ldots, e_m(t)]^T = [\hat{y}_1(t) - y_1(t), \hat{y}_2(t) - y_2(t), \ldots, \hat{y}_m(t) - y_m(t)]^T \in \mathbb{R}^m$ related to matrix W and parameter error $\phi(t) \in \mathbb{R}^n$ as follows

$$E(t) = W^T(t)\phi(t). \tag{31}$$

Applying gradient approach to the objective function $J = E^T E$, the following adaptive law is proposed

$$D^\alpha\phi(t) = -\gamma W(t)E(t) = -\gamma W(t)W^T(t)\phi(t). \tag{32}$$

Since the filters are linearly independents, the rows of W turns out to be linearly independent functions for any not null input, whereby $WW^T(t) > 0$ for all $t > 0$. Hence, by choosing $\gamma(t)$ properly, arbitrarily fast convergence rates can be achieved (theoretically, one can choose for instance $\gamma(t) = \epsilon/\lambda_m$ where λ_m is the minimum eigenvalue of WW^T, getting $WW^T \geq \epsilon I$). By setting $\gamma = 1$, a convergence condition of error E could be $WW^T > \epsilon I$.

Example 4. Let F_1, F_2 be scalar linear transfer functions with numerators of different orders n_1 and n_2 respectively and the same denominator. Consider u a non null scalar function. If y_1, y_2 are the respective scalar outputs for null initial conditions, then y_1, y_2 are linearly independent and therefore F_1, F_2 are linearly independents. In fact, if $ay_1 + by_2 \equiv 0$, we have in Laplace domain $aF_1 + bF_2 \equiv 0$. Since F_1, F_2 are rational functions of different orders and a, b are non zero real numbers, $aF1 + bF2$ has as numerator a polynomial of degree $n = n_1 + n_2$. Hence, the only way to achieve $ay_1 + by_2 \equiv 0$ is by setting $a = b = 0$. Recursively, it can be proved for m linear filters $(F_i)_{i=1}^m$ of pairwise different orders. Furthermore, by similar arguments, one can allow non zero initial conditions by using asymptotically stable linear filters.

ii) Matrix case

Let us consider now the case when the parameter is a matrix $\Theta \in \mathbb{R}^{n\times n}$, the output system is a vector $Y \in \mathbb{R}^n$ and the information signal is a vector $w \in \mathbb{R}^n$. They are linearly related through

$$Y(t) = \Theta w(t), \tag{33}$$

where $Y(t)$ and $w(t)$ are known signals. The estimated output of the system at instant of time t is denoted as $\hat{Y}(t) \in \mathbb{R}^n$ and will be given by

$$\hat{Y}(t) = \hat{\Theta}(t)w(t), \tag{34}$$

where $\hat{\Theta}(t) \in \mathbb{R}^{n\times n}$ is the estimated of Θ at time t. Subtracting Eq. (33) from (34), the relationship between the output estimation error $E(t) = \hat{Y}(t) - Y(t) \in \mathbb{R}^n$ and the parameter estimation error $\Phi(t) = \hat{\Theta}(t) - \Theta \in \mathbb{R}^{n\times n}$ is given through

$$E(t) = \Phi(t)w(t). \tag{35}$$

By applying the gradient approach and considering the objective function $J = E^T E$, we end up with the following adjustment law

$$D^\alpha\Phi(t) = -E(t)w^T(t),$$

or equivalently

$$D^\alpha\Phi = -\Phi(t)w(t)w^T(t). \tag{36}$$

By applying [24], Lemma 5 and choosing $V = trace(\Phi^T\Phi)$, it follows that

$$\begin{aligned} D^\alpha V(t) &\leq -2trace(\Phi^T\Phi ww^T)(t) = -2w^T(t)\Phi^T(t)\Phi(t)w(t) \\ &= -2E^T(t)E(t) \leq 0. \end{aligned} \tag{37}$$

Hence, the system has bounded trajectories and $\lim\inf_t E^T(t)E(e) = 0$. A sufficient condition for convergence of the error to zero is that for any positive semi-definite constant matrix A, there exists a scalar differentiable bounded function f, independent of A, whose fractional α-integral diverges such that $w^T(t)Aw(t) \geq f(t)trace(A)$.

