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

Parameter Estimation of Nonlinearly Parameterized Regressions: Application to System Identification and Adaptive Control

Romeo Ortega [*,**] **Vladislav Gromov** [**,***]
Emmanuel Nuño [****] **Anton Pyrkin** [**]
Jose Guadalupe Romero [†]

[*] *Laboratoire des Signaux et Systémes, Supélec, CNRS-SUPELEC, Gif-sur-Yvette, France (e-mail: ortega@lss.supelec.fr)*
[**] *Faculty of Control Systems and Robotics, ITMO University, Saint Petersburg, Russia, (email: gromov{pyrkin}@itmo.ru)*
[***] *Center for Technologies in Robotics and Mechatronics Components, Innopolis University, Innopolis, Russia*
[****] *Department of Computer Science, CUCEI, University of Guadalajara, Guadalajara, Mexico, (email: emmanuel.nuno@cucei.udg.mx)*
[†] *Department of Digital Systems, ITAM, Mexico City, Mexico,(e-mail: jose.romerovelazquez@itam.mx)*

Abstract: We propose a solution to the problem of parameter estimation of *nonlinearly parameterized regressions*—continuous or discrete time—and apply it for system identification and adaptive control. We restrict our attention to parameterizations that can be factorized as the product of two functions, a measurable one and a *nonlinear* function of the parameters to be estimated. Another feature of the proposed estimator is that parameter convergence is ensured *without* a persistency of excitation assumption. It is assumed that, after a coordinate change, *some of the elements* of the transformed function satisfy a *monotonicity* condition. The proposed estimators are applied to design identifiers and adaptive controllers for nonlinearly parameterized systems, which are traditionally tackled using overparameterization and assuming persistency of excitation.

Keywords: Adaptive control, systems identification, nonlinear control.

Caveat This is an abridged version of the full paper (Ortega et al., 2019) where all proofs and simulation results are omitted.

1. INTRODUCTION AND LITERATURE REVIEW

It is well known that nonlinear parameterizations are inevitable in any realistic practical problem (Ljung, 1987; Nelles, 2001; Ortega et al., 1998). Unfortunately, designing adaptive (identification or control) algorithms for nonlinearly parameterized systems is a difficult poorly understood problem (Pyrkin et al., 2014). See (Ortega et al., 2019) for an extensive review of the literature.

Some results for *gradient estimators* have been reported in the literature for *convexly or monotonically* parameterized continuous-time (CT) systems. A very important drawback of these approaches is that the monotonicity or convexity conditions are imposed on functions that depend, not only on the parameters, but *also* on external signals, *e.g.*, time or the system state. This renders the verification of the condition very hard to carry out. One of the main contributions of this paper is to overcome this limitation.

To the best of the authors' knowledge no developments—similar to the ones done for CT—have been reported for case of nonlinearly parameterized *discrete-time* (DT) regressions that, in spite of its great practical importance, have attracted less attention in the identification and adaptive control community. One of the objectives of our paper is to contribute, if modestly, towards the development of estimation algorithms for DT NPRE. In particular, we provide solutions to the, essentially open, problems of direct and indirect adaptive pole-placement control (APPC) without overparameterization nor persistency of excitation (PE) requirements. [1] It should be pointed out that a solution to the direct APPC problem using overparameterization, hence requiring some excitation conditions, has been reported in (Pyrkin et al., 2019).

The dependence of the convexity or monotonicity assumption on external signals mentioned above happens even

[1] We recall that a bounded vector signal $\Omega \in \mathbb{R}^q$ is said to be PE if there exist $\delta \in \mathbb{R}_{>0}$ such that $\int_t^{t+T} \Omega(\tau)\Omega^\top(\tau)d\tau \geq \delta I_q$ for some $T \in \mathbb{R}_{>0}$ and all $t \in \mathbb{R}_{\geq 0}$ in CT, or $\sum_{j=k+1}^{k+K} \Omega(j)\Omega^\top(j) \geq \delta I_m$, $\forall k \in \mathbb{N}_{\geq 0}$, for some $K \in \mathbb{N}_{>0}$, with $K \geq m$, in DT.

in the case when the uncertain terms appear as products of a function of the unknown parameters times a known function—the so-called, *factorized mappings*, that is NPRE of the form

$$y = \Omega \mathcal{S}(\theta),$$

with $\mathcal{S} : \mathbb{R}^q \to \mathbb{R}^p$, with $p > q$. Although in this case it is possible to define the extended parameter vector $\theta_a := \mathcal{S}(\theta)$ to obtain a linear parametrization, overparametrization suffers from the following well-known shortcomings (Ljung, 1987; Sastry and Bodson, 1989).

(S1) Performance degradation, *e.g.*, slower convergence, due to the need of a search in a larger parameter space.

(S2) More stringent conditions imposed on the reference signals to ensure the PE requirement needed for convergence of the new parameters.

(S3) Inability to recover the true parameters—except for injecting mappings.

(S4) Conservativeness introduced when incorporating prior knowledge in restricted parameter estimation.

In this paper we propose a parameter estimator for monotonic, factorized NPRE that achieves the following objectives.

(O1) It *does not* rely on overparameterization.

(O2) Imposes the monotonicity property *directly* to the function $\mathcal{S}(\theta)$.

(O3) Ensures parameter convergence *without* the stringent PE requirement.

CT estimators for NPRE with factorizable mappings that avoid overparameterization and rely on monotonicity have been reported in (Liu et al., 2011, Section 3) and (Aranovskiy et al., 2017, Section III). In (Liu et al., 2011) neither the second nor the third objectives above are achieved. On the other hand, in (Aranovskiy et al., 2017) these objectives are achieved, via the use of a dynamic regressor extension and mixing (DREM) estimator. As is well known, the main feature of DREM is that it generates, out of a q-dimensional regression equations, one *scalar* equation for each of the q unknown parameters. Another important feature of DREM is that parameter convergence is ensured without assuming PE.

In this paper, we also use DREM to derive both, CT and DT, parameter estimators. We obtain simpler and stronger results than (Aranovskiy et al., 2017) due to the following three key modifications.

(M1) Generate the extended regressor matrix using the linear *time-varying* (LTV) operators first introduced in (Kreisselmeier, 1977).

(M2) Directly apply the "mixing" operation—that is the multiplication by the adjugate of the extended matrix—to generate the scalar regressions.

