2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Parameter Estimation of Some Special Classes of Dynamical Nonlinear Systems with Non-separable Nonlinear Parameterizations*

Romeo Ortega[1], Alexey Bobtsov[2], Ramon Costa-Castelló[3], Nikolay Nikolaev[2] and Anton Pyrkin[2]

***Abstract*— In this paper we address the challenging problem of designing globally convergent estimators for the parameters of nonlinear systems containing a *non-separable* exponential nonlinearity. This class of terms appears in many practical applications, and none of the existing parameter estimators is able to deal with them in an efficient way. The proposed estimation procedure is illustrated with two modern applications: fuel cells and human musculoskeletal dynamics. The procedure does not assume that the parameters live in known compact sets, that the nonlinearities satisfy some Lipschitzian properties, nor rely on injection of high-gain or the use of complex, computationally demanding methodologies. Instead, we propose to design a classical on-line estimator whose dynamics is described by an ordinary differential equation given in a compact precise form. A further contribution of the paper is the proof that parameter convergence is guaranteed with the extremely weak interval excitation requirement.**

I. INTRODUCTION

To comply with the stringent monitoring and control requirements in modern applications an accurate model of the system is vital. It is well-known that *nonlinear parameterizations* (NLP) are inevitable in any realistic dynamic model of practical problems with complex dynamics. Constitutive relations and conservation equations used to characterize physical variables always involve NLP. Classical examples are friction dynamics [1], biochemical processes [2] and in more recent technological developments we can mention fuel cells [3], photovoltaic arrays [4], windmill generators [5] and biomechanics [6]. However, one of the assumptions that pervades almost all results in adaptive estimation and control is *linearity* in the unknown parameters and there are very few results available for NLP systems. Quite often, in practical problems, there are only few physical parameters that are uncertain and occur nonlinearly in the underlying dynamic model. In some cases, it is possible to use suitable transformations so as to convert it into a problem where the unknown parameters occur linearly, usually involving overparameterizations. This procedure, however, suffers from serious drawbacks including the enlarging of dimension of the parameter space, with the subsequent increase in the excitation requirements needed to ensure parameter convergence. The reader is referred to [7] for a thorough discussion on the drawbacks of overparameterization.

Some results for gradient estimators have been reported in the literature for *convexly* parameterized systems. It was first reported in [8] (see also [9]) that convexity is enough to ensure that the gradient search "goes in the right direction" in a *certain region* of the estimated parameter space. The idea is then to apply a standard adaptive scheme in this region, while in the "bad" region either the adaptation is frozen and a robust constant parameter controller is switched-on [10] or, as proposed in [11], the adaptation is running all the time and stability is ensured with a high-gain mechanism which is suitably adjusted incorporating prior knowledge on the parameters. In [12] *reparametrization* to convexify an otherwise non-convexly parameterized system is proposed. See also [13], [14] for some interesting results along these lines, where the controller and the estimator switch between over/underbounding convex/concave functions. Some calculations invoking computationally demanding set membership principles—similar to fuzzy systems—have recently been reported in [15].

Using the Immersion and Invariance adaptation laws proposed in [16], stronger results were obtained in [17], [18] invoking the property of *monotonicity*, see also [13], [14] for related results. The main advantage of using monotonicity, instead of convexity, is that in the former case the parameter search "goes in the right direction" in *all regions* of the estimated parameter space—this is in contrast to the convexity-based designs where, as pointed out above, this only happens in some regions of this space. See the recent work [19] where these results relying on monotonicity have been significantly extended. The reader is referred to [7], [19] for recent reviews of the literature on parameter estimation and adaptive control of NLP systems. Unfortunately, the monotonicity property can be exploited only for the case of *separable* NLP. That is for the case where we can factor the parameter dependent terms as $h_i(u,y,\theta) = \bar{h}_i(u,y)\psi_i(\theta_i)$, where u and y are measurable and θ_i is the unknown parameter. However, there are many practical application models where this factorization is not possible, we refer to this case as *non-separable* NLP. Two often encountered cases are $\cos(\theta_i \cdot h_i(u,y))$ or $e^{\theta_i \cdot h_i(u,y)}$. In particular, the last example appears in many physical processes including Arrenhius laws [20], biochemical reactors [2], friction models [1], windmill systems [21], fuel cell systems [22], photovoltaic arrays [23] and models of elastic moments [24], [25], [26]. This paper is

*This paper is supported by the Ministry of Science and Higher Education of Russian Federation, passport of goszadanie no. 2019-0898. This work is part of the project MAFALDA (PID2021-126001OB-C31) and MASHED (TED2021-129927B-I00) funded by MCIN/ AEI/10.13039/501100011033 and by the European Union Next GenerationEU/PRTR

[1]Romeo Ortega is with the Departamento Académico de Sistemas Digitales, ITAM, Ciudad de México, México romeo.ortega@itam.mx
[2]Alexey Bobtsov, Nikolay Nikolaev and Anton Pyrkin are with the Department of Control Systems and Robotics, ITMO University, Saint-Petersburg, Russia bobtsov@mail.ru, nikona@yandex.ru, pyrkin@itmo.ru
[3]Ramon Costa-Castelló is with the Universitat Politècnica de Catalunya (UPC), Spain ramon.costa@upc.edu