Also, in this context we can analyze the case when $\phi(t) \in \mathbb{R}$. Let $e \in \mathbb{R}^n$ be a vector whose components are given by $e_i = \phi w_i$ for $i = 1, \ldots, n$. Using the gradient approach with the objective function $J = \| e \|^2$, it is suggested the following adjustment

$$D^\alpha \phi(t) = -e^T(t) w(t) = -\phi(t) \left(\sum_{i=1}^{n} w_i(t) \right)^2. \tag{38}$$

Since $\phi(t) D^\alpha \phi(t) \le 0$ for all $t \ge 0$, the trajectories are bounded. Let us consider for example $\alpha = 1$ and $n=2$. The analytic solution is $\phi(t) = \phi(0) \exp(-\int_0^t (w_1(\tau) + w_2(\tau))^2) d\tau$, showing that ϕ always converges. If $\| w \|$ has bounded integral and it is uniformly continuous, then the error converges to zero. But it is not enough that $w_i \in PE(1, 1)$ to make ϕ converge to zero (for instance, if $w_1 = -w_2$ then neither ϕ nor e will converge to zero). The condition for convergence of error e when $\alpha \le 1$ is that the fractional α-integral of $(\sum_i w_i)^2$ be divergent.

4.5. Differential error model

In this section instead of having an algebraic error equation like (7) we will analyze the case when there is a differential equation relating the output error $e \in \mathbb{R}^n$ with the parameter error $\phi \in \mathbb{R}^n$ and the information vector $W \in \mathbb{R}^{n \times n}$. Let us consider the following fractional order equation

$$D^\beta e(t) = Ae(t) + W^T(t) \phi(t) \tag{39}$$

where $A \in \mathbb{R}^{n \times n}$ is a constant matrix such that when $W=0$, e converges to zero. From (39) we have $\partial e / \partial \phi(t) = -A^{-1} W(t)$. If A is negative definite, the gradient approach for $J = \| e \|^2$ suggests the following form for adaptive law

$$D^\alpha \phi(t) = -\gamma W(t) e(t). \tag{40}$$

For the case of $\alpha = \beta = 1$ and e scalar (A scalar and ϕ, W vector functions) it can be shown that e is uniformly continuous and that converges to zero. Therefore de/dt also converges to zero by Barbalat Lemma (see Lemma 2) and thus, in the scalar case, $\phi^T w$ converges to zero and ϕ becomes orthogonal to w, giving the same geometric motivation that in the Error Model I.

Considering now the vector case with $\gamma = 1$, A asymptotically stable matrix in the integer sense and $\alpha = \beta \le 1$. Defining $2V = \phi^T \phi + e^T e$, it follows that $D^\alpha V(t) \le e^T(t) A e(t) \le 0$, whereby trajectories $(e(t), \phi(t))$ have uniformly bounded norms. Moreover, by using Proposition 15 in [13] together with the fact that A is constant and negative definite, we have $\lambda_m(A) e^T e \le e^T A e \le \lambda_M(A) e^T e$, where $\lambda_{m,M}$ are the minimal and maximal eigenvalues, respectively, we have that $\lim\inf_{t \to \infty} \| e(t) \| = 0$.

There are some simple results derived from the previous analysis; (a) if W vanish for $t > T$ then ϕ converges to $\phi(0)$ (Property 4 [13]), whereby $D^\alpha e(t) = Ae(t) + f(t)$ with $f \in \mathcal{L}^2$, then e converges to zero (Theorem 5 of [14]). More general, if $I^\alpha W \to 0$ then ϕ converges to $\phi(0)$. (b) If W is scalar and constant, e converges to zero since ϕ converges to zero.

By considering γ variable with time, the following is the implicit condition on W for convergence of ϕ to zero,

$$\begin{bmatrix} -A & -W^T(t) \\ \gamma(t) W(t) & 0 \end{bmatrix} \ge f(t) I, \tag{41}$$

where f is a differentiable bounded function whose fractional α-integral diverges.

To get an explicit condition on the information vector w that assures convergence to zero of the error, an auxiliary error signal $\varepsilon(t) := D^\beta e(t) - Ae(t)$ is defined whereby $\varepsilon(t) = \phi^T(t) w(t)$ and one can adjust with the adaptive law $D^\alpha \phi(t) = -\varepsilon(t) w(t)$ for $0 < \alpha < 1$. Therefore, by applying Theorem 2, if $w \in SE(n, \beta)$ then $\phi, \varepsilon \to 0$. Since w is bounded by assumption, $\phi^T w \to 0$. By applying Theorem 5 of [14] to $D^\beta e = Ae + \varepsilon$, we conclude that $e \to 0$. If $\alpha = 1$ it is enough for $w(\cdot)$ to be a uniformly continuous function to prove that $e \to 0$.

5. Fractional adaptive indirect control

We give now a relevant application in indirect adaptive control of integer systems of the fractional adjustment of parameters. This application is numerically exemplified for the pole placement problem.

