2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Enhanced Parameter Convergence for Linear Systems Identification: The DREM Approach*

A. Belov[1], S. Aranovskiy[2], R. Ortega[3], N. Barabanov[4] and A. Bobtsov[5]

Abstract— **Dynamic regressor extension and mixing is a new technique for parameter estimation that has proven instrumental in the solution of several open problems in system identification and adaptive control. A key property of the estimator is that, for linear regression models, it guarantees monotonicity of *each* element of the parameter error vector that is a much stronger property than monotonicity of the vector *norm*, as ensured with classical gradient or least-squares estimators. On the other hand, the overall performance improvement of the estimator is strongly dependent on the suitable choice of certain operators that enter in the design. In this paper we investigate the impact of these operators on the convergence properties of the estimator in the context of identification of linear time-invariant systems. In particular, we give some guidelines for their selection to ensure convergence under the same (persistence of excitation) conditions as standard identification schemes.**

I. INTRODUCTION

A new procedure to design *parameter estimators* for linear and nonlinear regressions, called dynamic regressor extension and mixing (DREM), was recently proposed in [2]. The technique has been successfully applied in a variety of identification and adaptive control problems [3], [5], [6], [7], [15]. For linear regressions DREM estimators outperform classical gradient or least-squares estimators in the following precise aspect: independently of the excitation conditions, DREM guarantees monotonicity of *each* element of the parameter error vector that is much stronger than monotonicity of the vector *norm*, which is ensured with classical estimators. Another interesting property of DREM is that its convergence is established without the usual, restrictive requirement of regressor persistence of excitation (PE) [11], [16]. Instead of PE a non-square integrability condition on the determinant of a designer-dependent extended (square) regressor matrix is imposed. Similarly to instrumental variable methods [8] where the regression model is multiplied by some signals to generate new regressions, in DREM new regressions are created selecting a certain number of linear, stable operators, which act on the linear regression to create new regressors (with filtered signals), which are then pile up on the aforementioned matrix. Multiplying by the adjoint of this matrix generates a series of independent *scalar* regressions for each of the unknown parameters with the determinant of the matrix being the common regressor to all of them. The non-square integrability of this determinant is, then, the necessary and sufficient condition for parameter convergence. To make the paper self-contained a brief description of DREM as applied in identification problems is given in the next section—see [2] for a more general and detailed presentation of DREM and [13] for its reformulation as a functional Luenberger observer.

Clearly, the overall performance of the estimator is strongly dependent on the suitable choice of the aforementioned operators. Roughly speaking, they should be selected to generate new (filtered) regressors that are, as much as possible, linearly independent among them. Prior information on the spectral content of the regressor may then be used to select these operators, which may be selected as linear time-invariant (LTI) band-pass filters or simple delays. In this paper we investigate the impact of these operators on the convergence properties of the estimator in the context of parameter identification of linear time-invariant stable systems. In this case, the (original) regressor is generated applying some LTI filters to the systems input. It is well known, *cf.*, Theorem 2.7.3 of [16], that standard gradient and least-squares algorithms will generate a globally exponentially convergent estimate of the parameters if and only if the input signal contains a sufficient number of spectral

*This paper supported by the Ministry of Education and Science of Russian Federation (goszadanie no. 8.8885.2017/8.9) and Government of the Russian Federation (Grant 074-U01) through ITMO Postdoctoral Fellowship program

[1]A. Belov is with Department of Control Systems and Informatics, ITMO University, 197101 Saint Petersburg, Russia and with Laboratory of Dynamics of Control Systems, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow 117997, Russia

[2]S. Aranovskiy is with Equipe Automatique, CentaleSupélec – IETR, Avenue de la Boulaie - CS 47601, 35576 Cesson-Sévigné, France and with Department of Control Systems and Informatics, ITMO University, 197101 Saint Petersburg, Russia, stanislav.aranovskiy@centralesupelec.fr

[3]R. Ortega is with Laboratoire des Signaux et Systèmes, CNRS-CentraleSupélec, Plateau du Moulon, 91192 Gif-sur-Yvette, France

[4]N. Barabanov is Department of Mathematics, North Dakota State University, Fargo, ND 58105, USA and with Department of Control Systems and Informatics, ITMO University, Saint Petersburg, 197101, Russia

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

lines—a property called “sufficient richness” in [16]—where it is also shown to be equivalent to having a PE regressor.