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

devoted to the development of a systematic methodology for the parameter identification of systems containing this kind of exponential terms. More precisely, we consider systems of the form

$$\dot{x} = F_x(u, y, \theta), y = H_x(u, y, \theta)$$

with u and y measurable and θ a vector of unknown parameters, with some of its elements entering into the functions F_x and/or H_x via exponential terms of the form $e^{\theta_i \cdot h_i(u,y)}$. The objective is to design an on-line *estimator*

$$\dot{\chi} = F_\chi(\chi, u, y), \hat{\theta} = H_\chi(\chi, u, y)$$

with $\chi(t) \in \mathbb{R}^{n_\chi}$ such that we ensure *global exponential convergence* (GEC) of the estimated parameters. That is, for all $x(0) \in \mathbb{R}^n, \chi(0) \in \mathbb{R}^{n_\chi}$ and all continuous u that generates a bounded state trajectory x we ensure

$$\lim_{t\to\infty} |\tilde{\theta}(t)| = 0, \quad \text{(exp)}, \tag{1}$$

where $\tilde{\theta} := \hat{\theta} - \theta$ is the parameter estimation error, with all signals remaining bounded.

Notice that, in contrast with the existing approaches for non-separable NLP systems, we do not assume that the parameters live in known compact sets, that the nonlinearities satisfy some Lipschitzian properties, nor rely on injection of high-gain to dominate the nonlinearities or the use of complex, computationally demanding methodologies like min-max optimizations, parameter projections or set membership techniques. Instead, we propose to design a classical on-line estimator whose dynamics is described by an ordinary differential equation given in a compact precise form.

We identify in the paper two classes of systems for which the problem formulated above can be solved. The design procedure consists of the construction—from the non-separable NLP containing an exponential term—a new NLP regression equation (NLPRE) of the form $Y(u, y) = \phi^\top(u, y)\mathcal{G}(\theta)$, where the functions $Y(u, y)$ and $\phi(u, y)$ are known and $\mathcal{G}(\theta)$ is a nonlinear mapping. To estimate the parameters θ from the NLPRE we invoke the recent result of [19], where a least-squares plus dynamic regression equation and mixing [27] (LS+DREM) estimator applicable for this kind of NLPRE is reported. A key feature of the LS+DREM estimator is that it ensures GEC imposing an extremely weak *interval excitation* (IE) assumption [28], [29] of the regressor ϕ. On the other hand, this estimator requires that the mapping of the NLPRE satisfies a rather weak monotonizability property—that is captured by the verifiability of a linear matrix inequality (LMI) imposed on $\mathcal{G}(\theta)$. We give two practical examples of the application of the proposed estimation method and illustrate their performance with some simulations.

Notation. I_n is the $n \times n$ identity matrix and $0_{s\times r}$ is an $s \times r$ matrix of zeros. $\mathbb{R}_+$ and $\mathbb{Z}_+$ denote the positive real and integer numbers, respectively. For $q \in \mathbb{Z}_+$ we define the set $\bar{q} := \{1, 2, \dots, q\}$. For $a \in \mathbb{R}^n$, we denote $|a|^2 := a^\top a$, and for any matrix A its induced norm is $\|A\|$. All functions and mappings are assumed smooth and all dynamical systems are assumed to be forward complete. Given a function $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ we define its transposed gradient via the differential operator $\nabla_{(\cdot)} h(x, u) := \left[\frac{\partial h}{\partial(\cdot)}(x, u)\right]^\top$. For a mapping $\mathcal{G} : \mathbb{R}^{n_\eta} \to \mathbb{R}^{p_\eta}$ we denote its Jacobian by $\nabla\mathcal{G}(\eta) := \frac{\partial \mathcal{G}}{\partial \eta}(\eta)$. To simplify the notation, the arguments of all functions and mappings are written only when they are first defined and are omitted in the sequel.

II. FIRST CLASS OF SYSTEMS

In this section we consider NLP systems of the form

$$\dot{x} = f_1(x, u) + f_2(x, u)\,\mathcal{G}(\eta) \tag{2a}$$

$$y = \begin{bmatrix} y_1 \\ x \end{bmatrix} = \begin{bmatrix} h_1(x, u) + h_2(x, u)\,\theta_2 + h_3(x, u)\,e^{h_4(x)\theta_1} \\ x \end{bmatrix} \tag{2b}$$

with $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^{n+1}$ and $u(t) \in \mathbb{R}^m$ the systems state, output and control, respectively. The functions $f_i,\ i = 1, 2$, and $h_i,\ i = 1, \dots, 4$, are *known* nonlinear functions, $\mathcal{G} : \mathbb{R}^{n_\eta} \mapsto \mathbb{R}^{p_\eta},\ p_\eta > n_\eta$, is a *known* mapping of the unknown parameters $\eta \in \mathbb{R}^{n_\eta}$, and $\theta_i \in \mathbb{R},\ i = 1, 2$ are also unknown parameters. Hence, the overall vector of *unknown* parameters, which needs to be estimated on-line, consists of $\theta := \text{col}(\theta_1, \theta_2, \eta) \in \mathbb{R}^{\ell_I}$, where $\ell_I := 2 + n_\eta$.

We make the important observation that, in view of the presence of the *exponential term* in the signal y_1, the parameterization of the system is nonlinear and *non-separable*. As discussed in the Introduction none of the existing parameter estimators can deal with this difficult—but often encountered in practice—scenario.

A. Assumptions

We make the following assumptions on the system.

A1 [Sign definiteness] The scalar function h_3 is bounded away from *zero*. That is $|h_3| > 0$.