Consider an unknown SISO integer order linear time-invariant system with input u and output y, defined by the proper transfer function

$$y = \frac{n_p(s)}{d_p(s)} u = \frac{\sum_{i=0}^{n-1} b_i s^i}{\sum_{i=0}^{n} a_i s^i} u \tag{42}$$

where n_p,d_p are coprime polynomials. Without loss of generality, we assume $a_n=1$. The problem of adaptive control, namely to design u such that y has a desired behavior, can be solved by making an estimation of plant parameters and building a rational transfer function compensator $C(s)$ of order m given by the transfer function

$$C = \frac{\sum_{i=0}^{m} \bar{c}_i s^i}{\sum_{i=0}^{m} c_i s^i}, \tag{43}$$

with input $r - y$ and output u, based on these estimations, where r is a reference signal. Provided that the estimation converges asymptotically to the true values, the desired behavior is achieved (Lemma 2.1 in [25]).

The estimation can be done using fractional adaptive adjustment in the following way. To have initial condition terms decaying to zero and as a filter of noise, the new filtered variables $\bar{y} := y/\lambda$, $\bar{u} := u/\lambda$ with a Hurwitz polynomial λ of degree at least n are defined. Then,

$$\bar{y} = \frac{\sum_{i=0}^{n-1} b_i s^i}{\sum_{i=0}^{n} a_i s^i} \bar{u}.$$

This system can be realized in the time domain as ([25])

$$\sum_{i=0}^{n} a_i D^i \bar{y}(t) = \sum_{i=0}^{n-1} b_i D^i \bar{u}(t)$$

where D^i stand for integer derivative for i a natural number.

Putting $w := (\bar{y}, D\bar{y}, \ldots, D^{n-1}\bar{y}, \bar{u}, D\bar{u}, \ldots, D^{n-1}\bar{u})^T$ and $\theta := (-a_0, -a_1, \ldots, -a_{n-1}, b_0, \ldots, b_{n-1})^T$, the last equation can be written as

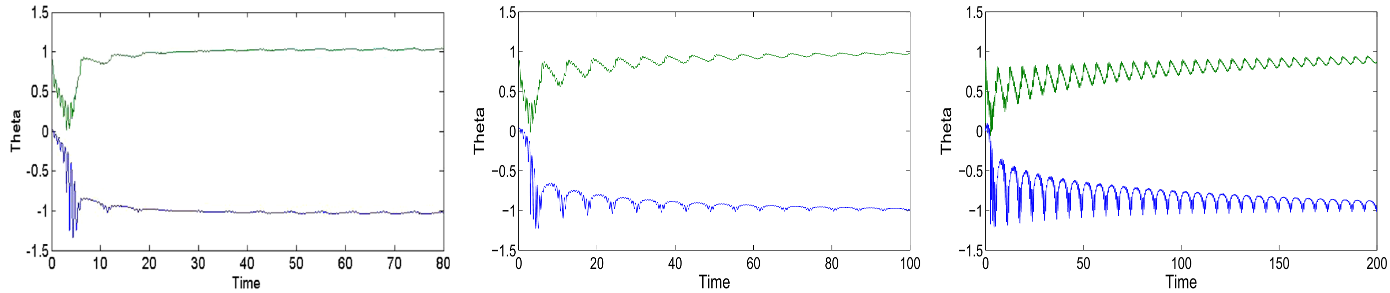
$$D^n \bar{y} = \theta^T w. \tag{44}$$

By defining the identification error as $e := \hat{\theta}^T w - D^n \bar{y}$ and the parametric error as $\phi := \hat{\theta} - \theta$, we have (up to exponentially decaying terms associated to initial conditions of $1/\lambda$ filter)

$$e(t) = \phi^T(t) w(t) \tag{45}$$

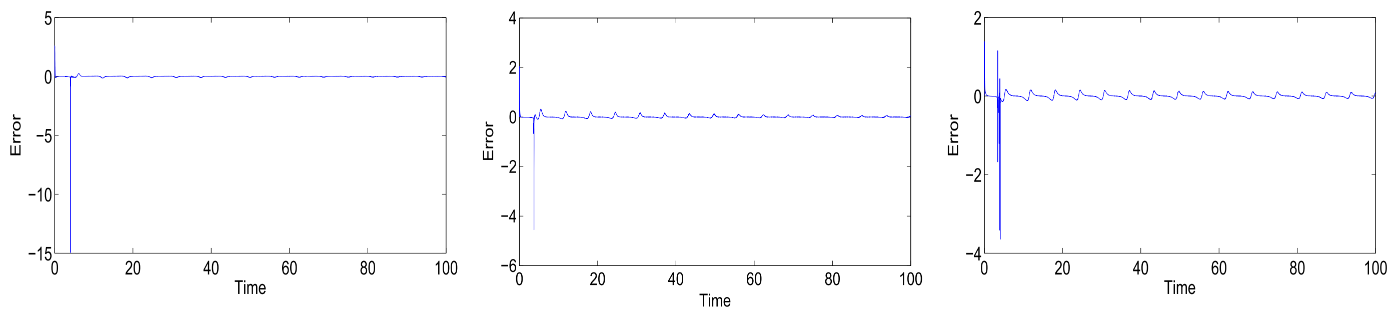
where e, w are known functions.

