Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

On key properties of the Lion's and Kreisselmeier's adaptation algorithms ⋆

Dmitry Gerasimov * **Vladimir Nikiforov** *

* *ITMO University, 197101, Kronverkskiy av. 49., Saint-Petersburg, Russia (e-mail: dngerasimov@corp.ifmo.ru, nikiforov_vo@itmo.ru).*

Abstract: The paper revises properties of two identification/adaptation algorithms proposed by Lion (1967) and Kreisselmeier (1977) more than 40 years ago to accelerate parametric convergence under regressor persistency of excitation (PE) condition. First, being motivated by paper Aranovskiy et al. (2017) it is demonstrated that these algorithms can provide asymptotic (not exponential) parametric convergence under simple condition which is weaker than requirement of PE. Second, it is shown that via some condition these schemes can be used for generating the high order time derivatives (HOTD) of the adjustable parameters that are necessary for solution of a wide range of problems of identification and adaptive control including backstepping design procedure.

Keywords: Adaptive identification, Adaptive Control, Persistent Excitation Condition.

1. INTRODUCTION

In spite of the fact that the identification/adaptation algorithms proposed by Lion (1967) and Kreisselmeier (1977) are well known and widely used during last decades, in the authors' opinion these algorithms have some unrevealed properties which can be used in various practical applications. The first property consists in asymptotic (not exponential) convergence under simple condition which is weaker than requirement of *persistency of excitation* (PE). The second property is the ability to calculate the high order time derivative (HOTD) of the adjustable parameters that can be used, for example, in procedures of adaptive backstepping.

The impetus for study of the first issue was inspired by the recently published paper of Aranovskiy et al. (2017), in which an identification algorithm called *dynamic regressor extension and mixing* (DREM) was proposed. For this algorithm the authors established a simple condition for asymptotic (not exponential) parametric convergence. This condition requires *square nonintegrability* of some signals and is weaker than usual requirement of PE. Moreover, under this condition DREM provides monotonic convergence of each element of parametric error vector. To demonstrate advantages of the proposed algorithm the authors of Aranovskiy et al. (2017) presented a numerical example in which asymptotic parametric convergence is provided for a regressor which is not only not PE, but also has components approaching zero. Effective applications of DREM to different identification and control problems were demonstrated in Ortega et al. (2018, 2019a); Gerasimov et al. (2018); Ortega et al. (2019b).

In the present paper we demonstrate that similar asymptotic parametric convergence under weak condition can be demonstrated for the Lion's and Kreisselmeier's algorithms (Lion (1967); Kreisselmeier (1977)). Initially the Lion's and Kreisselmeier's algorithms were proposed to improve parametric convergence under PE condition. Namely, it was proved that under PE condition the rate of parametric convergence can be made *arbitrary fast* by increasing adaptation gain. However, simulation evidence and some intuitive reasoning showed that the Lion's and Kreisselmeier's schemes improved parametric convergence not only under PE condition. Now we can rigorously explain this phenomenon by asymptotic (not exponential) parametric convergence under some weaker condition.

The second issue was motivated by an idea of model reference adaptive control with high order tuner presented in Itamiya et al. (1999). In accordance with this idea, analytical expressions for each HOTD of adjustable parameters are derived directly from the Kreisselmeier's scheme. Unfortunately, since these expressions are used for calculation of each HOTD independently, the general order of the adaptation algorithm increases dramatically. As a result, in Gerasimov et al. (2020a,b) the authors, in the framework of backstepping procedures, proposed the closed-loop modification of the Kreisselmeier's scheme without increasing the order. In this paper we demonstrate three modified algorithms based on the Kreisselmeier's and the Lion's schemes and used for calculation the HOTD of adjustable parameters.

The paper is organized as follows. In Section 2 we briefly summarize the main properties of the algorithms considered. Section 3 contains the main result concerning asymptotic (alas, not exponential) convergence of the considered algorithms and their new closed-loop representation generating the HOTD of adjustable parameters. Section 4 presents some simulation results. The conclusion is given in Section 5.

⋆ This work was financially supported by Goverment of Russian Federation (Grant 08-08).

2. PRELIMINARIES: BASIC ADAPTATION ALGORITHMS

2.1 Regression model

Consider the static regression model given by

$$\varepsilon = \omega^\top \tilde{\theta} + \sigma, \tag{1}$$

where $\varepsilon \in \mathbb{R}$ is the measurable control error, $\omega \in \mathbb{R}^n$ is the bounded regressor, $\tilde{\theta} = \hat{\theta} - \theta$ is the *parametric error*, $\hat{\theta}$ is a vector of adjustable control parameters, θ is a constant vector of unknown "true" values for the adjustable parameters, σ is an exponentially decaying term due to nonzero initial conditions [1]. For example, the regression (1) can represent the augmented error in the problem of MRAC or identification error in corresponding problems of identification (Ioannou and Sun (1996); Narendra and Annaswamy (1989); Sastry and Bodson (1989)).

