

# Finite-time parameter estimation without persistence of excitation

Jian Wang, Denis Efimov, Alexey Bobtsov

► **To cite this version:**

Jian Wang, Denis Efimov, Alexey Bobtsov. Finite-time parameter estimation without persistence of excitation. ECC 2019 - European Control Conference, Jun 2019, Naples, Italy. hal-02084983

**HAL Id: hal-02084983**

**<https://hal.inria.fr/hal-02084983>**

Submitted on 30 Mar 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Finite-time parameter estimation without persistence of excitation

J. Wang, D. Efimov, A.A. Bobtsov

**Abstract**—The problem of adaptive estimation of constant parameters in the linear regressor model is studied without the hypothesis that regressor is Persistently Excited (PE). First, the initial vector estimation problem is transformed to a series of the scalar ones using the method of Dynamic Regressor Extension and Mixing (DREM). Second, several adaptive estimation algorithms are proposed for the scalar scenario. In such a case, if the regressor may be nullified asymptotically or in a finite time, then the problem of estimation is also posed on a finite interval of time. The efficiency of the proposed algorithms is demonstrated in numeric experiments for an academic example.

## I. INTRODUCTION

Estimation and identification of parameters of models of industrial plant's is an important problem, whose solution forms a basis for posterior state estimation and control synthesis [1]. One of the most popular problem statements is based on the static linear regression model [1], [2] (the basic problem of on-line estimation of constant parameters of the  $n$ -dimensional linear regression):

$$\begin{aligned} x(t) &= \omega^\top(t)\theta, \quad t \in \mathbb{R}, \\ y(t) &= x(t) + w(t), \end{aligned} \quad (1)$$

where  $x(t) \in \mathbb{R}$  is the model output,  $\theta \in \mathbb{R}^n$  is the vector of unknown constant parameters that is necessary to estimate,  $\omega : \mathbb{R} \rightarrow \mathbb{R}^n$  is the regressor function (usually assumed bounded and known),  $y(t) \in \mathbb{R}$  is the signal available for measurements with a measurement noise  $w : \mathbb{R} \rightarrow \mathbb{R}$  (here  $\mathbb{R}$  denotes the set of real numbers). A conventional additional requirement, which is usually imposed on the regressor function  $\omega$ , consists in its persistent excitation [2], [3], *i.e.* there are  $\ell > 0$  and  $\vartheta > 0$  such that

$$\int_t^{t+\ell} \omega(s)\omega^\top(s)ds \geq \vartheta I_n$$

for any  $t \in \mathbb{R}$ , where  $I_n$  denotes the identity matrix of dimension  $n \times n$ . It is a well-known fact [4], that if  $\omega$  is PE, then the estimation algorithm

$$\dot{\hat{\theta}}(t) = \gamma \omega(t) \left( y(t) - \omega^\top(t)\hat{\theta}(t) \right), \quad \gamma > 0 \quad (2)$$

is globally exponentially stable at the origin for the estimation error variable  $e(t) = \theta - \hat{\theta}(t)$  provided that  $w(t) \equiv 0$  for all  $t \in \mathbb{R}$ ; and it is input-to-state stable (ISS) for any essentially bounded noise  $w$  [3], [5]. There are several recent results attempting to relax the requirement of persistence of excitation and imposing some nonuniform in initial time restrictions on  $\omega$ , which lead to a (nonuniform) global asymptotic stability of (2) [6], [7], and without

Jian Wang is with Hangzhou Dianzi University, Hangzhou, China.

Denis Efimov is with Inria, Univ. Lille, CNRS, UMR 9189 - CRISTAL, F-59000 Lille, France.

Alexey A. Bobtsov is with Department of Control Systems and Informatics, University ITMO, 197101 Saint Petersburg, Russia.

This work was partially supported by 111 project No. D17019, the Government of Russian Federation (Grant 08-08) and the Ministry of Education and Science of Russian Federation (Project 14.Z50.31.0031).

introducing additional constraints they guarantee only integral ISS property of  $e(t)$  with respect to  $w(t)$  [8]. The linear estimation algorithm (2) is one of the most popular methods to solve (1) and there are many others [2], [3], see for example a recent work [9] and references there, but almost all of them are based on the assumption that the regressor  $\omega(t)$  is PE or has an analogous property. For example, in the concurrent learning approach [10], [11], it is assumed that a kind of PE is satisfied on a finite interval of time only.

In this note we will consider the problem (1) with a similar assumption on an excitation of  $\omega$  during an initial time window. Thus, for our design we will implicitly assume that the norm  $|\omega(t)|$  of the regressor  $\omega(t)$  might be converging to zero (asymptotically or in a finite time), which implies that the norm of the model output  $|x(t)| \leq |\omega(t)||\theta|$  is also converging to zero that leads to a necessity of estimation of  $\theta$  during the time interval when  $|x(t)| \geq |w(t)|$ , *i.e.* when initially the measured output  $y(t)$  disposes the information about  $x(t)$  and it is not hidden completely by the noise  $w(t)$ . Therefore, we will consider the problem of finite-time estimation in (1) for  $t \in [0, T]$ , where  $T$  is fixed, without persistence of excitation. For this purpose, first, we will apply DREM [9] to (1) in order to decouple this vector  $\theta$  estimation problem on a series of independent problems of estimation of scalar parameters. And, second, several algorithms are proposed for the scalar estimation without the regressor excitation and with the convergence in finite-time. Efficiency of the proposed algorithms is demonstrated in simulations for an exponentially converging regressor and a bounded noise.

The outline of this note is as follows. Some preliminary results are introduced in section II. The problem statement is given in Section III. The estimation algorithms and convergence conditions are established in Section IV. A simple illustrating example is considered in Section V.

