

Parameter identification of linear time-invariant systems using dynamic regressor extension and mixing

Stanislav Aranovskiy, Alexey Belov, Romeo Ortega, Nikita Barabanov, Alexey Bobtsov

▶ **To cite this version:**

Stanislav Aranovskiy, Alexey Belov, Romeo Ortega, Nikita Barabanov, Alexey Bobtsov. Parameter identification of linear time-invariant systems using dynamic regressor extension and mixing. International Journal of Adaptive Control and Signal Processing, 2019, 33 (6), pp.1016-1030. 10.1002/acs.3006 . hal-02181842

HAL Id: hal-02181842

https://centralesupelec.hal.science/hal-02181842

Submitted on 17 Jun 2020

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.

Received <day> <Month>, <year>; | Revised <day> <Month>, <year>; | Accepted <day> <Month>, <year>

DOI: xxx/xxxx

ARTICLE TYPE

Parameter Identification of Linear Time-Invariant Systems Using Dynamic Regressor Extension and Mixing

Stanislav Aranovskiy*[1,2] | Alexey Belov[3,2] | Romeo Ortega[4] | Nikita Barabanov[5,2] | Alexey Bobtsov[2]

[1]Equipe Automatique, CentaleSupélec – IETR, Avenue de la Boulaie, 35576 Cesson-Sévigné, France

[2]Faculty of Control Systems and Robotics, ITMO University, 197101 Saint Petersburg, Russia

[3]Laboratory of Dynamics of Control Systems, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow 117997, Russia

[4]Laboratoire des Signaux et Systèmes, CNRS-CentraleSupélec, Plateau du Moulon, 91192 Gif-sur-Yvette, France

[5]Department of Mathematics, North Dakota State University, Fargo, ND 58105, USA

Correspondence
*Stanislav Aranovskiy. Email: stanislav.aranovskiy@centralesupelec.fr

Abstract

Dynamic regressor extension and mixing (DREM) 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, by generation of scalar 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 single-input single-output time-invariant systems with periodic excitation. The most important contribution is that the DREM (almost surely) converges under the same persistence of excitation (PE) conditions as the gradient estimator, while providing improved transient performance. In particular, we give some guidelines how to select the DREM operators to ensure convergence under the same PE conditions as standard identification schemes.

KEYWORDS:
DREM, transient performance, system identification, persistent excitation

1 | INTRODUCTION

A new procedure to design parameter estimators for linear and nonlinear regressions, called dynamic regressor extension and mixing (DREM), was recently proposed in [1]. The technique has been successfully applied in a variety of identification and adaptive control problems [2,3,4,5,6]. 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) [7,8]. Instead of PE a non-square integrability condition on the determinant of a designer-dependent extended (square) regressor matrix is imposed. Similarly to the dynamic regressor extension approach proposed in [9] and to instrumental variable methods [10] 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 [1] for a more general and detailed presentation of DREM and [11] 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 [8], 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 lines—a property called "sufficient richness" in [8]—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 [1] 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 remainder of the paper is organized as follows. The application of DREM for identification of an LTI system parameters is presented in Section 2. In Section 3 we give the answer to (Q1) while the answer to (Q2) is presented in Section 4. Some simulation results that illustrate our results and show the performance improvement of DREM, with respect to gradient estimators, are given in Section 5 The paper is wrapped–up with some conclusions and future work in Section 6. The proofs of some technical lemmas are presented in the appendix.

The preliminary results of this study were reported in [12]. The contribution of this paper includes the main claim (Proposition 3) that is stronger than the one reported in [12], as well as the proofs of technical lemmas and corollaries that were not presented before.

2 | 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 [7,13,8].

2.1 | 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 $t \geq 0$ is the time, $y(t)$, $u(t)$ are the plant output and input, respectively, where $A(p) = \sum_{i=0}^{n} a_i p^i$, $B(p) = \sum_{i=0}^{n-1} b_i p^i$, $p := \frac{d}{dt}$, $a_n = 1$, $A(p)$ 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 [8] 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

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

$\lambda(p) = \sum_{i=0}^{n} \lambda_i p^i$, $\lambda_n = 1$, is an arbitrary Hurwitz polynomial and ϵ_t is a generic 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

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

Also, it is well–known [14,8] 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

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

for some $T, \delta > 0$ and for all $t \geq 0$, 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 (see Theorems 2.7.2 and 2.7.3 of [8]).

Proposition 1. 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,

$$\phi(t) \in \text{PE} \ \Leftrightarrow\ N \geq n.$$

Remark 1. For ease of presentation we consider only a particular case of the more general result reported in [8]. In 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 [15] conditions on ϕ 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.

2.2 | 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

$$Y(t) := \mathcal{H}[y(t)]$$
$$\Phi(t) := \mathcal{H}[\phi^\top(t)].$$

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

$$Y(t) = \Phi(t)\theta + \epsilon_t. \tag{8}$$

[1] See [8] and Remark 3 in [1] where the effect of these term is rigorously analysed.

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

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

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 $d_i > 0$. See Section 4 of[11] for the case of general linear time-varying operators.

Pre-multiplying (8) by the adjunct matrix of $\Phi(t)$, denoted $\mathrm{adj}\{\Phi(t)\}$, we get $2n$ scalar regressors of the form

$$\mathcal{Y}_i(t) = \Delta(t)\theta_i, \tag{9}$$

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 (9) can be easily carried out via

$$\dot{\hat{\theta}}_i(t) = \gamma_i\Delta(t)(\mathcal{Y}_i(t) - \Delta(t)\hat{\theta}_i(t)), \tag{10}$$

with adaptation gains $\gamma_i > 0$. From (9) it is clear that the latter equations are equivalent to

$$\dot{\tilde{\theta}}_i(t) = -\gamma_i\Delta^2(t)\tilde{\theta}_i(t). \tag{11}$$

Solving the simple scalar differential equation (11) as

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

shows that

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

that is, parameter convergence is established without the restrictive PE assumption. Moreover, if $\Delta(t) \in$ PE, the convergence of DREM is exponential. An important advantage of DREM is that the individual parameter errors satisfy

$$|\tilde{\theta}_i(t)| \leq |\tilde{\theta}_i(0)|,\ \forall\, t \geq 0, \tag{12}$$

that is strictly stronger than the monotonicity property (6).