We also assume that we can measure the *output variable*

$$y = \omega^\top \theta \tag{2}$$

or can directly calculated it as $y = \omega^\top \hat{\theta} - \varepsilon$.

We will use the following definition.

Definition 1. *A bounded signal vector $\omega \in \mathbb{R}^n$ is persistently exciting ($\omega \in PE$) if there exist $T > 0$ and $\alpha > 0$ such that*

$$\int_t^{t+T} \omega(\tau)\omega^\top(\tau)d\tau \;\; \succeq \;\; \alpha I \;\; for \;\; \forall t \; \geq 0.$$

It is well known that standard gradient-based algorithm

$$\dot{\hat{\theta}} = -\gamma\omega\varepsilon, \tag{3}$$

where $\gamma > 0$ is adaptation gain, yields the following closed-loop *parametric error model*

$$\dot{\tilde{\theta}} = -\gamma\omega\omega^\top \tilde{\theta} \tag{4}$$

and ensures the following properties (Ioannou and Sun (1996); Narendra and Annaswamy (1989)).

Proposition 1. Consider the regression model (1) and the estimator (3).

- P1.1 For any bounded $\omega, \dot{\omega}$, the signals $\hat{\theta}$ and ε are bounded and $\varepsilon \to 0$ as $t \to \infty$.
- P1.2 If $\omega \in PE$ then the norm $||\tilde{\theta}||$ tends to zero exponentially and there exists an optimal value of γ for which the rate of convergence is maximum (see Subsection 4.2.1 in Narendra and Annaswamy (1987)).

It is worth noting that in the recent paper of Barabanov and Ortega (2017) new necessary conditions for global asymptotic stability of (4) as well as a new sufficient (but not necessary) condition, that is strictly weaker than the ones reported in literature, were established. Unfortunately, these conditions are highly technical and hard to verify in practical examples.

2.2 Lion's scheme

In Lion (1967) it was proposed the simple idea to extend the dimension of the error and regressor in order to improve parametric convergence. To derive such an algorithm, $n-1$ different LTI causal dynamic operators $\mathcal{H}_i\{\cdot\}$, $i = \overline{1, n-1}$ ($\mathcal{H}_i\{\cdot\} : \mathcal{L}_\infty \to \mathcal{L}_\infty$) are introduced. In particular case, the operators $\mathcal{H}_i\{\cdot\}$ can be selected as the first order filters:

$$\mathcal{H}_i\{\cdot\} = \frac{\alpha_i}{s + \beta_i}\{\cdot\}, \tag{5}$$

where $\beta_j > 0$, $\beta_i \neq \beta_j$ for $i \neq j$, and $s = d/dt$ is the differential operator.

Following the terminology proposed in Aranovskiy et al. (2017); Ortega et al. (2020), Lion's scheme generates a *dynamic regressor extension* (DRE) as

$$Y_d \triangleq [y,\ \mathcal{H}_1\{y\},\ \mathcal{H}_2\{y\}, \ldots,\ \mathcal{H}_{l-1}\{y\}]^\top, \tag{6}$$

$$W^\top \triangleq [\omega,\ \mathcal{H}_1\{\omega\},\ \mathcal{H}_2\{\omega\}, \ldots,\ \mathcal{H}_{l-1}\{\omega\}]^\top. \tag{7}$$

It is obvious that in this case

$$Y_d = W^\top \theta. \tag{8}$$

This motivates the following form of *estimation error*:

$$E \triangleq W^\top \hat{\theta} - Y_d = W^\top \tilde{\theta}.$$

Then, the gradient-based algorithm with dynamically extended regressor

$$\dot{\hat{\theta}} = -\gamma W E \tag{9}$$

gives the following closed-loop error model

$$\dot{\tilde{\theta}} = -\gamma W W^\top \tilde{\theta}. \tag{10}$$

It can be shown that the parametric error equation (10) has the following properties of convergence.

Proposition 2. Consider the regression model (2) and the adaptation algorithm (9).

- P2.1 For any bounded ω, $\dot{\omega}$, the signals $\hat{\theta}$ and $||E||$ are bounded and $||E|| \to 0$, $\varepsilon \to 0$ as $t \to \infty$.
- P2.2 If $\omega \in PE$, then the norm $||\tilde{\theta}||$ tends to zero exponentially.
- P2.3 If $\omega \in PE$, then additionally to P2.2 the rate of parametric convergence can be made arbitrarily fast by increasing γ.

It is worth noting that the property P2.3 cannot be ensured by the algorithm (3).

Remark 1. Strictly speaking, due to nonzero initial conditions exciting the exponentially decaying term σ in (1) *arbitrarily fast* convergence is achieved in theory after some transient period only.