Two natural questions arise in this respect.

(Q1) Can DREM relax the assumption of sufficiently rich input? More precisely, is there a suitable selection of the operators of DREM such that parameter convergence is ensured even if the PE assumption on the regressor is *not satisfied*?

(Q2) If the regressor is PE will DREM ensure parameter convergence for a well-defined class of operators? In [2] it is shown that there exists a “bad choice” of operators, in the sense that applied to a PE regressor generates an (asymptotically) singular extended regressor matrix and, consequently, DREM will not work. Therefore, the question is how to verify that the chosen operators are not “bad”.

In the paper we give answers to the previous questions. Unfortunately, the answer to (Q1) is negative even allowing for arbitrary linear, possibly time-varying, $\mathcal{L}_\infty$-stable operators. On the other hand, we give a positive answer to (Q2) for LTI filters and delay operators.

The remaining of the paper is organized as follows. The application of DREM for identification of an LTI system parameters is presented in Section II. In Section III we give the answer to (Q1) while the answer to (Q2) is presented in Section IV. Some simulation results that illustrate our results and show the performance improvement of DREM, with respect to gradient estimators, are given in Section V. The paper is wrapped–up with some conclusions and future work in Section VI. The proof of the main claim in Section IV, being notationally involved, is deferred to an appendix, where a preliminary lemma is also presented.

II. Parameter Identification of LTI Systems

In this section we briefly review the problem of parameter identification of LTI systems using the classical gradient algorithm and the new DREM estimator. For more details on system identification the reader is referred to [11], [12], [16].

A. Problem formulation and classical solution

We are interested in the classical problem of parameter identification of the scalar LTI continuous–time plant

$$A(p)y(t) = B(p)u(t) \tag{1}$$

where $y(t),\ u(t)$ are the plant output and input, respectively, $A(p) = \sum_{i=0}^{n} a_i p^i, \quad B(p) = \sum_{i=0}^{n-1} b_i p^i,\ p := \frac{d}{dt},\ a_n = 1$, $A(p)$ and $B(p)$ are coprime with unknown coefficients. We make the standard assumptions that $A(p)$ is a Hurwitz polynomial, $u(t)$ is regular and bounded and n is known.

In [16] it is shown that the system (1) can be represented in the linear regression form

$$y(t) = \phi^\top(t)\theta + \epsilon_t \tag{2}$$

where

$$F(p) := \frac{1}{\lambda(p)}\begin{bmatrix} 1 \\ p \\ \vdots \\ p^{n-1}\end{bmatrix},\ \theta := \begin{bmatrix} \lambda_0 - a_0 \\ \vdots \\ \lambda_{n-1} - a_{n-1} \\ b_0 \\ \vdots \\ b_{n-1}\end{bmatrix}, \tag{3}$$
$$\phi(t) := \begin{bmatrix} \frac{F(p)B(p)}{A(p)} \\ F(p)\end{bmatrix} u(t),$$

$\lambda(p) = \sum_{i=0}^{n} \lambda_i p^i,\ \ \lambda_n = 1$, is an arbitrary Hurwitz polynomial and ϵ_t is the generic notation for an exponentially decaying term due to the filters initial conditions that, without loss of generality, we neglect in the sequel.[1]

The standard gradient estimator

$$\dot{\hat{\theta}}(t) = \Gamma\phi(t)[y(t) - \phi^\top(t)\hat{\theta}(t)],\ \Gamma > 0, \tag{4}$$

yields the error equation

$$\dot{\tilde{\theta}}(t) = -\Gamma\phi(t)\phi^\top(t)\tilde{\theta}(t), \tag{5}$$

where $\tilde{\theta}(t) := \hat{\theta}(t) - \theta$ are the parameter estimation errors.

Evaluating the derivative of $|\tilde{\theta}(t)|^2$, with $|\cdot|$ the Euclidean norm, is easy to show that

$$|\tilde{\theta}(t)| \le |\tilde{\theta}(0)|,\ \forall t \ge 0. \tag{6}$$

Also, it is well–known [1], [16] that the zero equilibrium of the linear time–varying system (5) is globally exponentially stable if and only if the regressor vector $\phi(t)$ is PE, that is, if

$$\int_t^{t+T} \phi(s)\phi^\top(s)ds \ge \delta I,$$

for some $T, \delta > 0$ and for all $t \ge 0$, which will be denoted as $\phi(t) \in$ PE. The PE condition of $\phi(t)$ is translated to the input signal $u(t)$ via the following fundamental result.