The relationship between the condition $\Delta(t) \notin \mathcal{L}_2$ and $\phi(t) \in$ PE is far from obvious for arbitrary regressor vectors $\phi(t)$—see[1] 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 (12), this feature is instrumental to eliminate the need to overparameterise nonlinear regressions to obtain a linear one—a practice that, as is well-known[7,13,8], entails a serious performance degradation. This, and other advantages of DREM, have been discussed in a series of publications including[1,2,3,4,5,11,6]

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.

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

3 | DREM CANNOT RELAX THE PE CONDITION

In this section we give the answer, alas negative, to the question (Q1) of Section 1.

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

$$\Delta(t) = \det\{\mathcal{H}[\phi^\top(t)]\} \tag{13}$$

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

$$\phi(t) = \phi_{\rm ss}(t) + \epsilon_t, \tag{14}$$

where the components of the steady-state term $\phi_{\rm ss}(t)$ are given by

$$\phi_{{\rm ss},i}(t) = \sum_{k=1}^{N} A_{i,k} \cos\left(\omega_k t + \psi_{i,k}\right), \tag{15}$$

with $A_{i,k}$ and $\psi_{i,k}$ constants and the transient component ϵ_t tends to zero exponentially fast. We can write the steady-state vector in a compact form as

$$\phi_{\rm ss}(t) = X^\top \xi(t) \tag{16}$$

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

$$X := \begin{bmatrix} A_{1,1}\cos(\psi_{1,1}) & \cdots & A_{2n,1}\cos(\psi_{2n,1}) \\ \vdots & \ddots & \vdots \\ A_{1,N}\cos(\psi_{1,N}) & \cdots & A_{2n,N}\cos(\psi_{2n,N}) \\ -A_{1,1}\sin(\psi_{1,1}) & \cdots & -A_{2n,1}\sin(\psi_{2n,1}) \\ \vdots & \ddots & \vdots \\ -A_{1,N}\sin(\psi_{1,N}) & \cdots & -A_{2n,N}\sin(\psi_{2n,N}) \end{bmatrix}, \tag{17}$$

and

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

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

$$XC = 0. \tag{18}$$

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

$$\Phi(t) = \mathcal{H}[\phi^\top(t)] = \mathcal{H}[\phi_{\rm ss}^\top(t) + \epsilon_t^\top(t)] = \mathcal{H}[\phi_{\rm ss}^\top(t)] + \mathcal{H}[\epsilon_t^\top(t)].$$

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

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

where we invoked (16) to get the last equation. From (18) we then conclude that

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

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

4 | PE (GENERICALLY) IMPLIES DREM IS EXPONENTIALLY STABLE

In this section we address the question (Q2) of Section 1 and present a condition, under which, the equivalence

$$\phi(t) \in \mathrm{PE} \;\Leftrightarrow\; \Delta(t) \in \mathrm{PE} \tag{19}$$

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 (12). Moreover, we identify a class of operators $\mathcal{H}$ such that (19) holds.

This section is organized as follows. First we present a general result providing necessary and sufficient conditions for $\phi(t) \in$ PE. Then, we give corollaries illustrating how this result is applied to the LTI identification problem.

4.1 | Necessary and sufficient conditions

Consider the vector $\phi(t)$ generated via (3) with $u(t)$ given by (7) with $N \geq n$ and the function $\Delta(t)$ defined in (13). Let the elements of $\mathcal{H}$ used in (13) be $\mathcal{L}_\infty$-stable LTI operators, either rational minimum-phase transfer functions or constant time delays, such that the frequency response of the filter $\mathcal{H}_j$ at the frequency ω_k equals to $L_{j,k}e^{\mathrm{i}\alpha_{j,k}}$, where $j \in \{1, \dots, 2n\}$ and $k \in \{1, \dots, N\}$.

Define the vectors

$$w_k := \begin{bmatrix} A_{1,k}e^{\mathrm{i}\psi_{1,k}} \\ A_{2,k}e^{\mathrm{i}\psi_{2,k}} \\ \vdots \\ A_{2n,k}e^{\mathrm{i}\psi_{2n,k}} \end{bmatrix} \in \mathbb{C}^{2n}, \quad z_k := \begin{bmatrix} L_{1,k}e^{\mathrm{i}\alpha_{1,k}} \\ L_{2,k}e^{\mathrm{i}\alpha_{2,k}} \\ \vdots \\ L_{2n,k}e^{\mathrm{i}\alpha_{2n,k}} \end{bmatrix} \in \mathbb{C}^{2n}, \tag{20}$$

where $A_{i,k}$ and $\psi_{i,k}$ are the parameters of the signal $\phi_{\mathrm{ss}}(t)$ as defined in (15). Denote

$$\begin{aligned} Z &:= \begin{bmatrix} z_1 & z^*{}_1 & z_2 & z^*{}_2 & \cdots & z_N & z^*{}_N \end{bmatrix}, \\ W &:= \begin{bmatrix} w_1 & w^*{}_1 & w_2 & w^*{}_2 & \cdots & w_N & w^*{}_N \end{bmatrix}, \end{aligned} \tag{21}$$

where $(\cdot)^*$ is the complex conjugate of $(\cdot)$. Let $Z_{j_1,j_2,\dots,j_{2n}}$ and $W_{j_1,j_2,\dots,j_{2n}}$ be the square matrices consisting of columns j_1, j_2, …, j_{2n} of matrices Z and W, respectively, and define

$$C_{j_1,j_2,\dots,j_{2n}} := \det\left(Z_{j_1,j_2,\dots,j_{2n}}\right) \det\left(W_{j_1,j_2,\dots,j_{2n}}\right). \tag{22}$$

Define $2N$ values x_j as

$$x_{2k-1} := \omega_k, \quad x_{2k} := -\omega_k, \tag{23}$$

where $j = 1, \dots, 2N$, $k = 1, \dots, N$, and ω_k are the frequencies of sinusoidal components in (7), and denote by $\Omega_1, \Omega_2, \dots, \Omega_M$ all M distinct values of sums $x_{j_1} + \dots + x_{j_{2n}}$ with indices $1 \leq j_1 < j_2 < \dots < j_{2n} \leq 2N$.