2.3 Kreisselmeier's scheme

An alternative approach to improve the estimator convergence was proposed by Kreisselmeier initially for design of adaptive observers (see Kreisselmeier (1977)) and later for MRAC (see Kreisselmeier and Joos (1982)). In accordance with this approach we introduce a single LTI causal dynamic operator $\mathcal{L}\{\cdot\}$ such that $\mathcal{L}\{\omega\omega^\top\}$ is bounded for any bounded ω and is positive semidefinite. Then we apply $\mathcal{L}\{\cdot\}$ to obtain alternatively extended regression model

$$Y_m = \Omega\theta, \tag{11}$$

where $Y_m \triangleq \mathcal{L}\{\omega\omega^\top \hat{\theta} - \omega\varepsilon\}$ or $Y_m \triangleq \mathcal{L}\{\omega y\}$ is the extended output and

$$\Omega \triangleq \mathcal{L}\{\omega\omega^\top\}. \tag{12}$$

[1] If ω is bounded, the term σ being the result of nonzero initial conditions does not influence the stability of closed-loop system and is usually neglected. However, as it is discussed in Remark 1, this term can deteriorate overall transient performance.

Regression (11) motivates the following gradient-based adaptation algorithm

$$\dot{\hat{\theta}} = -\gamma \left(\Omega\hat{\theta} - Y_m \right), \tag{13}$$

where $\gamma > 0$ is the adaptation gain.

The main distinguishing feature of the algorithm is to use past history of matrix $\omega\omega^\top$ provided by the operator $\mathcal{L}\{\cdot\}$ to improve transient performance of $\hat{\theta}$. In this context, following the terminology proposed in Gerasimov et al. (2019); Ortega et al. (2020) we can say that $\mathcal{L}\{\cdot\}$ is the operator with "memory", and hence the Kreisselmeier's scheme provides *memory regressor extension* (MRE). In particular, the operator $\mathcal{L}\{\cdot\}$ can be selected as the simplest first order filter:

$$\mathcal{L}\{\cdot\} = \frac{1}{s+1}\{\cdot\}. \tag{14}$$

It can be shown that the estimator (13) yields the parametric error model

$$\dot{\tilde{\theta}} = -\gamma\Omega\tilde{\theta}. \tag{15}$$

It can be shown that the estimator (13) ensures the following convergence properties.

Proposition 3. For the error model (1) the algorithm (13) ensures the following properties:

P3.1 For any bounded ω the signals $\hat{\theta}$ and ε are bounded and $\varepsilon \to 0$ as $t \to \infty$.
P3.2 If $\omega \in PE$ then the norm $||\tilde{\theta}||$ tends to zero exponentially.
P3.3 If $\omega \in PE$ then additionally to P3.2 the rate of parametric convergence can be made arbitrary fast by increasing γ.

More detailed discussion on estimators with different LTI filters $\mathcal{L}\{\cdot\}$ can be found in Section 4.4.5 of Narendra and Annaswamy (1989).

2.4 DREM

Recently, new scheme with improved parametric convergence was proposed in Aranovskiy et al. (2017). This scheme can be obtained from (9) with the following choice of time-varying matrix adaptation gain γ:

$$\gamma = \gamma_0\, adj\{W^\top\} adj\{W\} \succeq O_{n\times n}, \tag{16}$$

where adj$\{\cdot\}$ is the adjunct matrix, $\gamma_0 \in \mathbb{R}_+$ is a constant. Replacing this coefficient in (9) with the use of definition of E and the identity $adj\{W\}W = det\{W\}I_m$ we get

$$\dot{\hat{\theta}} = \gamma_0\delta(\delta\hat{\theta} - Y), \tag{17}$$

where $Y \triangleq adj\{W^\top\}Y_d$, $\delta \triangleq \det\{W\}$. Since δ is a scalar, algorithm (17) can be rewritten in the element-wise form

$$\dot{\hat{\theta}}_i = -\gamma_0\delta \left(\delta\hat{\theta}_i - Y_i \right) \tag{18}$$

that yields the set of scalar error models

$$\dot{\tilde{\theta}}_i = -\gamma_0\delta^2\tilde{\theta}_i, \qquad i = \overline{1,n}. \tag{19}$$

The algorithm (18) is called DREM.

In Aranovskiy et al. (2017) based on the model (19) the following properties are proved.

Proposition 4. For regression (1) the estimator (18) ensures the following:

P4.1 $\delta(t) \notin \mathcal{L}_2 \quad \Leftrightarrow \quad \lim_{t\to\infty} \tilde{\theta}(t) = 0$ and $\tilde{\theta}_i(t)$ tends to zero monotonically.
P4.2 If $\omega \in PE$ the parametric errors $\tilde{\theta}_i$ tend to zero exponentially. Moreover, the rate of convergence can be made arbitrary fast by increasing γ_0.