Proposition 1 ([16], Theorems 2.7.2 and 2.7.3): Consider the vector $\phi(t)$ defined in (3) *with $u(t)$ given by*

$$u(t) = \sum_{k=1}^{N} A_k \sin(\omega_k t), \tag{7}$$

with $\omega_k \neq \omega_j, \forall k \neq j$ and $A_k \neq 0$. Then,

$$\phi(t) \in PE \ \Leftrightarrow\ N \ge n.$$

Remark 1: For ease of presentation we consider only a particular case of the more general result reported in [16]. In

[1] See [16] and Remark 3 in [2] where the effect of these term is rigorously analysed.

particular, the translation of the PE condition of the regressor to a suitable excitation of the input is established for all regular signals admitting a suitable spectral decomposition without assuming it is of the form (7).

Remark 2: In [4] conditions on $\phi(t)$ for global asymptotic (but not exponential) stability of (5), which are strictly weaker than PE, are given. It is not clear at this point how these conditions are related with the input signal in the present identification context.

B. Dynamic regressor extension and mixing estimator

To apply DREM in the identification problem the first step is to introduce a *linear, single-input* $2n$*-output,* $\mathcal{L}_\infty$*–stable* operator $\mathcal{H}:\mathcal{L}_\infty \to \mathcal{L}_\infty^{2n}$, and define the vector $Y \in \mathbb{R}^{2n}$ and the matrix $\Phi \in \mathbb{R}^{2n\times 2n}$ as

$$\begin{aligned} Y(t) &:= \mathcal{H}[y(t)] \\ \Phi(t) &:= \mathcal{H}[\phi^\top(t)]. \end{aligned} \tag{8}$$

Clearly, because of linearity and $\mathcal{L}_\infty$ stability, these signals satisfy

$$Y(t) = \Phi(t)\theta + \epsilon_t. \tag{9}$$

The elements of the operator $\mathcal{H}$ may be simple, exponentially stable *LTI filters* of the form[2]

$$\mathcal{H}_i(p) = \frac{\alpha_i}{p+\beta_i},\ i \in \{1,2,\ldots,2n\}$$

with $\alpha_i \neq 0$, $\beta_i > 0$. Another option of interest are *delay operators*, that is

$$[\mathcal{H}_i(\cdot)](t) := (\cdot)(t-d_i),$$

where $d_i \in \mathbb{R}_+$. See Section 4 of [13] for the case of general LTV operators.

Pre-multiplying (9) by the *adjunct matrix* of $\Phi(t)$, denoted $\mathrm{adj}\{\Phi(t)\}$, we get $2n$ *scalar* regressors of the form

$$\mathcal{Y}_i(t) = \Delta(t)\theta_i, \tag{10}$$

where we defined the scalar function $\Delta(t) \in \mathbb{R}$

$$\Delta(t) := \det\{\Phi(t)\},$$

and the vector $\mathcal{Y}(t) \in \mathbb{R}^{2n}$

$$\mathcal{Y}(t) := \mathrm{adj}\{\Phi(t)\}Y(t).$$

The estimation of the parameters θ_i from the scalar regression form (10) can be easily carried out via

$$\dot{\hat{\theta}}_i(t) = \gamma_i\Delta(t)(\mathcal{Y}_i(t) - \Delta(t)\hat{\theta}_i(t)), \tag{11}$$

with adaptation gains $\gamma_i > 0$. From (10) it is clear that the latter equations are equivalent to

$$\dot{\tilde{\theta}}_i(t) = -\gamma_i\Delta^2(t)\tilde{\theta}_i(t). \tag{12}$$

[2]In the sequel the clarification $i \in \{1,2,\ldots,2n\}$ is omitted for brevity.