(M3) Incorporate the possibility of adding a *change of coordinates* to the original parameters to satisfy the required monotonicity property.

Notation. $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, $\mathbb{N}_{>0}$ and $\mathbb{N}_{\geq 0}$ denote the positive and non-negative real and integer numbers, respectively. For $n \in \mathbb{N}_{>0}$ we define the set $\bar{n} := \{1, 2, \ldots, n\}$. For $x \in \mathbb{R}^n$, we denote $|x|^2 := x^\top x$. CT signals $s : \mathbb{R}_{\geq 0} \to \mathbb{R}$ are denoted $s(t)$, while for DT sequences $s : \mathbb{N}_{\geq 0} \to \mathbb{R}$ we use $s(k) := s(kT_s)$, with $T_s \in \mathbb{R}_{>0}$ the sampling time. When a formula is applicable to CT signals and DT sequences the time argument is *omitted.* The action of an operator $H : \mathcal{L}_\infty \to \mathcal{L}_\infty$ is denoted $H[u](t)$, while for an operator $H : \ell_\infty \to \ell_\infty$ we use $H[u](k)$. With $i \in \mathbb{N}_{>0}$ we define $q^{\pm i}u(k) := u(k \pm i)$, and $p^i[u](t) := \frac{d^i u}{dt^i}$. All mappings and reference signals are assumed *smooth.* Given a function $F : \mathbb{R}^n \to \mathbb{R}$ we define $\nabla F := \left(\frac{\partial F}{\partial x}\right)^\top$.

2. MONOTONIZABLE NONLINEARLY PARAMETERIZED FACTORIZED REGRESSIONS

2.1 Problem formulation

In many system identification and adaptive control applications one is confronted with the problem of estimation of the parameters appearing in a NPRE of the form

$$y = \Omega \mathcal{S}(\theta) + \varepsilon \tag{1}$$

where $y \in \mathbb{R}^n$, $\Omega \in \mathbb{R}^{n \times p}$ are *measurable* signals, $\theta \in \mathbb{R}^q$ is a constant vector of *unknown* parameters, $\mathcal{S} : \mathbb{R}^q \to \mathbb{R}^p$, with

$$p > q, \tag{2}$$

and ε is a (generic) exponentially decaying term. The task is to identify on-line the parameters θ, out of the measurements of y and Ω.

Remark 2.1. The NPRE (1) is a particular case of the more general regression $y(\cdot) = \mathcal{H}(\cdot, \theta)$, with $(\cdot) = t$ in CT or k in DT. Although in the factorized case it is possible to introduce extra parameters to obtain a linear parametrization, overparametrization suffers from the well-known shortcomings S1-S5 mentioned in the Introduction.

2.2 Key monotonicity assumption

Similarly to (Aranovskiy et al., 2017; Liu et al., 2011) the key property of the parameterization that we will exploit is P-monotonicity, which is defined a follows.

Definition 2.1. Given a positive definite *matrix* $P \in \mathbb{R}^{q \times q}$, *a mapping* $\mathcal{L} : \mathbb{R}^q \to \mathbb{R}^q$ *is* strongly P-monotone *if and only if there exists a constant* $\rho \in \mathbb{R}_{>0}$ *such that*

$$(a-b)^\top P\,[\mathcal{L}(a) - \mathcal{L}(b)] \geq \rho|a-b|^2 > 0,\ \forall a, b \in \mathbb{R}^q,\ a \neq b. \tag{3}$$

Lemma 2.1. A sufficient condition for a differentiable mapping $\mathcal{L} : \mathbb{R}^q \to \mathbb{R}^q$ to be strictly P–monotone is

$$P\nabla\mathcal{L} + (\nabla\mathcal{L})^\top P \geq \rho I_q > 0. \tag{4}$$

Assumption 2.1. Consider the mapping $\mathcal{S}(\theta)$. There exists:

(i) a bijective mapping $\mathcal{D} : \mathbb{R}^q \to \mathbb{R}^q, \theta \mapsto \eta$ with right inverse $\mathcal{D}^I : \mathbb{R}^q \to \mathbb{R}^q, \eta \mapsto \theta$;
(ii) a permutation matrix $T \in \mathbb{R}^{p \times p}$ and;
(iii) a positive definite matrix $P \in \mathbb{R}^{q \times q}$

such that

$$P\nabla\mathcal{W}(\eta)C^\top + C[\nabla\mathcal{W}(\eta)]^\top P \geq \rho I_q > 0, \tag{5}$$

where

$$\mathcal{W}(\eta) := \mathcal{S}(\mathcal{D}^I(\eta)),\ C := \left[I_q \mid 0_{q\times(p-q)}\right] T. \tag{6}$$

□□□

In words, the construction associated with Assumption 2.1 proceeds as follows. First, introduce a bijective coordinate change for the parameters θ, namely $\eta = \mathcal{D}(\theta)$, with inverse $\theta = \mathcal{D}^I(\eta)$. Second, write the original mapping $\mathcal{S}(\theta)$ in terms of the parameters η via the definition of the new mapping $\mathcal{W}(\eta) := \mathcal{S}(\mathcal{D}^I(\eta))$. Third, assuming that these mapping contains q elements that are "good"—term to be defined below—place them at the top with the permutation matrix T and select them with the fat matrix $\left[I_q \mid 0_{q\times(p-q)}\right]$. Whence, define the new "good" mapping $\mathcal{G} : \mathbb{R}^q \to \mathbb{R}^q$ as

$$\mathcal{G}(\eta) := C\mathcal{W}(\eta). \tag{7}$$

Observing that $\nabla\mathcal{G} = \nabla\mathcal{W}C^\top$, and invoking Lemma 2.1, the condition (5) ensures that this "good" mapping is strongly P-monotonic. For future reference we rewrite this condition in terms of the "good" mapping as

$$P\nabla\mathcal{G}(\eta) + [\nabla\mathcal{G}(\eta)]^\top P \geq \rho I_q > 0. \tag{8}$$

