On preserving-excitation properties of a dynamic regressor extension scheme

Stanislav Aranovskiy, Rosane Ushirobira, Marina Korotina, Alexey Vedyakov

On preserving-excitation properties of Kreisselmeier's regressor extension scheme

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

***Abstract*—In this work, we consider the excitation preservation problem of Kreisselmeier's regressor extension scheme. We analyze this problem within the context of the Dynamic Regressor Extension and Mixing procedure. The well-known qualitative result is that such a scheme preserves excitation. We perform a quantitative analysis and derive lower bounds on the resulting regressor signal considering both persistent and interval excitation cases. We also show that the resulting signal is excited if and only if the original regressor is. Studying the dynamics of the novel regressor, we provide a lower bound on its derivative. Illustrative simulations support our theoretical results.**

I. INTRODUCTION

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]–[4] and adaptive pole-placement [5], in filtering and prediction [6], in reinforcement learning [7], and several other 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.

There exist various approaches to tackle the parameter estimation problem, where two traditional strategies are the least-squares method and the gradient estimator. The convergence of parameter estimation schemes relies on the essential premise of a good richness of the regressor ϕ, realized by the definition of the *persistence of excitation* (PE) condition, see Definition 1 in Section II. In recent years, various efforts to ease the PE requirement have been suggested, such as concurrent [8], [9], or composite learning [10], [11]. Within this methodology, a dynamic data stack is built to record online historical data discretely, and the convergence of parameter estimation is managed following the interval excitation (IE) condition, a weaker requirement than the persistence of excitation, see Definition 2 in Section II. In [12], several fixed-time convergence algorithms have been proposed under the IE assumption on the regressor, and the PE relaxation has also been examined in [13] 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.

The Dynamic Regressor Extension and Mixing (DREM) procedure has been recently introduced in [14], [15]. This procedure allows for translating the LRE (1) to a set of n scalar LREs for each element of the vector θ separately. 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]. The DREM procedure also ensures the 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, which does not affect transients for other elements making the gain tuning more straightforward and transparent.

Two fundamental steps are involved in the DREM method: the *dynamic regressor extension* (DRE) and *mixing* steps. In the first step, the original LRE (1) is linearly extended to obtain an extended LRE with a square 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, an estimator is applied, e.g., a standard gradient estimator or an estimator with the fixed-time convergence [12].

Linearly extending the dynamics in the first step of the DREM procedure is a key point. The question is how to perform such an extension so that the excitation level, either PE or IE, of the regressor ϕ is preserved. A poor choice 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 preserving the excitation, widely used in adaptive control, is Kreisselmeier's regressor extension introduced in [23]. For the LRE (1), Kreisselmeier's regressor extension scheme generates extended matrices Φ and Y as solutions of

$$\dot{\Phi}(t) = -a\Phi(t) + \phi(t)\phi^\top(t), \ \forall t \in \mathbb{R}_+, \tag{2}$$

$$\dot{Y}(t) = -aY(t) + \phi(t)y(t), \tag{3}$$

*This work was supported by Russian Science Foundation grant (project No. 17-79-20341), and by Rennes Metropole, France (project AIS-19C0326). The corresponding author is S. Aranovskiy.

[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.

where $\Phi(0) = \Phi_0 \geq 0$, $Y(0) = Y_0$, and $a > 0$ is a scalar tuning parameter. The PE preservation properties of (2) are well-known and are summarized in the following implication:

$$\phi \text{ is PE } \Rightarrow \Phi(t) > 0 \tag{4}$$

for all $t \geq T$, where T is the excitation interval of ϕ, see Definition 1 in Section II. This PE preservation property of (2) motivates its use in the DREM scheme; the authors in [24], [25] proposed such a choice referring to it as Memory Regressor Extension.

Specifically, within the DREM context, we are interested in the scalar signal Δ that is the determinant of the extended matrix Φ,

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

and for which (4) implies the positiveness and thus the PE property:

$$\phi \text{ is PE } \Rightarrow \Delta(t) > 0, \ \forall t > T \ \Rightarrow \Delta \text{ is PE.} \tag{6}$$

The proof of the implication (4) and thus (6) is well-known and can be found, e.g., in [3, Theorem 4.3.3], where the integral cost gradient adaptation algorithm is considered. However, to the best of our knowledge, only the *qualitative* results (4), (6) are available nowadays, and there is no *quantitative* result for estimating the lower bound of Δ based on the excitation characteristics of ϕ and parameters of (2).