## Notation

- $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ , where  $\mathbb{R}$  is the set of real number.
- $|x|$  denotes the absolute value for  $x \in \mathbb{R}$  or a vector norm for  $x \in \mathbb{R}^n$ , and the corresponding induced matrix norm for a matrix  $A \in \mathbb{R}^{n \times n}$  is denoted by  $\|A\|$ .
- For a Lebesgue measurable and essentially bounded function  $x : \mathbb{R} \rightarrow \mathbb{R}^n$  denote  $\|x\|_\infty = \sup_{t \in \mathbb{R}} |x(t)|$ , and define by  $\mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$  the set of all such functions with finite norms  $\| \cdot \|_\infty$ ; if

$$\int_{-\infty}^{+\infty} |x(t)|^2 dt < +\infty$$

then this class of functions is denoted by  $\mathcal{L}_2(\mathbb{R}, \mathbb{R}^n)$ .

- A continuous function  $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  belongs to the class  $\mathcal{K}$  if  $\alpha(0) = 0$  and the function is strictly increasing, a function  $\alpha \in \mathcal{K}$  belongs to the class  $\mathcal{K}_\infty$  if it is increasing to infinity.
- The identity matrix of dimension  $n \times n$  is denoted as  $I_n$ .

- A sequence of integers  $1, 2, \dots, n$  is denoted by  $\overline{1, n}$ .
- Define  $\mathbf{e} = \exp(1)$ .

## II. PRELIMINARIES

Consider a time-dependent differential equation [4]:

$$dx(t)/dt = f(t, x(t)), \quad t \geq t_0, \quad t_0 \in \mathbb{R}, \quad (3)$$

where  $x(t) \in \mathbb{R}^n$  is the state vector;  $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$  is a continuous function with respect to  $x$  and piecewise continuous with respect to  $t$ ,  $f(t, 0) = 0$  for all  $t \in \mathbb{R}$ . A solution of the system (3) for an initial condition  $x_0 \in \mathbb{R}^n$  at time instant  $t_0 \in \mathbb{R}$  is denoted as  $X(t, t_0, x_0)$ , and we assume that  $f$  ensures definiteness and uniqueness of solutions  $X(t, t_0, x_0)$  in forward time at least on some finite time interval  $[t_0, t_0 + T)$ .

### A. Stability definitions

Let  $\Omega, \Xi$  be open neighborhoods of the origin in  $\mathbb{R}^n$ ,  $0 \in \Omega \subset \Xi$ .

**Definition 1.** [4], [12] At the steady state  $x = 0$  the system (3) is said to be

(a) *uniformly stable* if for any  $\epsilon > 0$  there is  $\delta(\epsilon)$  such that for any  $x_0 \in \Omega$  and  $t_0 \in \mathbb{R}$ , if  $|x_0| \leq \delta(\epsilon)$  then  $|X(t, t_0, x_0)| \leq \epsilon$  for all  $t \geq t_0$ ;

(b) *uniformly asymptotically stable* if it is uniformly stable and for any  $\kappa > 0$  and  $\epsilon > 0$  there exists  $T(\kappa, \epsilon) \geq 0$  such that for any  $x_0 \in \Omega$  and  $t_0 \in \mathbb{R}$ , if  $|x_0| \leq \kappa$  then  $|X(t, t_0, x_0)| \leq \epsilon$  for all  $t \geq t_0 + T(\kappa, \epsilon)$ ;

(c) *uniformly finite-time stable* if it is uniformly stable and *finite-time converging from  $\Omega$* , i.e. for any  $x_0 \in \Omega$  and  $t_0 \in \mathbb{R}$  there exists  $0 \leq T^{t_0, x_0} < +\infty$  such that  $X(t, t_0, x_0) = 0$  for all  $t \geq T^{t_0, x_0}$ . The function  $T_0(t_0, x_0) = \inf\{T^{t_0, x_0} \geq 0 : X(t, t_0, x_0) = 0 \forall t \geq T^{t_0, x_0}\}$  is called the settling time of the system (3).

If  $\Omega = \mathbb{R}^n$ , then the corresponding properties are called global uniform stability/asymptotic stability/finite-time stability of  $x = 0$ .

Another version of uniform finite-time stability has also been proposed in [13].

In this work we will also be interested in a special stability notion defined not for all  $t_0 \in \mathbb{R}$  as in Definition 1, but for a compact interval of initial times  $t_0$  and only on a fixed interval of time [14], [15], [16], [17]:

**Definition 2.** [18] At the steady state  $x = 0$  the system (3) is said to be

(a) *short-time stable* with respect to  $(\Omega, \Xi, T^0, T_f)$  if for any  $x_0 \in \Omega$  and  $t_0 \in [-T^0, T^0]$ ,  $X(t, t_0, x_0) \in \Xi$  for all  $t \in [t_0, T_f]$ ;

(b) *short-finite-time stable* with respect to  $(\Omega, \Xi, T^0, T_f)$  if it is short-time stable with respect to  $(\Omega, \Xi, T^0, T_f)$  and finite-time converging from  $\Omega$  with the convergence time  $T^{t_0, x_0} \leq T_f$  for all  $x_0 \in \Omega$  and  $t_0 \in [-T^0, T^0]$ ;

(c) *globally short-finite-time stable* if for any bounded set  $\Omega \subset \mathbb{R}^n$  containing the origin there exist a bounded set  $\Xi \subset \mathbb{R}^n$ ,  $\Omega \subset \Xi$  and  $T_f > 0$  such that the system is short-finite-time stable with respect to  $(\Omega, \Xi, T^0, T_f)$  for any  $T^0$ .