Define the value B_k as a sum of the values $C_{j_1,j_2,\dots,j_{2n}}$ computed for all sets of indices $1 \leq j_1 < j_2 < \dots < j_{2n} \leq 2N$ such that the sum $x_{j_1} + \dots + x_{j_{2n}}$ equals Ω_k:

$$B_k := \sum_{1 \leq j_1 < j_2 < \dots < j_{2n} \leq 2N \,:\, x_{j_1} + \dots + x_{j_{2n}} = \Omega_k} C_{j_1,j_2,\dots,j_{2n}}. \tag{24}$$

With these definitions, the main result is formulated as follows.

Proposition 3. The signal $\det \Phi(t)$ is not persistently exciting if and only if $B_k = 0$ for all $k = 1, \dots, M$. In other words, the following equivalence holds

$$\exists k \in \{1, 2, \dots, M\} \,:\, B_k \neq 0 \Leftrightarrow \det \Phi(t) \in \mathrm{PE}.$$

Proof. Note that the matrix $\Phi(t)$ can be written as

$$\Phi(t) = \Phi_{\mathrm{ss}}(t) + \Phi_\varepsilon(t),$$

where $\Phi_{\mathrm{ss}}(t)$ denotes the steady-state response and all elements of $\Phi_\varepsilon(t)$ decay exponentially. Then (see Lemma 1)

$$\det \Phi(t) = \det \Phi_{\mathrm{ss}}(t) + \epsilon_t,$$

where ϵ_t decays exponentially. It is known, see[8], that a sum of a PE signal and an exponentially decaying term is PE; hence the following equivalence holds:

$$\det \Phi(t) \in \mathrm{PE} \Leftrightarrow \det \Phi_{\mathrm{ss}}(t) \in \mathrm{PE}.$$

The element of $\Phi_{\mathrm{ss}}(t)$ with the indices j, i is the steady-state response of $\mathcal{H}_j[\phi_i(t)]$ and is given by

$$\Phi_{\mathrm{ss}_{j,i}}(t) = \sum_{k=1}^{N} L_{j,k} A_{i,k} \cos(\omega_k t + \psi_{i,k} + \alpha_{j,k}). \tag{25}$$

Recalling Leibnitz formula, it follows that $\det \Phi_{\mathrm{ss}}(t)$ is a sum of products of sinusoidal functions. Thus, it is an *almost-periodic* function, see[16], and it is bounded, continuous, and differentiable; the first derivative of $\det \Phi_{\mathrm{ss}}(t)$ is also bounded. As it is shown in Lemma 2, if an almost periodic function is not identically zero, then it is persistently exciting, i.e. the following implication holds:

$$\exists t_0 \,:\, \det \Phi_{\mathrm{ss}}(t_0) \neq 0 \Leftrightarrow \det \Phi_{\mathrm{ss}}(t) \in \mathrm{PE}.$$

Note that the signals (15) can be also rewritten as

$$\phi_{\mathrm{ss},i}(t) = \sum_{k=1}^{N} \mathrm{Re}\left(A_{i,k} e^{\mathrm{i}\psi_{i,k}} e^{\mathrm{i}\omega_k t}\right),$$

and (25) can be written as

$$\Phi_{\mathrm{ss}_{j,i}}(t) = \sum_{k=1}^{N} \mathrm{Re}\left(L_{j,k} e^{\mathrm{i}\alpha_{j,k}} A_{i,k} e^{\mathrm{i}\psi_{i,k}} e^{\mathrm{i}\omega_k t}\right).$$

Recalling definition (20), it holds

$$\Phi_{\mathrm{ss}}(t) = \sum_{k=1}^{N} \mathrm{Re}\left(z_k w_k^{\top} e^{\mathrm{i}\omega_k t}\right)$$

or

$$\begin{aligned} 2\Phi_{\mathrm{ss}}(t) &= \sum_{k=1}^{N} \begin{bmatrix} z_k & z^*{}_k \end{bmatrix} \begin{bmatrix} e^{\mathrm{i}\omega_k t} & 0 \\ 0 & e^{-\mathrm{i}\omega_k t} \end{bmatrix} \begin{bmatrix} w_k^{\top} \\ w^{*\top}_k \end{bmatrix} \\ &= \begin{bmatrix} z_1 & z^*{}_1 & z_2 & z^*{}_2 & \cdots & z_N & z^*{}_N \end{bmatrix} \times \\ &\begin{bmatrix} \begin{bmatrix} e^{\mathrm{i}\omega_1 t} & 0 \\ 0 & e^{-\mathrm{i}\omega_1 t} \end{bmatrix} & 0 & \cdots & 0 \\ 0 & \begin{bmatrix} e^{\mathrm{i}\omega_2 t} & 0 \\ 0 & e^{-\mathrm{i}\omega_2 t} \end{bmatrix} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \begin{bmatrix} e^{\mathrm{i}\omega_N t} & 0 \\ 0 & e^{-\mathrm{i}\omega_N t} \end{bmatrix} \end{bmatrix} \begin{bmatrix} w_1^{\top} \\ w^{*\top}_1 \\ w_2^{\top} \\ w^{*\top}_2 \\ \vdots \\ w_N^{\top} \\ w^{*\top}_N \end{bmatrix}. \end{aligned}$$

Then

$$2\Phi_{\mathrm{ss}}(t) = Z \,\mathrm{diag}\left(\left\{e^{\mathrm{i}x_j t}\right\}_{j=1}^{2N}\right) W^{\top},$$

where Z and W are defined in (21).