This paper aims to close this gap by providing a *quantitative* analysis of excitation preservation via Kreisselmeier's scheme.

Novelty and Contribution. This paper's contribution is the analysis of Kreisselmeier's scheme's properties in the context of the DREM procedure. Specifically, besides the discussed implication ϕ is PE $\Rightarrow \Delta$ is PE, we estimate the lower bound of Δ as a function of the excitation characteristics of ϕ and the gain a in (2). Moreover, we show that the inverse implication holds and Δ is PE only if ϕ is PE; to the best of our knowledge, this intuitive observation was not rigorously proven before. We also study the Interval Excitation property of ϕ and show that it is preserved as well; we also provide the quantitative analysis of the resulting interval excitation of Δ.

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 $\mathbb{N}$, and the set of reals is denoted by $\mathbb{R}$. For $m, n \in \mathbb{N}$, $\overline{m,n} := \{p \in \mathbb{N} \mid m \leq p \leq n\}$ if $m \leq n$ and $\emptyset$, otherwise. The matrix I_n is the $n \times n$ identity matrix, for all $n \in \mathbb{N}$. For a vector $x \in \mathbb{R}^n$, $|x|$ denotes the Euclidean norm of x. For a function $f : \mathbb{R}_+ \to \mathbb{R}$, we say that $f \in \mathcal{L}_2$ if $\int_0^t f^2(s)ds$ 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 excited *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 [8], [9].

Definition 2. *A bounded signal $\phi : \mathbb{R}_+ \to \mathbb{R}^{n \times \ell}$ is (t_1, T, μ)-*interval excited *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 [10].

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

$$\dot{\hat{\theta}}(t) = \Gamma\phi(t)\left(y(t) - \phi^\top(t)\hat{\theta}(t)\right), \tag{7}$$

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. Otherwise, the gradient estimator is 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 slightly technical and can hardly be applied in practice.

The DREM procedure [14]: To apply the DREM procedure, we start by performing the *dynamic regressor extension* step, where the goal is to generate an LRE for the same vector of unknown parameters θ as in (1) but with a square $n \times n$ matrix regressor. To this end, we apply Kreisselmeier's regressor extension scheme given by (2), (3) with zero initial conditions; see [14] for other possible schemes. Due to the linearity, the matrices Φ and Y generated by (2) and (3), respectively, satisfy the extended LRE

$$Y(t) = \Phi(t)\theta + W(t), \forall t \in \mathbb{R}_+, \tag{8}$$

where $W:\mathbb{R}_+\to\mathbb{R}^n$ is the result of the distortion w propagation in (3).

Next, the *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 $\mathrm{adj}(A)A=\det(A)I_n$, where $\mathrm{adj}(\cdot)$ is the *adjoint* (also called *adjugate*) matrix. Multiplying (8) by $\mathrm{adj}(\Phi(t))$ on the left and setting $\mathcal{Y}(t):=\mathrm{adj}(\Phi(t))Y(t)$, $\mathcal{W}(t):=\mathrm{adj}(\Phi(t))W(t)$, $\forall t\in\mathbb{R}_+$, we get

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

where $\mathcal{Y}_i$, θ_i, and $\mathcal{W}_i$ are i-th elements of the vectors $\mathcal{Y}$, θ, and $\mathcal{W}$, respectively, $i\in\overline{1,n}$, and the scalar function $\Delta:\mathbb{R}_+\to\mathbb{R}$ is defined in (5). 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 (9) sharing the same bounded scalar regressor Δ is the main result of the DREM procedure. Applying the gradient estimator to (9) as

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

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 $t_b\geq t_a\geq 0$ it holds $|\tilde{\theta}_i(t_b)|\leq|\tilde{\theta}_i(t_a)|$;
- 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 (10) 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.

To get the exponential convergence with the DREM procedure, the PE property of Δ must be satisfied. As discussed above, it motivates the use of Kreisselmeier's regressor extension scheme (2), (3) yielding the implication (6). However, to evaluate the exponential convergence rate or the input-to-state gain with respect to $\mathcal{W}_i$, one has to estimate the characteristics of the new regressor Δ. These characteristics depend on the excitation of ϕ and the parameter a in (2); in the following section, we analyze this dependence.