In [14], [15], [16], [17] the short-time stability is considered for a fixed initial time instant  $t_0$  only.

*Remark 1.* In the literature, short-time stability [16] is frequently called stability over a finite interval of time [14], [15], [17], but following [18], we prefer here the former notion to avoid a confusion with finite-time stability from [19], [20], since both concepts of stability are used in the paper.

**Lemma 1.** [18] Let the system in (3) with  $d = 0$  possess a Lyapunov function  $V : \mathbb{R} \times \Omega \rightarrow \mathbb{R}_+$ , where  $\Omega \subset \mathbb{R}^n$  is an open neighborhood of the origin, such that for all  $x \in \Omega$  and  $t \in \mathbb{R}$

$$\begin{aligned} \alpha_1(|x|) \leq V(t, x) \leq \alpha_2(|x|), \quad \alpha_1, \alpha_2 \in \mathcal{K}_\infty; \\ \dot{V}(t, x) \leq -\alpha V^\eta(t, x) + k(\varpi t) V^\eta(t, x), \\ \alpha > 0, \quad \varpi \in \mathbb{R}, \quad \eta \in (0, 1) \end{aligned}$$

for a continuous  $k : \mathbb{R} \rightarrow \mathbb{R}$ ,  $k(0) = 0$ . Then there exist  $\varpi \in \mathbb{R}$  and  $T^0 > 0$  such that the system (3) is short-finite-time stable with respect to  $(\Omega', \Xi, T^0, T_f)$  for some  $\Xi \subset \mathbb{R}^n$  with  $\Omega' \subseteq \Omega \subset \Xi$  and  $T_f > 0$ .

### B. Dynamic regressor extension and mixing method

Consider the estimation problem of the vector  $\theta \in \mathbb{R}^n$  in (1) under the following hypothesis:

**Assumption 1.** Let  $\omega \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$  and  $w \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$ .

As it has been proposed in [9], in order to overcome the limitations imposed by the condition that  $\omega$  is PE, the DREM procedure transforms (1) to  $n$  new one-dimensional regression models, which allows independent estimates of  $\theta_i$ ,  $i = \overline{1, n}$  to be computed under a condition on the regressor  $\omega$  that differs from the persistent excitation.

For this purpose  $n - 1$  linear operators  $H_j : \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$  are introduced for  $j = \overline{1, n-1}$  (for instance an operator  $H_j$  can be chosen as a stable linear time-invariant filter with the transfer function  $W_j(s) = \frac{\alpha_j}{s + \beta_j}$ , where  $s \in \mathbb{C}$  is a complex variable and  $\alpha_j \neq 0$ ,  $\beta_j > 0$  are selected to filter the noise  $w$  in (1); or it can realize the delay operation with the transfer function  $W_j(s) = e^{-\tau_j s}$  for  $\tau_j > 0$ ). Note that  $y \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$  under Assumption 1, then these operators are applied to the measured output  $y(t)$  of (1), and using the superposition principles (the operators  $H_j$  are linear) we obtain:

$$\tilde{y}_j(t) = H_j(y(t)) = \tilde{\omega}_j^\top(t)\theta + \tilde{w}_j(t), \quad j = \overline{1, n-1},$$

where  $\tilde{y}_j(t) \in \mathbb{R}$  is the  $j^{\text{th}}$  operator output,  $\tilde{\omega}_j : \mathbb{R} \rightarrow \mathbb{R}^n$  is the  $j^{\text{th}}$  filtered regressor function and  $\tilde{w}_j(t) : \mathbb{R} \rightarrow \mathbb{R}$  is the new  $j^{\text{th}}$  noise signal, which is composed by the transformation of the noise  $w(t)$  by  $H_j$  and other exponentially converging components related to the initial conditions of the filters. By construction  $\tilde{\omega}_j \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$  and  $\tilde{w}_j \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$  for all  $j = \overline{1, n-1}$ . Define new vector

$$\begin{aligned} \tilde{Y}(t) &= [y(t) \tilde{y}_1(t) \dots \tilde{y}_{n-1}(t)]^\top \in \mathbb{R}^n, \\ \tilde{W}(t) &= [w(t) \tilde{w}_1(t) \dots \tilde{w}_{n-1}(t)]^\top \in \mathbb{R}^n \end{aligned}$$

and matrix

$$M(t) = [\omega(t) \tilde{\omega}_1(t) \dots \tilde{\omega}_{n-1}(t)]^\top \in \mathbb{R}^{n \times n}$$

variables, then stacking the original equation (1) with the  $n - 1$  filtered regressor models we design an extended regressor system:

$$\tilde{Y}(t) = M(t)\theta + \tilde{W}(t).$$

For any matrix  $M(t) \in \mathbb{R}^{n \times n}$  the following equality is true [21]:

$$\text{adj}(M(t)) M(t) = \det(M(t)) I_n,$$

even if  $M(t)$  is singular, where  $\text{adj}(M(t))$  is the adjugate matrix of  $M(t)$  and  $\det(M(t))$  is its determinant, and each element

$$\text{adj}(M(t))_{k,s} = (-1)^{k+s} \mathbf{M}_{k,s}(t)$$

for all  $k, s = \overline{1, n}$ , where  $\mathbf{M}_{k,s}(t)$  is the  $(k, s)$  minor of  $M(t)$ , i.e. it is the determinant of the  $(n-1) \times (n-1)$  matrix that results from deleting the  $k^{\text{th}}$  row and the  $s^{\text{th}}$  column of  $M(t)$ . Define

$$Y(t) = \text{adj}(M(t)) \tilde{Y}(t), \quad W(t) = \text{adj}(M(t)) \tilde{W}(t), \\ \phi(t) = \det(M(t)),$$