Using the definitions above in the NPRE (1) we obtain the new NPRE in terms of the parameters η as

$$y = \Omega\mathcal{W}(\eta) \tag{9}$$

3. GENERATION OF SCALAR NPRE VIA DREM

3.1 Continuous-time case

Proposition 1. Consider the NPRE (9) for CT signal. Fix $\lambda > 0$ and define the signals

$$\begin{aligned}
\dot{Y}(t) &= -\lambda Y(t) + \Omega^\top(t) y(t)\\
\dot{\Phi}(t) &= -\lambda\Phi(t) + \Omega^\top(t)\Omega(t)\\
\mathcal{Y}(t) &= C\mathrm{adj}\{\Phi(t)\}Y(t)\\
\Delta(t) &= \det\{\Phi(t)\}.
\end{aligned} \tag{10}$$

The q scalar NPRE hold

$$\mathcal{Y}_i(t) = \Delta(t)\mathcal{G}_i(\eta),\ i \in \bar{q}\ \Leftrightarrow\ \mathcal{Y}(t) = \Delta(t)\mathcal{G}(\eta). \tag{11}$$

3.2 Discrete-time case

Proposition 2. Consider the NPRE (9) for DT sequences. Fix $0 < \alpha < 1$ and define the signals

$$\begin{aligned}
Y(k) &= -\alpha Y(k-1) + \Omega^\top(k-1) y_p(k-1)\\
\Phi(k) &= -\alpha\Phi(k-1) + \Omega^\top(k-1)\Omega(k-1)\\
\mathcal{Y}(k) &= C\mathrm{adj}\{\Phi(k)\}Y(k)\\
\Delta(k) &= \det\{\Phi(k)\}.
\end{aligned} \tag{12}$$

The q scalar NPRE hold

$$\mathcal{Y}_i(k) = \Delta(k)\mathcal{G}_i(\eta),\ i \in \bar{q}\ \Leftrightarrow\ \mathcal{Y}(k) = \Delta(k)\mathcal{G}(\eta). \tag{13}$$

4. PARAMETER ESTIMATORS CONVERGENCE ANALYSIS

4.1 Continuous-time case

Proposition 3. Consider the NPRE (11) satisfying (8) of Assumption 2.1. Propose the parameter estimator

$$\dot{\hat{\eta}}(t) = \Gamma P\Delta(t)[\mathcal{Y}(t) - \Delta(t)\mathcal{G}(\hat{\eta}(t))], \tag{14}$$

with $\Gamma \in \mathbb{R}^{q\times q}$, $\Gamma > 0$ the adaptation gain.

(i) The *norm* of the parameter estimation vector $\tilde{\eta}(t) := \hat{\eta}(t) - \eta$ is monotonically non-increasing, that is,

$$|\tilde{\eta}(t_2) \leq |\tilde{\eta}(t_1)|, \forall t_2 \geq t_1 \in \mathbb{R}_{\geq 0}. \tag{15}$$

(ii) The following implication holds

$$\Delta(t) \notin \mathcal{L}_2\ \Rightarrow\ \lim_{t\to\infty} |\tilde{\eta}(t)| = 0.$$

Remark 4.1. Convergence in *all* parameter estimators—as well as in state observers—can only be ensured under some kind of excitation conditions (Ljung, 1987). In particular, for standard gradient and least-squares estimators this property is encrypted in the well known PE requirement of the regressor. As it has been shown in (Aranovskiy et al., 2017) convergence of DREM estimators can be ensured without requiring PE and replacing it, instead, by the assumption $\Delta(t) \notin \mathcal{L}_2$, which is necessary and sufficient for parameter convergence for *linear* regression equation. Proposition 3 shows that this condition is sufficient for NPRE of the form (1) with a P-"monotonizable" regressor $\mathcal{S}(\theta)$.

Remark 4.2. We have assumed that the mapping $\mathcal{G}(\eta)$ is *strongly* P-monotonic. This can be relaxed to *strictly* P-monotonic adding some further assumptions on $\Delta(t)$.

4.2 Discrete-time case

Assumption 4.1. The mapping $\mathcal{G}(\eta)$ satisfies the *Lipschitz condition*

$$|\mathcal{G}(a) - \mathcal{G}(b)| \leq \nu|a - b|,\ \forall a, b \in \mathbb{R}^q, \tag{16}$$

for some $\nu > 0$.

Proposition 4. Consider the DT NPRE (13) with $\mathcal{G}(\eta)$ satisfying (8) of Assumption 2.1 and Assumption 4.1. Propose the DT parameter estimator

$$\hat{\eta}(k+1) = \hat{\eta}(k) + \gamma P\frac{\Delta(k)}{1 + \kappa\Delta^2(k)}[\mathcal{Y}(k) - \Delta(k)\mathcal{G}(\hat{\eta})], \tag{17}$$

with $\gamma > 0$ the adaptation gain selected such that the constant [2]

$$\sigma := 2\gamma\rho - \gamma^2\nu^2\lambda^2_{\max}\{P\} > 0, \tag{18}$$

and the constant κ verifying

$$\kappa \geq \max\{1, \sigma\}. \tag{19}$$

P1 The *norm* of the parameter estimation error $\tilde{\eta}(k) := \hat{\eta}(k) - \eta$ is monotonically non-increasing, that is,

$$|\tilde{\eta}(k_2) \leq |\tilde{\eta}(k_1)|, \forall k_2 \geq k_1 \in \mathbb{N}_{\geq 0}. \tag{20}$$

P2 The following implication is true

$$\prod_{i=0}^{\infty} \frac{1 + (\kappa - \sigma)\Delta^2(k)}{1 + \kappa\Delta^2(k)} = 0\ \Rightarrow\ \lim_{k\to\infty} |\tilde{\eta}(k)| = 0.$$

P3 The following implication is true

$$\lim_{k\to\infty} \Delta(k) =: \Delta(\infty) \neq 0\ \Rightarrow\ \lim_{k\to\infty} |\tilde{\eta}(k)| = 0.$$

Remark 4.3. Similarly to the observation made in Remark 4.1 the sufficient conditions for parameter convergence of Properties P2 and P3 should be interpreted as *excitation requirements* imposed on $\Delta(k)$. Notice that the condition of Property P3 is *sufficient* to ensure $\Delta(k) \notin \ell_2$ and *necessary* for it to be PE.

[2] Clearly, for any positive ρ, ν and $P > 0$, this condition is satisfied with $\gamma < \frac{2\rho}{\nu^2\lambda^2_{\max}\{P\}}$.

