

On preserving-excitation properties of a dynamic regressor extension scheme

Stanislav Aranovskiy, Rosane Ushirobira, Marina Korotina, Alexey Vedyakov

▶ **To cite this version:**

Stanislav Aranovskiy, Rosane Ushirobira, Marina Korotina, Alexey Vedyakov. On preserving-excitation properties of a dynamic regressor extension scheme. 2021. hal-03245139v1

HAL Id: hal-03245139

https://centralesupelec.hal.science/hal-03245139v1

Preprint submitted on 1 Jun 2021 (v1), last revised 5 May 2022 (v2)

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.

On preserving-excitation properties of a dynamic regressor extension scheme

Stanislav Aranovskiy[1,2], Rosane Ushirobira[3], Marina Korotina[2], Alexey Vedyakov[2]

***Abstract*—In this work, we consider the excitation preservation problem within the context of the Dynamic Regressor Extension and Mixing procedure. To ensure that the input excitation is not lost, we apply the Kreisselmeier's regressor extension scheme and prove that this choice always preserves both the persistent and the interval input excitations. We also provide a lower bound on the resulting excitation level and study the dynamics of the novel regressor. Illustrative simulations support our theoretical results.**

I. Introduction

The investigation of physical phenomena by mathematical modeling leads frequently to the problem of estimating model parameters. Indeed, the differential equations appearing in the model under consideration may contain parameters that are difficult to determine in advance. Through the years, numerous research works, in several disciplines, have been dedicated to this fundamental problem.

The linear regression equation (LRE) plays a central role in adaptive parameter estimation and adaptive control. It can be found in system identification [1], in model-reference adaptive control [2], [3] and adaptive pole-placement [4], in filtering and prediction [5], in reinforcement learning [6], and several others areas. The linear regression model is given by

$$y(t) = \phi^\top(t)\theta + w(t), \tag{1}$$

where $y(t) \in \mathbb{R}^\ell$ is the output signal, $\phi(t) \in \mathbb{R}^{n\times\ell}$ is the regressor, $w(t) \in \mathbb{R}^\ell$ is an additive distortion, *e.g.*, a measurement noise, and $\theta \in \mathbb{R}^n$ is the vector of unknown constant parameters. The signals y and ϕ are known (*e.g.*, they are measured), and the distortion signal w is unknown. The purpose is to estimate the vector of parameters θ using the measurements y and ϕ.

The convergence of various parameter estimation schemes relies on the essential premise of an adequate richness of the regressor ϕ, realized by the definition of the persistence of excitation (PE) condition, see Section II. In recent years, various efforts to ease the PE requirement have been suggested, such as concurrent [7], [8], or composite learning [9], [10]. Within this methodology, a dynamic data stack is built to discretely record online historical data, and the convergence of parameter estimation is managed following the interval excitation (IE) condition, a weaker requirement than the persistence of excitation. In [11], several fixed-time convergence algorithms have been proposed under the IE assumption on the regressor, and the PE relaxation has been also examined in [12] in the context of model reference control. To summarize, the persistence of excitation and interval (or sufficient) excitation are nowadays two crucial components in convergence analysis of parameter estimation algorithms.

Among many approaches to tackle parameter estimation problems, two traditional strategies can be used: the least-squares method and the gradient method. Many different versions of these procedures are known, such as the least-squares estimator with forgetting factor, the normalized gradient estimator, to mention a few [2], [3]. A disadvantage of these techniques is that even if there is a guarantee for the weak monotonicity of a (weighted) norm of the estimation errors, the estimation transients for each component of the vector θ may be rather unpredictable, presenting notable oscillations and peaking phenomena. Furthermore, usual tuning procedures for these estimators involve the tuning of a gain matrix, and they can be delicate, involving many trial-and-error attempts. For instance, it has been proved in [13] that the amplification of the gains for gradient estimators does not always produce accelerated transients, and it provokes pikes augmentation.

The Dynamic Regressor Extension and Mixing (DREM) procedure has been recently introduced in [14], and in [15], the authors propose an interpretation of this procedure as a functional Luenberger observer. The DREM approach has been successfully applied to a variety of adaptive control problems and applications, such as direct model reference control [16], time-varying frequency estimation [17], electrical drives velocity estimation [18], and power system applications [19], [20]. Among many advantages, the DREM procedure ensures an element-wise transient monotonicity preventing oscillations and peaking, independently of the excitation conditions. Moreover, each element of the estimate of θ is tuned with a separate scalar gain, that does not affect transients for other elements making the gain tuning simpler and transparent.

Two fundamental steps are involved in the DREM method, namely the *dynamic regressor extension* (DRE) and *mixing* steps. In the first step, a linear dynamic operator is introduced to extend the original linear regression equation (1) and to obtain an extended matrix regressor. In the second step, a nonlinear transformation is employed on the previously generated data to obtain a set of n scalar independent LRE for each component of the vector θ sharing the same new scalar regressor. Finally, the gradient estimator is applied to each of these scalar equations.

To choose the linear operator for extending the dynamics in the first step of the DREM procedure is a key point. The

*This work was supported by the Ministry of Science and Higher Education of Russian Federation, passport of goszadanie no. 2019-0898, and by Rennes Métropole (France), project AIS-19C0326. The corresponding author is S. Aranovskiy, contact email: stanislav.aranovskiy@centralesupelec.fr.

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

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

[3]Inria, Univ. Lille, CNRS, UMR 9189 - CRIStAL, F-59000 Lille, France.

question is how to chose such an operator that the excitation level of the regressor ϕ is preserved. A poor choice of this operator can compromise the convergence even if the original regressor ϕ is PE, as it has been shown in [14]. In the discrete-time domain, a summation over a fixed window has been proposed in [21], but *a priori* knowledge about the original regressor ϕ must be available for choosing the window size. This choice problem has also been considered in [22] for a particular class of LRE, where the regressor consists of a finite sum of sinusoidal signals, and the upper frequency bound is known.