then multiplying from the left the extended regressor system by the adjugate matrix  $\text{adj}(M(t))$  we get  $n$  scalar regressor models of the form:

$$Y_i(t) = \phi(t)\theta_i + W_i(t) \quad (4)$$

for  $i = \overline{1, n}$ . Again, by construction  $Y \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$ ,  $W \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$  and  $\phi \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$ . For the scalar linear regression model (4) the estimation algorithm (2) takes the form:

$$\dot{\hat{\theta}}_i(t) = \gamma_i \phi(t) \left( Y_i(t) - \phi(t) \hat{\theta}_i(t) \right), \quad \gamma_i > 0 \quad (5)$$

for all  $i = \overline{1, n}$ , where now, contrarily to (2), the estimation processes for all components of  $\theta$  are decoupled, and the adaptation gain  $\gamma_i$  can be adjusted separately for each element of  $\theta$ . However, all these estimation algorithms are dependent on the same regressor  $\phi(t)$  (determinant of  $M(t)$ ).

Define the parameter estimation error as  $e(t) = \theta - \hat{\theta}(t)$ , then its dynamics admits the differential equation:

$$\dot{e}_i(t) = -\gamma_i \phi(t) (\phi(t)e_i(t) + W_i(t)), \quad i = \overline{1, n} \quad (6)$$

and the following result can be proven for the DREM method:

**Proposition 1.** *Consider the linear regression (1) under Assumption 1. Assume that for any  $t_0 \in \mathbb{R}$*

$$\int_{t_0}^{+\infty} \phi^2(t) dt = +\infty \quad (7)$$

for the selected operators  $H_j : \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$ ,  $j = \overline{1, n-1}$ , then for the estimation algorithm (5)

(A) *If  $\|W\|_\infty = 0$ , the system (6) is globally asymptotically stable at the origin iff (7) is valid.*

(B) *For all  $W \in \mathcal{L}_2(\mathbb{R}, \mathbb{R}^n)$  we have  $e \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$ , in addition, for all  $t_0 \in \mathbb{R}$  and  $e_i(t_0) \in \mathbb{R}$ :*

$$|e_i(t)| \leq e^{-\gamma_i \int_{t_0}^t \phi^2(\tau) d\tau} |e_i(t_0)| \\ + \sqrt{\frac{\gamma_i}{2}} \sqrt{1 - e^{-2\gamma_i \int_{t_0}^t \phi^2(\tau) d\tau}} \sqrt{\int_{t_0}^t W_i^2(s) ds}$$

for all  $t \geq t_0$  and  $i = \overline{1, n}$ .

*Proof.* If  $\|W\|_\infty = 0$ , then the system (6) can be rewritten as follows:

$$\dot{e}_i(t) = -\gamma_i \phi^2(t) e_i(t), \quad i = \overline{1, n},$$

and its global uniform stability can be established considering a Lyapunov function  $V(e) = e^T \text{diag}\{\gamma_i^{-1}\}_{i=1}^n e$ . The equivalence of global convergence of  $e(t)$  to zero and (7) has been established in [9]. Thus, the part (A) is proven.

If  $\|W\|_\infty \neq 0$ , then solutions of (6) can be calculated analytically for all  $t \geq t_0$  and  $e_i(t_0) \in \mathbb{R}$ :

$$e_i(t) = e^{-\gamma_i \int_{t_0}^t \phi^2(\tau) d\tau} e_i(t_0) - \gamma_i \int_{t_0}^t e^{-\gamma_i \int_s^t \phi^2(\tau) d\tau} \phi(s) W_i(s) ds$$

for all  $i = \overline{1, n}$ . Further, for all  $t \geq t_0$  and  $e_i(t_0) \in \mathbb{R}$ :

$$|e_i(t)| \leq e^{-\gamma_i \int_{t_0}^t \phi^2(\tau) d\tau} |e_i(t_0)| + \gamma_i \int_{t_0}^t e^{-\gamma_i \int_s^t \phi^2(\tau) d\tau} |\phi(s)| |W_i(s)| ds \\ \leq e^{-\gamma_i \int_{t_0}^t \phi^2(\tau) d\tau} |e_i(t_0)| + \gamma_i \sqrt{\int_{t_0}^t e^{-2\gamma_i \int_s^t \phi^2(\tau) d\tau} \phi^2(s) ds} \sqrt{\int_{t_0}^t W_i^2(s) ds},$$

where the Cauchy–Schwarz inequality was used on the last step. Note that

$$\frac{\partial}{\partial s} e^{-2\gamma_i \int_s^t \phi^2(\tau) d\tau} = 2\gamma_i e^{-2\gamma_i \int_s^t \phi^2(\tau) d\tau} \phi^2(s),$$

hence,

$$2\gamma_i \int_{t_0}^t e^{-2\gamma_i \int_s^t \phi^2(\tau) d\tau} \phi^2(s) ds = 1 - e^{-2\gamma_i \int_{t_0}^t \phi^2(\tau) d\tau},$$

then the desired estimate of the part (B) follows.  $\square$

Obviously, if the signal  $\phi(t)$  is PE, then the error dynamics is ISS with respect to  $W \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$  and an exponential convergence rate can be guaranteed [3], [5], [8].

An interested reader is directed to [9] for a comparison of the condition (7) imposed for  $\phi(t)$  and the requirement that  $\omega(t)$  is PE, and also for the discussion about a possibility to select the operators  $H_j$ ,  $j = \overline{1, n-1}$  in a way enforcing the condition (7) for  $\phi(t)$  while initially  $\omega(t)$  does not admit a persistent excitation. Also an inverse question can be posed: assume that  $\omega(t)$  is PE, is there a guarantee or restrictions to be imposed for the operators  $H_j$ ,  $j = \overline{1, n-1}$  that the condition (7) is satisfied? In other words, can additional filtering, which leads to a decoupled scalar regressor model, destroys good estimation abilities in (1)? But since in this work we do not need the conditions of persistence of excitation, then there is no obstruction for us to use DREM, as it is stated in the problem statement below.