From Lemma 3 it follows that $\det \Phi_{\mathrm{ss}}(t)$ is equal to a linear combination with constant coefficients of functions $e^{\mathrm{i}(x_{j_1}+\ldots+x_{j_{2n}})t}$:

$$2^{2n} \det \Phi_{\mathrm{ss}}(t) = \sum_{1 \leq j_1 < j_2 < \ldots < j_{2n} \leq 2N} e^{\mathrm{i}(x_{j_1}+\ldots+x_{j_{2n}})t} C_{j_1,j_2,\ldots,j_{2n}},$$

or

$$2^{2n} \det \Phi_{\mathrm{ss}}(t) = \sum_{k=1}^{M} e^{\mathrm{i}\Omega_k t} B_k, \tag{26}$$

where $C_{j_1,j_2,\ldots,j_{2n}}$ and B_k are defined in (22) and (24), respectively. Thus, function $\det \Phi_{\mathrm{ss}}(t)$ is equal to zero for all t if and only if $B_k = 0$ for all $k = 1, \ldots, M$.

The proof is summarized as follows:

$$\exists k \in \{1, 2, \ldots, M\} \,:\, B_k \neq 0 \Leftrightarrow \det \Phi_{\mathrm{ss}}(t) \not\equiv 0 \Leftrightarrow \det \Phi_{\mathrm{ss}}(t) \in \mathrm{PE} \Leftrightarrow \det \Phi(t) \in \mathrm{PE}.$$

□

Remark 5. It can be counterintuitive that the left-hand side of (26) is always real while the right-hand side is a sum of complex values. Note that due to the choice of x_j given by (23) it follows that the set $\{\Omega_j\}_{j=1}^{M}$ of all distinct values of sums $x_{j_1} + \ldots + x_{j_{2n}}$ with indices $1 \le j_1 < j_2 < \ldots < j_{2n} \le 2N$ has odd number of elements, contains zero as an element, and for each positive element of the set there exists a corresponding negative element with the same absolute value. Without loss of generality, the elements of this set can be rearranged as

$$\{\Omega_{-k}, \ldots, \Omega_{-1}, \Omega_0, \Omega_1, \ldots, \Omega_k\}, \tag{27}$$

where $k = 1, 2, \ldots, \frac{M-1}{2}$, elements with positive indices are positive, $\Omega_k > 0$, elements with negative indices are defined as $\Omega_{-k} = -\Omega_k$, and $\Omega_0 = 0$. Then equation (26) can be rewritten as

$$2^{2n} \det \Phi_{\mathrm{ss}}(t) = B_0 + \sum_{k=1}^{\frac{M-1}{2}} \left(e^{\mathrm{i}\Omega_k t} B_k + e^{-\mathrm{i}\Omega_k t} B_{-k} \right). \tag{28}$$

Recalling that for any complex square matrix Z it holds

$$\det(Z^*) = (\det(Z))^*,$$

it follows from (24) and (22) that $B_{-k} = B_k^*$ and

$$e^{\mathrm{i}\Omega_k t} B_k + e^{-\mathrm{i}\Omega_k t} B_{-k} = 2 \operatorname{Re}\left(e^{\mathrm{i}\Omega_k t} B_k \right).$$

Hence the right-hand side of (26) is a real-valued function of t.

4.2 | Corollaries for LTI identification problem

From Proposition 3 we obtain the following corollary.

Corollary 1. Under the conditions of Proposition 3 consider the case when $N = n$ and the matrices Z and W defined in (21) are square. Recall that the elements of $\mathcal{H}$ used in (13) are $\mathcal{L}_\infty$-stable LTI operators, such that the frequency response of the filter $\mathcal{H}_j$ at the frequency ω_k equals to $L_{j,k} e^{\mathrm{i}\alpha_{j,k}}$, where $j \in \{1, \ldots, 2n\}$ and $k \in \{1, \ldots, N\}$. Define the $2N \times 2n$ square matrix

$$H := \begin{bmatrix} L_{1,1}\cos(\alpha_{1,1}) & L_{2,1}\cos(\alpha_{2,1}) & \cdots & L_{2n,1}\cos(\alpha_{2n,1}) \\ \vdots & \vdots & \ddots & \vdots \\ L_{1,N}\cos(\alpha_{1,N}) & L_{2,N}\cos(\alpha_{2,N}) & \cdots & L_{2n,N}\cos(\alpha_{2n,N}) \\ -L_{1,1}\sin(\alpha_{1,1}) & -L_{2,1}\sin(\alpha_{2,1}) & \cdots & -L_{2n,1}\sin(\alpha_{2n,1}) \\ \vdots & \vdots & \ddots & \vdots \\ -L_{1,N}\sin(\alpha_{1,N}) & -L_{2,N}\sin(\alpha_{2,N}) & \cdots & -L_{2n,N}\sin(\alpha_{2n,N}) \end{bmatrix}. \tag{29}$$

Then $\det \Phi(t) \in \mathrm{PE}$ if and only if the matrix H is of full rank.

Proof. First we note that for the case $N = n$ the set (27) contains only the element $\Omega_0 = 0$ and (28) can be rewritten as

$$2^{2n} \det \Phi_{\mathrm{ss}}(t) = B_0 = \det(Z) \det(W).$$

From (21) it follows that $\det(Z)$ and $\det(W)$ are real numbers. With the definitions (20) the matrix H defined as (29) and the matrix X defined as (17) can be rewritten as

$$X^\top = \left[\operatorname{Re}(w_1) \ \ldots \ \operatorname{Re}(w_N) \ -\operatorname{Im}(w_1) \ \ldots \ -\operatorname{Im}(w_N) \right],$$
$$H^\top = \left[\operatorname{Re}(z_1) \ \ldots \ \operatorname{Re}(z_N) \ -\operatorname{Im}(z_1) \ \ldots \ -\operatorname{Im}(z_N) \right],$$

and the following implications hold

$$\det(X) \neq 0 \Leftrightarrow \det(W) \neq 0, \quad \det(H) \neq 0 \Leftrightarrow \det(Z) \neq 0.$$

Conditions of Proposition 3 are satisfied if both $\det(Z)$ and $\det(W)$ are not zero. Then, to complete the proof we have to show that the matrix X is of the full rank under the conditions of Proposition 3.