Remark 1. *Let us remark that the computation of the adjoint matrix* $\mathrm{adj}(\Phi)$ *can be avoided. So the elements* $\mathcal{Y}_i$ *in* (9) *can be computed using Cramer's rule as*

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

where $\Phi_{Y,i}$ *is the matrix* Φ *where the ith column is replaced by the vector* Y, *and* $i\in\overline{1,n}$. *This can be beneficial when only some elements of the parameter vector* θ *are estimated.*

III. The excitation-preserving regressor extension

A. Excitation propagation

Consider Kreisselmeier's regressor extension (2), (3). The applicability of (2), (3) for the DREM procedure and the implication (6) can be derived from the proof of Theorem 4.3.3 in [3]. However, in that theorem, only the lower bound's positiveness on the smallest eigenvalue of the matrix Φ is established. Extending that result, we present the following theorem providing precise lower bounds for the determinant of the matrix Φ. We also show that the inverse implication in (6) holds supporting a somewhat intuitive observation that the dynamic extension (2) does not create new excitation.

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* (2) *for some initial value* $\Phi(0)=\Phi_0\geq 0$. *Let* $\Delta:\mathbb{R}_+\to\mathbb{R}$ *be the determinant of* Φ. *Then if* ϕ *is* (T,μ)*-PE, then for any positive integer* $q\geq 1$ *and for all* $t\geq qT$, *it holds*

$$\Delta(t)\geq\mu^n\left(\sum_{k=1}^{q}e^{-akT}\right)^n \tag{11}$$

and

$$\liminf_{t\to\infty}\Delta(t)\geq\left(\frac{\mu}{e^{aT}-1}\right)^n. \tag{12}$$

Moreover, the following implication holds

$$\phi\in\mathrm{PE}\Leftrightarrow\Delta\in\mathrm{PE}. \tag{13}$$

Proof: The proof consists of two parts. First, we show that if ϕ is (T,μ)-PE, then the inequalities (11) and (12) hold proving the direct implication in (13). Next, we show that the inverse implication in (13) also holds.

Part 1: $\phi\in$ PE implies the inequalities (11), (12) and $\Delta\in$ PE.

The solution of (2) is given by

$$\Phi(t)=e^{-at}\Phi(0)+\int_0^t\psi(t,s)ds,\ \forall t\in\mathbb{R}_+, \tag{14}$$

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 (14) 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

$$\Phi(t)\geq\mu\sum_{k=1}^{q}e^{-akT}I_n+\int_0^{t-qT}\psi(t,s)ds+e^{-at}\Phi(0). \tag{15}$$

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 (15) 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 (11) 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 (12), 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 (12) for $\Delta(t)$ follows.

Part 2: $\Delta \in$ PE $\Rightarrow \phi \in$ 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 (14) 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

$$\int_t^{t+T} \lambda_m(\Phi(s))\, ds \geq \frac{\mu}{c}. \tag{16}$$

From (14) 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 (16), 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

$$\int_t^{t+kT} \lambda_m \left(\int_t^s e^{-a(s-\tau)} \phi(\tau)\phi^\top(\tau) d\tau \right) ds \geq k\frac{\mu}{c} - c_0. \tag{17}$$

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

$$\begin{aligned} & \int_t^{t+kT} \lambda_m \left(\int_t^s e^{-a(s-\tau)} \phi(\tau)\phi^\top(\tau) d\tau \right) ds \\ \leq\ & \int_t^{t+kT} \lambda_m \left(\int_t^s \phi(\tau)\phi^\top(\tau) d\tau \right) ds \\ \leq\ & \int_t^{t+kT} \lambda_m \left(\int_t^{t+kT} \phi(\tau)\phi^\top(\tau) d\tau \right) ds \\ =\ & kT\lambda_m \left(\int_t^{t+kT} \phi(\tau)\phi^\top(\tau) d\tau \right). \end{aligned} \tag{18}$$

Finally, combining (17) and (18) 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 $k > \frac{c_0 c}{\mu}$. Then $\phi(t)$ is PE with $L = kT$ and $\alpha = \frac{\mu}{cT} - \frac{c_0}{kT} > 0$. ■

B. Interval Excitation

In the vein of Theorem 1, it can be also shown that the dynamic extension (2) 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* (2) *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 excited for some* $t_1 \geq 0$, $T > 0$, *and* $\mu > 0$, *then the signal* Δ *is* (t_1, T, α)*-interval excited for some* $\alpha > 0$.

Proof: As it is discussed in the proof of Theorem 1, the solution of (2) is given by (14). 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

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

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

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

[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.