### III. PROBLEM STATEMENT

Consider the static linear regression model (1) under Assumption 1, and assume that the DREM method has been applied in order to reduce the initial problem of vector estimation to  $n$  scalar regressor models in the form (4). Note that  $Y \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$ ,  $W \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R}^n)$  and  $\phi \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$  under Assumption 1 and due to properties of the DREM approach. Instead of the condition (7), which is imposed on an infinite interval of time, assume that  $\omega(t)$  may not admit PE, and that  $\phi(t)$  is a converging function of time:

**Assumption 2.** *There exist  $\rho > 0$ ,  $\kappa > 0$  and  $T > 0$  such that for any  $t_0 \in \mathbb{R}$*

$$|\phi(t)| \geq \kappa e^{-\rho(t-t_0)} \quad \forall t \in [t_0, t_0 + T].$$

It is supposed that the values of  $\rho$ ,  $\kappa$  and  $T$  are known for a designer, denote  $\beta = \kappa e^{-\rho T}$ . It is also assumed that a constant  $\bar{\theta} > 0$  is given such that  $\theta \in \Omega = [-\bar{\theta}, \bar{\theta}]^n$ .

It is necessary to propose an algorithm generating an estimate  $\hat{\theta}(t) \in \mathbb{R}^n$  of the vector of unknown parameters  $\theta \in \mathbb{R}^n$  and providing the property of short-finite-time stability with respect to  $(\Omega, \Omega, T^0, T^0 + T)$  (see Definition 2) of the estimation error  $e(t) = \theta - \hat{\theta}(t)$  dynamics under assumptions 1 and 2 for some suitably defined  $T^0$ .

Since by applying DREM method the problem is decoupled on  $n$  independent ones, for brevity of notation, we will further omit the index  $i$  in (4) by assuming that  $n = 1$ :

$$Y(t) = \phi(t)\theta + W(t), \quad (8)$$

then  $\theta \in \mathbb{R}$ ,  $Y \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$ ,  $W \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$  and  $\phi \in \mathcal{L}_\infty(\mathbb{R}, \mathbb{R})$ . To simplify the notation suppose, without losing generality under Assumption 2, that  $\phi(t) > 0$  for all  $t \in [t_0, t_0 + T]$ .

#### IV. ESTIMATION ALGORITHMS CONVERGING IN SHORT-FINITE-TIME

In this section we will assume that there is no noise and  $\|W\|_\infty = 0$ . Denote  $\lceil s \rceil^\alpha = |s|^\alpha \text{sign}(s)$  for any  $s \in \mathbb{R}$  and  $\alpha \in \mathbb{R}_+$ .

Consider an adaptive estimation algorithm proposed in [22], [23]:

$$\dot{\hat{\theta}}(t) = \gamma \phi(t) \left[ Y(t) - \phi(t) \hat{\theta}(t) \right]^\alpha, \quad \gamma > 0, \quad \alpha \in [0, 1), \quad (9)$$

which admits the following properties:

**Lemma 2.** *Let assumptions 1 and 2 be satisfied. For any  $\alpha \in [0, 1)$  take*

$$\gamma \geq \frac{\bar{\theta}^{1-\alpha}}{T\beta^{1+\alpha}(1-\alpha)},$$

then the estimation error  $e(t) = \theta - \hat{\theta}(t)$  dynamics of (9) with  $\hat{\theta}(t_0) = 0$  is short-finite-time stable with respect to  $(\Omega, \Omega, T^0, T^0 + T)$  for any  $T^0 \in \mathbb{R}_+$ .

All proofs are excluded due to space limitations.

An appealing idea to design a finite- or fixed-time converging system using time-varying feedback is presented in [24]. This idea in our context has an interpretation in the form of the following estimation algorithm:

$$\dot{\hat{\theta}}(t) = \gamma \phi(t) (|\phi(t)| - \beta)^{-\eta} \left( Y(t) - \phi(t) \hat{\theta}(t) \right), \quad \gamma > 0, \quad \eta > 1 \quad (10)$$

for  $\hat{\theta}(t_0) = 0$ , which has well-defined solutions while  $\phi(t)$  is separated with  $\beta$ .

**Lemma 3.** *Let assumptions 1 and 2 be satisfied, then the estimation error  $e(t) = \theta - \hat{\theta}(t)$  dynamics of (10) is short-time stable with respect to  $(\Omega, \Omega, T^0, T^0 + T)$  for any  $T^0 \in \mathbb{R}_+$  and exponentially converging. In addition, if  $|\phi(t_0 + T)| = \beta$ , then the estimation error  $e(t)$  of (10) is short-finite-time stable with respect to  $(\Omega, \Omega, T^0, T^0 + T)$  for any  $T^0 \in \mathbb{R}_+$ .*

For implementation of (10) it is enough that the condition  $|\phi(t_0 + T)| = \beta$  is verified approximately, i.e.  $|\phi(t_0 + T)| = \beta + \varepsilon$  for some  $\varepsilon > 0$  being close to machine computation precision.