One possible dynamics extension, widely used in adaptive control, is the Kreisselmeier's regressor extension introduced in [23], see also the Integral Cost Gradient Adaptation algorithm [3, Section 4.3]. The authors in [24], [25] have used the Kreisselmeier's regressor extension in the DREM procedure, calling it the memory regressor extension; however, these papers do not address the excitation preservation properties of such a choice.

This paper aims to close this gap and to study in detail the applicability of the Kreisselmeier's scheme. To this end, we investigate under which conditions it preserves the PE and IE properties of the original regressor.

Novelty and contribution. The contribution of this paper is the analysis of the properties of the Kreisselmeier's scheme in the context of the DREM procedure. We show that both PE and IE properties of the original regressor are *always preserved*, and that the novel regressor is PE *if and only if* the original regressor is PE. We derive the lower bound of the resulting excitation level and show its dependence on the parameters of the Kreisselmeier's scheme. Finally, we study the dynamics of the novel regressor and provide some bounds on it.

A preliminary version of the present article was published in [26]. The novelties in the paper are its several extensions, such as the output y is in $\mathbb{R}^\ell$, the equivalence of the PE condition in Theorem 1, and the proof of Theorem 1. Also, an interval excitation analysis of the dynamic extension is given in Proposition 1. Also, compared to [26], the proof of Proposition 2 is added, and the simulations are revised.

The organization of this paper is as follows. In Section II, some basic material is provided, together with a brief description of the DREM method. The main results on PE and IE conditions for the extension of the dynamics are provided in Section III, and the regressor dynamics is studied in Section IV. In Section V, the numerical simulations illustrate our results.

Notation. The set of positive integers is denoted by $\mathbb{N}$. For $m,n \in \mathbb{N}$, $\overline{m,n}$ is the set of elements of $\mathbb{N}$ greater or equal than m and smaller or equal than n. The matrix I_n is the $n\times n$ identity matrix, for all $n \in \mathbb{N}\setminus\{0\}$. For a signal $x:\mathbb{R}_+\to\mathbb{R}$ and a linear operator $\mathcal{H}$, we denote the action of this operator on the signal x as $\mathcal{H}[x]$. For a function $x:\mathbb{R}_+\to\mathbb{R}$, we say that $x\in\mathcal{L}_2$ if $\int_0^t x^2(s)ds$ converges to a constant as t tends to $+\infty$. If the integral does not converge, we write $x\notin\mathcal{L}_2$.

II. Background material

The persistence of excitation (PE): First, we present the definition of the (T,μ)-PE property.

Definition 1. *A bounded signal $\phi:\mathbb{R}_+\to\mathbb{R}^{n\times\ell}$ is (T,μ)-*persistently exciting *if there exist $T>0$ and $\mu>0$ such that for all $t\in\mathbb{R}_+$,*

$$\int_t^{t+T}\phi(s)\phi^\top(s)ds\geq\mu I_n.$$

This property is further denoted as $\phi\in$ PE, or ϕ is PE.

The persistence of excitation property and its connection with the exponential convergence in various estimation schemes are widely known. One relaxation of this condition is the *interval* (or *sufficient*) excitation that is used in, e.g., concurrent and composite learning algorithms [7], [8].

Definition 2. *A bounded signal $\phi:\mathbb{R}_+\to\mathbb{R}^{n\times\ell}$ is (t_1,T,μ)-*interval exciting *if there exist $t_1\geq 0$, $T>0$, and $\mu>0$ such that*

$$\int_{t_1}^{t_1+T}\phi(s)\phi^\top(s)ds\geq\mu I_n.$$

The fundamental difference is that the persistence of excitation is uniform in time, whereas the interval excitation holds for the particular time interval starting at t_1. If $t_1=0$, then the interval excitation is also called the *initial excitation*, see [9].

The gradient estimators [2]: For all $t\in\mathbb{R}_+$, the gradient estimator for the LRE (1) is given by

$$\dot{\hat\theta}(t)=\Gamma\phi(t)\left(y(t)-\phi^\top(t)\hat\theta(t)\right),\tag{2}$$

where $\hat\theta$ denotes the estimate of θ and $\Gamma>0$ is the gain matrix. Define the estimation error $\tilde\theta(t):=\hat\theta(t)-\theta$, $\forall t\in\mathbb{R}_+$. Then the error dynamics is given by

$$\dot{\tilde\theta}(t)=-\Gamma\phi(t)\phi^\top(t)\tilde\theta(t)+\Gamma\phi(t)w(t),\forall t\in\mathbb{R}_+.$$

In the noise-free scenario, i.e., assuming $w\equiv 0$, the gradient estimator ensures exponential convergence to zero of the error $\tilde\theta$ if and only if the regressor ϕ is PE. In this case, the gradient estimator is also input-to-state stable with respect to the noise w. Some sufficient (but not necessary) and necessary (but not sufficient) conditions for asymptotic convergence when ϕ is not PE have been discussed in [27] for $\ell=1$. However, they are somewhat technical and can be hardly applied in practice.

The DREM procedure [14]: To apply the DREM procedure, we start by performing the *dynamic regressor extension* step. For that, we introduce a linear, ℓ-input n-output, bounded-input bounded-output (BIBO)–stable operator $\mathcal{H}$ and define the vector $Y:\mathbb{R}_+\to\mathbb{R}^n$ and the matrix $\Phi:\mathbb{R}_+\to\mathbb{R}^{n\times n}$ by:

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

Due to the linearity of the operator $\mathcal{H}$ and the BIBO stability, these signals satisfy

$$Y(t)=\Phi(t)\theta+W(t),\forall t\in\mathbb{R}_+,\tag{3}$$

where $W:=\mathcal{H}[w]$. For example, the operator $\mathcal{H}$ can produce an LTI system or can be chosen as a delay operator, as proposed in [14].