If the support of the spectrum of u has no less than $2n$ points, the following adjustment assures the convergence of $\phi = \hat{\theta} - \theta$ to zero,



(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 1. Speed of parameter convergence: $\hat{\theta}$ vs time for the stable plant.



(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 2. Identification error function: stable plant.

$$D^\alpha \phi(t) = D^\alpha \hat{\theta}(t) = -e(t)w(t) \tag{46}$$

where $0 < \alpha \le 1$. The reason is that for such u, we have $w \in EP(2n, 1)$ and then Theorem 2 can be applied. By using that $u(s) = C(s)(r(s) - y(s))$, the above condition for u holds, if the support of the spectrum of r has no less than $3n + m$ points, since w is related to r by a filter $w = Hr$ which has order at most $3n + m$ (see the proof of [25], Lemma 2.2 for the specific form of H).

The above general procedure is illustrated in how to choose $C = n_c/d_c$ for the pole placement problem. Consider d^* the characteristic polynomial of the desired poles of order $2n - 1$. Suppose first that n_p,d_p are known. The transfer function relating y with r is given $\frac{PC}{1+PC}$ where $P = n_p/d_p$. By equating the denominator of the closed loop transfer function with d^* we obtain

$$n_c n_p + d_c d_p = d^*. \tag{47}$$

Since n_p and d_p are coprimes, there exist n_c and d_c satisfying (47). For implementation, $n_c(t)$, $d_c(t)$ must be modified whenever the estimates of n_p and d_p will be coprimes, otherwise it remain in its constant value.

Remark 6. We make the following comments to the above procedure:

(i) Filter $\frac{1}{\lambda}$ can be chosen stable in the fractional sense.

(ii) The error function e is related to the model error $e_i := \hat{y} - y$ in Laplace domain by $e = (s^n/\lambda)y$.

(iii) Since Theorem 2 requires w to be bounded, when unstable systems are considered w can be normalized by employing the factor $\gamma(t) := 1/\sqrt{1 + w^T(t)w(t)}$.

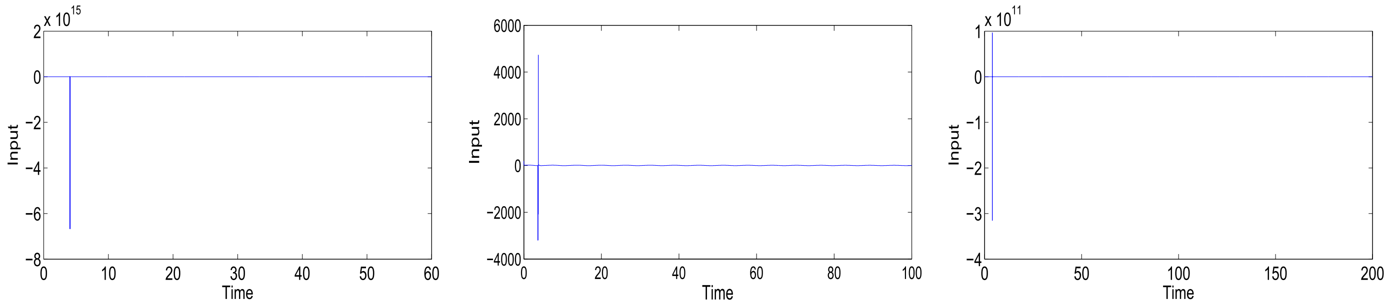
(iv) Provided that the decaying to zero initial condition term associated to $1/\lambda$ filter has bounded fractional α-integral, one can show rigorously that the convergence conditions are not affected (see for instance Proposition 2).

Let us now consider a numerical example of this application.

Example 5. Consider $n=1$ and let $d^* = s + 2$ be the desired denominator. Applying Eq. (47), the compensator is given by $C = (2 - a_0)/b_0$. The parameters are estimated using Eq. (46). Then, considering an instant $t = t_i$, if $\hat{b}_0(t_i) \neq 0$, $C(t_i) = (2 - \hat{a}_0(t_i))/\hat{b}_0(t_i)$, otherwise the last modification is kept $C(t_i) = (2 - \hat{a}_0(t_{i-1}))/\hat{b}_0(t_{i-1})$.