A2 [Monotonicity] There exists a matrix $T_\mathcal{G} \in \mathbb{R}^{n_\eta \times p_\eta}$ such that the mapping $\mathcal{G}(\eta)$ satisfies the *LMI*

$$T_\mathcal{G}\nabla\mathcal{G}(\eta) + [\nabla\mathcal{G}(\eta)]^\top T_\mathcal{G}^\top \geq \rho_\mathcal{G} I_{n_\eta}, \tag{3}$$

for some $\rho_\mathcal{G} > 0$.

Discussion on the assumptions: **D1** In [7, Proposition 1] it is shown that (7) ensures the mapping $T_\eta\mathcal{G}(\eta)$ is *strictly monotonic* [30]. That is, it satisfies

$$(a - b)^\top [T_\eta\mathcal{G}(a) - T_\eta\mathcal{G}(b)] \geq \rho_\eta |a - b|^2,\ \forall a, b \in \mathbb{R}^{n_\eta},\ a \neq b. \tag{4}$$

This is the fundamental property that is required by the LS+DREM estimator used in the next section.

D2 The assumption that the state trajectories of (2) are bounded is standard in parameter estimation theory [31], [32]. Similarly, the assumption that the dimension n_η of the unknown parameters vector η is smaller than p_η is reasonable, otherwise we could redefine a new vector of unknown parameters $\bar{\eta} := \mathcal{G}(\eta) \in \mathbb{R}^{n_\eta}$ without overparameterization and get a LRE.

B. Regression Equation for Parameter Estimation

In this section we derive the regression equation that will be used to estimate the unknown parameters θ. As expected, this regressor equation is nonlinearly parameterized, which hampers the application of standard estimation techniques. Therefore, we are compelled to appeal—in Section IV—to the LS+DREM parameter estimator recently reported in [19], [33].

Lemma 1: Consider the system (2) *verifying Assumptions* **A1, A2**. *There exist measurable, scalar signals*

$Y_I(x,u,y),\ \phi_{I,i}(x,u,y),\ i=1,\ldots,s_I,\ s_I := 3+2p_\eta$, *such that the following NLPRE holds:*

$$Y_I(x,u,y) = \phi_I^\top(x,u,y)\mathcal{W}_I(\theta), \tag{5}$$

where we defined the mapping $\mathcal{W}_I : \mathbb{R}^{\ell_I} \to \mathbb{R}^{s_I}$

$$\mathcal{W}_I(\theta) := \begin{bmatrix}\theta_1 & \theta_2 & \theta_1\theta_2 & \theta_1\mathcal{G}^\top(\eta) & \theta_1\theta_2\mathcal{G}^\top(\eta)\end{bmatrix}^\top. \tag{6}$$

Discussion on the regressor equation: **D3** It is possible to construct another NLPRE proceeding as follows. First, exploiting the monotonicity property of Assumption **A2** and using the LS+DREM algorithm estimate the parameters η filtering (2a). Then, use this estimate in the (approximate) calculation of $\dot{h}_4$, yielding

$$\dot{\hat{h}}_4 = \nabla^\top h_4[f_1 + f_2\mathcal{G}(\hat{\eta})].$$

Applying the certainty equivalent principle, and replacing this expression in the chain of implications of the proof of Lemma 1 in the full paper version [34] would then yield a simpler NLPRE where only the terms $(\theta_1,\theta_2,\theta_1\theta_2)$ will appear. Of course, the drawback of this approach is that we rely on the fast convergence of $\tilde{\eta} := \hat{\eta} - \eta$ to zero.

D4 In the system (2) the function h_4 appearing in the exponential does not depend on u. It is possible to adapt the result of Lemma 1 to consider that case in the following way. The expression for $\dot{h}_4$ given in (**??**) would need to be replaced by

$$\dot{h}_4 = \nabla_x^\top h_4[f_1 + f_2\mathcal{G}(\eta)] + \nabla_u^\top h_4\dot{u}.$$

To construct the NPLRE as in Lemma 1 for this case it is clearly necessary to know $\dot{u}$. However, in many practical applications the control law contains an *integral action*—*e.g.*, in PID control—therefore this signal is available for measurement.

C. Construction of a Strictly Monotonic Mapping

To estimate the parameters θ from the NLPRE (5) we invoke the recent result of [19], where the LS+DREM estimator proposed in [33], which is applicable for linear regression equations, was *extended* to deal with NLPRE. However, this estimator requires that the mapping of the NLPRE satisfies a *monotonicity* property, which is not verified by $\mathcal{W}_I(\theta)$ given in (6). Therefore, in this section we construct a new mapping verifying the required monotonicity condition.

Lemma 2: Consider the mapping $\mathcal{W}(\theta)$ *given in* (6) *with* $\mathcal{G}(\eta)$ *verifying Assumption* **A2**. *There exists a constant* $\alpha_m > 0$ *such that for all* $\alpha \geq \alpha_m$ *the mapping* $\mathcal{W}_I(\theta)$ *satisfies the* LMI

$$T_{\mathcal{W}_I}\nabla\mathcal{W}_I(\theta) + [\nabla\mathcal{W}_I(\theta)]^\top T_{\mathcal{W}_I}^\top \geq \rho_{\mathcal{W}_I} I_{\ell_I}, \tag{7}$$

for some $\rho_{\mathcal{W}_I} > 0$, *with the matrix*

$$T_{\mathcal{W}_I} := \begin{bmatrix}\alpha & 0 & 0 & 0_{1\times p_\eta} & 0_{1\times p_\eta}\\ 0 & \alpha & 0 & 0_{1\times p_\eta} & 0_{1\times p_\eta}\\ & 0_{n_\eta\times 3} & & \mathrm{sign}(\theta_1)T_\mathcal{G} & 0_{n_\eta\times p_\eta}\end{bmatrix} \in \mathbb{R}^{\ell_I\times s_I}.$$

Discussion on the mapping: **D5** Notice that the only prior knowledge needed to construct the matrix $T_{\mathcal{W}_I}$ is sign(θ_1). On the other hand, to select the value of α some prior knowledge on the parameters θ is required. Specifically, as shown in the proof of Lemma 2 in the full paper version [34], it is necessary to know an upper bound on $\|T_\mathcal{G}\mathcal{G}(\eta)\|$.