A first important advantage of DREM is that the *individual* parameter errors satisfy

$$|\tilde{\theta}_i(t)| \leq |\tilde{\theta}_i(0)|,\ \forall\, t \geq 0, \tag{13}$$

that is *strictly stronger* than the monotonicity property (6). Moreover, solving the simple scalar differential equation (12) as

$$\tilde{\theta}_i(t) = e^{-\gamma_i\int_0^t \Delta^2(s)ds}\tilde{\theta}_i(0).$$

shows that

$$\lim_{t\to\infty}\tilde{\theta}_i(t) = 0 \quad \iff \quad \Delta(t) \notin \mathcal{L}_2,$$

that is, parameter convergence is established without the restrictive PE assumption. Moreover, if $\Delta(t) \in \mathrm{PE}$, the convergence of DREM is exponential.

The relationship between the condition $\Delta(t) \notin \mathcal{L}_2$ and $\phi(t) \in \mathrm{PE}$ is far from obvious for *arbitrary* regressor vectors $\phi(t)$—see [2] for examples that show that neither one of the conditions is stronger than the other. However, for the particular case of identification, when $\phi(t)$ is generated via (3), the relation between these assumptions can be clarified, which constitutes the main contribution of this paper.

Remark 3: The importance of having established scalar regressor models for each of the unknown parameters can hardly be overestimated. Besides the important element-by-element monotonicity property of the parameter errors captured by (13), this feature is instrumental to eliminate the need to overparameterise nonlinear regressions to obtain a linear one—a practice that, as is well-known [11], [12], [16], entails a serious performance degradation. This, and other advantages of DREM, have been discussed in a series of publications including [2], [3], [5], [6], [7], [13], [15]

Remark 4: It is well-known that non-square integrability and PE of a signal are *not equivalent* properties—even in the scalar case. For instance, the signal $\frac{1}{\sqrt{1+t}}$ is not in $\mathcal{L}_2$ but it is not PE, on the other hand, all PE signals are not in $\mathcal{L}_2$. Besides this issue, the comparison of the convergence conditions of gradient and DREM estimators is further complicated by the fact that $\Delta(t)$ and $\phi(t)$ are related via, not just the action of the linear operator $\mathcal{H}$, but also by the nonlinear operation of the determinant computation.

III. DREM Cannot Relax the PE Condition

In this section we give the answer, alas negative, to the question (Q1) of Section I.

Proposition 2: Consider the vector $\phi(t)$ *generated via* (3) *with* $u(t)$ *given by* (7)*. Define the function* $\Delta(t)$ *as*

$$\Delta(t) = \det\{\mathcal{H}[\phi^\top(t)]\} \tag{14}$$

where $\mathcal{H}$ is an arbitrary *linear, single-input $2n$-output, $\mathcal{L}_\infty$–stable operator. Then,*

$$N < n \;\Rightarrow\; \Delta(t) \in \mathcal{L}_2.$$

In other words, independently of the choice of the operator $\mathcal{H}$, a necessary *condition for DREM to ensure global convergence of the parameter error is $\phi(t) \in$ PE.*

Proof: From (3) and (7) it is clear that

$$\phi_i(t) = \phi_i^{\mathtt{ss}}(t) + \phi_i^{\mathtt{tr}}(t),$$

where the steady-state component is given by

$$\phi_i^{\mathtt{ss}}(t) := \sum_{k=1}^{N} A_{i,k} \sin(\omega_k t + \psi_{i,k}),$$

with $A_{i,k}$ and $\psi_{i,k}$ constants and the transient component $\phi_i^{\mathtt{tr}}(t)$ tends to zero exponentially fast. Piling up all the components we can write the steady-state vector in a compact form as

$$\phi_{\mathtt{ss}}(t) := \mathrm{col}(\phi_1^{\mathtt{ss}}(t), \ldots, \phi_{2n}^{\mathtt{ss}}(t)) = X^\top \xi(t) \tag{15}$$

where $X \in \mathbb{R}^{2N \times 2n}$ is given by

$$X := \begin{bmatrix} A_{1,1}\cos(\psi_{1,1}) & \cdots & A_{2n,1}\cos(\psi_{2n,1}) \\ \vdots & \ddots & \vdots \\ A_{1,N}\cos(\psi_{1,N}) & \cdots & A_{2n,N}\cos(\psi_{2n,N}) \\ A_{1,1}\sin(\psi_{1,1}) & \cdots & A_{2n,1}\sin(\psi_{2n,1}) \\ \vdots & \ddots & \vdots \\ A_{1,N}\sin(\psi_{1,N}) & \cdots & A_{2n,N}\cos(\psi_{2n,N}) \end{bmatrix} \tag{16}$$

and

$$\xi(t) := \begin{bmatrix} \sin(\omega_1 t) \\ \cdots \\ \sin(\omega_N t) \\ \cos(\omega_1 t) \\ \cdots \\ \cos(\omega_N t) \end{bmatrix} \in \mathbb{R}^{2N \times 1}.$$