Next, a *mixing* step is applied to obtain a set of n scalar equations. Recall that for any square and possibly singular $n\times n$ matrix A, we have $\operatorname{adj}(A)A=\det(A)I_n$, where $\operatorname{adj}(\cdot)$ is

the *adjoint* (also called *adjugate*) matrix. Multiplying (3) by $\operatorname{adj}(\Phi(t))$ from the left, we get

$$\mathcal{Y}_i(t) = \Delta(t)\theta_i + \mathcal{W}_i(t), \tag{4}$$

where $i \in \overline{1,n}$, the scalar function $\Delta : \mathbb{R}_+ \to \mathbb{R}$ is defined as

$$\Delta(t) := \det \Phi(t), \forall t \in \mathbb{R}_+, \tag{5}$$

and $\mathcal{Y}(t) := \operatorname{adj}(\Phi(t))\, Y(t)$, $\mathcal{W}(t) := \operatorname{adj}(\Phi(t))\, W(t)$, $\forall t \in \mathbb{R}_+$. It is worth noting that for a bounded regressor ϕ, the vector $\mathcal{W}$ is also bounded, and $w \equiv 0$ implies $\mathcal{W} \equiv 0$.

The set of n scalar LRE (4) sharing the same bounded scalar regressor Δ is the main result of the DREM procedure. Applying the gradient estimator to (4) as

$$\dot{\hat{\theta}}_i(t) = \gamma_i\ \Delta(t)\left(\mathcal{Y}_i(t) - \Delta(t)\hat{\theta}_i(t)\right), \tag{6}$$

where $\gamma_i > 0$ is a scalar tuning parameter, we obtain

$$\dot{\tilde{\theta}}_i(t) = -\gamma_i \Delta^2(t)\tilde{\theta}_i(t) + \gamma_i \Delta(t)\mathcal{W}_i(t),$$

and thus

$$\tilde{\theta}_i(t) = e^{-\gamma_i \int_0^t \Delta^2(\tau)d\tau}\theta_i(0) + \gamma_i \int_0^t e^{-\gamma_i \int_s^t \Delta^2(\tau)d\tau}\Delta(s)\mathcal{W}_i(s)ds.$$

Assuming the noise-free scenario $w \equiv 0$, the following properties hold:

P1: $\Delta \notin \mathcal{L}_2 \Leftrightarrow |\tilde{\theta}| \to 0$ asymptotically;
P2: Δ is PE $\Leftrightarrow |\tilde{\theta}| \to 0$ exponentially fast;
P3: (the element-wise monotonicity) for all $i \in \overline{1,n}$ for $t_a \le t_b$ it holds $|\tilde{\theta}_i(t_a)| \le |\tilde{\theta}_i(t_b)|$;
P4: (the element-wise tuning) variations in the gain γ_i affect the transients for $\hat{\theta}_i$ only.

Concerning the case $w \not\equiv 0$, the estimator (6) is input-to-state stable with respect to $\mathcal{W}_i$ if $\Delta \in$ PE, which is a similar result as for the standard gradient estimator discussed above. Moreover, as it has been shown in [28], if $\mathcal{W}_i \in \mathcal{L}_2$ and $\Delta \notin \mathcal{L}_2$, then $\tilde{\theta}_i$ is bounded.

The property P1 illustrates the new convergence condition, namely the non-square-integrability of Δ. As shown in [14], this condition is weaker than PE for ϕ, where the price paid is the asymptotic convergence instead of the exponential one. To get the exponential convergence with the DREM procedure, the PE property of Δ must be satisfied. Therefore, the main design question when applying the DREM procedure is to choose an operator $\mathcal{H}$ such that the PE property of the original regressor ϕ is preserved. Such a choice is discussed in the following section.

Remark 1. *Let us remark that the computation of the adjoint matrix* $\operatorname{adj}(\Phi)$ *can be avoided in numerical implementations of the DREM estimator. So the elements* $\mathcal{Y}_i$ *in* (4) *can be computed using the Cramer's rule as*

$$\mathcal{Y}_i(t) = \det \Phi_{Y,i}(t), \forall t \in \mathbb{R}_+,$$

where $\Phi_{Y,i}$ *is the matrix* Φ *where the i-th column is replaced with the vector Y, and* $i \in \overline{1,n}$.

III. The excitation-preserving regressor extension

A. Excitation propagation

Consider the Kreisselmeier's regressor extension [23] given by

$$\dot{\Phi}(t) = -a\Phi(t) + \phi(t)\phi^\top(t), \tag{7}$$
$$\dot{Y}(t) = -aY(t) + \phi(t)y(t) \tag{8}$$

for some initial values $\Phi(0) = \Phi_0 \ge 0$ and $Y(0) = Y_0$, where the scalar $a > 0$ is the tuning parameter. As it is discussed in [21], [25], the scheme (7), (8) can be seen either as an LTI filter applied to the signals $\phi\phi^\top$ and ϕy, or as an LTV operator $\mathcal{H}_\phi$ chosen such that the relationship $y_u = \mathcal{H}_\phi[u]$ for an input signal $u(t) \in \mathbb{R}^\ell$ and an output signal $y_u(t) \in \mathbb{R}^n$ have the following linear time-varying state-space representation:

$$\dot{x}(t) = -ax(t) + \phi(t)u(t),$$
$$y_u(t) = x(t),$$

where $x(t) \in \mathbb{R}^n$ is the internal state vector. Then the Kreisselmeier's regressor extension (7), (8) corresponds to $\Phi = \mathcal{H}_\phi[\phi^\top]$ and $Y = \mathcal{H}_\phi[y]$, where the use of the ℓ-input n-output operator $\mathcal{H}_\phi$ for the $\ell \times n$ matrix $\phi^\top$ implies that the operator is applied to the each column of $\phi^\top$, and the resulting vectors are then collected to the $n \times n$ matrix Φ. Thus, the scheme (7), (8) is a valid choice for the *dynamic regressor extension* step of the DREM procedure.