2.5 Memory regressor extension and mixing

In the recent paper of Ortega et al. (2020) inspired by the idea of the algorithm with DREM it was proposed to apply "regressor mixing" to Kreisselmeier's algorithm (13). That is, to select for (13) time-varying adaptation gain

$$\gamma(t) = \gamma_0\, adj\{\Omega\} \succeq O_{n\times n}. \tag{20}$$

By applying identity $adj\{\Omega\}\Omega = det\{\Omega\}I_n$ we get the adaptation algorithm

$$\dot{\hat{\theta}} = -\gamma\, \text{adj}\{\Omega\} \left(\Omega\hat{\theta} - Y \right) \tag{21}$$

that yields the following closed-loop error model:

$$\dot{\tilde{\theta}}_i = -\gamma\delta_\Omega\tilde{\theta}_i, \qquad i = 1, \ldots, n, \tag{22}$$

where $\delta_\Omega \triangleq \det\{\Omega\}$. Notice that, due to the definition of the matrix Ω in (12), $\delta_\Omega(t) \geq 0$.

Using terminology proposed in Ortega et al. (2020) we refer to the estimator (21) as *memory regressor extension and mixing* (MREM).

The convergence properties of MREM are established by Proposition 4 with replacement of condition $\delta(t) \notin \mathcal{L}_2$ by $\delta_\Omega(t) \notin \mathcal{L}_1$.

3. MAIN RESULT: PROPERTIES OF DRE AND MRE

In this section we present and discuss *asymptotic* properties of the schemes with MRE and DRE (i.e., Kreisselmeir's and Lion's schemes) and show that with some choice of operators $\mathcal{L}\{\cdot\}$ and $\mathcal{H}_i\{\cdot\}$ these schemes can generate the HOTD of adjustable parameters $\hat{\theta}$.

3.1 Asymptotic convergence

One interesting contribution of the paper of Aranovskiy et al. (2017) is the answer to the question: is the condition P4.1 ($\delta(t) \notin \mathcal{L}_2$) of DREM weaker or stronger than $\omega(t) \in PE$? To answer this question the authors give an example of ω satisfying P4.1 and at the same time is not PE. They consider particular case with $n = 2$ and

$$\omega = [1 \;\; g + \dot{g}]^\top, \tag{23}$$

where g is a continuous function such that $g, \dot{g} \in \mathcal{L}_\infty$, $\dot{g} \notin \mathcal{L}_2$ and

$$\lim_{t\to\infty} g(t) = \lim_{t\to\infty} \dot{g}(t) = 0. \tag{24}$$

It is obvious that in this case $\omega \notin PE$.

For example, function $g(t) = \sin(t)(1+t)^{-\frac{1}{2}}$ satisfies these conditions and yields the following regressor

$$\omega(t) = \left[1 \qquad \frac{\sin(t) + \cos(t)}{(1+t)^{\frac{1}{2}}} - \frac{\sin(t)}{2(1+t)^{\frac{3}{2}}} \right]^\top. \tag{25}$$

It was proved that if $\mathcal{H}_1\{\cdot\} = \frac{1}{s+1}\{\cdot\}$ and $W^\top = [\omega,\ \mathcal{H}_1\{\omega\}]^\top$, then

$$\delta(t) = -\dot{g}(t) \notin \mathcal{L}_2.$$

In this section, we use this numerical examples to illustrate the asymptotic properties of Lion's and Kreisselmeier's algorithms.

Kreisselmeier's scheme (MRE). Before we consider a particular example, we note that parametric error model (15) yields the following differential inequality:

$$\frac{d}{dt}||\tilde{\theta}||^2 = -2\gamma\tilde{\theta}^\top\Omega\tilde{\theta} \leq -2\gamma\lambda_{min}(t)||\tilde{\theta}||^2, \qquad (26)$$

where $\lambda_{min}(t) \geq 0$ is the minimum eigenvalue of matrix $\Omega(t)$. Solution of (26) gives the upper bound of the squared norm $||\tilde{\theta}(t)||$:

$$||\tilde{\theta}(t)||^2 \leq e^{-2\gamma\int_0^t \lambda_{min}(\tau)d\tau}||\tilde{\theta}(0)||^2. \qquad (27)$$

Later gives the condition for asymptotic parametric convergence.

Proposition 5. Additionally to Proposition 3, algorithm (13) provides

P3.4 Asymptotic convergence $||\tilde{\theta}|| \to 0$ and $\varepsilon \to 0$ as $t \to \infty$, if $\lambda_{min}(t) \notin \mathcal{L}_1$.