III. SECOND CLASS OF SYSTEMS

In this section we consider second-order systems of the form

$$\ddot{x} = f_1(x) + f_2^\top(x,\dot{x})\mathcal{G}(\eta) + h_3(x)e^{\theta_1 h_4(x)} + u \tag{8a}$$

$$y = \begin{bmatrix}x\\ \dot{x}\end{bmatrix} \tag{8b}$$

with $x(t) \in \mathbb{R}$ and $u(t) \in \mathbb{R}$. The functions $f_i,\ i = 1,2$, and $h_i,\ i = 1,3$, are *known* nonlinear functions, $\mathcal{G} : \mathbb{R}^{n_\eta} \mapsto \mathbb{R}^{p_\eta},\ p_\eta > n_\eta$, is a *known* mapping of the unknown parameters $\eta \in \mathbb{R}^{n_\eta}$, and $\theta_1 \in \mathbb{R}$ is also an unknown parameter. Hence, the overall vector of unknown parameters, which needs to be estimated on-line, consists of $\theta := \mathrm{col}(\theta_1,\eta) \in \mathbb{R}^{\ell_{II}}$, where $\ell_{II} := 1 + n_\eta$.

Notice that, in contrast to system (2), in this case the dynamics is second order and the nasty exponential term enters into the state equation instead of the readout map. Moreover, note that the control signal is scalar and enters linearly in the state equation. In particular, observe that the function h_3 appearing in the exponential *does not* depend on u now.[1]

To simplify the calculations, in the model (8) we do not include unknown parameters multiplying the function h_3 or the control u. As explained in Discussion **D7** below, this can be easily added redefining $h_3(x) := \theta_2\bar{h}_3(x)$ and $u := \theta_3\bar{u}$, where the functions $\bar{h}_3$ and $\bar{u}$ are known but θ_2 and θ_3 are unknown parameters.

A. Assumptions

We make on the system (8) Assumptions **A1, A2** together with the following.

A3 [Separability] The function $f_2(x,\dot{x})$ verifies

$$\nabla_{\dot{x}} f_2(x,\dot{x}) = \psi_a(x)\psi_b(\dot{x}),$$

for some functions $\psi_a(x)$ and $\psi_b(\dot{x})$.

Discussion on Assumption **A3***:* **D6** As shown in the proof of Lemma 3 given in the full paper version [34], Assumption **A3** is needed to be able to generate—via LTI filtering—a measurable regressor in the NLPRE. We observe that the function $\nabla_{\dot{x}} f_2 \in \mathbb{R}^{p_\eta}$ hence, for $p_\eta > 1$, this is a vector function. However, there is no restriction on the dimensions of the functions ψ_a and ψ_b, as long as they comply with the dimensionality requirement $\psi_a\psi_b \in \mathbb{R}^{p_\eta}$. This degree of freedom relaxes the condition of the assumption.

B. Regression Equation for Parameter Estimation

As in Subsection II-B we derive here the NLPRE that will be used to estimate the unknown parameters θ.

Lemma 3: Consider the system (8) *verifying Assumptions* **A1-A3**. *There exist measurable, scalar signals* $Y_{II}(x,u,y),\ \phi_{II,i}(x,u,y),\ i=1,\ldots,s_{II},\ s_{II} := 1+2p_\eta$, *such that the following NLPRE holds:*

$$Y_{II}(x,u,y) = \phi_{II}^\top(x,u,y)\mathcal{W}_{II}(\theta), \tag{9}$$

where we defined the mapping $\mathcal{W}_{II} : \mathbb{R}^{\ell_{II}} \to \mathbb{R}^{s_{II}}$

$$\mathcal{W}_{II}(\theta) := \begin{bmatrix}\theta_1 & \mathcal{G}^\top(\eta) & \theta_1\mathcal{G}^\top(\eta)\end{bmatrix}^\top. \tag{10}$$

[1] To simplify the presentation, but with an obvious abuse of notation, we keep the same symbol for both functions.

Discussion on regression equation: **D7** To include an unknown multiplicative parameter in the function h_3 or the control u we proceed as follows. Define $h_3(x) = \theta_2 \bar{h}_3(x)$ and $u = \theta_3 \bar{u}$, where the functions $\bar{h}_3$ and $\bar{u}$ are known but θ_2 and θ_3 are unknown parameters. Tracing back the proof of Lemma 3 given in the full paper version [34], in the first step where we divide the model equation by h_3 we divide instead by $\bar{h}_3$. Then, the parameter θ_2 appears multiplying the exponential in the term in parenthesis and it is removed in the next line. That is, the first three lines of the proof become

$$\frac{1}{\bar{h}_3}\ddot{x} = \theta_2 e^{h_4\theta_1} + \bar{f}_3 + \bar{f}_4^\top \mathcal{G}(\eta) + \theta_3 \frac{\bar{u}}{\bar{h}_3}$$