And, finally, let us introduce a united version of the algorithms (9) and (10), which borrows the nonlinear paradigm of the former and uses the time-varying feedback of the latter:

$$\dot{\hat{\theta}}(t) = \gamma \phi^{\frac{\phi(t)-\beta}{\phi_{\max}-0.5\beta}}(t) \left[ Y(t) - \phi(t) \hat{\theta}(t) \right]^{\frac{\phi(t)-\beta}{\phi_{\max}-0.5\beta}}, \quad \gamma > 0, \quad (11)$$

where  $\phi_{\max} = \max_{t_0 \in [-T^0, T^0]} \max_{t \in [t_0, t_0+T]} |\phi(t)|$  for some given  $T^0 \in \mathbb{R}_+$  and  $\hat{\theta}(t_0) = 0$ .

**Lemma 4.** *Let assumptions 1 and 2 be satisfied. Take*

$$\gamma \geq \frac{\bar{\theta}^{1-\alpha}(1-\alpha)^{-1}T^{-1}}{\min\{1, \bar{\theta}^{-\alpha-1}\} \min\{1, \beta^{2\alpha}\}},$$

where  $\alpha = \frac{\phi_{\max}-\beta}{\phi_{\max}-0.5\beta}$ , then the estimation error  $e(t) = \theta - \hat{\theta}(t)$  dynamics of (11) is short-finite-time stable with respect to  $(\Omega, \Omega, T^0, T^0 + T)$ . In addition, if  $\phi(t_0 + T) = \beta$ , then for any  $\phi_{\max} > 0.5\beta$  there exists  $\gamma > 0$  such that the estimation error  $e(t)$  of (11) is short-finite-time stable with respect to  $(\Omega, \Omega, T^0, T^0 + T)$  for any  $T^0 \in \mathbb{R}_+$ .

Thus, the idea of the algorithm (11) consists in the utilization of a nonlinearity such that for  $\phi(t) < \min\{1, \beta + 1\}$  the function  $\phi^{\frac{\phi(t)-\beta}{\phi_{\max}-0.5\beta}}(t)$  becomes growing overcoming the absence of excitation in the system. The price for that is a possible discontinuity

at  $t_0 + T$ . However, if  $\phi(t_0 + T) > \beta$  and stay very small, then the algorithm stays continuous being independent in the excitation of (1).

*Remark 2.* If the regressor  $\phi(t)$  is just asymptotically converging, then all these algorithms, (9), (10) and (11), can be applied for any finite  $T > 0$  and  $t_0 \in \mathbb{R}$ .

*Remark 3.* Following the concurrent learning approach [10], [11], in order to relax the dependence of the algorithm (5) on persistence of excitation, it can be extended by an auxiliary static feedback (a kind of  $\sigma$ -modification [10]) whose gains are updated at isolated instants of time  $t_i \in [t_0, t_0 + T]$ ,  $i = 1, 2, \dots$ :

$$\begin{aligned} \dot{\hat{\theta}}(t) &= \gamma \phi(t) \left( Y(t) - \phi(t) \hat{\theta}(t) \right) \\ &\quad + \gamma_i \phi(t_i) \left( Y(t_i) - \phi(t_i) \hat{\theta}(t) \right), \end{aligned} \quad (12)$$

where  $\gamma_i > 0$  are tuning gains. The instants  $t_i$  can be selected after a dwell time  $\tau > 0$  under a restriction on sufficiency of excitation, e.g.:

$$t_i = \arg \inf_{t \geq t_{i-1} + \tau} |\phi(t)| \geq \bar{\varrho}$$

for all  $i = 1, 2, \dots$  and some  $\bar{\varrho} > 0$ . The gain  $\gamma_i$  can be chosen to ensure a uniform influence of the static feedback:

$$\gamma_i \phi^2(t_i) \approx \gamma_{i-1} \phi^2(t_{i-1}).$$

It is worth to highlight that the same extension can also be applied to any algorithm proposed in this note. For example, for (9) we will obtain:

$$\begin{aligned} \dot{\hat{\theta}}(t) &= \gamma \phi(t) \left[ Y(t) - \phi(t) \hat{\theta}(t) \right]^\alpha \\ &\quad + \gamma_i \phi(t_i) \left[ Y(t_i) - \phi(t_i) \hat{\theta}(t) \right]^\alpha, \end{aligned} \quad (13)$$

where all parameters save their meaning.

*Remark 4.* Note also that in this scalar case the multiplicative terms  $\phi(t)$  and  $\phi^{\frac{\phi(t)-\beta}{\phi_{\max}-0.5\beta}}(t)$  in the algorithms (9) and (11), respectively, can be replaced with  $\text{sign}(\phi(t))$ . Then the algorithms are optimized for the convergence rate. Inversely, if the term  $\phi(t)$  is kept in both estimators, e.g. (11) is transformed to

$$\dot{\hat{\theta}}(t) = \gamma \phi(t) \left[ Y(t) - \phi(t) \hat{\theta}(t) \right]^{\frac{\phi(t)-\beta}{\phi_{\max}-0.5\beta}}, \quad (14)$$

then the noise filtering is improved (since for big values of the time in such a case the derivative of  $\hat{\theta}(t)$  is proportional to  $\phi(t)W(t)$  for a converging regressor  $\phi(t)$ ).

*Remark 5.* One of the most important features of estimation algorithms, after estimation error convergence in the ideal case, is their robustness with respect to measurement noises. In our case, since the regressor  $\phi(t)$  may converge to zero, the appearance of  $W(t) \neq 0$  additionally limit the time of convergence, since it is reasonable to use the output  $Y(t)$  for estimation with  $t \in [t_0, t_0 + T]$  only while

$$|Y(t)| > |W(t)| + \varepsilon$$

for some  $\varepsilon > 0$ . If  $|Y(t)| \leq |W(t)| + \varepsilon$  (or  $|Y(t)|$  is almost equal to  $|W(t)|$  for a sufficiently small  $\varepsilon$ ), then the measured output mainly contains the measurement noise, and it is ambiguous to ask an algorithm to estimate  $\theta$  due to a bad ratio between the signal and the noise. Therefore, anyway it is reasonable to stop the estimation process when

$$|\phi(t)| \leq \underline{\varrho} \quad (15)$$

for some  $\underline{\varrho} > 0$  related with available upper bound on the noise  $W(t)$ . The proposed algorithms (9), (10) and (11) have different robust stability properties with respect to  $W$ , as we will demonstrate in the next section for an example, while the analytical investigations we will keep for a future research (the main issue is a necessity of introduction of an analogue of ISS property for short-finite-time stability in the noise-free case).