The regressor extension (7), (8) is widely used in adaptive control. Particularly, it was recently used in [12], [29] to obtain the matrix equation $Y(t) = \Phi(t)\theta$, $\forall t \in \mathbb{R}_+$. Applicability of (7), (8) for the DREM procedure can be derived from the proof of Theorem 4.3.3 in [3], where the integral cost gradient adaptation algorithm is considered. However, in that Theorem, only the positiveness of the lower bound on the smallest eigenvalue of the matrix Φ is established. Extending that the result, we present the following theorem showing that (7) preserves the persistence of excitation and the determinant of Φ is PE *if and only if* ϕ is PE and providing a more precise asymptotic lower bound for the determinant of the matrix Φ.

Theorem 1. *Consider the bounded signal* $\phi : \mathbb{R}_+ \to \mathbb{R}^{n\times\ell}$ *and let* $\Phi : \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ *be a solution of* (7) *for some initial value* $\Phi(0) = \Phi_0 \ge 0$. *Then the following implication holds*

$$\phi \in \text{PE} \Leftrightarrow \Delta \in \text{PE},$$

where $\Delta : \mathbb{R}_+ \to \mathbb{R}$ *is the determinant of* Φ. *Moreover, if* ϕ *is* (T,μ)*-PE, then for any positive integer* $q \ge 1$ *and for all* $t \ge qT$, *it holds*

$$\Delta(t) \ge \mu^n \left(\sum_{k=1}^{q} e^{-akT}\right)^n \tag{9}$$

and

$$\liminf_{t\to\infty} \Delta(t) \ge \left(\frac{\mu}{e^{aT} - 1}\right)^n. \tag{10}$$

Proof: The proof consists of two parts. First, we prove that $\phi \in$ PE implies $\Delta \in$ PE, and then we prove that the inverse implication also holds.

Part 1: $\phi \in \text{PE} \Rightarrow \Delta \in \text{PE}$. First we show that if ϕ is (T,μ)-PE, then Δ is also PE and the inequalities (9) and (10) hold.

The solution of (7) is given by

$$\Phi(t) = e^{-at}\Phi(0) + \int_0^t \psi(t,s)ds, \ \forall t \in \mathbb{R}_+, \tag{11}$$

where

$$\psi(t,s) := e^{-a(t-s)}\phi(s)\phi^\top(s).$$

Consider $t \geq T$ and let $q \geq 1$ be a positive integer number such that $t \geq qT$. The integral term in (11) can be rewritten as

$$\int_0^t \psi(t,s)ds = \int_0^{t-qT} \psi(t,s)ds + \sum_{k=1}^{q} \int_{t-kT}^{t-kT+T} \psi(t,s)ds.$$

For any positive integer $k \leq q$ it holds

$$\begin{aligned} \int_{t-kT}^{t-kT+T} \psi(t,s)ds &= e^{-at}\int_{t-kT}^{t-kT+T} e^{as}\phi(s)\phi^\top(s)ds \\ &\geq e^{-at}e^{a(t-kT)}\mu I_n = \mu e^{-akT} I_n. \end{aligned}$$

Then

$$\Phi(t) \geq \mu \sum_{k=1}^{q} e^{-akT} I_n + \int_0^{t-qT} \psi(t,s)ds + e^{-at}\Phi(0). \tag{12}$$

For $\Phi(0) \geq 0$, the sum of the last two terms in the right-hand side of this inequality is a semi positive-definite matrix,

$$\int_0^{t-qT} \psi(t,s)ds + e^{-at}\Phi(0) \geq 0.$$

Then from (12) it follows that for all $t \geq qT$ the smallest eigenvalue of $\Phi(t)$ is not less than $\mu\sum_{k=1}^{q} e^{-akT}$. Thus

$$\det\Phi(t) \geq \mu^n \left(\sum_{k=1}^{q} e^{-akT}\right)^n,$$

and (9) follows.

The PE property follows by noting that for $t \geq T$ it holds

$$\Delta(t) \geq \mu^n e^{-anT} > 0,$$

and Δ is strictly separated from zero for all $t \geq T$.

To get the inequality (10), we choose q as the largest integer such that $t \geq qT$. Then $q \to \infty$ as $t \to \infty$. Since

$$\lim_{q\to\infty} \sum_{k=1}^{q} e^{-akT} = \frac{1}{e^{aT}-1},$$

the asymptotic lower bound (10) for $\Delta(t)$ follows.

Part 2: $\Delta \in \text{PE} \Rightarrow \phi \in \text{PE}$. Now we will show that if Δ is PE, then ϕ is also PE. More precisely, we will show that if ϕ is bounded and there exist $T > 0$ and $\mu > 0$ such that for all $t \in \mathbb{R}_+$

$$\int_t^{t+T} \Delta(s)^2 ds \geq \mu,$$

then there exist $L > 0$ and $\alpha > 0$ such that for all $t \in \mathbb{R}_+$

$$\int_t^{t+L} \phi(s)\phi^\top(s)ds \geq \alpha I_n.$$

Since the matrix Φ given by (11) is bounded for bounded ϕ, it follows that all its eigenvalues are non-negative and also bounded, and there exists a constant $c > 0$ such that[1]

$$c\lambda_m(\Phi(t)) \geq \Delta^2(t), \forall t \in \mathbb{R}_+,$$

where $\lambda_m(\Phi)$ is the smallest eigenvalue of Φ. Then

$$\int_t^{t+T} \lambda_m(\Phi(s))\,ds \geq \frac{\mu}{c}. \tag{13}$$

From (11) it follows that for $s \geq t$

$$\Phi(s) = e^{-a(s-t)}\Phi(t) + \int_t^s e^{-a(s-\tau)}\phi(\tau)\phi^\top(\tau)d\tau.$$