We make now the key observation that since $N < n$ the matrix X is flat hence there exists a nonzero vector $C \in \mathbb{R}^{2n}$ such that

$$XC = 0. \tag{17}$$

Now, because of linearity of the operator $\mathcal{H}$, the extended regressor matrix $\Phi(t)$ can be written as

$$\begin{aligned}\Phi(t) = \mathcal{H}[\phi^\top(t)] &= \mathcal{H}[\phi_{\mathtt{ss}}^\top(t) + \phi_{\mathtt{tr}}^\top(t)] \\ &= \mathcal{H}[\phi_{\mathtt{ss}}^\top(t)] + \mathcal{H}[\phi_{\mathtt{tr}}^\top(t)],\end{aligned}$$

where we defined the vector

$$\phi_{\mathtt{tr}}(t) := \mathrm{col}(\phi_1^{\mathtt{tr}}(t), \ldots, \phi_{2n}^{\mathtt{tr}}(t)).$$

From stability of the operator $\mathcal{H}$ we have that $\mathcal{H}[\phi_{\mathtt{tr}}^\top(t)]$ converges to zero exponentially. Therefore, invoking Lemma 1 in Appendix A, we can concentrate our attention on the steady-state term $\mathcal{H}[\phi_{\mathtt{ss}}^\top(t)]$, which can be written as

$$\mathcal{H}[\phi_{\mathtt{ss}}^\top(t)] = \begin{bmatrix} \mathcal{H}_1[\phi_{\mathtt{ss}}^\top(t)] \\ \vdots \\ \mathcal{H}_{2n}[\phi_{\mathtt{ss}}^\top(t)] \end{bmatrix} = \begin{bmatrix} \mathcal{H}_1[\xi^\top(t)] \\ \vdots \\ \mathcal{H}_{2n}[\xi^\top(t)] \end{bmatrix} X,$$

where we invoked (15) to get the last equation. From (17) we then conclude that

$$\mathcal{H}[\phi_{\mathtt{ss}}^\top(t)]C = 0,$$

which implies that $\det\{\Phi(t)\}$ converges to zero exponentially and, consequently, $\Delta(t) \in \mathcal{L}_2$. ■

Remark 5: Instrumental to establish the proof of Proposition 2 is the assumption that the input signal consists of a sum of sinusoids of different frequencies, *i.e.*, given as (7). As indicated in Remark 1 the fundamental result of Proposition 1 is applicable to much wider class of input signals. Current investigation is under way to see whether the claim of Proposition 2 is still applicable in that case.

IV. PE (Generically) Implies DREM is Exponentially Stable

In this section we address the question (Q2) of Section I and present a condition, under which, the equivalence

$$\phi(t) \in \text{PE} \;\Leftrightarrow\; \Delta(t) \in \text{PE} \tag{18}$$

holds true for a given choice of operators $\mathcal{H}$—consisting of LTI filters and delay operators. In other words, under suitable excitation conditions, the asymptotic behaviour of DREM will (generically) be as good as the one of standard gradient estimators, with the additional advantage of an improved transient performance due to the monotonicity property (13). Moreover, we identify a class of operators $\mathcal{H}$ such that (18) holds.