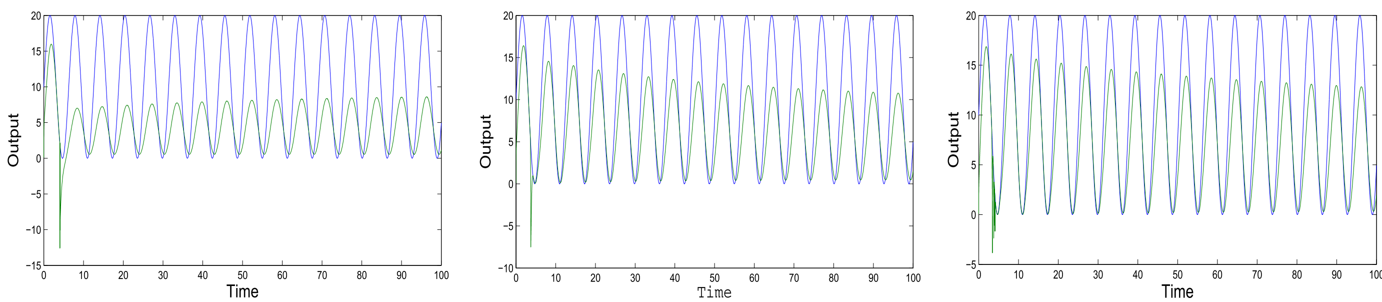
For the first experiment an stable plant with $a_0 = 1$ and $b_0 = 1$ is controlled. As reference function, it is used $r = 10 + 10 sint$ –so that r has $3 \ge 3n + m = 3$ spectral frequencies– and fixed initial conditions ($y(0) = 0.1$, $\hat{\theta}(0) = (0, 0.1)^T$). Recall that $-a_0 = \theta_1$, $b_0 = \theta_2$. Fig. 1 shows that the speed of parameter's convergence is related to the order of derivation, in such a way that no advantage of fractional adjustment is observed in comparison with the integer case. This is not a crucial point since one has the γ parameter to handle the speed of convergence (see Section 4.1). Similar experiments indicate that the amplitude (which can be subsumed in the γ factor) and the number of spectral frequencies of r is also related to the speed of convergence. However, Fig. 2 shows that, in absolute value, the error's maximum amplitude increases with the order of derivation (worst for $\alpha = 1$) and Fig. 3 shows that the scale of the maximum amplitude in absolute value of the input is 10^{15} in the integer case whereas in both fractional controllers are of order lesser that 10^{11}. Fig. 4 shows the controlled output; note that, since the filter gain of the closed loop transfer function is 1/2, to tracking the reference one must include a proportional factor depending on the reference frequency.

In the second experiment, an unstable plant with $a_0 = -0.1$, $b_0 = 1$ and the same reference as before was employed. Recall that $-a_0 = \theta_1$, $b_0 = \theta_2$. Fig. 5 shows again the order of derivation dependence of the convergence speed, but there is no clear advantage of the integer case in comparison with the fractional adjustment of $\alpha = 0.8$. Fig. 6 shows the error function where again the maximum amplitude is larger in the integer case. Fig. 7 shows no significant difference in fractional or integer adjustment regarding to the input amplitude. Fig. 8 shows the controlled output; note that although for tracking reference one must add an



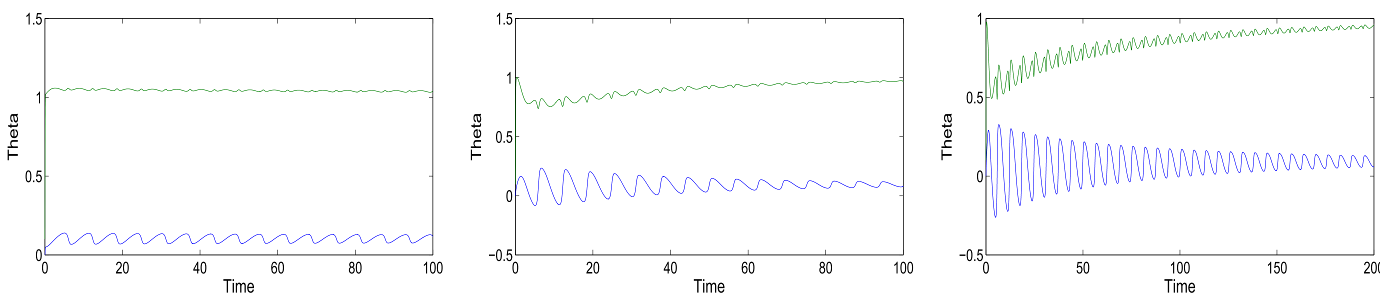
(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 3. Control input: stable plant.