Then recalling (13), for any positive integer k it holds

$$\begin{aligned} &\int_t^{t+kT} \lambda_m(\Phi(s))\,ds \\ &= \int_t^{t+T} \lambda_m\left(e^{-a(s-t)}\Phi(t) + \int_t^s e^{-a(s-\tau)}\phi(\tau)\phi^\top(\tau)d\tau\right)ds \\ &\geq \int_t^{t+kT} \lambda_m\left(e^{-a(s-t)}\Phi(t)\right)ds \\ &+ \int_t^{t+kT} \lambda_m\left(\int_t^s e^{-a(s-\tau)}\phi(\tau)\phi^\top(\tau)d\tau\right)ds \geq k\frac{\mu}{c}. \end{aligned}$$

Note that

$$\begin{aligned} &\int_t^{t+kT} \lambda_m\left(e^{-a(s-t)}\Phi(t)\right)ds \\ &= \lambda_m(\Phi(t))\int_t^{t+kT} e^{-a(s-t)}ds \leq \frac{1}{a}\lambda_m(\Phi(t)) \end{aligned}$$

for $a > 0$. Choose c_0 as $c_0 := \frac{1}{a}\sup_t \lambda_m(\Phi(t))$. Then we have that for any k

$$\int_t^{t+kT} \lambda_m\left(\int_t^s e^{-a(s-\tau)}\phi(\tau)\phi^\top(\tau)d\tau\right)ds \geq k\frac{\mu}{c} - c_0. \tag{14}$$

Note that

$$\lambda_m\left(\int_t^s e^{-a(s-\tau)}\phi(\tau)\phi^\top(\tau)d\tau\right) \leq \lambda_m\left(\int_t^s \phi(\tau)\phi^\top(\tau)d\tau\right),$$

and for all s satisfying $t \leq s \leq t+kT$ it holds

$$\lambda_m\left(\int_t^s \phi(\tau)\phi^\top(\tau)d\tau\right) \leq \lambda_m\left(\int_t^{t+kT} \phi(\tau)\phi^\top(\tau)d\tau\right).$$

Thus

$$\begin{aligned} &\int_t^{t+kT} \lambda_m\left(\int_t^s e^{-a(s-\tau)}\phi(\tau)\phi^\top(\tau)d\tau\right)ds \\ &\leq \int_t^{t+kT} \lambda_m\left(\int_t^s \phi(\tau)\phi^\top(\tau)d\tau\right)ds \\ &\leq \int_t^{t+kT} \lambda_m\left(\int_t^{t+kT} \phi(\tau)\phi^\top(\tau)d\tau\right)ds \\ &= kT\lambda_m\left(\int_t^{t+kT} \phi(\tau)\phi^\top(\tau)d\tau\right). \end{aligned} \tag{15}$$

[1] A conservative estimate of c is $c = \left(\sup_t \lambda_M(\Phi(t))\right)^{2n-1}$, where $\lambda_M(\Phi)$ is the largest eigenvalue of Φ.

Finally, combining (14) and (15) yields

$$\lambda_m\left(\int_t^{t+kT}\phi(\tau)\phi^\top(\tau)d\tau\right)\geq\frac{\mu}{cT}-\frac{c_0}{kT}.$$

Since c and c_0 are constants and do not depend on k, we can choose the positive integer $k\geq 1$ such that $k>\frac{c_0c}{\mu}$. Then $\phi(t)$ is PE with $L=kT$ and $\alpha=\frac{\mu}{cT}-\frac{c_0}{kT}>0$. ■

In the vein of Theorem 1, it can be also shown that the dynamic extension (7) preserves the interval excitation (in the sense of Definition 2) as well.

Proposition 1. *Consider the bounded signal $\phi:\mathbb{R}_+\to\mathbb{R}^{n\times\ell}$ and let $\Phi:\mathbb{R}_+\to\mathbb{R}^{n\times n}$ be a solution of (7) for some initial value $\Phi(0)=\Phi_0\geq 0$. Let the signal $\Delta:\mathbb{R}_+\to\mathbb{R}$ be the determinant of Φ. If the signal ϕ is (t_1,T,μ)-interval exciting for some $t_1\geq 0$, $T>0$, and $\mu>0$, then the signal Δ is (t_1,T,α)-interval exciting for some $\alpha>0$.*

Proof: As it is discussed in the proof of Theorem 1, the solution of (7) is given by (11). Then for $t_1\leq t\leq t_1+T$

$$\Phi(t)=e^{-a(t-t_1)}\Phi(t_1)+\int_{t_1}^{t}e^{-a(t-s)}\phi(s)\phi^\top(s)ds,$$

and

$$\begin{aligned}\int_{t_1}^{t_1+T}\lambda_m\left(\Phi(t)\right)dt&\geq\int_{t_1}^{t_1+T}\lambda_m\left(\int_{t_1}^{t}e^{-a(t-s)}\phi(s)\phi^\top(s)ds\right)dt\\&\geq\int_{t_1}^{t_1+T}e^{-a(t-t_1)}f(t)dt,\end{aligned}$$

where $f:\mathbb{R}_+\to\mathbb{R}$ is defined as

$$f(t)=\lambda_m\left(\int_{t_1}^{t}\phi(s)\phi^\top(s)ds\right).$$

The function f is continuous and non-decreasing, $f(t_1)=0$ and (due to the interval excitation) $f(t_1+T)\geq\mu$. The function f is not necessarily differentiable; however, since ϕ is bounded it admits a Lipschitz constant, i.e., there exists a positive constant $\rho>0$ such that for any $s\geq t_1$ and $h\geq 0$ it holds