Proposition 3: Consider the vector $\phi(t)$ generated via (3) *with $u(t)$ given by* (7) *with $N = n$ and the function $\Delta(t)$ defined in* (14). *Let the elements of $\mathcal{H}$ be $\mathcal{L}_\infty$-stable LTI operators, either rational minimum-phase transfer functions or constant time delays, such that*

$$\mathcal{H}_i(\mathrm{i}\omega_k) = M_{i,k}\exp(\mathrm{i}\alpha_{i,k}), \tag{19}$$

where i *is the imaginary unit. Define the $2N \times 2n$ matrix*

$$H := \begin{bmatrix} M_{1,1}\cos(\alpha_{1,1}) & \cdots & M_{2n,1}\cos(\alpha_{2n,1}) \\ \vdots & \ddots & \vdots \\ M_{1,N}\cos(\alpha_{1,N}) & \cdots & M_{2n,N}\cos(\alpha_{2n,N}) \\ M_{1,1}\sin(\alpha_{1,1}) & \cdots & M_{2n,1}\sin(\alpha_{2n,1}) \\ \vdots & \ddots & \vdots \\ M_{1,N}\sin(\alpha_{1,N}) & \cdots & M_{2n,N}\sin(\alpha_{2n,N}) \end{bmatrix}. \tag{20}$$

The following equivalence is true

$$\textit{rank}\,\{H\} = 2n \;\Leftrightarrow\; \Delta(t) \in \mathrm{PE}.$$

From Claim (C1) of Proposition 3 we immediately obtain the following corollary.

Corollary 1: Under the conditions of Proposition 3 if the elements of $\mathcal{H}$ are delay operators of the form

$$\mathcal{H}_i[x(t)] = x(t - d_i),\; d_i = d_c + (i-1)\,d_0, \qquad (21)$$

where $d_c \geq 0$ and $d_0 > 0$ is such that $\max_k \omega_k\, d_0 < \pi$, then $\Delta(t) \in \mathrm{PE}$ and DREM is exponentially convergent.

The proofs of Proposition 3 and Corollary 1 are based on complex-domain representation of the signals and linear operators and are omitted due to space restrictions.

Remark 6: Corollary 1 shows that the simple choice (21) will always ensure that the PE property of the regressor will be preserved for $\Delta(t)$. Clearly, to design this operators it is sufficient to know and upper bound on the bandwidth of the systems input signal, which is a reasonable assumption in most applications. However, increasing the size of the operators delays will adversely affect the transient performance of the DREM estimator.

V. Numerical example

Consider the following system

$$W(p) = \frac{b_1 p + b_0}{p^2 + a_1 p + a_0}$$

with $a_0 = 2$, $a_1 = 1$, $b_0 = 2$, $b_1 = 1$. The regression model (2) is constructed following (2) with $\lambda_1 = 10$, $\lambda_0 = 20$. The input signal is considered as $u(t) = 3\sin(2t) + 10\sin(5t)$. Initial conditions on the estimated parameters $\hat{\theta}(0)$ and the DREM filters were taken as zeros and initial conditions for the parameters estimation errors are

$$\tilde{\theta}(0) = -\begin{bmatrix} \lambda_0 - a_0 & \lambda_1 - a_1 & b_0 & b_1 \end{bmatrix}^\top.$$

We consider two cases of the gains: $\Gamma_1 = 10^2 \times I_4$ and $\Gamma_2 = 10^3 \times I_4$ for the gradient estimator (4), and $\Gamma_1 = \mathrm{diag}\{\gamma_{11}, \gamma_{12}\} = 10^3 \times I_4$, $\Gamma_2 = \mathrm{diag}\{\gamma_{21}, \gamma_{22}\} = 10^4 \times I_4$ for the DREM estimator (11).

According to Corollary 1, we choose the filters $\mathcal{H}$ as (21) with $d_c = 0$ and $d_0 \in (0; \frac{\pi}{5})$. The value of d_0 is chosen sovling the following optimization problem:

$$\text{find:} \max_{d_0 \in (0; \frac{\pi}{5})} \det(H(d_0)),$$

where

$$\begin{bmatrix} 1 & \cos(\omega_1 d_0) & \cos(2\omega_1 d_0) & \cos(3\omega_1 d_0) \\ 1 & \cos(\omega_2 d_0) & \cos(2\omega_2 d_0) & \cos(3\omega_2 d_0) \\ 0 & -\sin(\omega_1 d_0) & -\sin(2\omega_1 d_0) & -\sin(3\omega_1 d_0) \\ 0 & -\sin(\omega_2 d_0) & -\sin(2\omega_2 d_0) & -\sin(3\omega_2 d_0) \end{bmatrix} =: H(d_0).$$