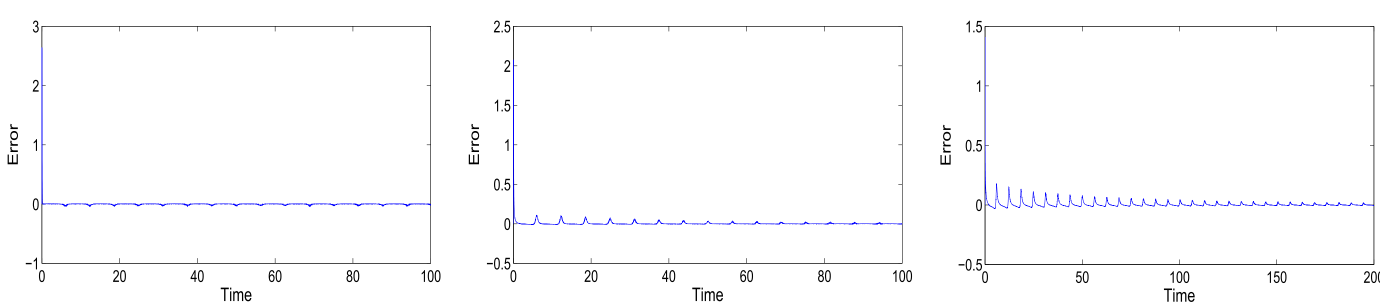
(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 4. Controlled output: stable plant (blue line is the reference signal). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 5. Speed of parameter convergence: $\hat{\theta}$ vs time for the unstable plant.

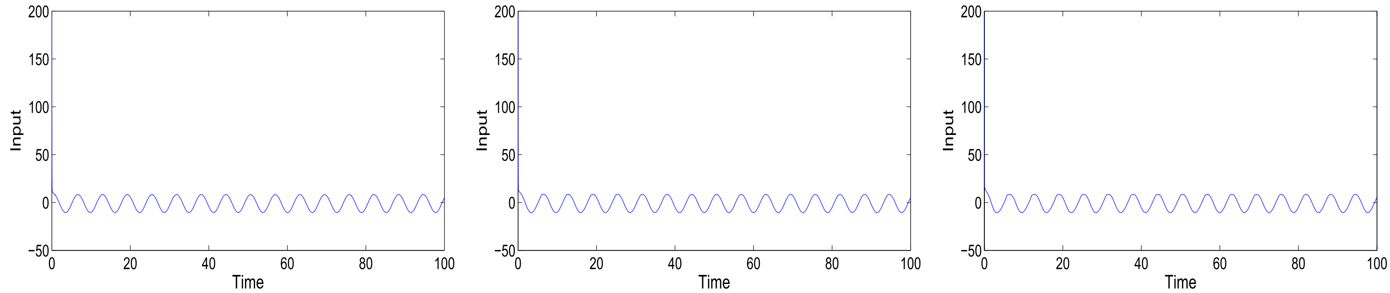


(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 6. Identification error function: unstable plant.

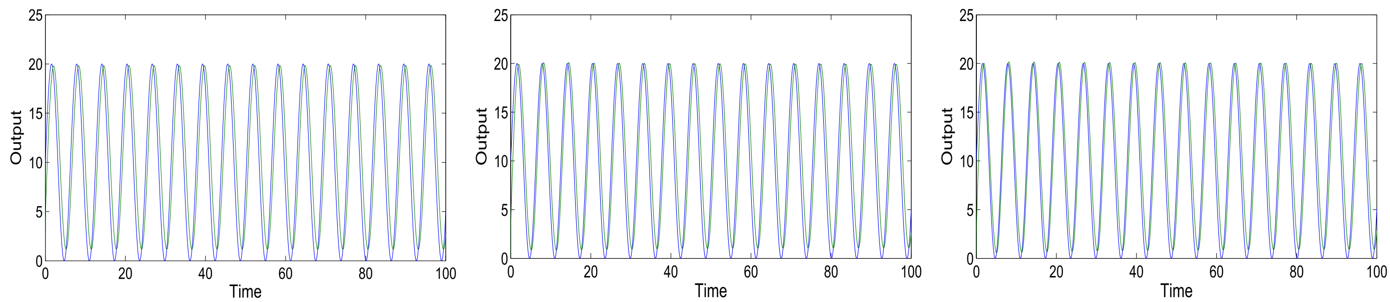
amplification factor depending on the reference frequency, the gain of the closed loop transfer function is now 2.1/2 and a nearly perfect tracking occurs for the unstable plant.