## V. EXAMPLE

Select

$$\begin{aligned} \theta &= 1, \bar{\theta} = 1, t_0 = 0, T = 5, \\ \phi'(t) &= 2\beta + \cos^2(t)e^{-t}, \beta = 0.01, \end{aligned}$$

where  $\phi'(t)$  is the ideal regressor. We assume that the regressor and the output are measured with a noise:

$$\begin{aligned} Y(t) &= \phi'(t)\theta + \delta (\cos(10t) + 5\nu_1(t)e^{-t}), \\ \phi(t) &= \phi'(t) + \delta\nu_2(t), \end{aligned}$$

where  $\nu_j(t) \in [-1, 1]$ ,  $j = 1, 2$  are uniformly distributed stochastic signals and  $\delta \geq 0$  determines the noise amplitude, then

$$W(t) = \delta (\cos(10t) + 5\nu_1(t)e^{-t} - \theta\nu_2(t)).$$

Let  $\alpha = 0.5$  for the algorithm (9),  $\eta = 1.5$  for the algorithm (10),  $\gamma = 1$  and  $\phi_{\max} = \phi(0)$  for the algorithm (11). For simplicity,  $\gamma = 5$  for all algorithms, and the rule (15) is used to anticipate the noise presence with  $\underline{\varrho} = 0.02$ . The algorithms (5) and (9) are augmented by by concurrent learning terms at  $t_1 = 0.1T$  with  $\gamma_1 = 0.5$  in (12) and (13), respectively. The version (14) of the algorithm (11) is used to improve its transients in noisy scenarios.

For the case  $\delta = 0$  ( $\|W\|_\infty = 0$ ) the results of simulation are shown in Fig. 1, for  $\delta = 0.01$  the results are given in Fig. 2 and for  $\delta = 0.1$  in Fig. 3, where the blue line represents the linear algorithm (12), the magenta line stays for (13), the green line corresponds to (10) and the red line is for (14). As we can conclude from these results, the linear algorithm (12) suffers from the absence of excitation (due to presence of the concurrent learning term it is converging to the right value, but not on the simulated time interval; then increasing of  $\gamma$  or  $\gamma_1$  may solve the convergence rate issue, but at the price of quality degradation in the presence of noise, as usual), all the rest converge in a finite time to the ideal value of  $\theta$  in the absence of noise, however, in the presence of noise the algorithm (10) starts to oscillate (since the perturbation  $W(t)$  is divided by the time-varying gain approaching 0), while the algorithms (13) and (14) generate bounded trajectories, but (14) becomes more sensitive to the noise since it is approaching discontinuity. All these results confirm the theoretical findings of this note.

## VI. CONCLUSIONS

The problem of adaptive estimation of a vector of constant parameters in the linear regressor model is studied without the hypothesis that regressor is PE. For this purpose, the initial estimation problem is transformed to a series of the scalar ones using the DREM approach. Several adaptive estimation algorithms are proposed for the scalar scenario converging in a short-finite-time. Future directions of research have to include robustness analysis, relaxation of the introduced hypotheses and tuning guidelines development for the proposed estimation algorithms.