5. APPLICATION TO CT NONLINEARLY PARAMETERIZED NONLINEAR SYSTEMS

5.1 Direct adaptive control of a general class of CT nonlinear systems

Consider CT systems described by the state equations

$$\dot{x}(t) = F(x(t), u(t)) + R(x(t))\mathcal{S}(\theta) \tag{21}$$

where $x(t) \in \mathbb{R}^n$ is the measurable state, $u(t) \in \mathbb{R}^m$, with $n \geq m$, is the control signal, the mappings $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $R : \mathbb{R}^n \to \mathbb{R}^{n\times p}$ and $\mathcal{S} : \mathbb{R}^q \to \mathbb{R}^p$ are *known* with $p > q$, and $\theta \in \mathbb{R}^q$ is a constant vector of *unknown parameters.*

Assumption 5.1. There exists a mapping $\beta : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^m$, such that the system

$$\dot{x}(t) = F(x(t), \beta(x(t), \theta)) + R(x(t))\mathcal{S}(\theta) =: f_\star(x(t)) \tag{22}$$

has a *globally exponentially stable* (GES) equilibrium at a desired value $x_\star \in \mathbb{R}^n$.

The *control objective* is then to design a parameter estimator such that the (certainty-equivalent) adaptive control $u = \beta(x(t), \hat{\theta}(t))$ ensures the asymptotic convergence

$$\liminf x(t) = x_\star, \tag{23}$$

with all signals bounded. A fist step in the design is the derivation of the NPRE (1) for the system (21). This is easily obtained applying to (21) the stable, LTI filter $H(p) = \frac{1}{p+\lambda}$ and defining

$$\begin{aligned} y(t) &:= pH(p)[x](t) - H(p)[F(x, u)](t) \\ \Omega &:= H(p)[R(x)](t), \end{aligned} \tag{24}$$

and $\varepsilon(t)$ is the solution of $H(p)[\varepsilon](t) = 0$. A state-space realization of (24) is given by

$$\begin{aligned} \dot{z}(t) &= -\lambda(z(t) + x(t)) - F(x(t), u(t)) \\ \dot{\Omega}(t) &= -\lambda\Omega(t) + R(x(t)) \\ y(t) &= z(t) + x(t). \end{aligned} \tag{25}$$

Proposition 5. Consider the nonlinearly parameterized, nonlinear system (21) satisfying Assumptions 2.1 and 5.1. Let the adaptive control be given by

$$u(t) = \beta(x(t), \mathcal{D}^I(\hat{\eta}(t))),$$

together with the parameter estimator (10), (14) and (25). If $\Delta(t) \notin \mathcal{L}_2$ (23) holds with all signals bounded.

5.2 Adaptive tracking of Euler-Lagrange systems

System dynamics and adaptive control problem formulation. We consider n_q degrees-of-freedom (dof), possibly underactuated, EL systems with generalized coordinates $q(t) \in \mathbb{R}^{n_q}$ and control vector $u(t) \in \mathbb{R}^m$, $m \leq n_q$, whose dynamics is described by

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + \nabla\mathbb{U}(q) = G(q)u, \tag{26}$$

where $M : \mathbb{R}^{n_q} \to \mathbb{R}^{n_q\times n_q}$ is the generalized inertia matrix, which is positive definite and assumed to be *bounded*, $\mathbb{U} : \mathbb{R}^{n_q} \to \mathbb{R}$ the potential energy function, $G : \mathbb{R}^{n_q} \to \mathbb{R}^{n_q\times m}$ the input matrix and $C : \mathbb{R}^{n_q} \times \mathbb{R}^{n_q} \to \mathbb{R}^{n_q\times n_q}$ represents the Coriolis and centrifugal forces matrix, which satisfies the key skew-symmetry property

$$z^\top[\dot{M}(q) - 2C(q, \dot{q})]z,\ \forall z \in \mathbb{R}^{n_q}, \tag{27}$$

holds. See Ortega et al. (1998) for additional details on this model and many practical examples.

Assumption 5.2. Given a desired bounded trajectory for the state vector $(q_\star(t), \dot{q}_\star(t)) \in \mathbb{R}^{n_q} \times \mathbb{R}^{n_q}$. Define the state tracking error $\mathrm{col}(\tilde{q}, \dot{\tilde{q}}) := \mathrm{col}(q - q_\star, \dot{q} - \dot{q}_\star)$. There exists a mapping $\beta : \mathbb{R}^{n_q} \times \mathbb{R}^{n_q} \times \mathbb{R}^q \times \mathbb{R}_{\geq 0} \to \mathbb{R}^m$, such that the system

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + \nabla\mathbb{U}(q) = G(q)\beta(q, \dot{q}, \theta, t),$$

has an *error dynamics*

$$\begin{bmatrix} \dot{\tilde{q}} \\ \ddot{\tilde{q}} \end{bmatrix} = f_\star(\tilde{q}, \dot{\tilde{q}}, t)$$

whose origin is GES.

The *control objective* is then to design a parameter estimator such that the (certainty-equivalent) adaptive control $u = \beta(q, \dot{q}, \hat{\theta}, t)$ ensures global asymptotic tracking, that is,

$$\liminf \mathrm{col}(\tilde{q}(t), \dot{\tilde{q}}(t)) = 0, \tag{28}$$

with all signals bounded.

Derivation of the regression equation. A fist step in the design is the derivation of the NPRE (1) for the system (26)—which was already reported in (Slotine and Li, 1989). Towards this end, we introduce the following parameterization of the inertia matrix $M(q)$ and the potential energy $\mathbb{U}(q)$

$$M(q) = \sum_{i=1}^{\ell} m_i(q)\mathcal{S}_i^m(\theta), \qquad \mathbb{U}(q) = \sum_{j=1}^{r} \mathbb{U}_j(q)\mathcal{S}_j^{\mathbb{U}}(\theta) \tag{29}$$

with *known* matrices $m_i : \mathbb{R}^{n_q} \to \mathbb{R}^{n_q\times n_q}$ and functions $\mathbb{U}_j : \mathbb{R}^{n_q} \to \mathbb{R}$ and known functions $\mathcal{S}_i^m(\theta), \mathcal{S}_j^{\mathbb{U}}(\theta) : \mathbb{R}^q \to \mathbb{R}$ of the *unknown* physical parameters $\theta \in \mathbb{R}^q$. We group together all functions $\mathcal{S}_i^m(\theta), \mathcal{S}_j^{\mathbb{U}}(\theta)$ in a single vector mapping $\mathcal{S} : \mathbb{R}^q \to \mathbb{R}^p$ as

$$\mathcal{S}(\theta) := \mathrm{col}(\mathcal{S}_1^m(\theta), \cdots, \mathcal{S}_\ell^m(\theta), \mathcal{S}_1^{\mathbb{U}}(\theta), \cdots, \mathcal{S}_r^{\mathbb{U}}(\theta)) \in \mathbb{R}^p, \tag{30}$$

where $p := \ell + r > q$. We are in position to present the following.