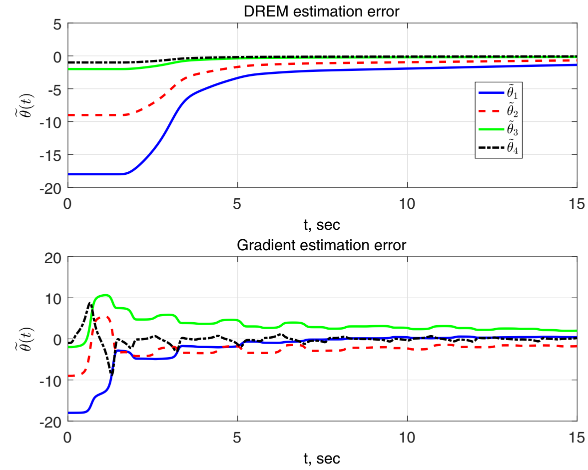

Fig. 1: Parameters estimation errors with the gain Γ_1.

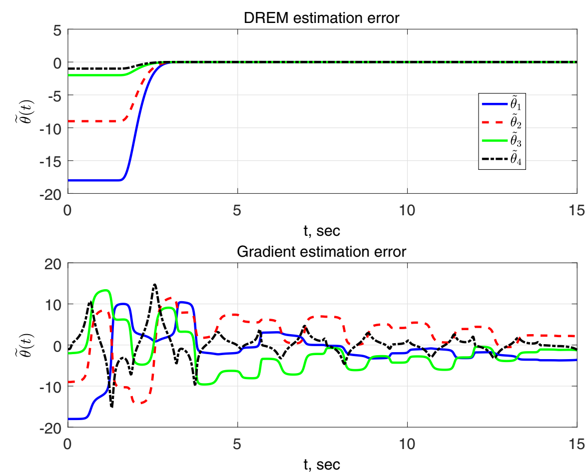

Fig. 2: Parameters estimation errors with the gain Γ_2.

It can be shown that increase of $\det(H(d_0))$ yields faster convergence of the DREM estimator. The optimal delay value is $d_0^* = 0.4647$ sec.

The simulation results with gain Γ_1 are given in Fig. 1 and show that both methods yield consistent estimations. Notice, however, that the gradient method is slower and has oscillating transients while DREM ensures the monotonicity.

To evaluate the effect on the transient performance of the adaptation gains we provide simulation results with gain Γ_2 shown in Fig. 2. As expected the transient performance of the gradient estimator significantly degrades and, although the parameters converge, the transient need more time to decay. On the other hand, increasing Γ for the DREM-based estimator leads to faster parameter convergence.

VI. Conclusions and Future Work

We have addressed in this paper the critical question of selection of the operators $\mathcal{H}$ introduced in DREM estimators to generate the extended regressor matrix $\Phi(t)$. As it has been widely documented in the publications [2], [3], [5], [6], [7], [13], [15], a suitable choice of these operators is essential to guarantee a good transient performance of the DREM estimator. It has been shown that, for the particular task of identification of LTI systems, the PE condition for exponential convergence of the parameter errors of gradient (or least-squares) estimators cannot be relaxed by DREM. On the other hand, we have proven that this convergence property is preserved in DREM for almost all choices of the operators, and some simple selection rules for them have been reported.

Within the context of identification we are currently exploring the use of DREM for some practical problems where under-excitation is prevalent. For instance, for identification of the Thevenin equivalent of the power network for synchronisation [9] or adaptive active damping in power converters [14], or for the estimation of a power system inertia and active power imbalance [18]. In these kind of applications it is not expected to achieve consistent estimation, being sufficient to ensure fast convergence to a neighbourhood of the true parameters, a feature that due to its monotonicity property can be guaranteed by DREM.

A far reaching objective is the use of DREM in classical model reference adaptive control problems. Unfortunately, some preliminary results reveal that the fundamental self-tuning property—required in these applications to ensure global tracking of the reference model output without PE—is lost with the use of DREM. On the other hand, some interesting robustness properties, conspicuous by their absence in gradient based schemes, have been established for DREM-based controllers. In particular, the instability mechanisms revealed by the widely known Rohrs' counterexamples, do not appear with DREM.

References

[1] B. D. O. Anderson, Exponential stability of linear equations arising in adaptive identification, *IEEE Trans. Autom. Control*, Vol. 22, No. 1, pp. 83-88, 1977.