$$0\leq f(s+h)-f(s)\leq\rho h,$$

and $\rho T\geq\mu$. Define $\tau:=\frac{\mu}{\rho}\leq T$. Then for all $t\in[t_1+T-\tau,t_1+T]$ it holds

$$f(t)\geq\rho\left(t-(t_1+T-\tau)\right).$$

Thus

$$\begin{aligned}&\int_{t_1}^{t_1+T}e^{-a(t-t_1)}f(t)dt\\\geq&\int_{t_1+T-\tau}^{t_1+T}e^{-a(t-t_1)}\rho\left(t-(t_1+T-\tau)\right)dt\\=&\frac{\rho}{a^2}e^{-aT}\left(e^{a\tau}-1-a\tau\right).\end{aligned}$$

Finally, recalling that $\Delta(t)\geq(\lambda_m(\Phi(t)))^n$, the (t_1,T,α)-interval excitation follows:

$$\int_{t_1}^{t_1+T}\Delta^2(t)dt\geq\alpha,$$

where

$$\alpha:=\left(\frac{\rho}{a^2}e^{-aT}\left(e^{\frac{a\mu}{\rho}}-1-\frac{a\mu}{\rho}\right)\right)^{2n}>0.$$

■

It is worth noting that the inverse implication in Proposition (1) does not hold. I.e., if Δ is (t_1,T,α)-interval exciting for some $t_1>0$, it does not imply that ϕ is (t_1,T,μ)-interval exciting for the same t_1 and T. It is reasonable to expect that (given zero initial condition in (7)) ϕ is $(0,t_1+T,\mu)$-interval exciting, however, this claim is not rigorously proven here.

Theorem 1 motivates the use of (7), (8) as a reasonable choice for the dynamic regressor extension step of the DREM procedure. Under this choice, the persistence of excitation and the interval excitation properties of the original regressor are always preserved and no prior knowledge about the regressor ϕ (e.g., the excitation period length T) is required. Moreover, after the first excitation interval, the new regressor Δ remains strictly positive for all t, where the lower asymptotic bound on Δ can be computed as in (10).

IV. Dynamics of the regressor Δ

Theorem 1 provides the lower bounds (9), (10) for the dynamic regressor extension (7). However, it is also of interest to study the dynamics of the DREM-generated new regressor Δ. In this section, we present such a study for the particular case $\ell=1$, i.e., for the scalar output case. Extension of the results presented in this section to a more general case remains an open question.

Let us consider a time evaluation of Δ that is the determinant of the time-dependent matrix Φ. The matrix Φ is a solution of (7), so Δ obeys Jacobi's formula (see Theorem 8.1 in [30]):

$$\dot\Delta(t)=\operatorname{tr}\left(\operatorname{adj}\left(\Phi(t)\right)\dot\Phi(t)\right),\ \forall t\in\mathbb{R}_+,$$

where $\Delta(0)=\det\Phi(0)$ and tr denotes the matrix trace.

Substituting (7), we obtain for all $t\in\mathbb{R}_+$:

$$\begin{aligned}\dot\Delta(t)&=\operatorname{tr}\left(-a\operatorname{adj}\left(\Phi(t)\right)\Phi(t)+\operatorname{adj}\left(\Phi(t)\right)\phi(t)\phi(t)^\top\right)\\&=-an\Delta(t)+\operatorname{tr}\left(\operatorname{adj}\left(\Phi(t)\right)\phi(t)\phi(t)^\top\right),\end{aligned}$$

where n is the dimension of ϕ. Next, due to elementary properties of the matrix trace function, it follows for all $t\geq 0$,

$$\operatorname{tr}\left(\operatorname{adj}\left(\Phi(t)\right)\phi(t)\phi(t)^\top\right)=\phi(t)^\top\operatorname{adj}\left(\Phi(t)\right)\phi(t),$$

and we obtain

$$\dot\Delta(t)=-an\Delta(t)+\phi(t)^\top\operatorname{adj}\left(\Phi(t)\right)\phi(t).\tag{16}$$

Remark 2. *Note that (16) holds only in the case $\ell=1$ when $\phi(t)$ is a vector and not a matrix.*

Recall that the eigenvalues of an adjoint matrix can be estimated as follows. Let $\lambda_{1,\Phi},\dots,\lambda_{n,\Phi}$ denote the eigenvalues of Φ. Applying Schur's Lemma, it is then straightforward to show that the eigenvalues of $\operatorname{adj}(\Phi)$ are given by

$$\lambda_{i,\operatorname{adj}(\Phi)}=\Pi_{j\neq i}\lambda_{j,\Phi},\ \forall i=1,\dots,n,$$

and for all $i=1,\dots,n$ it holds

$$\lambda_{i,\operatorname{adj}(\Phi)}\lambda_{i,\Phi}=\det\left(\Phi\right).$$

In particular

$$\min_i\lambda_{i,\operatorname{adj}(\Phi)}\max_i\lambda_{i,\Phi}=\det\left(\Phi\right).\tag{17}$$

Let λ_M denote the maximum eigenvalue of Φ. Since for $t \geq 0$, $\Phi(t) \geq 0$, then $\lambda_M(t) = 0$ implies that all eigenvalues of $\Phi(t)$ are zeros, and so are the eigenvalues of $\operatorname{adj}(\Phi(t))$. That implies for all $t \geq 0$,

$$\phi(t)^\top \operatorname{adj}(\Phi(t))\,\phi(t) = 0.$$

On the other hand, if $\lambda_M(t) > 0$, then due to (17)

$$\|\phi(t)\|^2 \frac{\Delta(t)}{\lambda_M(t)} \leq \phi^\top(t) \operatorname{adj}(\Phi(t))\,\phi(t), \ \forall t \in \mathbb{R}_+.$$

The derivation above yields the following proposition.