Proposition 6. The regressor matrix $\Omega : \mathbb{R}^{n_q} \times \mathbb{R}^{n_q} \to \mathbb{R}^{n_q\times p}$

$$\Omega(q, \dot{q}) := H(p) \begin{bmatrix} pm_1(q)\dot{q} - \frac{1}{2}\nabla_q(\dot{q}^\top m_1(q)\dot{q}) \\ \vdots \\ pm_\ell(q) - \frac{1}{2}\nabla_q(\dot{q}^\top m_\ell(q)\dot{q}) \\ \nabla\mathbb{U}_1(q) \\ \vdots \\ \nabla\mathbb{U}_r(q) \end{bmatrix}^\top, \tag{31}$$

is such that the EL system (26) satisfies the NPRE

$$y = \Omega(q, \dot{q})\mathcal{S}(\theta) \tag{32}$$

where

$$y := H(p)\left[G(q)u\right], \tag{33}$$

with θ and $\mathcal{S}(\theta)$ defined via (29) and (30).

Main stabilization result. We are now in position of present the main result of this subsection.

Proposition 7. Consider the EL system (26) with NPRE (32) verifying Assumptions 2.1 and 5.2. Let the adaptive control be given by

$$u = \beta(q, \dot{q}, \mathcal{D}^I(\hat{\eta}), t) \tag{34}$$

together with the parameter estimator (10), (14), (33) and (31). If $\Delta \notin \mathcal{L}_2$ we have that (28) holds with all signals bounded.

In what follows we present two well-known choices of $\beta(q,\dot q,\theta,t)$ for *fully actuated* systems, *i.e.*, $m = n_q$, and prove that they satisfy the key GES Assumption 5.2

- The *Computed Torque Controller* in the known parameter case is given by

$$\beta(q,\dot q,\theta,t) = M(q)[\ddot q_\star - K_1\dot{\tilde q} - K_2\tilde q] + C(q,\dot q)\dot q + g(q),$$

resulting in the LTI closed-loop system $\ddot{\tilde q} + K_1\dot{\tilde q} + K_2\tilde q = 0$, that, obviously, has a GES equilibrium at the origin for all positive definite control gains $K_1, K_2 \in \mathbb{R}^{n_q\times n_q}$.

- The *Slotine-Li Controller* in the known parameter case is given by (Slotine and Li, 1988)

$$\beta(q,\dot q,\theta,t) = M(q)\ddot q_r + C(q,\dot q)\dot q_r + g(q) + K_1 s, \quad (35)$$

where we defined the signals

$$\dot q_r := \dot q_\star - K_2\tilde q, \quad s := \dot{\tilde q} + K_2\tilde q. \quad (36)$$

The closed-loop system is then $M(q)\dot s + [C(q,\dot q) + K_1]s = 0$, $\dot{\tilde q} + K_2\tilde q = s$, that—as indicated in (Ortega et al., 1998, Remark 4.5)—has an GES equilibrium at the origin.

Remark 5.1. To the best of our knowledge, the proof of global stability of the adaptive version of the computed torque scheme proposed above is the first one reported in the literature.

Verifying Assumption 2.1 on a 2-DOF robot manipulator. The equations of motion of the robot are given by (26) with

$$M(q) = \begin{bmatrix} \mathcal{S}_1(\theta) + 2\mathcal{S}_2(\theta)\cos(q_2) & \mathcal{S}_3(\theta) + \mathcal{S}_2(\theta)\cos(q_2) \\ \mathcal{S}_3(\theta) + \mathcal{S}_2(\theta)\cos(q_2) & \mathcal{S}_3(\theta) \end{bmatrix}$$
$$\mathbb{U}(q) = \mathcal{S}_4(\theta)g\,(1+\sin(q_1+q_2)) + \mathcal{S}_5(\theta)g\,(1+\sin(q_1)), \quad (37)$$

with g the gravitational constant, the physical parameters $\theta := \mathrm{col}(l_1, l_2, m_1, m_2)$, where $l_i > 0$ is the the length of the link i with mass $m_i > 0$ for $i = 1, 2$, and the mappings

$$\mathcal{S}^m(\theta) := \begin{bmatrix} \theta_2^2\theta_4 + \theta_1^2(\theta_3+\theta_4) \\ \theta_1\theta_2\theta_4 \\ \theta_2^2\theta_4 \end{bmatrix}, \; \mathcal{S}^{\mathbb{U}}(\theta) := \begin{bmatrix} \theta_2\theta_4 \\ \theta_1(\theta_3+\theta_4) \end{bmatrix}, \quad (38)$$

and $\mathcal{S}(\theta) := \begin{bmatrix} \mathcal{S}^m(\theta) \\ \mathcal{S}^{\mathbb{U}}(\theta)) \end{bmatrix}$. In the following lemma we verify Assumption 2.1 for the mapping $\mathcal{S}(\theta)$.

Lemma 5.1. Consider the vector $\theta \in \mathbb{R}^4_{>0}$ and the mapping $\mathcal{S} : \mathbb{R}^4_{>0} \to \mathbb{R}^5_{>0}$ given by (38). Assume the bounds

$$\theta_1 \le \theta_1^M, \; \theta_2^m \le \theta_2 \le \theta_2^M, \; \theta_4^m \le \theta_4. \quad (39)$$

The mapping $\mathcal{D} : \mathbb{R}^4_{>0} \to \mathbb{R}^4_{>0}$

$$\eta = \mathcal{D}(\theta) = \mathrm{col}(\theta_1, \theta_2, \theta_2\theta_4, \theta_1(\theta_3+\theta_4)),$$

with right inverse $\mathcal{D}^I : \mathbb{R}^4_{>0} \to \mathbb{R}^4$

$$\theta = \mathcal{D}^I(\eta) = \mathrm{col}(\eta_1, \eta_2, \frac{\eta_4}{\eta_1} - \frac{\eta_3}{\eta_2}, \frac{\eta_3}{\eta_2}), \quad (40)$$

verifies Assumption 2.1 with

$$T = \begin{bmatrix} 0&1&0&0&0 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \\ 1&0&0&0&0 \end{bmatrix}, \; P = \mathrm{diag}\{a, a, 1, 1\}$$

for any $a \ge \frac{1}{4\theta_4^m}\left[\theta_2^M + \frac{(\theta_1^M)^2}{\theta_2^m}\right]$.