From Proposition 1 it follows that $\phi(t) \in$ PE for $N = n$. Then from (14) it follows that $\phi_{ss} \in$ PE as well and there exist $T_\phi > 0$ and $\delta_\phi > 0$ such that for all $t > 0$ and for all $z \in \mathbb{R}^{2n}$, $|z| = 1$,

$$z^\top \left(\int_t^{t+T_\phi} \phi_{ss}(s)\phi_{ss}^\top(s) ds \right) z \geq \delta_\phi.$$

Recalling (16), the latter implies

$$z^\top X^\top \left(\int_t^{t+T_\phi} \xi(s)\xi^\top(s) ds \right) X z \geq \delta_\phi$$

for all z. Hence X is of full rank and $\det W \neq 0$. □

Based on Proposition 3 and Corollary 1 a systematic procedure to construct the filters $\mathcal{H}$ can be proposed as described in the following corollary.

Corollary 2. Consider the case $N = n$. Under the conditions of Proposition 3 if the elements of $\mathcal{H}$ are delay operators of the form

$$\mathcal{H}_i[x(t)] = x(t - d_i), \;\; d_i = d_c + (i-1)\, d_0, \tag{30}$$

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

Proof. As it was shown in the proof of Corollary 1, for the case $N = n$ to prove that $\Delta(t) \in$ PE it is sufficient to show that the square matrix Z defined in (21) is of full rank. Let us rearrange the elements of Z and consider

$$Z_e := \begin{bmatrix} z_1 & \cdots & z_N & z^*{}_1 & \cdots & z^*{}_N \end{bmatrix}^\top,$$

where $\det(Z) = \det(Z_e)$. Clearly for (30)

$$\mathcal{H}_j(\mathrm{i}\omega_k) = \exp\left(-\mathrm{i}\omega_k\, d_c\right) \exp\left(-\mathrm{i}\omega_k\, (j-1) d_0\right)$$

and

$$\begin{aligned} Z_e = \operatorname{diag}\left\{ e^{-\mathrm{i}\omega_1 d_c}, \ldots, e^{-\mathrm{i}\omega_N d_c}, e^{\mathrm{i}\omega_1 d_c}, \ldots, e^{\mathrm{i}\omega_N d_c} \right\} \\ \times \begin{bmatrix} 1 & e^{-\mathrm{i}\omega_1 d_0} & \left(e^{-\mathrm{i}\omega_1 d_0}\right)^2 & \cdots & \left(e^{-\mathrm{i}\omega_1 d_0}\right)^{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-\mathrm{i}\omega_N d_0} & \left(e^{-\mathrm{i}\omega_N d_0}\right)^2 & \cdots & \left(e^{-\mathrm{i}\omega_N d_0}\right)^{2n} \\ 1 & e^{\mathrm{i}\omega_1 d_0} & \left(e^{\mathrm{i}\omega_1 d_0}\right)^2 & \cdots & \left(e^{-\mathrm{i}\omega_1 d_0}\right)^{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{\mathrm{i}\omega_N d_0} & \left(e^{\mathrm{i}\omega_N d_0}\right)^2 & \cdots & \left(e^{\mathrm{i}\omega_N d_0}\right)^{2n} \end{bmatrix}. \end{aligned} \tag{31}$$

The first term in (31) is a nonsingular diagonal matrix. To obtain the full-rank conditions, we note that the second term in (31) is the Vandermond matrix with complex entries. Recalling the properties of the determinant of a Vandermond matrix [17], we conclude that this matrix is singular if and only if there exist two elements of the second column such that their difference equals zero. Hence the matrix Z_e is of full rank if and only if the following conditions hold:

$$\begin{aligned} e^{\mathrm{i}\omega_i d_0} - e^{-\mathrm{i}\omega_i d_0} &\neq 0 \quad \text{for all } i, \\ e^{\mathrm{i}\omega_i d_0} - e^{\mathrm{i}\omega_j d_0} &\neq 0 \quad \text{for all } i \neq j, \\ e^{\mathrm{i}\omega_i d_0} - e^{-\mathrm{i}\omega_j d_0} &\neq 0 \quad \text{for all } i \neq j, \end{aligned}$$

where $i, j \in \{1, \ldots, N\}$. These conditions are satisfied since $\omega_i \neq \omega_j$ for $i \neq j$, and $\omega_i d_0 < \pi$ for all i. □

Remark 6. Corollary 2 shows that the simple choice (30) will always ensure that the PE property of the regressor will be preserved by $\Delta(t)$. 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.

TABLE 1 Standard deviation σ of the parameter estimation errors

	$\tilde{\theta}_1$	$\tilde{\theta}_2$	$\tilde{\theta}_3$	$\tilde{\theta}_4$
Gradient estimator (4)	1.67	1.25	0.39	0.59
DREM estimator (10)	0.1	0.11	0.27	0.02

Remark 7. Instrumental to establish the proofs of Proposition 2 and Proposition 3 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 claims of Proposition 2 and Proposition 3 are still applicable in that case.

5 | SIMULATION RESULTS

Consider the system of the form (1), where

$$\frac{B(p)}{A(p)} = \frac{b_1 p + b_0}{p^2 + a_1 p + a_0} = \frac{2p+1}{p^2+p+2}.$$

The regression model (2) is constructed following (3) where the choice $\lambda_1 = 20$, $\lambda_0 = 100$ yields

$$\theta = \begin{bmatrix} 98 & 19 & 1 & 2 \end{bmatrix}^\top.$$

The estimators are initialized with zero initial conditions, thus $\tilde{\theta}(0) = -\theta$. The input signal is chosen as $u(t) = \sin(2\pi t) + \cos(3t)$. The DREM operators are chosen as time-delay elements following Corollary 2 and taking $d_c = 0$ and $d_0 = 0.2$.