$$\overset{\frac{d}{dt}}{\Longrightarrow} -\frac{\dot{\bar{h}}_3}{\bar{h}_3^2}\ddot{x} + \frac{1}{\bar{h}_3}\frac{d^3x}{dt^3} = \theta_1 \dot{h}_4\Big(\theta_2 e^{h_4\theta_1}\Big) + \dot{\bar{f}}_3 + \dot{\bar{f}}_4^\top \mathcal{G}(\eta) - \theta_3\Big(\frac{\dot{\bar{h}}_3}{\bar{h}_3^2}\bar{u} - \frac{\dot{\bar{u}}}{\bar{h}_3}\Big)$$

$$\Longleftrightarrow -\frac{\dot{\bar{h}}_3}{\bar{h}_3^2}\ddot{x} + \frac{1}{\bar{h}_3}\frac{d^3x}{dt^3} = \theta_1 \dot{h}_4\Big(\frac{1}{\bar{h}_3}\ddot{x} - \bar{f}_3 - \bar{f}_4^\top \mathcal{G}(\eta) - \theta_3\frac{\bar{u}}{\bar{h}_3}\Big) + \dot{\bar{f}}_3 + \dot{\bar{f}}_4^\top \mathcal{G}(\eta) - \theta_3\Big(\frac{\dot{\bar{h}}_3}{\bar{h}_3^2}\bar{u} - \frac{\dot{\bar{u}}}{\bar{h}_3}\Big),$$

with the new definitions

$$\bar{f}_3 := \frac{f_1}{\bar{h}_3},\ \bar{f}_4 := \frac{1}{\bar{h}_3} f_2.$$

The remaining part of the proof remains unchanged leading to a NLPRE similar to (9), with the new $\bar{(\cdot)}$ terms and adding to the parameter vector θ_3 and $\theta_1\theta_3$. As proven in Proposition 1, from this NLPRE we can estimate exponentially fast $(\theta_1, \theta_3, \eta)$. Therefore, we can replace their estimates in the model (8) leading to the system

$$\ddot{z} = f_1(z) + f_2^\top(z,\dot{z})\mathcal{G}(\hat{\eta}) + \theta_2 \bar{h}_3(z) e^{\hat{\theta}_1 h_4(z)} + \hat{\theta}_3\bar{u},$$

which is a classical linearly parameterized system from which we can estimate θ_2 with standard filtering plus gradient descent techniques.

C. Construction of a Strictly Monotonic Mapping

Similarly to the calculations presented in Subsection II-C we present here the matrix $T_{\mathcal{W}_{II}} \in \mathbb{R}^{\mathcal{L}_{II}\times s_{II}}$ that defines the new monotonic mapping. The proof of this lemma is trivial, therefore it is omitted for brevity.

Lemma 4: Consider the mapping $\mathcal{W}_{II}(\theta)$ given in (10) *with $\mathcal{G}(\eta)$ verifying Assumption* **A2**. *The mapping $\mathcal{W}_{II}(\theta)$ satisfies the* LMI

$$T_{\mathcal{W}_{II}}\nabla\mathcal{W}_{II}(\theta) + [\nabla\mathcal{W}_{II}(\theta)]^\top T_{\mathcal{W}_{II}}^\top \geq \rho_{\mathcal{W}_{II}} I_{\ell_{II}}, \quad (11)$$

with the matrix

$$T_{\mathcal{W}_{II}} := \begin{bmatrix} 1 & 0_{1\times p_\eta} & 0_{1\times p_\eta} \\ 0_{n_\eta\times 1} & T_{\mathcal{G}} & 0_{n_\eta\times p_\eta}\end{bmatrix} \in \mathbb{R}^{\ell_{II}\times s_{II}}.$$

Discussion on the mapping: **D8** Notice that, in contrast with the construction of Subsection II-C, here there is no requirement of prior knowledge on the parameter θ_1. This stems from the fact that, as seen in (10), the mapping $\mathcal{G}(\eta)$ appears once without multiplying this parameter—compare with (6). Therefore, Assumption **A2** is sufficient to construct the new monotonic mapping.

IV. A GLOBALLY EXPONENTIALLY CONVERGENT ESTIMATOR OF θ

In this section we present the main result of the paper, that is, an estimator of the parameters θ that achieves GEC of the parameter error. We proceed from the NLPREs constructed in Lemmata 1 and 3 and, as explained in Subsection II-C, we propose to use the LS+DREM estimator recently reported in [19]. Towards this end, we use the new mappings identified in Lemmata 2 and 4 that verify the monotonicty conditions required by the LS+DREM estimator. To simplify the notation we avoid the subindices $(\cdot)_I$ and $(\cdot)_{II}$ of the various terms appearing in previous sections and present a single proposition applicable to both classes of systems.

Therefore, we consider a general scalar NLPRE of the form

$$Y(t) = \phi^\top(t)\mathcal{W}(\theta) \quad (12)$$

with $\mathcal{W}: \mathbb{R}^\ell \to \mathbb{R}^s$. The main feature of the LS+DREM estimator is that it ensures GEC imposing the following extremely weak *IE* assumption [28], [29] of the regressor ϕ.

A4 [Excitation] The regressor vector ϕ is IE. That is, there exist constants $C_c > 0$ and $t_c > 0$ such that

$$\int_0^{t_c} \phi(s)\phi^\top(s)ds \geq C_c I_s.$$

The proof of the proposition below is given in [19, Proposition 1], therefore it is omitted here.