[2] S. Aranovskiy, A. Bobtsov, R. Ortega and A. Pyrkin, Performance enhancement of parameter estimators via dynamic regressor extension and mixing, *IEEE Trans. Automatic Control*, Vol. 62, No. 2, pp. 3546-3550, 2017. (See also `arXiv preprint arXiv:1509.02763`.)

[3] S. Aranovskiy, A. Bobtsov, R. Ortega and A. Pyrkin, Improved transients in multiple frequencies estimation via dynamic regressor extension and mixing, *12th IFAC International Workshop on Adaptation and Learning in Control and Signal Processing*, Eindhoven, The Netherlands, 29/06–01/07, 2016.

[4] 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*, 2017, (to appear).

[5] A. Bobtsov, A. Pyrkin, R. Ortega, S. Vukosavic, A. M. Stankovic and E. V. Panteley, A robust globally convergent position observer for the permanent magnet synchronous motor, *Automatica*, Vol. 61, pp. 47-54, 2015.

[6] A. Bobtsov, D. Bazylev, A. Pyrkin, S. Aranovskiy and R. Ortega, A robust nonlinear position observer for synchronous motors with relaxed excitation conditions, *Int. J. of Control*, Vol. 90, No. 4, pp. 813-824, 2017.

[7] J. Choi, K. Nam, A. Bobtsov, A. Pyrkin and R. Ortega, Robust adaptive sensorless control for permanent magnet synchronous motors, *IEEE Transactions on Power Electronics*, Vol. 32, No. 5, pp. 3989-3997, 2017.

[8] M. Gilson, What has Instrumental Variable method to offer for system identification?, *8th IFAC International Conference on Mathematical Modelling, MATHMOD 2015*, Vienna, Austria, February 2015. (See also `hal-01242758`).

[9] N. Hoffmann and F. Fuchs, Minimal invasive equivalent grid impedance estimation in inductive-resistive power networks using extended Kalman filter, *IEEE Trans. Power Electron.*, Vol. 29, No. 2, pp. 631-641, Feb. 2014.

[10] R. Horn and C. Johnson, *Matrix Analysis* (2nd ed.), Cambridge University Press, 2013.

[11] P. Ioannou and J. Sun, *Robust Adaptive Control*, Prentice–Hall, NJ, 1996.

[12] L. Ljung, *System Identification: Theory for the User*, New Jersey: Prentice Hall, 1987.

[13] R. Ortega, L. Praly, S. Aranovskiy, B. Yi and W. Zhang, On dynamic regressor extension and mixing parameter estimators: Two Luenberger observers interpretations, *Automatica*, 2017, (to be published).

[14] R. Peña-Alzola, M. Liserre, F. Blaabjerg, M. Ordoñez, and T. Kerekes, A self-commissioning notch filter for active damping in a three-phase LCL-filter-based grid-tie converter, *IEEE Transactions on Power Electronics*, Vol. 29, No. 12, pp. 6754-6761, 2014.

[15] A. Pyrkin, F. Mancilla, R. Ortega, A. Bobtsov and S. Aranovskiy, Identification of photovoltaic arrays' maximum power extraction point via dynamic regressor extension and mixing, *Int. J. on Adaptive Control and Signal Processing*, Vol. 31, No. 9, pp. 1337-1349, 2017.

[16] S. Sastry and M. Bodson, *Adaptive Control; Stability, Convergence and Robustness*, Prentice-Hall, New Jersey, 1989.

[17] R. Horn A and C. R. Johnson, *Matrix analysis*, Cambridge university press, 2012.

[18] P. Wall and V. Terzija, Simultaneous estimation of the time of disturbance and inertia in power systems, *IEEE Transactions on Power Delivery*, Vol. 29, No. 4, pp. 2018-2013, August 2014.

Appendix

Appendix A: A Preliminary Lemma

Lemma 1: Consider matrix functions $A, B : \mathbb{R}_+ \to \mathbb{R}^{q\times q}$ with $A(t)$ bounded, and each entry of $B(t)$ tending to zero exponentially fast. The following implication is true:

$$\det\{A(t)\} \equiv 0 \;\Rightarrow\; \lim_{t\to\infty}\{\det(A(t)+B(t))\} = 0, \textit{ (exp).}$$

The proof is omitted due to space restrictions.