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

$$\alpha:=\left(\frac{\rho}{a^2}e^{-aT}\left(e^{\frac{a\mu}{\rho}}-1-\frac{a\mu}{\rho}\right)\right)^{2n}>0. \tag{19}$$

■

Theorem 1 and Proposition 1 justify using (2), (3) 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. The obtained bounds (11), (12), and (19) allow for performance evaluation of the DREM-enhanced estimation algorithms, e.g., the convergence rate estimation.

IV. Dynamics of the regressor Δ

In Theorem 1, we provide useful lower bounds (11), (12) for the DREM-generated new regressor Δ when Kreisselmeier's regressor extension scheme (2) is used. However, it is also of interest to study the dynamics of Δ. In this section, we present such a study.

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 (2), so Δ obeys Jacobi's formula (see Theorem 8.1 in [29]):

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

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

Substituting (2), we obtain for all $t\in\mathbb{R}_+$:

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

where we recall that the dimension of ϕ is $n\times\ell$. Let $\phi_k\in\mathbb{R}^n$ be the k-th column of ϕ, $k\in\overline{1,\ell}$. Due to elementary properties of the matrix trace function, it follows for all $t\geq 0$,

$$\operatorname{tr}\big(\operatorname{adj}(\Phi(t))\,\phi(t)\phi(t)^\top\big)=\sum_{k=1}^{\ell}\phi_k(t)^\top\operatorname{adj}(\Phi(t))\,\phi_k(t),$$

and we obtain

$$\dot{\Delta}(t)=-an\Delta(t)+\sum_{k=1}^{\ell}\phi_k(t)^\top\operatorname{adj}(\Phi(t))\,\phi_k(t). \tag{20}$$

Recall that the eigenvalues of an adjoint matrix can be estimated as follows. Let $\lambda_{1,\Phi},\ldots,\lambda_{n,\Phi}$ 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

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

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

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

In particular

$$\min_i\lambda_{i,\operatorname{adj}(\Phi)}\max_i\lambda_{i,\Phi}=\det(\Phi). \tag{21}$$

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 $t\geq 0$ and all $k\in\overline{1,\ell}$,

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

On the other hand, if $\lambda_M(t)>0$, then due to (21)

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

Recall that for the induced matrix norm $\|\phi\|$ it holds $\sum_{k=1}^{\ell}|\phi_k|^2\geq\|\phi\|^2$. Then the derivation above yields the following proposition.

Proposition 2. *Let Φ be a solution of* (2) *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

$$\begin{aligned}y_1(t)&=B+A\sin(t+\psi)+w_1(t),\\ y_2(t)&=A\cos(2t+\psi)+w_2(t)\end{aligned}$$

where the scalars B, $A>0$, and $\psi\in[-\pi,\pi)$ are the unknown parameters, and w_1, w_2 are the measurement distortions. These signals can be rewritten as the LRE (1) with

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

$$\phi(t)=\begin{bmatrix}1 & \sin(t) & \cos(t)\\ 0 & \cos(2t) & -\sin(2t)\end{bmatrix}^\top,\ w=\begin{bmatrix}w_1\\ w_2\end{bmatrix}.$$

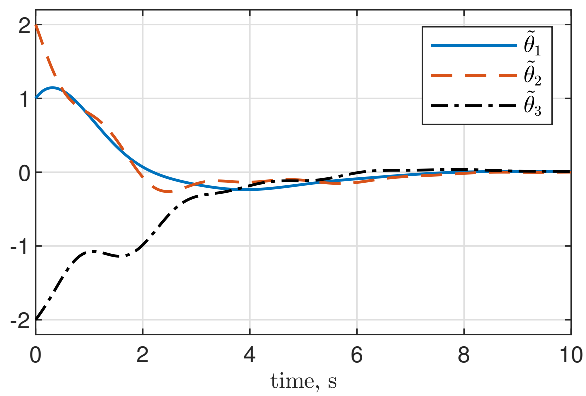

Fig. 1. The estimation error $\tilde{\theta}$ under the standard gradient algorithm (7).