Proposition 2. *Let Φ be a solution of (7) and let λ_M denote the maximum eigenvalue of Φ and $\Delta = \det\Phi$. Then:*

- *if $\lambda_M = 0$, then $\dot{\Delta} = 0$;*
- *if $\lambda_M > 0$, then*

$$\dot{\Delta}(t) \geq \left(-an + \frac{\|\phi(t)\|^2}{\lambda_M(t)}\right)\Delta(t), \ \forall t \in \mathbb{R}_+.$$

It is also worth noting that an upper bound of the maximum eigenvalue λ_M can be estimated given an upper bound of ϕ.

V. Simulations

To illustrate the results obtained in Section III, we consider the problem of magnitude and phase estimation for sinusoidal signals with known frequencies. To illustrate that the presented results hold not only for scalar linear regression models, we choose $\ell = 2$ and $n = 3$. Consider

$$y_1(t) = B + A\sin(t+\psi),$$
$$y_2(t) = A\cos(2t+\psi)$$

where B, $A > 0$, and $\psi \in [-\pi, \pi)$ are the unknown parameters. These signals can be rewritten as the LRE (1) with

$$y(t) = \begin{bmatrix} y_1(t) & y_2(t) \end{bmatrix}^\top, \ \theta = \begin{bmatrix} B & A\cos(\psi) & A\sin(\psi) \end{bmatrix}^\top,$$
$$\phi(t) = \begin{bmatrix} 1 & \sin(t) & \cos(t) \\ 0 & \cos(2t) & -\sin(2t) \end{bmatrix}^\top,$$

and $w \equiv 0$. Obviously, the values A, B, and ψ can be reconstructed given θ.

It is straightforward to verify that the regressor ϕ is $(2\pi, 2\pi)$-PE, i.e., for all $t \geq 0$

$$\int_t^{t+2\pi} \phi(s)\phi^\top(s)ds \geq 2\pi I_3.$$

For simulations, we set $B = -1$, $A = 2\sqrt{2}$, and $\psi = \frac{3}{4}\pi$ providing

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

First, we apply the standard gradient algorithm (2) with $\Gamma = I_3$ and $\hat{\theta}(0) = 0$; the simulation results are depicted in Fig. 1. Since ϕ is PE, the estimation error $\tilde{\theta}$ converges to zero; however, the transients exhibit oscillations.

Then we apply the dynamic regressor extension (7), (8), where the only tuning parameter is chosen as $a = 0.1$. The new regressor Δ computed as (5) is depicted in Fig. 2 with the lower bound (9) and with the asymptotic lower bound (10).

Finally, the estimation error transient $\tilde{\theta}$ for the DREM procedure with the gradient estimator (6) with $\gamma_i = 1$, $i = 1, 2, 3$, are depicted in Fig. 3 and illustrate performance improvement for the standard gradient estimator; note the difference in the time scale and the monotonicity of the transients.

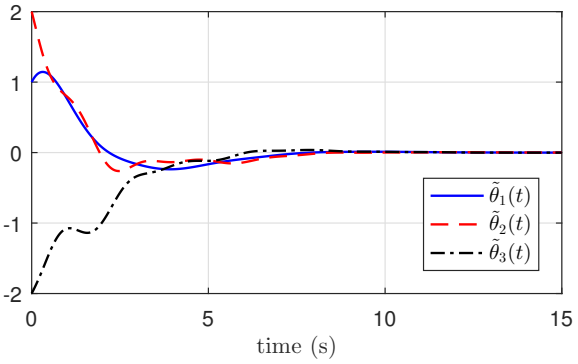


Fig. 1. The estimation error $\tilde{\theta}$ under the standard gradient algorithm (2).

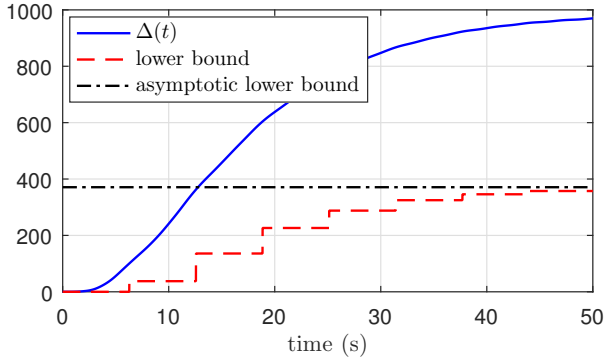


Fig. 2. The new regressor Δ, the lower bound (9), and the asymptotic lower bound (10).

VI. Conclusion

The recently proposed DREM procedure provides significant performance improvement in linear regression parameter estimation, where the main degree of freedom is the choice of the dynamic extension operator. In this paper, we have studied a special choice of this operator as the Kreisselmeier's regressor extension (7), (8) that ensures the preservation of the persistence of excitation property. In particular, we have proved that the determinant of the extended matrix is persistently exciting if and only if the original regressor is, where the asymptotic lower bound of the excitation constant is also provided. Moreover, we have shown that the suggested choice

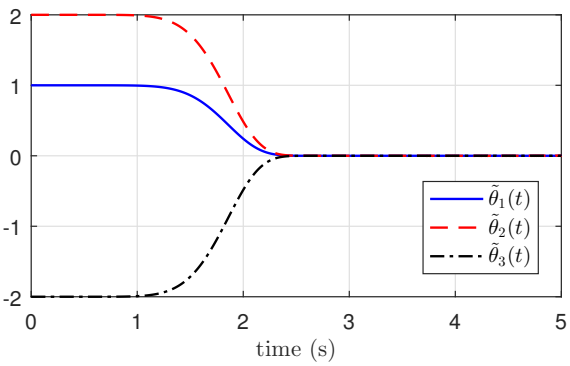


Fig. 3. The estimation error $\tilde{\theta}$ under the DREM procedure (7), (8) and the gradient estimator (6).

also preserves the interval excitation property. This result alleviates the main design question of the DREM procedure.