In the third experiment additive band-limited withe noise was introduced in the measured output y. By employing the RMS value of the error function e as the optimization function, for a simulation time of 70, Table 1 shows that minimal values do not occurs at $\alpha = 1$ but at $\alpha = 0.8$ for the stable plant. In Table 2, for the unstable plant, the minimum depends on the noise power and can be fractional.



(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 7. Control input: unstable plant.



(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 8. Controlled output: unstable plant (blue line is the reference signal). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Finally, in Fig. 9 shows the output for an unstable system $a_0 = -0.1$ with r as a step function of amplitude 1 (without enough spectral frequency). Additive band-limited withe noise of power 0.001 was added in the measured output contributing with the necessaries spectral frequencies. This experiment suggests that under noise conditions, reference input tracking can be undertaken even if r has not enough spectral frequencies. Since we are estimating simultaneously the parameters of the plant, the zero frequency gain can be adjusted, to have perfect matching. Note, however, that a nearly perfect tracking occur for the unstable plant when $\alpha = 0.5$ (Fig. 8(c)).

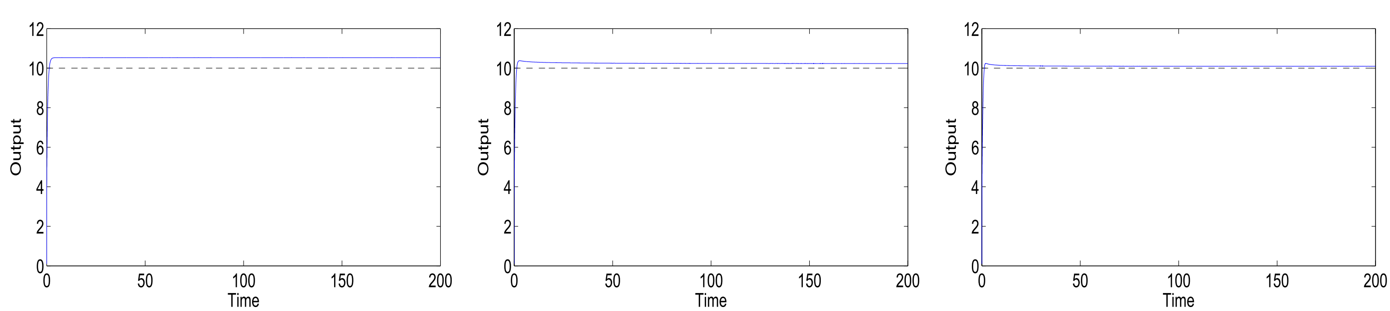
Table 1
RMS error: stable plant.

Power	$\alpha = 1$	$\alpha = 0.8$	$\alpha = 0.5$
0.001	0.00286	0.00202	0.0098
0.01	0.0055	0.00637	0.012
0.1	0.0032	0.0202	0.026

Table 2
RMS error: unstable plant.

Power	$\alpha = 1$	$\alpha = 0.8$	$\alpha = 0.5$
0.001	0.001155	0.0008	0.011
0.01	0.009659	0.0016	0.014
0.1	0.02596	0.0044	0.023

The experiments show that a fractional optimal order can be attained when the objective function considers for instance minimal control amplitude and/or minimal RMS error of the plant output around the reference. In order to find the optimal order for the controllers, the standard methodology to tune parameters used in adaptive control system applications can be followed. First, an objective function J should be selected. This function J will depend on the tune parameters – in particular, on the derivation order coming from the adaptive adjustment of parameters of the controller. If a rough plant model and operation conditions are available the optimization procedure can be initially solved off-line by a computationally efficient method such as particle swarm optimization or genetic algorithms. In this way, an optimal order to adjust control parameters is obtained. If a



(a) $\alpha = 1$ (b) $\alpha = 0.8$ (c) $\alpha = 0.5$

Fig. 9. Controlled output: dashed line is the reference.

rough plant model is not available, an on-line approach can be followed. To that extent, measures of the available signals can be collected during certain period of time. Based on these collected samples, a rough model of the plant is estimated and the optimization proceeds as before. In both cases, the tuning process can be repeated many times until a satisfactory fine tuning of parameters is achieved. The following reference can be consulted for further details on this standard methodology [26].

6. Conclusions

The bounded and convergence of some fractional adaptive systems (including the Fractional Error Model of Type I) employing fractional adaptive laws based on the gradient approach has been presented and analyzed in this paper. It has been demonstrated that convergence for the output error and the parameter error is achieved provided certain conditions related to the newly introduced concept of sufficiently exciting signals are satisfied.

The results obtained indicate that fractional order parameter adjustment is an alternative to the standard integer order adaptive laws to be used in adaptive systems. Simulations performed on adaptive systems reveal that the fractional order of the adaptive laws is a relevant variable of optimization, in particular in transient and robust behavior aspects.