Proposition 1: Consider the NLPRE (12) *with ϕ verifying Assumption* **A4** *and $\mathcal{W}$ satisfying the LMI*

$$T_{\mathcal{W}}\nabla\mathcal{W}(\theta) + [\nabla\mathcal{W}(\theta)]^\top T_{\mathcal{W}}^\top \geq \rho_{\mathcal{W}} I_\ell$$

for some matrix $T_{\mathcal{W}} \in \mathbb{R}^{\ell\times s}$ and $\rho_{\mathcal{W}} > 0$. Define the LS+DREM interlaced estimator

$$\dot{\hat{\mathcal{W}}} = \gamma_{\mathcal{W}} F\phi(Y - \phi^\top\hat{\mathcal{W}}),\ \hat{\mathcal{W}}(0) = \mathcal{W}_0 \in \mathbb{R}^s$$

$$\dot{F} = -\gamma_{\mathcal{W}} F\phi\phi^\top F,\ F(0) = \frac{1}{f_0} I_s$$

$$\dot{\hat{\theta}} = \Delta\Gamma T_{\mathcal{W}}[\mathcal{Y} - \Delta\mathcal{W}(\hat{\theta})],\ \hat{\theta}(0) = \theta_0 \in \mathbb{R}^\ell,$$

with tuning gains the scalars $\gamma_{\mathcal{W}} > 0$, $f_0 > 0$ and the positive definite matrix $\Gamma \in \mathbb{R}^{\ell\times\ell}$, and we defined the signals

$$\Delta := \det\{I_s - f_0F\}, \mathcal{Y} := \text{adj}\{I_s - f_0F\}(\hat{\mathcal{W}} - f_0F\mathcal{W}_0),$$

where adj$\{\cdot\}$ *denotes the adjugate matrix. For all initial conditions $\mathcal{W}_0 \in \mathbb{R}^s$ and $\theta_0 \in \mathbb{R}^\ell$. The estimation errors of the parameters $\tilde{\theta}$ verify* (1) *with all signals* bounded.

V. TWO PRACTICAL EXAMPLES

A. Proton Exchange Membrane Fuel Cell

Parameter estimation is vital for modeling and control of fuel cell systems. However, an accurate description of the fuel cell dynamics implies the use of models with nonlinear parameterizations [3]. The interested reader is refered to [22] where a detailed review of the literature is reported.

Verification of the conditions from the general result: A widely accepted mathematical model of a *Proton Exchange Membrane Fuel Cell* (PEMFC) is given in [22, Section II.B]. It can be shown that this model can be written in the form (2) with $n = m = n_\eta = p_\eta = 1$ and the scalar linear map $\mathcal{G}(\eta) = \eta$.

We make the observation that function h_3 is bounded away from zero, hence verifying Assumption **A1**.

Since $\mathcal{G} = \theta_3$ the mapping $\mathcal{W}_I : \mathbb{R}^3 \to \mathbb{R}^5$ defined in (6) is simpler and given as

$$\mathcal{W}_I(\theta) := \begin{bmatrix}\theta_1 & \theta_2 & \theta_1\theta_2 & \theta_1\theta_3 & \theta_1\theta_2\theta_3\end{bmatrix}^\top.$$

Some simple calculations give us terms Y_I and $\phi_I^\top$ for the NLPRE (5). And the matrix $T_{\mathcal{W}_I}$ of Lemma 2 is given as

$$T_{\mathcal{W}_I} := \begin{bmatrix}\alpha & 0 & 0 & 0 & 0\\ 0 & \alpha & 0 & 0 & 0\\ 0 & 0 & 0 & \mathrm{sign}(\theta_1) & 0\end{bmatrix},$$

and the minimum value for α is $\alpha_m = \frac{\theta_3^2}{|\theta_1|}$.

B. Human Shank Dynamics

Neuromuscular electrical stimulation is an active research area that aims at restoring functionality to human limbs with motor neuron disorders. Control of these systems is a challenging problem because the musculoskeletal dynamics are nonlinear and highly uncertain [6]. In this subsection we are interested in the mechanical dynamics of the human shank motion where the input is the joint torque produced by electrode stimulation of the shank muscles. We consider the scenario described in detail in [26], see also [24], [25] and concentrate our attention on the problem of estimating the parameters of a widely accepted mathematical model of this system. Namely, the system described by equations (11) to (14) of [26], that we repeat here for ease of reference

$$J\ddot{x} + b_1\dot{q} + b_2\mathrm{sign}(\dot{x}) + k_1 e^{-k_2 x}(x - q_0) + mg\ell\sin(x) = u, \quad (13)$$

where $(x, \dot{x})$ are assumed measurable and all the parameters are assumed unknown. The reader is referred to this reference for further details on the model, in particular, the physical interpretation of the different terms in the model, and the overall formulation of the neuromuscular electrical stimulation problem.

Verification of the conditions from the general result: The following clarifications regarding our formulation of the parameter estimation problem are in order.

C1 As indicated in [26], the term $\mathrm{sign}(\dot{x})$ of our model (13) is replaced in equation (12) of [26] by the function $\tanh(b_3\dot{x})$, with a large value for $b_3 > 0$, which is a smooth *approximation* of the sign function. This approximation is made for mathematical convenience of their calculations that rely on a smoothness assumption, but is not required in our approach that can deal with discontinuous nonlinearities.

C2 In this paper we assume that the term q_0, which is the constant resting knee angle, and the constant inertia J are *known*. Therefore the uncertain parameters in our case are $\mathrm{col}(b_1, b_2, k_1, k_2, m\ell)$. The assumption of known J is not too restrictive because the inertia can be predicted from the subject's anthropometric data [6].