REFERENCES

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

[2] S. Sastry and M. Bodson, *Adaptive control: Stability, convergence, and robustness*. New Jersey: Prentice-Hall, 1989.

[3] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. New Jersey: Prentice–Hall, 1996.

[4] M. M'Saad, R. Ortega, and I. Landau, "Adaptive controllers for discrete-time systems with arbitrary zeros: An overview," *Automatica*, vol. 21, no. 4, pp. 413–423, jul 1985.

[5] G. C. Goodwin and K. S. Sin, *Adaptive filtering prediction and control*. Courier Corporation, 2014.

[6] F. Lewis, D. Vrabie, and K. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," *IEEE Control Systems*, vol. 32, no. 6, pp. 76–105, 2012.

[7] 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, apr 2013.

[8] R. Kamalapurkar, B. Reish, G. Chowdhary, and W. E. Dixon, "Concurrent learning for parameter estimation using dynamic state-derivative estimators," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3594–3601, jul 2017.

[9] Y. Pan and H. Yu, "Composite learning robot control with guaranteed parameter convergence," *Automatica*, vol. 89, pp. 398–406, mar 2018.

[10] Y. Pan, S. Aranovskiy, A. Bobtsov, and H. Yu, "Efficient learning from adaptive control under sufficient excitation," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 10, pp. 3111–3124, jul 2019.

[11] J. Wang, D. Efimov, S. Aranovskiy, and A. A. Bobtsov, "Fixed-time estimation of parameters for non-persistent excitation," *European Journal of Control*, 2019.

[12] N. Cho, H.-S. Shin, Y. Kim, and A. Tsourdos, "Composite model reference adaptive control with parameter convergence under finite excitation," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 811–818, mar 2018.

[13] D. Efimov and A. Fradkov, "Design of impulsive adaptive observers for improvement of persistency of excitation," *International Journal of Adaptive Control and Signal Processing*, vol. 29, no. 6, pp. 765–782, jun 2015.

[14] 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, jul 2017.

[15] R. Ortega, L. Praly, S. Aranovskiy, B. Yi, and W. Zhang, "On dynamic regressor extension and mixing parameter estimators: Two luenberger observers interpretations," *Automatica*, vol. 95, pp. 548–551, sep 2018.

[16] D. N. Gerasimov, R. Ortega, and V. O. Nikiforov, "Relaxing the high-frequency gain sign assumption in direct model reference adaptive control," *European Journal of Control*, vol. 43, pp. 12–19, sep 2018.

[17] A. A. Vedyakov, A. O. Vediakova, A. A. Bobtsov, A. A. Pyrkin, and S. V. Aranovskiy, "A globally convergent frequency estimator of a sinusoidal signal with a time-varying amplitude," *European Journal of Control*, vol. 38, pp. 32–38, nov 2017.

[18] A. Bobtsov, D. Bazylev, A. Pyrkin, S. Aranovskiy, and R. Ortega, "A robust nonlinear position observer for synchronous motors with relaxed excitation conditions," *International Journal of Control*, vol. 90, no. 4, pp. 813–824, apr 2017.

[19] J. Schiffer, P. Aristidou, and R. Ortega, "Online estimation of power system inertia using dynamic regressor extension and mixing," *IEEE Transactions on Power Systems*, pp. 1–1, 2019.

[20] A. Pyrkin, F. Mancilla-David, R. Ortega, A. Bobtsov, and S. Aranovskiy, "Identification of photovoltaic arrays' maximum power extraction point via dynamic regressor extension and mixing," *International Journal of Adaptive Control and Signal Processing*, vol. 31, no. 9, pp. 1337–1349, sep 2017.

[21] R. Ortega, S. Aranovskiy, A. Pyrkin, A. Astolfi, and A. Bobtsov, "New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases," *IEEE Transactions on Automatic Control*, 2020, accepted.

[22] S. Aranovskiy, A. Belov, R. Ortega, N. Barabanov, and A. Bobtsov, "Parameter identification of linear time-invariant systems using dynamic regressor extension and mixing," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 6, pp. 1016–1030, jun 2019.

[23] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Transactions on Automatic Control*, vol. 22, no. 1, pp. 2–8, feb 1977.

[24] D. N. Gerasimov and V. O. Nikiforov, "On key properties of the lion's and kreisselmeier's adaptation algorithms," in *2020 IFAC World Congress*. IFAC, 2020, accepted.

[25] R. Ortega, V. O. Nikiforov, and D. N. Gerasimov, "On modied parameter estimators for identication and adaptive control. a unied framework and some new schemes," *Annual Reviews in Control*, 2020, to appear.

[26] M. Korotina, S. Aranovskiy, R. Ushirobira, and A. Vedyakov, "On parameter tuning and convergence properties of the drem procedure," in *2020 19th European Control Conference (ECC)*. IEEE, 2020, pp. 53–58.

[27] N. Barabanov and R. Ortega, "On global asymptotic stability of $\dot{x} = -\phi(t)\phi^\top(t)x$ with $\phi(t)$ not persistently exciting," *Systems & Control Letters*, vol. 109, pp. 24–29, nov 2017.

[28] J. Wang, D. Efimov, and A. Bobtsov, "On robust parameter estimation in finite-time without persistence of excitation," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. –, to appear in 2020, avaliable online.

[29] M. Noack, J. G. Rueda-Escobedo, J. Reger, and J. A. Moreno, "Fixed-time parameter estimation in polynomial systems through modulating functions," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 2067–2072.

[30] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, ser. Probabilistics and Statistics. Wiley, 1999.