Does the regressor (25) ensure this condition? In view of (23) we have

$$\omega\omega^\top = \begin{bmatrix} 1 & g+\dot{g} \\ g+\dot{g} & (g+\dot{g})^2 \end{bmatrix}.$$

Then invoking operator (14) and neglecting exponentially vanishing terms we calculate each element of the matrix Ω:

$$\Omega_{11} = \mathcal{L}\{1\} = \frac{1}{s+1}\{1\} = 1,$$

$$\Omega_{21} = \Omega_{12} = \mathcal{L}\{g+\dot{g}\} = \frac{1}{s+1}\{g+\dot{g}\} = \frac{1}{s+1}(s+1)\{g\} = g,$$

$$\Omega_{22} = \mathcal{L}\{(g+\dot{g})^2\} = \frac{1}{s+1}\{g^2+2g\dot{g}+\dot{g}^2\} = \frac{1}{s+1}(s+1)\{g^2\} + \frac{1}{s+1}\{\dot{g}^2\} = g^2 + \dot{g}_H^2,$$

where

$$\dot{g}_H^2 = \frac{1}{s+1}\{\dot{g}^2\}.$$

Thus, we obtain that

$$\Omega = \mathcal{L}\{\omega\omega^\top\} = \begin{bmatrix} 1 & g \\ g & g^2+\dot{g}_H^2 \end{bmatrix}.$$

Consider the maximum eigenvalue and the determinant of Ω given by

$$\lambda_{max}(t) = \frac{1}{2}\left(1+g^2(t)+\dot{g}_H^2(t)\right) + \frac{1}{2}\sqrt{\left(1+g^2(t)+\dot{g}_H^2(t)\right)^2 - 4\dot{g}_H^2(t)}$$

and

$$\delta_\Omega(t) = \det\Omega(t) = \lambda_{min}(t)\lambda_{max}(t) = \dot{g}_H^2(t),$$

respectively.

In order to verify the property $\lambda_{min}(t) \notin \mathcal{L}_1$ we establish the following facts:

- Since $g^2(t)$ and $\dot{g}_H^2(t)$ are nonnegative functions,
$$\lambda_{max} = \frac{1}{2}\left(1+g^2(t)+\dot{g}_H^2(t)\right) + \frac{1}{2}\sqrt{\left(1+g^2(t)+\dot{g}_H^2(t)\right)^2 - 4\dot{g}_H^2(t)} \leq \frac{1}{2}\left(1+g^2(t)+\dot{g}_H^2(t)\right) + \frac{1}{2}\sqrt{\left(1+g^2(t)+\dot{g}_H^2(t)\right)^2} = 1+g^2(t)+\dot{g}_H^2(t).$$
- From the last inequality we get
$$\lambda_{min} = \frac{\delta}{\lambda_{max}} = \frac{\dot{g}_H^2}{\lambda_{max}} \geq \frac{\dot{g}_H^2}{1+g^2(t)+\dot{g}_H^2(t)} \geq \frac{\dot{g}_H^2}{1+sup(g^2(t))+sup(\dot{g}_H^2(t))} = \frac{\dot{g}_H^2}{C},$$
where $C > 0$ is a constant, $sup(\cdot)$ is the supremum.
- Since $\dot{g} \notin \mathcal{L}_2$ and the operator $\mathcal{L}\{\cdot\}$ is stable ($\mathcal{L}_\infty \to \mathcal{L}_\infty$), it is seen from the last inequality and the definition of $\dot{g}_H^2$ that the eigenvalue $\lambda_{min} \notin \mathcal{L}_1$, i.e.
$$\lim_{t\to\infty}\int_0^t \lambda_{min}(\tau)d\tau = \infty.$$

Thus, we have demonstrated that the regressor (25) used in Aranovskiy et al. (2017) to illustrate asymptotic convergence of DREM without PE also provides asymptotic convergence for the Kreisselmeier's scheme.

Lion's scheme (DRE) Now consider the Lion's algorithm (9). It is easy to show that algorithm (9) ensures asymptotic convergence of the norm of parametric error if $\lambda_W \notin \mathcal{L}_1$, where $\lambda_W(t)$ is the minimum eigenvalue of $WW^\top$.

Proposition 6. For regression (1) algorithm (9) provides:

P2.4 Asymptotic convergence $||\tilde{\theta}|| \to 0$ as $t \to \infty$ if $\lambda_W(t) \notin \mathcal{L}_1$.

This proposition can be illustrated again by the example above taken from Aranovskiy et al. (2017) and applied to the algorithm (9). For the regressor (23) and $\mathcal{H}_1\{\cdot\} = \frac{1}{1+s}\{\cdot\}$, we (neglecting exponentially vanishing transiences) can show that

$$W^\top = \begin{bmatrix} 1 & g+\dot{g} \\ 1 & g \end{bmatrix}, \qquad \lambda_W(t) \geq \dot{g}^2, \qquad \det WW^\top = \dot{g}^2.$$

Since $\dot{g} \notin \mathcal{L}_2$, we have $\lambda_W \notin \mathcal{L}_1$.