C3 In [26] there is an additional, bounded, unstructured, additive term in (13) that is omitted here for brevity. As shown in Proposition 1 we achieve GEC of the parameter estimates, therefore this term could be easily accommodated in our analysis to ensure practical stability.

The dynamics (13) belongs to the second class of systems given by (8) with $n_\eta = p_\eta = 3$, and the following definitions for the functions

$$f_1(x) = 0,\; f_2(x, \dot{x}) = \frac{1}{J}\mathrm{col}(-\dot{x}, -\mathrm{sign}(\dot{x}), g\sin(x)),$$

$$h_3(x) = k_1\bar{h}_3(x) := k_1\frac{1}{J}(x - q_0),\; h_4(x) = -x,$$

and the parameters

$$\theta_1 = k_2,\; \mathcal{G}(\eta) = \eta = \mathrm{col}(b_1, b_2, m\ell),\; \theta := \mathrm{col}(\theta_1, \eta^\top).$$

We bring to the readers attention the fact that the model (13) has a parameter k_1 multiplying the exponential term. Therefore, it is necessary to invoke the two-stage certainty-equivalent based procedure described in Discussion **D7**. That is, we estimate with the NLPRE (9) the parameters $(k_2, b_1, b_2, m\ell)$ and then estimate, *e.g.*, with some filtering and a standard gradient, the remaining parameter k_1.

To comply with Assumption **A1**, we assume that $|x - q_0| > 0$.[2] Clearly, since $\mathcal{G}(\eta) = \eta$, Assumption **A2** is satisfied with $T_\mathcal{G} = \frac{\rho_\mathcal{G}}{2}I_3$, with any $\rho_\mathcal{G} > 0$. Finally Assumption **A3** is satisfied with the functions

$$\psi_a(x) := \begin{bmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & g\sin(x)\end{bmatrix},\; \psi_b(\dot{x}) := \begin{bmatrix}\dot{x}\\ \mathrm{sign}(\dot{x})\\ 1\end{bmatrix}.$$

The mapping $\mathcal{W}_{II} : \mathbb{R}^4 \to \mathbb{R}^7$ is given as

$$\mathcal{W}_{II}(\theta) := \begin{bmatrix}k_2 & b_1 & b_2 & m\ell & k_2 b_1 & k_2 b_2 & k_2 m\ell\end{bmatrix}^\top.$$

Some simple calculations give us terms Y_{II} and $\phi_{II}^\top$ for the NLPRE (9).

Finally, the matrix $T_{\mathcal{W}_{II}} \in \mathbb{R}^{4\times 7}$ of Lemma 4 is given as

$$T_{\mathcal{W}_{II}} := \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & \frac{\rho_\mathcal{G}}{2} & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & \frac{\rho_\mathcal{G}}{2} & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{\rho_\mathcal{G}}{2} & 0 & 0 & 0\end{bmatrix}.$$

VI. SIMULATION RESULTS

Simulation results you can find in the full paper version [34].

VII. CONCLUSIONS

We have presented in this paper a constructive procedure to design GEC estimators for the parameters of two classes of nonlinear, NLP systems containing nonseparable nonlinearities of the form $e^{\theta_i h_i(u,y)}$. Although this class of nonlinearities seems to be very particular, as discussed in the Introduction, it appears in many practical applications, including the two thoroughly studied in the paper, and is

[2] Adding a simple logic and a discontinuous function we can easily avoid the singularity points and replace this assumption by the knowledge of a set such that $q_0 \in [q_0^m, q_0^M]$.

not amenable for the application of the existing parameter estimation techniques. The design procedure consists of the construction—from the non-separable NLP containing the exponential term—a new separable NLPRE, for which we can apply the LS+DREM estimator of [19]. It is important to underscore that, to the best of our knowledge, only this estimator is capable of dealing with this kind of NLPREs. Moreover, the excitation requirement needed to ensure GEC is the very weak condition of IE defined in Assumption **A4**.

We would like to bring to the readers attention that techniques similar to the ones proposed here have been recently applied by the authors to solve two currently very relevant practical applications. Indeed, in [21] we solve the problem of estimation of the parameters of the power coefficient of windmill generators in off-grid operation. The mathematical model of this system is of the form

$$\dot{y} = -y^3(\theta_1 y - \theta_2)e^{-\theta_3 y},$$

with $\theta \in \mathbb{R}^3$ unknown parameters. Also, in [23] we proposed a GEC parameter estimator for photo-voltaic arrays, whose dynamic model is of the form

$$\dot{x} = -\theta_1 x - \theta_2 e^{bx} + \theta_3 - \theta_4 u, y = x - \theta_5 u,$$

with $\theta \in \mathbb{R}^5$ unknown parameters, and the state $x(t) \in \mathbb{R}$ *unmeasurable*. Notice that none of these applications fits into the class of systems considered in the paper.

References

[1] B. Armstrong-Hélouvry, P. Dupont, and C. C. De Wit, "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, no. 7, pp. 1083–1138, 1994.

[2] D. Dochain, "State and parameter estimation in chemical and biochemical processes: a tutorial," *Journal of Process Control*, vol. 13, no. 8, pp. 801–818, 2003.

[3] J. T. Pukrushpan, A. G. Stefanopoulou, and H. Peng, *Control of fuel cell power systems: principles, modeling, analysis and feedback design*. Springer Science & Business Media, 2004.

[4] G. M. Masters, *Renewable and efficient electric power systems*. John Wiley & Sons, 2013.

[5] S. Heier, *Grid integration of wind energy: onshore and offshore conversion systems*. John Wiley & Sons, 2014.