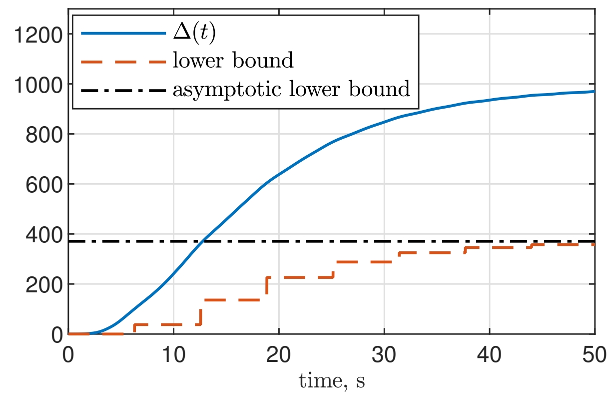

Fig. 2. The new regressor Δ, the lower bound (11), and the asymptotic lower bound (12).

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 consider the noise-free scenario, i.e., $w \equiv 0$. We apply the standard gradient algorithm (7) 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 (2), (3), 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 (11) and with the asymptotic lower bound (12). Finally, the estimation error transient $\tilde{\theta}$ for the DREM procedure with the gradient estimator (10) with $\gamma_i = 1$, $i = 1, 2, 3$, are depicted in Fig. 3 and illustrate performance improvement for the standard gradient estimator; note the monotonicity of the transients.

Next, to evaluate the measurement noise sensitivity, we let both w_1 and w_2 be random uniformly distributed between -1 and 1. For results comparison, we compute the mean squared error value,

$$MSE(\tilde{\theta}_i) = \frac{1}{50}\int_{50}^{100} \tilde{\theta}_i^2(s)ds, \ i = 1, 2, 3.$$

The results are summarized in Table I and illustrate the filtering properties of Kreisselmeier's regressor extension scheme.

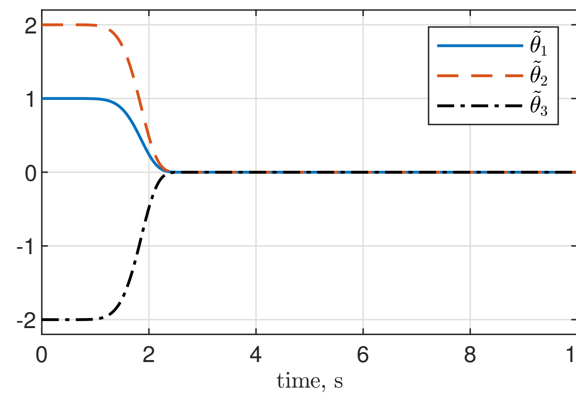

Fig. 3. The estimation error $\tilde{\theta}$ under the DREM procedure (2), (3) and the gradient estimator (10).

TABLE I
MSE OF ESTIMATION ERROR $\tilde{\theta}_i$, $i = 1, 2, 3$, FOR THE STANDARD GRADIENT ALGORITHM (7) AND THE DREM (2), (3), (10).

	$\tilde{\theta}_1$	$\tilde{\theta}_2$	$\tilde{\theta}_3$
MSE$\cdot 10^5$ for (7)	12.6	18.8	8.3
MSE$\cdot 10^5$ for (2), (3), (10)	0.85	0.44	0.44

VI. Conclusion

The recently proposed DREM procedure provides significant performance improvement in linear regression parameter estimation, where the principal degree of freedom is the choice of the dynamic extension. In this paper, we have studied a particular choice of this procedure as Kreisselmeier's regressor extension (2), (3) that ensures the preservation of the persistence of excitation property. We have analyzed the excitation propagation properties and estimated the lower bound of the new regressor Δ. We have also shown that the interval excitation property is also preserved and provided quantitative analysis of the new regressor's interval excitation. The obtained results are valuable for further use of the DREM procedure for linear regression estimation.

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] E. Lavretsky, "Combined/composite model reference adaptive control," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2692–2697, 2009.

[5] 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.

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

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

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

[13] 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.

[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. A. Pyrkin, A. Astolfi, and A. A. Bobtsov, "New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2265–2272, 2020.

[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 *Proc. 21st IFAC World Congress*, Berlin, Germany, 2020.

[25] R. Ortega, V. Nikiforov, and D. Gerasimov, "On modified parameter estimators for identification and adaptive control. a unified framework and some new schemes," *Annual Reviews in Control*, vol. 50, pp. 278–293, 2020.

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

[27] D. Efimov, N. Barabanov, and R. Ortega, "Robustness of linear time-varying systems with relaxed excitation," *Int. J. Adapt. Control Signal Process*, vol. 33, pp. 1885–1900, 2019.

[28] J. Wang, D. Efimov, and A. A. Bobtsov, "On robust parameter estimation in finite-time without persistence of excitation," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. 1731–1738, 2019.

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