Adaptive Slotine-Li control of the 2-DOF robot manipulator. In this subsubsection we present in detail the adaptive controller of Proposition 7 with the Slotine-Li scheme for the 2-DOF robot manipulator. Simulation results showing that the proposed scheme *largely outperforms* the classical one relying on overparameterization may be found in (Ortega et al., 2019).

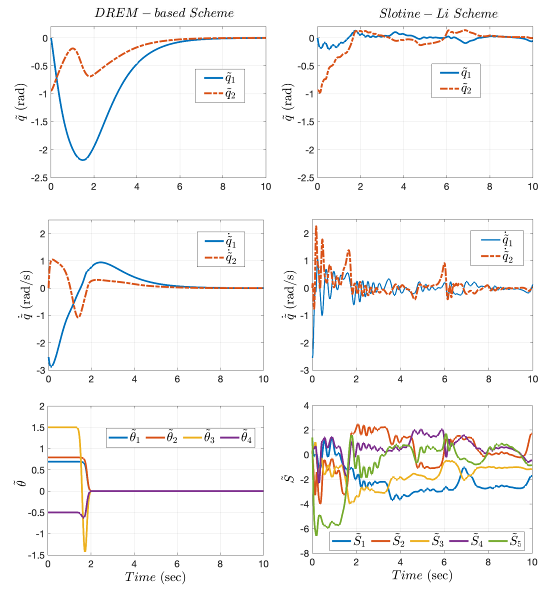

Fig. 1. Simulation results for the DREM-based adaptive scheme and the classical adaptive Slotine-Li.

To derive the NPRE (32) we invoke (29) and (37) and define

$$m_1 := \begin{bmatrix} 1&0 \\ 0&0 \end{bmatrix}, \; m_2(q_2) := \cos(q_2)\begin{bmatrix} 2&1 \\ 1&0 \end{bmatrix}, \; m_3 := \begin{bmatrix} 0&1 \\ 1&1 \end{bmatrix}$$
$$\mathbb{U}_1(q) := g[1+\sin(q_1+q_2)], \; \mathbb{U}_2(q_1) := g[1+\sin(q_1)].$$

Thus, the regressor matrix (32) takes the form

$$\Omega(q,\dot q) = H(p)\begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} & \Omega_{14} & \Omega_{15} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} & \Omega_{24} & \Omega_{25} \end{bmatrix}$$

where $\Omega_{11} = p\dot q_1$, $\Omega_{12} = p\cos(q_2)(2\dot q_1 + \dot q_2)$, $\Omega_{13} = p\dot q_2$, $\Omega_{14} = \Omega_{24} = g\cos(q_1+q_2)$, $\Omega_{15} = g\cos(q_1)$, $\Omega_{21} = \Omega_{25} = 0$, $\Omega_{22} = p\cos(q_2)\dot q_1 + \sin(q_2)(\dot q_1^2 + \dot q_1\dot q_2)$ and $\Omega_{23} = p(\dot q_1 + \dot q_2)$.

The known parameter version of the Slotine-Li controller (35) may be parameterized as $\beta(q,\dot q,\theta,t) := W(q,\dot q,t)\mathcal{S}(\theta) + K_1 s$, with the matrix

$$W(q,\dot q,t) := \begin{bmatrix} W_{11} & W_{12} & W_{13} & W_{14} & W_{15} \\ W_{21} & W_{22} & W_{23} & W_{24} & W_{25} \end{bmatrix},$$

with $W_{11} = \ddot q_{r1}$, $W_{12} = \cos(q_2)(2\ddot q_{r1} + \ddot q_{r2}) - \sin(q_2)(\dot q_2\dot q_{r1} + (\dot q_1 + \dot q_2)\dot q_{r2})$, $W_{13} = \ddot q_{r2}$, $W_{14} = W_{24} = \Omega_{14}$, $W_{15} = \Omega_{15}$, $W_{21} = W_{25} = \Omega_{21}$, $W_{22} = \cos(q_2)\ddot q_{r1} + \sin(q_2)\dot q_1\dot q_{r1}$ and $W_{23} = \ddot q_{r1} + \ddot q_{r2}$, where $\dot q_r$ and s are defined in (36). In its standard version (Slotine and Li, 1988), to get a linear parametrization, the adaptive implementation is obtained estimating the vector $S := \mathcal{S}(\eta)$, yielding $\beta(q,\dot q,\hat S,t) := W(q,\dot q,\dot q_r,\ddot q_r)\hat S + K_1 s$. The parameter estimator is given as $\dot{\hat S} := -\Gamma W^\top(q,\dot q,t)s$, that, as shown in (Ortega et al., 1998), yields a globally stable closed-loop system and ensures global tracking of the desired references.

In the proposed approach we estimate directly θ, that is, the adaptive control is $\beta(q,\dot q,\hat\theta,t) := W(q,\dot q,\dot q_r,\ddot q_r)\mathcal{S}(\hat\theta) + K_1 s$, with the parameter estimator (10), (14), (33) and

(31), combined with $\hat{\theta} = \mathcal{D}^I(\hat{\eta})$, where the mapping $\mathcal{D}^I(\cdot)$ is given in (40).

Figure 1 shows the results of the simulations of the DREM-based and the standard schemes, from which we can observe that the trajectory tracking and the parameter estimation capabilities of our proposal clearly *outperforms* those of the classical adaptive controller.