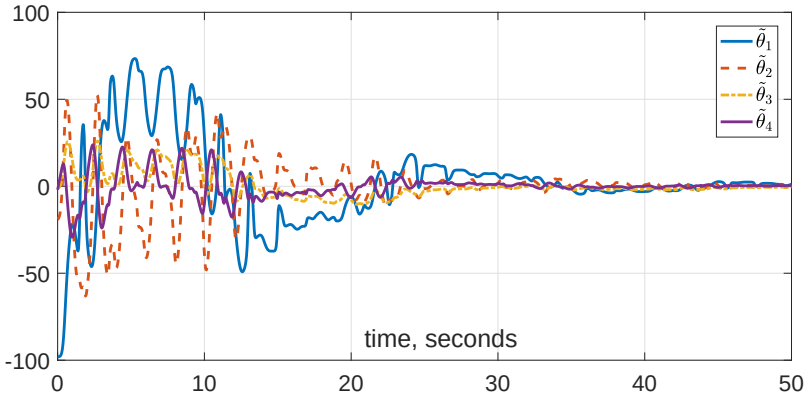
The gains tuning was performed as follows. First, for the standard gradient estimator (4) it was found that with the gain matrix $\Gamma = \mathrm{diag}(100, 50, 30, 10)$ parameter estimation errors converge to a neighborhood of zero in approximately 30 seconds, see Fig. 1 a. Next, the gains of the DREM estimator were tuned to have a comparable transient time. It was found that this goal is achieved choosing in (10) $\gamma_i = 1$, where $i = 1, 2, 3, 4$, see Fig. 1 b. The figures illustrate that the DREM estimator significantly outperforms the standard gradient solution since it provides monotonic transients without compromising the rate of convergence.

It can be argued that the choice of a diagonal gain matrix for the gradient estimator is restrictive since any positive-definite matrix Γ can be chosen. However, the tuning procedure for the gradient estimator is not straightforward and requires multiple trials. Thus, in practice the gain matrix is often restricted to be diagonal. In contrast, note that for the DREM estimator the tuning procedure is rather simple and transparent since each gain γ_i affects only the estimation error $\tilde{\theta}_i$ and does not affect other elements of the vector $\tilde{\theta}$.

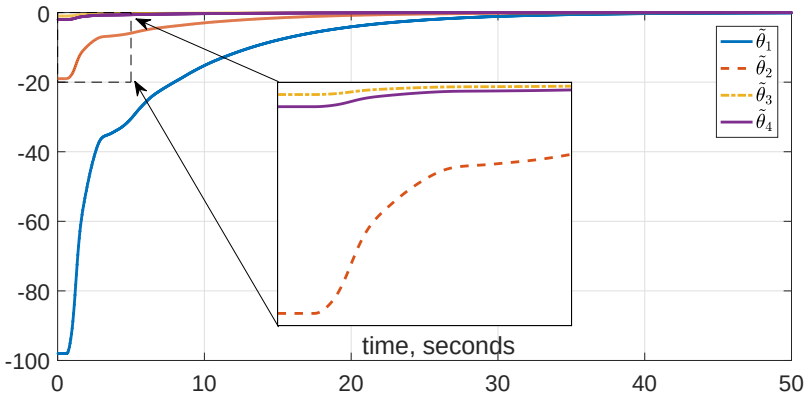
Next, the estimators that are tuned as described above are compared for the case when the measurements of the plant output signal $y(t)$ are distorted. The noise in measurements is modeled as a uniform random signal in the range from -0.1 to 0.1, while the signal $y(t)$ varies from -1.2 to 1.2 in the noise-free scenario. In this experiment we evaluate the standard deviation σ of the estimation errors $\tilde{\theta}$ after decay of the transients. Motivated by the results presented in Fig. 1 , the evaluation was performed on the interval from 50 to 100 seconds. The results of evaluation are given in Table 1 . It can be seen that the DREM estimator provides smaller deviation of the estimation errors; this can be partially explained noting that the DREM estimator requires smaller gains to achieve the same transient time than the gradient estimator. However, it should be highlighted that the DREM procedure involves a *nonlinear* signal transformation given by the calculation of the adjoint matrix. Thus, rigor analysis of noise propagation for the DREM scheme is not straightforward and represents a possible direction of future research.

6 | 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 [1,2,3,4,5,11,6], 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.



(a) The gradient estimator (4)



(b) The DREM estimator (10)

FIGURE 1 Parameter estimation errors of the gradient and DREM estimators.

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 [18] or adaptive active damping in power converters [19], or for the estimation of a power system inertia and active power imbalance [20]. In these kind of applications it is not expected to achieve consistent estimation, being sufficient to ensure fast convergence to a neighborhood 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 (MRAC) of LTI systems. Unfortunately, it has been shown 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, with DREM-based controllers it has been possible in [21] to remove the key assumptions of known sign of the high frequency gain in direct MRAC and the use of projections in indirect MRAC for multivariable systems. Also, in [22], the monotonicity property of the parameter estimation errors is used in [22] to propose a globally stable DREM-based MRAC for SISO systems with the only knowledge of a lower bound in the high-frequency gain.

FINANCIAL DISCLOSURE

This paper supported by the Ministry of Education and Science of the Russian Federation (goszadanie no. 8.8885.2017/8.9) and by the Government of the Russian Federation (Grant 08-08).

□

APPENDIX

APPENDIX A: TECHNICAL LEMMAS

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. Then

$$\lim_{t\to\infty} |\det\{A(t)+B(t)\} - \det\{A(t)\}| = 0.$$

Proof. First, we recall Leibnitz formula for the determinant of a square matrix $C(t) \in \mathbb{R}^{q\times q}$, which is given by

$$\det\{C(t)\} = \sum_{\delta\in S_q}\left(\operatorname{sign}(\delta)\prod_{i=1}^{q} c_{i,\delta_i}(t)\right),$$

where S_q is the set of permutations of the set $\{1, 2, \dots, q\}$, $\operatorname{sign}(\delta)$ is the signature of the permutation δ and $c_{i,\delta_i}(t)$ is the (i,δ_i)-th element of the matrix $C(t)$, see [17] for further details.