3.2 High order tuners generating the HOTD of estimates

Now, we show that the Lion's and Kreisselmeier's algorithms can be represented in the special forms allowing one to obtain the HOTD of the adjustable parameters $\hat{\theta}$ up to some prescribed order.

Kreisselmeier's scheme. For the algorithm (13) we introduce the operator

$$\mathcal{L}\{\cdot\} \triangleq L(s)\{\cdot\} = \frac{1}{d(s)}\{\cdot\}, \qquad (28)$$

where $d(s) = s^q + d_{q-1}s^{q-1} + \cdots + d_0$ is a Hurwitz polynomial of the order q. Now, taking into account the structure of this operator we demonstrate three ways of calculation of the HOTD of $\hat{\theta}$.

1. *Scheme with direct differentiation* is obtained by differentiation of (13):

$$\hat{\theta}^{(k+1)} = -\gamma\left(\sum_{j=0}^{k} C_j^k \Omega^{(k-j)}\hat{\theta}^{(j)} - Y_m^{(k)}\right),\ k=\overline{0,q}, \quad (29)$$

where C_j^k are the binomial coefficients,

$$\Omega^{(k-j)} = \frac{s^{k-j}}{d(s)}\{\omega\omega^\top\}, Y_m^{(k)} = \frac{s^k}{d(s)}\{\omega y\}.$$

2. *Closed-loop scheme #1 with swapping.* We represent the term $\Omega\hat{\theta}$ in (13) by applying the swapping lemma (see *Appendix A* of Ioannou and Sun (1996)):

$$\Omega\hat{\theta} = L(s)\left\{\omega\omega^\top\right\}\hat{\theta} = \quad (30)$$
$$L(s)\left\{\omega\omega^\top\hat{\theta}\right\} + I_n \otimes L_c(s)\left\{I_n \otimes L_b(s)\{\omega\omega^\top\}\dot{\hat{\theta}}\right\},$$

where $L_c(s) = \overline{L}_c(s)/d(s)$,

$$\overline{L}_c(s) = c_L^\top adj\left(I_q s - A_L\right), L_b(s) = \left(I_q s - A_L\right)^{-1} b_L,$$

A_L, b_L, c_L are the matrices implementing the minimal realization of the transfer function $L(s) = 1/d(s) = c_L^\top(I_q s - A_L)^{-1} b_L$, $\otimes$ is the Kronecker product operator. Replacing (30) in (13) with the use of definition of Y_m and applying the operator $d(s)$ to the both parts of expression obtained we get:

$$\hat{\theta}^{(q+1)} + d_{q-1}\hat{\theta}^{(q)} + \cdots + d_0\dot{\hat{\theta}} + \quad (31)$$
$$\gamma\bar{L}_c(s)\left\{L_b(s)\{\omega\omega^\top\}\dot{\hat{\theta}}\right\} = -\gamma\omega\varepsilon.$$

3. *Closed-loop scheme #2 with differentiation.* The scheme is obtained by application of the operator $d(s)$ to the both parts of (13):

$$d(s)\{\dot{\hat{\theta}} + \gamma\Omega\hat{\theta}\} = \gamma\omega y, \quad (32)$$

Proposition 7. If the operator $\mathcal{L}\{\cdot\}$ is defined by (28), then additionally to Proposition 3.

P3.5 Algorithm (13) can be used for calculation of the HOTD of $\hat{\theta}$ up to $(q+1)$th order by one of three schemes given by (29), (31) or (32).

Lion's scheme. For the algorithm (9) we define Y_d and W as

$$Y_d = L(s)\mathcal{H}(s)\{y\}, \quad (33)$$
$$W^\top = L(s)\mathcal{H}(s)\{\omega^\top\} \quad (34)$$

with the transfer function $L(s)$ given by (28) and the $l \times 1$ transfer matrix $\mathcal{H}(s) = col(1, \mathcal{H}_1(s), \ldots, \mathcal{H}_{l-1}(s))$.

Proposition 8. For extended output (33) and extended regressor matrix (34), additionally to Proposition 2

P2.5 Algorithm (9) can be used for calculation of the HOTD of $\hat{\theta}$ up to $(q+1)$th order by one of the following three schemes:

1. *Scheme with direct differentiation :*

$$\hat{\theta}^{(k+1)} = -\gamma\left(\sum_{j=0}^{k} C_j^k W^{(k-j)} E^{(j)}\right),\ k=\overline{0,q}, \quad (35)$$

where C_j^k are the binomial coefficients,

$$(W^{(k-j)})^\top = \frac{s^{(k-j)}}{d(s)}[\omega,\ \mathcal{H}_1\{\omega\},\ \mathcal{H}_2\{\omega\}, \ldots,\ \mathcal{H}_{l-1}\{\omega\}]^\top,$$
$$E^{(j)} = \sum_{l=0}^{j} C_j^j (W^{(j-l)})^\top \hat{\theta}^{(l)} - Y_d^{(j)}, \quad (36)$$
$$Y_d^{(j)} = \frac{s^{(j)}}{d(s)}[y,\ \mathcal{H}_1\{y\},\ \mathcal{H}_2\{y\}, \ldots,\ \mathcal{H}_{l-1}\{y\}]^\top.$$