6. APPLICATION TO NONLINEARLY PARAMETERIZED DT SYSTEMS

6.1 Adaptive Pole Placement Control of LTI Systems

We are interested in this subsection in the problem of APPC of LTI DT system represented by it pulse transfer function

$$A(q^{-1})y_p(k) = B(q^{-1})u(k), \tag{41}$$

where the polynomials

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{n_A} q^{-n_A}, \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \cdots + b_{n_B} q^{-n_B}, \end{aligned}$$

are coprime, with a known *upperbound* on their order, say v, but with *unknown* coefficients a_i, b_i. The pole-placement problem consists of designing a controller

$$L(q^{-1})u(k) + P(q^{-1})y_p(k) = r(k) \tag{42}$$

such that the closed-loop system takes the form $y_p(k) = \frac{B(q^{-1})}{A_m(q^{-1})}r(k)$, where $r(k)$ is a bounded external signal and $A_m(q^{-1}) = 1 + a_1^m q^{-1} + \cdots + a^m_{n_{A_m}} q^{-n_{A_m}}$, is a desired closed-loop polynomial whose roots are inside the unit circle. That is, the controller relocates the poles of the system in a desired position but preserves the open-loop zeros. For a lucid exposition of this problem see (Goodwin and Sin, 1984, Section 5.3) and (Pyrkin et al., 2019) for a review of the recent literature.

6.2 Obstacles for the adaptive implementation.

Computing (41) in closed-loop with (42) we get

$$y_p(k) = \frac{B(q^{-1})}{A(q^{-1})L(q^{-1}) + B(q^{-1})P(q^{-1})} r(k). \tag{43}$$

Hence, to achieve the objective, we need to verify the Bezout equation

$$A(q^{-1})L(q^{-1}) + B(q^{-1})P(q^{-1}) = A_m(q^{-1}). \tag{44}$$

As is well-known (Goodwin and Sin, 1984, Theorem 5.3.1), selecting $n_{A_m} := 2v - 1$, there exists unique polynomials $L(q^{-1})$ and $P(q^{-1})$, both of order $(v-1)$, solutions of (44). Indeed, it is possible to show that (44) admits a matrix representation

$$S(a_i, b_i)\eta = \mathrm{col}(a_0^m, a_1^m, \ldots, a_{2v-1}^m), \tag{45}$$

where

$$\eta := \mathrm{col}(l_0, l_1, \ldots, l_{v-1}, p_0, p_1, \ldots, p_{v-1}) \tag{46}$$

and $S(a_i, b_i) \in \mathbb{R}^{2v\times 2v}$—called the Sylvester matrix—is linearly dependent on the coefficients a_i, b_i, and is *full rank* if and only if $A(q^{-1})$ and $B(q^{-1})$ are coprime.

It is well-known that the adaptive version of the previous controller, called APPC, suffers from serious drawbacks (Goodwin and Sin, 1984; Pyrkin et al., 2019). In its *indirect* version—that is when we estimate the parameters of the plant a_i, b_i and then compute from them, via the solution of (45), the parameters of the controller l_i, p_i—the problem is that the Sylvester matrix with the estimated parameters $\hat{a}_i(k), \hat{b}_i(k)$ may loose rank during the transient behavior. Although this phenomenon can be avoided adding parameter projections, the prior knowledge required to implement this efficiently is never available in practice and relies on the availability of PE, see (Pyrkin et al., 2019, Section 1).

On the other hand, in its *direct* version the estimation of the *controller* parameters involves a NPRE. Indeed, applying (44) to the output of the plant $y_p(k)$ we get

$$\begin{aligned} &L(q^{-1})A(q^{-1})y_p(k) + P(q^{-1})B(q^{-1})y_p(k) = A_m(q^{-1})y_p(k) \\ &\Leftrightarrow B(q^{-1})[L(q^{-1})u(k) + P(q^{-1})y_p(k)] = A_m(q^{-1})y_p(k). \end{aligned} \tag{47}$$

The known parameter version of the direct pole-placement controller may be written in the LRE form $u(k) + \eta^\top \psi(k) = r(k)$ where we have used the fact that $L(q^{-1})$ is monic and defined $\psi(k) := \mathrm{col}(y_p(k), \ldots, y_p(k-v+1), u(k-1), \ldots, u(k-v+1)) \in \mathbb{R}^{2v-1}$, with η, as defined in (46), contains the unknown coefficients of the polynomials $L(q^{-1})$ and $P(q^{-1})$. A direct adaptive implementation of this controller takes then the form $u(k) + \hat{\eta}^\top(k)\psi(k) = r(k)$, where $\hat{\eta}(k)$ denotes the estimates of η. The difficulty of designing an estimator for the controller parameters η is due to the fact that, in terms of η, (47) defines a parameterization of the form

$$B(q^{-1})[u(k) + \eta^\top \psi(k)] = A_m(q^{-1})y_p(k) =: y(k), \tag{48}$$

which is *bilinear* because the polynomial $B(q^{-1})$ is *unknown.*

In the next two subsubsections we show that using the results reported in the paper it is possible to overcome the two obstacles mentioned above. To simplify the presentation we illustrate this fact with simple representative examples, that can be easily extended to the general case.

6.3 DREM-based indirect APPC.

Consider the LTI DT system

$$y_p(k+1) + \theta y_p(k) = u(k) + \theta^3 u(k-1), \tag{49}$$

where, to ensure the coprimeness assumption, $\theta \neq \pm 1$. Fixing a dead-beat objective, *e.g.*, $A_m(q^{-1}) = 1$, and selecting $L(q^{-1}) = l_0 + l_1 q^{-1}$ and $P(q^{-1}) = p_0 + p_1 q^{-1}$ the Bezout equation (44) takes the form

$$(1 + \theta q^{-1})(l_0 + l_1 q^{-1}) + q^{-1}(1 + \theta^3 q^{-1})(p_0 + p_1 q^{-1}) = 1. \tag{50}$$