[6] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[7] R. Ortega, V. Gromov, E. Nuño, A. Pyrkin, and J. G. Romero, "Parameter estimation of nonlinearly parameterized regressions: Application to system identification and adaptive control," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 1206–1212, 2020, 21st IFAC World Congress.

[8] V. Fomin, A. Fradkov, and V. Yakubovich, *Adaptive control of dynamical systems*. Nauka, Moskow, in Russian, 1981.

[9] R. Ortega, "Some remarks on adaptive neuro-fuzzy systems," *International Journal of Adaptive Control and Signal Processing*, vol. 10, no. 1, pp. 79–83, 1995.

[10] A. Fradkov, R. Ortega, and G. Bastin, "Semi-adaptive control of convexly parametrized systems with application to temperature regulation of chemical reactors," *International Journal of Adaptive Control and Signal Processing*, vol. 15, no. 4, pp. 415–426, 2001.

[11] A. M. Annaswamy, F. P. Skantze, and A.-P. Loh, "Adaptive control of continuous time systems with convex/concave parametrization," *Automatica*, vol. 34, no. 1, pp. 33–49, 1998.

[12] M. S. Netto, A. M. Annaswamy, R. Ortega, and P. Moya, "Adaptive control of a class of non-linearly parametrized systems using convexification," *International Journal of Control*, vol. 73, no. 14, pp. 1312–1321, 2000.

[13] I. Tyukin, D. Prokhorov, and V. Terekhov, "Adaptive control with nonconvex parameterization," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 554–567, 2003.

[14] I. Y. Tyukin, D. V. Prokhorov, and C. van Leeuwen, "Adaptation and parameter estimation in systems with unstable target dynamics and nonlinear parametrization," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1543–1559, 2007.

[15] V. Adetola, M. Guay, and D. Lehrer, "Adaptive estimation for a class of nonlinearly parameterized dynamical systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2818–2824, 2014.

[16] A. Astolfi, D. Karagiannis, and R. Ortega, *Nonlinear and adaptive control with applications*. Springer, 2008, vol. 187.

[17] X. Liu, R. Ortega, H. Su, and J. Chu, "Immersion and invariance adaptive control of nonlinearly parameterized nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2209–2214, 2010.

[18] L. Xiangbin, O. Romeo, S. Hongye, and C. Jian, "On adaptive control of nonlinearly parameterized nonlinear systems: Towards a constructive procedure," *Systems & Control Letters*, vol. 60, no. 1, pp. 36–43, 2011.

[19] R. Ortega, J. G. Romero, and S. Aranovskiy, "A new least squares parameter estimator for nonlinear regression equations with relaxed excitation conditions and forgetting factor," *Systems* & *Control Letters*, vol. 169, p. 105377, 2022.

[20] M. S. Silberberg, *Chemistry: The Molecular Nature of Matter and Change*. McGraw-Hill, 2006.

[21] A. Bobtsov, R. Ortega, S. Aranovskiy, and R. Cisneros, "On-line estimation of the parameters of the windmill power coefficient," *Systems & Control Letters*, vol. 164, p. 105242, 2022.

[22] Y. Xing, J. Na, M. Chen, R. Costa-Castelló, and V. Roda, "Adaptive nonlinear parameter estimation for a proton exchange membrane fuel cell," *IEEE Transactions on Power Electronics*, vol. 37, no. 8, pp. 9012–9023, 2022.

[23] A. Bobtsov, F. Mancilla-David, S. Aranovskiy, and R. Ortega, "On-line identification of photovoltaic arrays' dynamic model parameters," *Automatica, (submitted), arXiv preprint arXiv:2209.07246*, 2022.

[24] T. Schauer, N.-O. Negård, F. Previdi, K. Hunt, M. Fraser, E. Ferchland, and J. Raisch, "Online identification and nonlinear control of the electrically stimulated quadriceps muscle," *Control Engineering Practice*, vol. 13, no. 9, pp. 1207–1219, 2005, modelling and Control of Biomedical Systems.

[25] N. Sharma, C. M. Gregory, M. Johnson, and W. E. Dixon, "Closed-loop neural network-based nmes control for human limb tracking," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 712–725, 2012.

[26] R. Yang and M. de Queiroz, "Robust adaptive control of the nonlinearly parameterized human shank dynamics for electrical stimulation applications," *Journal of Dynamic Systems, Measurement, and Control*, vol. 140, no. 8, 2018.

[27] 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, 2017.

[28] G. Kreisselmeier and G. Rietze-Augst, "Richness and excitation on an interval-with application to continuous-time adaptive control," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 165–171, 1990.

[29] G. Tao, *Adaptive control design and analysis*. John Wiley & Sons, 2003, vol. 37.

[30] A. Pavlov, A. Pogromsky, N. van de Wouw, and H. Nijmeijer, "Convergent dynamics, a tribute to boris pavlovich demidovich," *Systems & Control Letters*, vol. 52, no. 3-4, pp. 257–261, 2004.

[31] L. Ljung, *System Identification: Theory for the user*. Prentice Hall, New Jersey, 1987.

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

[33] A. Pyrkin, A. Bobtsov, R. Ortega, and A. Isidori, "An adaptive observer for uncertain linear time-varying systems with unknown additive perturbations," *Automatica*, vol. 147, p. 110677, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0005109822005416

[34] R. Ortega, A. Bobtsov, R. Costa-Castelló, and N. Nikolaev, "Parameter estimation of two classes of nonlinear systems with non-separable nonlinear parameterizations," *http://arxiv.org/abs/2211.06455*, 2022.