Define $C(t) := A(t) + B(t)$. Then,

$$\det\{C(t)\} = \sum_{\delta\in S_n}\left(\operatorname{sign}(\delta)\prod_{i=1}^{n}\left(a_{i,\delta_i}(t)+b_{i,\delta_i}(t)\right)\right).$$

But

$$\operatorname{sign}(\delta)\prod_{i=1}^{n}\left(a_{i,\delta_i}(t)+b_{i,\delta_i}(t)\right) = \left(\operatorname{sign}(\delta)\prod_{i=1}^{n} a_{i,\delta_i}(t)\right) + f_\delta(t),$$

where $f_\delta(t)$ is equal to a product of $\operatorname{sign}(\delta)$ and of a finite sum of products of elements $a_{i,\delta_i}(t)$ and $b_{i,\delta_i}(t)$ such that at least one of the multipliers in each product is an entry of the matrix $B(t)$. Since all elements of matrix function $A(t)$ are bounded, the function $f_\delta(t)$ tends to zero exponentially for all $\delta \in S_n$. Besides,

$$\sum_{\delta\in S_n}\left(\operatorname{sign}(\delta)\prod_{i=1}^{n} a_{i,\delta_i}(t)\right) = \det A(t).$$

Therefore, the function

$$\det\{C(t)\} = \sum_{\delta\in S_n}\left[\left(\operatorname{sign}(\delta)\prod_{i=1}^{n} a_{i,\delta_i}(t)\right) + f_\delta(t)\right] = \det A(t) + \sum_{\delta\in S_n} f_\delta(t)$$

and

$$\lim_{t\to\infty} |\det\{A(t)+B(t)\} - \det\{A(t)\}| = \lim_{t\to\infty}\left|\sum_{\delta\in S_n} f_\delta(t)\right| = 0.$$

□

Lemma 2. Let $\Delta : \mathbb{R}_+ \to \mathbb{R}$ be a differentiable almost-periodic function with a bounded first derivative. If $\Delta(t)$ is not identically zero, then $\Delta(t)$ is PE.

Proof. Define $g(t) := (\Delta(t))^2$ and note that $g(t)$ is an almost-periodic as well. Since $\Delta(t)$ is not identically zero, it follows that there exits t_0 such that $g(t_0) > 0$. Since the derivative of $g(t)$ is bounded there exist $a_0 > 0$, $b_0 > a_0$, such that $t_0 \in [a_0,\ b_0]$ and

$$\int_{a_0}^{b_0} g(s)ds \geq \delta_g$$

for some $\delta_g > 0$.

Function $g(t)$ enjoys the following property of almost-periodic functions [16]: for any $\epsilon > 0$ there exists $L_\epsilon > 0$ such that for any $t^\star$ there exists (an ϵ-translation) $\tau \in [t^\star,\ t^\star + L_\epsilon]$ such that for all t it holds

$$|g(t) - g(t+\tau)| < \epsilon. \tag{1}$$

In other words, for any interval of length L_ϵ we can find a value τ belonging to this interval, such that τ is an (alomost-)period[3] of $g(t)$ up to accuracy ϵ.

Define $\epsilon = \frac{\delta_g}{2(b_0-a_0)}$ and choose L_ϵ correspondingly. Consider the interval $[t^\star, t^\star + L_\epsilon]$ for some $t^\star$, and choose τ as an ϵ-translation of $g(t)$ belonging to this interval. Then

$$t^\star \le a_0 + \tau < b_0 + \tau \le t^\star + L_\epsilon + b_0,$$

and recalling (1)

$$\int_{t^\star}^{t^\star+L_\epsilon+b_0} g(s)ds \ge \int_{a_0+\tau}^{b_0+\tau} g(s)ds \ge \int_{a_0}^{b_0} g(s)ds - \epsilon(b_0-a_0) \ge \frac{\delta_g}{2}.$$

Since $t^\star$ can be chosen arbitrary, it follows that there exist $T_\Delta = L_\epsilon + b_0$ and $\delta_\Delta = \frac{\delta_g}{2}$ such that

$$\int_{t}^{t+T_\Delta} (\Delta(s))^2\, ds \ge \delta_\Delta$$

for all t, i.e. $\Delta(t)$ is PE. □

Lemma 3. Let $N \ge n$ and consider the matrix function

$$F_{\rm ss}(t) = Z\ \mathrm{diag}(\{\alpha_j\}_{j=1}^{2N})W^\top,$$

where Z and W are complex matrices of size $(2n)\times(2N)$, and α_j are complex numbers. Denote $Z_{j_1,\dots,j_{2n}}$, $W_{j_1,\dots,j_{2n}}$ square matrices composed of columns $j_1,\dots,j_{2n}$ of matrices Z and W respectively, and $C_{j_1,\dots,j_{2n}} = \det(Z_{j_1,\dots,j_{2n}})\ \det(W_{j_1,\dots,j_{2n}})$. Then

$$\det(F_{\rm ss}(t)) = \sum_{1\le j_1<j_2<\dots<j_{2n}\le 2N} C_{j_1,\dots,j_{2n}} \left(\prod_{i=1}^{2n}\alpha_{j_i}\right).$$

Proof. Every element of matrix $F_{\rm ss}(t)$ has a form $F_{k,m}(t) = \sum_{j=1}^{2N}\alpha_j\beta_{k,m,j}$, where $\beta_{k,m,j}$ are complex numbers. Therefore the determinant of $F_{\rm ss}(t)$ is a sum of products of $2n$ such elements, and it has the following form

$$\det(F_{\rm ss}(t)) = \sum_{j_1\in\{1,\dots,2N\},\dots,j_{2n}\in\{1,\dots,2N\}} b_{j_1,\dots,j_{2n}} \left(\prod_{i=1}^{2n}\alpha_{j_i}\right),$$

where $b_{j_1,\dots,j_{2n}}$ are complex numbers, and all sets $\{j_1,\dots,j_{2n}\}$ are distinct.