The latter can be rewritten as

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ \theta & 1 & 1 & 0 \\ 0 & \theta & \theta^3 & 1 \\ 0 & 0 & 0 & \theta^3 \end{bmatrix} \begin{bmatrix} l_0 \\ l_1 \\ p_0 \\ p_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{51}$$

whose solution is $l_0 = 1$, $p_1 = 0$ and

$$\begin{bmatrix} 1 & 1 \\ \theta & \theta^3 \end{bmatrix} \begin{bmatrix} l_1 \\ p_0 \end{bmatrix} = \begin{bmatrix} -\theta \\ 0 \end{bmatrix}, \tag{52}$$

which corresponds to $\begin{bmatrix} l_1 \\ p_0 \end{bmatrix} = \frac{1}{\theta^3 - \theta} \begin{bmatrix} -\theta^4 \\ \theta^2 \end{bmatrix}$. Hence, the known-parameter controller (42) takes the form

$$u(k) = -\frac{1}{\theta^3 - \theta}\left[\theta^2 y_p(k) - \theta^4 u(k-1)\right] + r(k) \tag{53}$$

and yields the desired closed-loop system

$$y_p(k) = q^{-1}(1+\theta^3 q^{-1})r(k).$$

Obviously, the system admits an NPRE of the form (1) with

$$y_p(k) := y_p(k) - u(k-1),\ \Omega(k) := \begin{bmatrix} -y_p(k-1) \\ u(k-2) \end{bmatrix} \quad (54)$$

and $\mathcal{S}(\theta) := \begin{bmatrix} \theta \\ \theta^3 \end{bmatrix}$. If we overparametrize the NPRE and estimate the vector $\mathcal{S} \in \mathbb{R}^2$ the controller parameters are computed from

$$\begin{bmatrix} 1 & 1 \\ \hat{\mathcal{S}}_1(k) & \hat{\mathcal{S}}_2(k) \end{bmatrix} \begin{bmatrix} \hat{l}_1(k) \\ \hat{p}_0(k) \end{bmatrix} = \begin{bmatrix} -\hat{\mathcal{S}}_1(k) \\ 0 \end{bmatrix}, \quad (55)$$

which yields the adaptive controller

$$u(k) = -\frac{1}{\hat{\mathcal{S}}_2(k) - \hat{\mathcal{S}}_1(k)} \left[\hat{\mathcal{S}}_1^2(k) y_p(k) - \hat{\mathcal{S}}_1(k)\hat{\mathcal{S}}_2(k) u(k-1)\right] + r(k). \quad (56)$$

Clearly, the controller computation has a singularity on the line $\hat{\mathcal{S}}_1(k) = \hat{\mathcal{S}}_2(k)$. On the other hand, if we estimate θ, the adaptive version of (53) has a singularity only at the points $\hat{\theta}(k) = \pm 1$. Simulation results illustrating these facts may be found in (Ortega et al., 2019).

7. CONCLUSIONS

It has been shown that the DREM procedure can be used to estimate the parameters of a CT or DT NPRE of the form (1), provided the "monotonizability" Assumption 2.1 holds and some weak excitation conditions—encrypted in the scalar signal Δ—are satisfied. The applicability of the method has been illustrated with several classical examples.

We are currently pursuing the following research avenues.

(R1) The highly attractive parameterization of EL systems proposed in (Slotine and Li, 1989) seems to yield a non-identifiable NPRE. A rigorous proof of this claim is yet to be established.

(R2) Although the DREM estimator has a few tuning gains, *e.g.*, the filter constants (λ for CT, and α for DT) and the adaptation gain γ, their impact on the transient behavior is hard to predict—see Subsubsection 5.2.5 in (Ortega et al., 2019). A more thorough analysis of the sensitivity of the design *vis-à-vis* these coefficients is yet to be derived.

(R3) Although avoiding overparameterization to handle NLPRE seems, in principle, a sensible objective, it is not clear under which conditions this approach is really more convenient. Particularly considering that this is, until now, only applicable to "monotonizable" NLPRE. The results presented in the paper shows that there are practically important examples where this is the case.

(R4) The verification of the conditions of Proposition 1 is carried out in our examples via direct inspection. A deeper understanding of the underlying structural features of the mapping $\mathcal{S}(\theta)$ under which this is possible would be highly desirable. It seems that such a study should appeal to principles of differential algebra.

ACKNOWLEDGMENT

This work is supported by the Russian Science Foundation grant (project №19-19-00403).

REFERENCES

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* , *62*, pp. 3546–3550, 2017. (See also `arXiv:1509.02763` for an extended version.)

A. Astolfi, D. Karagiannis and R. Ortega, *Nonlinear and Adaptive Control Design with Applications.* Springer-Verlag, London, 2007.

G. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control.* Prentice Hall, Leban, Indiana, U.S.A, 1984.

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

X. Liu, R. Ortega, H. Su and J. Chu. On adaptive control of nonlinearly parameterized nonlinear systems: Towards a constructive procedure, *Systems and Control Letters*, *10*, pp. 36–43, 2011.

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

O. Nelles, *Nonlinear System Identification.* Springer-Verlag, Berlin, 2001.

R. Ortega, A. Loria, P. J. Nicklasson and H. Sira-Ramirez, *Passivity–Based Control of Euler–Lagrange Systems.* Springer-Verlag, Berlin, Communications and Control Engineering, 1998.

R. Ortega, V. Gromov, E. Nuño, A. Pyrkin and J. G. Romero. Parameter Estimation of Nonlinearly Parameterized Regressions without Overparameterization nor Persistent Excitation: Application to System Identification and Adaptive Control, *arXiv:1910.08016*, 2019.

A. Pavlov, A. Pogromsky, N. van de Wouw and H. Nijmeijer. Convergence dynamics, a tribute to Boris Pavlovich Demidovich, *System Control Letters*, *52*, pp. 257–261, 2004.

A. Pyrkin, A. Bobtsov, S. Kolyubin, O. Borisov, V. Gromov. Output controller for quadcopters based on mathematical model decomposition, *22nd Mediterranean Conference on Control and Automation*, pp. 1281–1286, 2014.

A. Pyrkin, A. Bobtsov, S. Kolyubin, O. Borisov, V. Gromov, S. Aranovskiy. Output controller for quadcopters with wind disturbance cancellation, *2014 IEEE Conference on Control Applications (CCA)*, pp. 166–170, 2014.

A. Pyrkin, R. Ortega, V. Gromov, A. Bobtsov and A. Vedyakov. A Globally convergent direct adaptive pole-placement controller for nonminimum phase systems with relaxed excitation assumptions, *Int. J. on Adaptive Control and Signal Processing*, *33*(10), pp.1491–1505, 2019.

S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence and Robustness.* Prentice Hall, Englewood Cliffs, N.J, 1989.

J.J. E. Slotine and W. Li. Adaptive manipulator control: a case study, *IEEE Trans. Automatic Control*, *33*(11), pp. 995–1003, 1988.

J.J. E. Slotine and W. Li. Composite adaptive control of robot manipulators, *Automatica*, *25* (4), pp. 509–519, 1989.