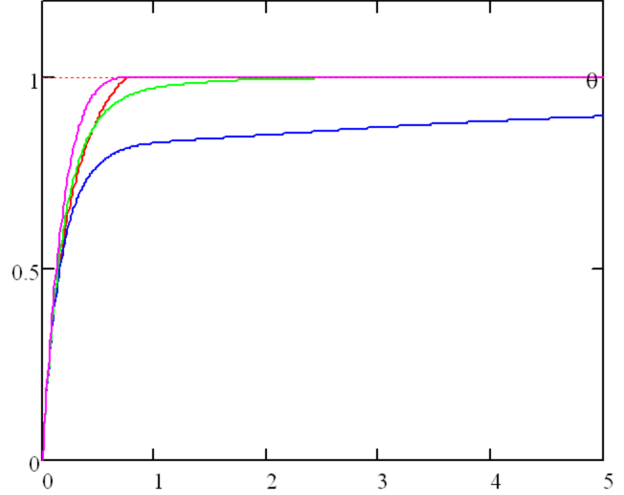


Figure 1. The results of simulation without noise

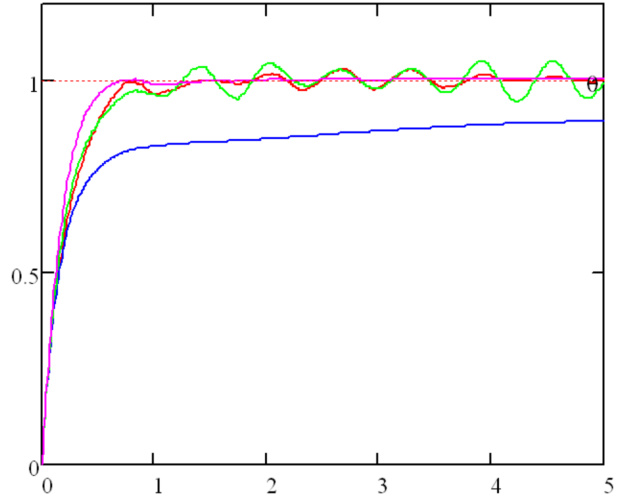


Figure 2. The results of simulation with a small noise

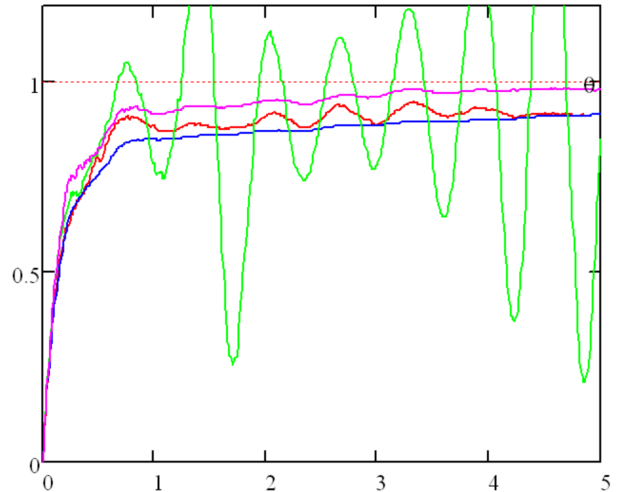


Figure 3. The results of simulation with a big noise

## REFERENCES

- [1] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice-Hall, 1987.
- [2] S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence and Robustness*. London: Prentice-Hall, 1989.
- [3] K. Narendra and A. Annaswamy, "Persistent excitation in adaptive systems," *International Journal of Control*, vol. 45, no. 1, pp. 127–160, 1987.
- [4] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, New Jersey: Prentice Hall, 2002.
- [5] D. Efimov and A. Fradkov, "Design of impulsive adaptive observers for improvement of persistency of excitation," *Int. J. Adaptive Control and Signal Processing*, vol. 29, no. 6, pp. 765–782, 2015.
- [6] L. Praly, "Convergence of the gradient algorithm for linear regression models in the continuous and discrete time cases," December 26 2016, int. Rep. MINES ParisTech, Centre Automatique et Systèmes. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01423048>
- [7] N. Barabanov and R. Ortega, "On global asymptotic stability of  $\dot{x} = -\phi(t)\phi^\top(t)x$  with  $\phi(t)$  bounded and not persistently exciting," *Systems & Control Letters*, vol. 109, pp. 24–27, 2017.
- [8] D. Efimov, N. Barabanov, and R. Ortega, "Robust stability under relaxed persistent excitation conditions," in *Proc. IEEE CDC*, 2018, submitted.
- [9] S. Aranovskiy, A. Bobtsov, R. Ortega, and A. Pyrkin, "Performance enhancement of parameter estimators via dynamic regressor extension and mixing," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3546–3550, 2017.
- [10] G. Chowdhary, T. Yucelen, M. Mühlegg, and E. N. Johnson, "Concurrent learning adaptive control of linear systems with exponentially convergent bounds," *International Journal of Adaptive Control and Signal Processing*, vol. 27, no. 4, pp. 280–301, 2012.
- [11] R. Kamalapurkar, P. Walters, and W. E. Dixon, "Model-based reinforcement learning for approximate optimal regulation," *Automatica*, vol. 64, pp. 94–104, 2016.
- [12] W. M. Haddad, S. G. Nersisov, and L. Du, "Finite-time stability for time-varying nonlinear dynamical systems," in *Proc. ACC*, Seattle, 2008, pp. 4135–4139.
- [13] Y. Orlov, "Finite time stability and robust control synthesis of uncertain switched systems," *SIAM J. Control Optim.*, vol. 43, no. 4, pp. 1253–1271, 2005.
- [14] G. Kamenkov, "On stability of motion over a finite interval of time," *Journal of Applied Math. and Mechanics (PMM)*, vol. 17, pp. 529–540, 1953.
- [15] A. Lebedev, "The problem of stability in a finite interval of time," *Journal of Applied Math. and Mechanics (PMM)*, vol. 18, pp. 75–94, 1954.
- [16] P. Dorato, "Short-time stability in linear time-varying systems," Ph.D. dissertation, Polytechnic Institute of Brooklyn, New York, 1961.
- [17] L. Weiss and E. Infante, "On the stability of systems defined over a finite time interval," *Proc. of the National Academy of Sciences*, vol. 54, pp. 440–448, 1965.
- [18] H. Ríos, D. Efimov, A. Polyakov, and W. Perruquetti, "Homogeneous time-varying systems: Robustness analysis," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4075–4080, 2016.
- [19] E. Roxin, "On finite stability in control systems," *Rendiconti del Circolo Matematico di Palermo*, vol. 15, pp. 273–283, 1966.
- [20] S. P. Bhat and D. S. Bernstein, "Geometric homogeneity with applications to finite-time stability," *Mathematics of Control, Signals and Systems*, vol. 17, pp. 101–127, 2005.
- [21] F. Gantmacher, *The Theory of Matrices*. New York: Chelsea, 1960.
- [22] H. Ríos, D. Efimov, J. A. Moreno, W. Perruquetti, and J. G. Rueda-Escobedo, "Time-varying parameter identification algorithms: Finite and fixed-time convergence," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3671–3678, 2017.
- [23] H. Ríos, D. Efimov, and W. Perruquetti, "An adaptive sliding-mode observer for a class of uncertain nonlinear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 32, no. 3, pp. 511–527, 2018.
- [24] Y. Song, Y. Wang, J. Holloway, and M. Krstic, "Time-varying feedback for regulation of normal-form nonlinear systems in prescribed finite time," *Automatica*, vol. 83, pp. 243–251, 2017.