Now we show that summation over the sets $\{j_1,\dots,j_{2n}\}$ where there are equal numbers among numbers $j_1,\dots,j_{2n}$ is zero. Indeed, choose a set $\{j_1,\dots,j_{2n}\}$ and assume that at least two numbers are equal. Allow $\alpha_j = 0$ for all $j\notin\{j_1,\dots,j_{2n}\}$. Since there are equal numbers among numbers $j_1,\dots,j_{2n}$, there are at most $2n-1$ nonzero elements α_j and rank $\left(\mathrm{diag}\{\alpha_j\}\right) < 2n$, and therefore $\det(F_{\rm ss}(t)) = 0$. On the other hand it is equal to

$$\det(F_{\rm ss}(t)) = \sum_{k_1\in\{j_1,\dots,j_{2n}\},\dots,k_{2n}\in\{j_1,\dots,j_{2n}\}} b_{k_1,\dots,k_{2n}} \left(\prod_{i=1}^{2n}\alpha_{k_i}\right).$$

It follows that summation over any set $\{j_1,\dots,j_{2n}\}$ where there are at least two equal numbers is zero and without loss of generality we can assume that

$$\det(F_{\rm ss}(t)) = \sum_{1\le j_1<j_2<\dots<j_{2n}\le 2N} b_{j_1,\dots,j_{2n}} \left(\prod_{i=1}^{2n}\alpha_{j_i}\right). \tag{2}$$

Now fix numbers $j_1<\dots<j_{2n}$ and let the rest numbers α_j be zero. Then

$$F_{\rm ss}(t) = Z_{j_1,\dots,j_{2n}}\ \mathrm{diag}(\{\alpha_{j_k}\}_{k=1}^{2n})W^\top_{j_1,\dots,j_{2n}}.$$

All matrices in the right hand side are square. Therefore $\det(F_{\rm ss}(t)) = C_{j_1,\dots,j_{2n}}\left(\prod_{i=1}^{2n}\alpha_{j_i}\right)$. But (2) implies $\det(F_{\rm ss}(t)) = b_{j_1,\dots,j_{2n}}\left(\prod_{i=1}^{2n}\alpha_{j_i}\right)$. Hence, $b_{j_1,\dots,j_{2n}} = C_{j_1,\dots,j_{2n}}$, and the claim is proven.

□

[3] If $g(t)$ is a periodic function with period T, then we have $L_0 = T$, and for any interval of length T we can choose τ, e.g. $2T$, $3T$ and so on, such that $g(t) = g(t+\tau)$ for all t.

References

1. Aranovskiy S, Bobtsov A, Ortega R, Pyrkin A. Performance enhancement of parameter estimators via dynamic regressor extension and mixing. *IEEE Transactions on Automatic Control.* 2017;62(7):3546–3550.
2. Aranovskiy S, Bobtsov A, Ortega R, Pyrkin A. Improved transients in multiple frequencies estimation via dynamic regressor extension and mixing. *IFAC-PapersOnLine.* 2016;49(13):99–104.
3. Bobtsov A, Pyrkin A, Ortega R, Vukosavic SN, Stankovic AM, Panteley EV. A robust globally convergent position observer for the permanent magnet synchronous motor. *Automatica.* 2015;61:47–54.
4. Bobtsov A, Bazylev D, Pyrkin A, Aranovskiy S, Ortega R. A robust nonlinear position observer for synchronous motors with relaxed excitation conditions. *International Journal of Control.* 2017;90(4):813–824.
5. Choi J, Nam K, Bobtsov A, Pyrkin A, Ortega R. Robust adaptive sensorless control for permanent-magnet synchronous motors. *IEEE Transactions on Power Electronics.* 2017;32(5):3989–3997.
6. Pyrkin A, Mancilla-David F, Ortega R, Bobtsov A, Aranovskiy S. Identification of photovoltaic arrays' maximum power extraction point via dynamic regressor extension and mixing. *International Journal of Adaptive Control and Signal Processing.* 2017;31(9):1337–1349.
7. Ioannou PA, Sun J. *Robust adaptive control.* New Jersey: Prentice–Hall; 1996.
8. Sastry S, Bodson M. *Adaptive control: stability, convergence, and robustness.* New Jersey: Prentice-Hall; 1989.
9. Kreisselmeier G. Adaptive observers with exponential rate of convergence. *IEEE Transactions on Automatic Control.* 1977;22(1):2–8.
10. Gilson M. What has Instrumental Variable method to offer for system identification?. *IFAC-PapersOnLine.* 2015;48(1):354–359.
11. Ortega R, Praly L, Aranovskiy S, Yi B, Zhang W. On dynamic regressor extension and mixing parameter estimators: Two Luenberger observers interpretations. *Automatica.* 2018;95:548–551.
12. Belov A, Aranovskiy S, Ortega R, Barabanov N, Bobtsov A. Enhanced parameter convergence for linear systems identification : the DREM approach. In: :2794–2799; 2018; Limassol. 2018 European Control Conference.
13. Ljung L. *System identification: theory for user.* New Jersey: Prentice Hall; 1987.
14. Anderson BDO. Exponential stability of linear equations arising in adaptive identification. *IEEE Transactions on Automatic Control.* 1977;22(1):83–88.
15. Barabanov N, Ortega R. On global asymptotic stability of $\dot{x} = \phi(t)\phi^{\top}(t)x$ with $\phi(t)$ not persistently exciting. *Systems & Control Letters.* 2017;109:24–29.
16. Besicovitoh AS. *Almost periodic functions.* Cambridge: Cambridge University Press; 1932.
17. Horn RA, Johnson CR. *Matrix analysis.* Cambridge University Press; 2013.
18. Hoffmann N, Fuchs FW. Minimal invasive equivalent grid impedance estimation in inductive-resistive power networks using extended Kalman filter. *IEEE Transactions on Power Electronics.* 2014;29(2):631–641.
19. Pena-Alzola R, Liserre M, Blaabjerg F, Ordonez M, Kerekes T. A self-commissioning notch filter for active damping in a three-phase LCL -filter-based grid-tie converter. *IEEE Transactions on Power Electronics.* 2014;29(12):6754–6761.
20. Wall P, Terzija V. Simultaneous estimation of the time of disturbance and inertia in power systems. *IEEE Transactions on Power Delivery.* 2014;29(4):2018–2031.
21. Ortega R, Gerasimov D, Barabanov N, Nikiforov V. Relaxing the assumptions on adaptive control of linear multivariable systems using dynamic regressor extension and mixing estimators. *Automatica.* 2018;. submitted.
22. Gerasimov D, Ortega R, Nikiforov V. Relaxing the high-frequency gain sign assumption in direct model reference adaptive control. *European Journal of Control.* 2018;43:12–19.