2. *Closed-loop scheme #1 with swapping:*

$$\hat{\theta}^{(q+1)} + d_{q-1}\hat{\theta}^{(q)} + \cdots + d_0\dot{\hat{\theta}} = -\gamma\left(\overline{W}E + \quad (37)\right.$$
$$\left.(I_n \otimes \overline{L}_c(s))\left\{I_n \otimes L_b(s)\left\{\overline{W}\right\}\dot{E}\right\}\right),$$

where $\overline{W}^\top = \mathcal{H}(s)\{\omega^\top\}$, $\dot{E}$ is calculated using (36) (assuming $j=1$).

3. *Closed-loop scheme #2 with differentiation:*

$$d(s)\left\{\dot{\hat{\theta}} + \gamma W W^\top\hat{\theta}\right\} = -d(s)\left\{WY_d\right\} \quad (38)$$

The proposition is proved by differentiating (9) (item (1)), applying the swapping lemma to the term WE (item (2)), applying $d(s)$ to the both parts of (9) (item(3)).

4. SIMULATION RESULTS

Consider simulation results illustrating asymptotic properties of the algorithms with DRE (9) and MRE (13). Fig.1a demonstrates the plots of $\int_0^t det\{\Omega\}(\tau)d\tau$ and $\int_0^t \lambda_{min}(\tau)d\tau$ (λ_{min} is the minimum eigenvalue of Ω). Fig.1b demonstrates $\int_0^t det^2\{W\}(\tau)d\tau$, $\int_0^t \lambda_W(\tau)d\tau$ (λ_W is the minimum eigenvalue of $WW^\top$). Matrices Ω and W are calculated for $\omega \notin PE$ in accordance with (25).

As seen from the plots of Fig.1, inspite of the fact that $\omega \notin PE$ we have: $\lambda_{min}, det\{\Omega\} \notin \mathcal{L}_1$ and $\lambda_W \notin \mathcal{L}_1, det\{W\} \notin \mathcal{L}_2$. Thus, we can assume that both the algorithms provide asymptotic parametric convergence even in the case of "bad" regressor (25).

For all the experiments we select $\theta^\top = [-3, 3]$ and $\hat{\theta}(0) = 0$. For comparison purposes Fig.3 demonstrates the function $||\tilde{\theta}(t)||$ calculated from the conventional gradient-based algorithm (3) with $\gamma = 3$ and $\gamma = 10$ (adaptation gains and initial conditions are equal to those ones used in Aranovskiy et al. (2017)).

Fig.3 shows the norm $||\tilde{\theta}(t)||$ provided by MRE algorithm (13) with operator (14), while Fig.4 shows this function given by DRE algorithm (9) with $W = \left[\omega,\ \frac{1}{s+1}\{\omega\}\right]$. In spite of the fact that $\omega \notin PE$, both figures demonstrate asymptotic parametric convergence and potentials for convergence acceleration by increasing γ.

5. CONCLUSION

Thus, in the paper we: 1) prove and demonstrate via a numerical example that the Lion's and Kreisselmeier's adaptation algorithms can ensure asymptotic (not exponential) parametric convergence under conditions weaker than the PE one; 2) show that the Lion's and Kreisselmeier's algorithms can be represented in the forms generating HOTD of adjustable parameters.

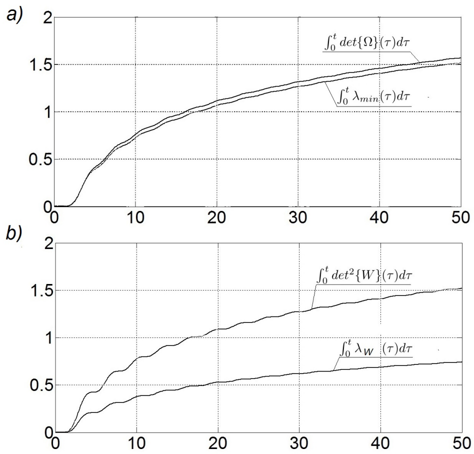

Fig. 1. Evolution of functions: a) $\int_0^t \lambda_{min} d\tau$, $\int_0^t det\{\Omega\} d\tau$; b) $\int_0^t \lambda_W d\tau$, $\int_0^t det^2\{W\} d\tau$.