It has yet to be analytically proved that fractional order adaptive laws present advantages over classic integer order adaptive laws.

Acknowledgements

This work has been supported by CONICYT- Chile, under the Grants FB0809 "Centro de Tecnología para la Minería" and FONDECYT 1150488 "Fractional Error Models in Adaptive Control and Applications".

References

[1] Narendra KS, Annaswamy AM. Stable adaptive systems. Englewood Cliffs, NJ: Dover Publications; 1989.

[2] Sastry S, Bodson M. Adaptive control: stability, convergence and robustness. Upper Saddle River, NJ: Prentice Hall; 1989.

[3] Vinagre B, et al. Using fractional order adjustment rules and fractional order reference models in model-reference adaptive control. Nonlinear Dyn 2002;29 (1–4):269–79.

[4] Aguila-Camacho N, Duarte-Mermoud MA. Boundedness of the solutions for certain classes of fractional differential equations with application to adaptive systems. ISA Trans 2016;60:82–8.

[5] Ladaci S, Bensafia Y. Indirect fractional order pole assignment based adaptive control. Eng Sci Technol 2016;19(1):518–30.

[6] Pu Y, et al. Fractional extreme value adaptive training method: fractional steepest descent approach. IEEE Trans Neural Netw Learn Syst 2015;26 (4):653–62.

[7] Yin C, et al. Adaptive minimum energy cognitive lighting control: integer order vs fractional order strategies in sliding mode based extremum seeking. Mechatronics 2013;23(7):863–72.

[8] Shi B, Yuan J, Dong C. On Fractional model reference adaptive control. Sci World J, vol. 2014, 8 pages.

[9] Kilbas A, Srivastava H, Trujillo J. Theory and applications of fractional differential equations. Amsterdam: Elsevier; 2006.

[10] Aguila-Camacho N, Duarte-Mermoud MA, Gallegos JA. Lyapunov functions for fractional order systems. Commun Nonlinear Sci Numer Simul 2014;19 (9):2951–7.

[11] Alikhanov A, Priori A. Estimates for solutions of boundary value problems for fractional-order equations. Differ Equations 2010;46(5):660–6.

[12] Li Y, Chen YQ, Podlubny I. Mittag-Leffler stability of fractional order nonlinear dynamic systems. Automatica 2009;45(8):1965–9.

[13] Gallegos JA, Duarte-Mermoud MA, Aguila-Camacho N, Castro-Linares R. On fractional extensions of Barbalat Lemma. Syst Control Lett 2015;84(1):7–12.

[14] Gallegos JA, Duarte-Mermoud MA. Boundedness and convergence on fractional order systems. J Comput Appl Math 2016;296(C):815–26.

[15] Álvarez F, Cabot A. Steepest descent with curvature dynamical system. J Optim Theory Appl 2004;120(2):247–73.

[16] Gallegos JA, Duarte-Mermoud M. On Lyapunov theory for fractional system. Appl Math Comput 2016;287:161–70.

[17] Morgan A, Narendra K. On the uniform asymptotic stability of certain non-autonomous linear differential equations. SIAM J Control Optim 1977;15(1):5–24.

[18] Narendra KS, Annaswamy AM. Persistent excitation of adaptive systems. Int J Control 1987;45(1):127–60.

[19] Boyd S, Sastry S. Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control Automatica, 22(1986). 1986, p. 629–39.

[20] Malti R. A note on Lp-norms of fractional systems. Autom (J IFAC) 2013;49 (9):2923–7.

[21] Boyd S, Sastry S. On parameter convergence in adaptive control. Syst Control Lett 1983;3(6):311–9.

[22] Anderson B. Exponential stability of linear equations arising in adaptive identification. IEEE Trans Autom Control AC 1977;22:83–8.

[23] Kreisselmeier G. Adaptive observers with exponential rate of convergence. IEEE Trans Autom Control 1977;22(1):2–8.

[24] Duarte-Mermoud MA, Aguila-Camacho N, Gallegos JA, Castro-Linares R. Using general quadratic Lyapunov functions to prove Lyapunov uniform stability for fractional order systems. CNSNS 2014;22(1):650–9.

[25] Bai E, Sastry S. Global stability proofs for continuous-time indirect adaptive control schemes. IEEE Trans Autom Control 1987;32(6):537–43.

[26] Ortiz-Quisbert M, Duarte-Mermoud MA, Milla F, Castro-Linares R, Lefranc G. Optimal Fractional Order Adaptive Controllers for AVR Applications. Electrical Engineering. [Accepted December 1, 2016. Published First Online December 21, 2016]. 2016. http://dx.doi.org/10.1007/s00202-016-0502-2.