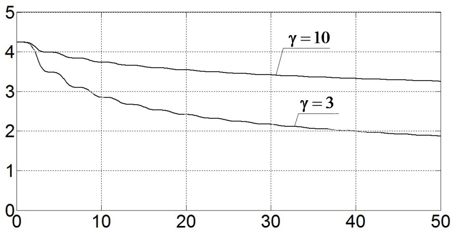

Fig. 2. The norms $||\tilde{\theta}(t)||$ provided by gradient-based algorithm (3).

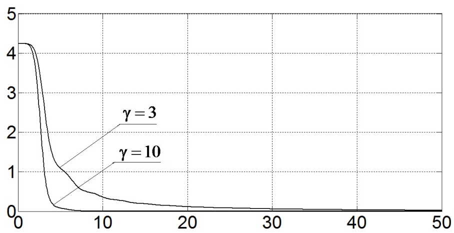

Fig. 3. The norms $||\tilde{\theta}(t)||$ provided by MRE algorithm (13).

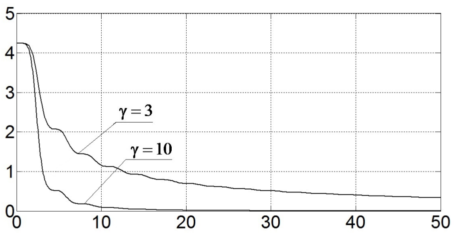

Fig. 4. The norms $||\tilde{\theta}(t)||$ provided by DRE algorithm (9).

REFERENCES

Aranovskiy, S., Bobtsov, A., Ortega, R., and Pyrkin, A. (2017). Performance enhancement of parameter estimator via dynamic regressor extension and mixing. *IEEE Trans. on Automatic Control*, 62(7), 3546 – 3550.

Barabanov, N. and Ortega, R. (2017). On global asymptotic stability of $\dot{x} = \phi(t)\phi^\top(t)x$ with $\phi(t)$ bounded and not persistently exciting. *Systems and Control Letters*, 109, 24–27.

Gerasimov, D., Ortega, R., and Nikiforov, V. (2018). Relaxing the high-frequency gain sign assumption in direct model reference adaptive control. *European J. of Control*, 12–19.

Gerasimov, D., Belyaev, M., and Nikiforov, V. (2019). Performance improvement of discrete mrac by dynamic and memory regressor extension. *18th European Control Conf. (ECC)*, Naples, Italy, 2950–2956.

Gerasimov, D., Pashenko, A., and Nikiforov, V. (2020a). Improved adaptive compensation of unmatched multi-sinusoidal disturbances in uncertain nonlinear plants. *2020 American Control Conf. (ACC)*, (accepted).

Gerasimov, D., Pashenko, A., and Nikiforov, V. (2020b). Improved adaptive servotracking for a class of nonlinear plants with unmatched uncertainties. *IFAC-PapersOnLine*, (accepted).

Ioannou, P. and Sun, J. (1996). *Robust Adaptive Control*. Prentice–Hall, NJ.

Itamiya, E., Sawada, M., and Suzuki, T. (1999). Model reference adaptive control system based on surrogate model control using plant parameter estimates. *Proc. of the 38th IEEE Conf. on Decision and Control*, 4, 3327–3328.

Kreisselmeier, G. (1977). Adaptive observers with exponential rate of convergence. *IEEE Trans. of Automatic Control*, 22(1), 2–8.

Kreisselmeier, G. and Joos, D. (1982). Rate of convergence in model reference adaptive control. *IEEE Trans. of Automatic Control*, 27, 710–713.

Lion, P. (1967). Rapid identification of linear and nonlinear systems. *AIAA J.*, 5, 1835–1842.

Narendra, K. and Annaswamy, A. (1987). Persistent excitation in adaptive systems. *Int. J. Control*, 45(1), 127–160.

Narendra, K. and Annaswamy, A. (1989). *Stable Adaptive Systems*. Prentice Hall, NJ.

Ortega, R., Gerasimov, D., Barabanov N.E., S., and Nikiforov, V. (2019a). Adaptive control of linear multivariable systems using dynamic regressor extension and mixing estimators: Removing the high-frequency gain assumptions. *Automatica*, 110, 108589. doi: https://doi.org/10.1016/j.automatica.2019.108589.

Ortega, R., Nikiforov, V., and Gerasimov, D. (2020). On modified parameter estimators for identification and adaptive control. a unified framework and some new schemes. *Annual Reviews in Control*, accepted.

Ortega, R., Praly, L., Aranovskiy, S., Yi, B., and Zhang, W. (2018). On dynamic regressor extension and mixing parameter estimators: Two Luenberger observers interpretations. *Automatica*, 95, 548–551.

Ortega, R., Aranovskiy, S., Pyrkin, A.A., Astolfi, A., and Bobtsov, A.A. (2019b). New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases. *ArXiv*, abs/1908.05125.

Sastry, S. and Bodson, M. (1989). *Adaptive Control: Stability, Convergence and Robustness*. Prentice Hall, NJ.