This is the post peer-review accepted manuscript of: M. Bin and L. Marconi, ""Class-Type" Identification-Based Internal Models in Multivariable Nonlinear Output Regulation," accepted for publication in IEEE Transaction on Automatic Control. The published version is available online at: https://doi.org/10.1109/TAC.2019.2955668

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

"Class-type" Identification-Based Internal Models in Multivariable Nonlinear Output Regulation

Michelangelo Bin and Lorenzo Marconi, *Fellow, IEEE*

***Abstract*— The paper deals with the problem of output regulation in a "non-equilibrium" context for a special class of multivariable nonlinear systems stabilizable by high-gain feedback. A post-processing internal model design suitable for the multivariable nature of the system, which might have more inputs than regulation errors, is proposed. Uncertainties in the system and exosystem are dealt with by assuming that the ideal steady state input belongs to a certain "class of signals" by which an appropriate model set for the internal model can be derived. The adaptation mechanism for the internal model is then cast as an identification problem and a least square solution is specifically developed. In line with recent developments in the field, the vision that emerges from the paper is that approximate, possibly asymptotic, regulation is the appropriate way of approaching the problem in a multivariable and uncertain context. New insights about the use of identification tools in the design of adaptive internal models are also presented.**

***Index Terms*— Output Regulation, Internal Model, Adaptive Systems, Identification for Control**

I. Introduction

We consider nonlinear systems of the form

$$\dot{x} = f(w, x, u), \quad y = h(w, x), \quad e = h_e(w, x) \tag{1}$$

with state $x \in \mathbb{R}^{n_x}$, control input $u \in \mathbb{R}^{n_u}$, measured outputs $y \in \mathbb{R}^{n_y}$, "regulation error" $e \in \mathbb{R}^{n_e}$, and with $w \in \mathbb{R}^{n_w}$ an exogenous signal generated by the "exosystem"

$$\dot{w} = s(w)\,. \tag{2}$$

The problem of *approximate* output regulation pertains the design of an output feedback regulator of the form

$$\dot{x}_c = f_c(x_c, y)\,, \quad u = k_c(x_c, y)$$

achieving the regulation objective $\limsup_{t\to\infty} |e(t)| \leq \epsilon$, with $\epsilon \geq 0$ possibly a "small" number measuring the regulator's asymptotic performance. If $\epsilon = 0$, then the regulator is said to achieve *asymptotic regulation*. If ϵ can be reduced arbitrarily by opportunely tuning the regulator parameters, the regulator is said to achieve *practical regulation*. If the regulation properties are obtained in spite of possible uncertainties in the system (1), the problem is referred to as *robust output regulation* [1], while the terminology *adaptive output regulation* is typically used in presence of uncertainties in the exosystem (2). An anchor point in the solution of the problem is represented by the steady-state trajectories $(x^\star(t), u^\star(t))$ solution of the so-called *regulator equations*

$$\dot{w} = s(w)\,, \quad \dot{x}^\star = f(w, x^\star, u^\star)\,, \quad 0 = h_e(w, x^\star)\,, \tag{3}$$

with $x^\star$ representing the ideal state trajectory associated with a zero regulation error and $u^\star$ the associated input (often referred to as "the friend" of $x^\star$). As shown in [2], indeed, solvability of (3) is a necessary condition for the problem at hand.

Michelangelo Bin (m.bin@imperial.ac.uk) is with the Department of Electrical and Electronic Engineering, Imperial College London, UK. Lorenzo Marconi (lorenzo.marconi@unibo.it) is with the Department of Electrical, Electronic, and Information Engineering, University of Bologna, Italy. Corresponding author: Michelangelo Bin.

This work was supported by the European project AirBorne (ICT 780960).

Regulator structures proposed in the nonlinear context are typically composed by two units, an *internal model unit* and a *stabilising unit*, with a neat, albeit limiting in many contexts, "role" conferred on the two at the design stage: the former is designed to generate the steady state input $u^\star(t)$ required to keep the error at zero in steady state, while the latter is designed to steer the closed-loop trajectories of the system to $x^\star(t)$. What makes the design problem particularly challenging is, of course, the fact that $(x^\star, u^\star)$ are unknown as the initial conditions of (3) and (2) are such and, in the robust/adaptive case, uncertainties in (1) and/or (2) strongly affect the solution of (3). The majority of the current literature on the subject have some limiting aspects that is worth pointing out to better frame the contribution of this paper.

Non-equilibrium context. Current frameworks typically assume that the solutions of (3) depend on time through $w(t)$, namely $(x^\star, u^\star) = (\pi(w), c(w))$ for some π and c. Moreover, further restrictions are usually imposed limiting the class of friends that can be dealt with, as for instance the so-called "immersion assumption" (the latter even more weakened over the years, see [3], [4], [5], [6]). This assumption, far to be necessary, leads to design principles of the internal model unit just driven by the exosystem dynamics and some appropriate "distortions" that, however, do not completely capture the full nonlinear context. A formal framework to overcome this limitation was given in [2], where a "non-equilibrium theory" for nonlinear output regulation was laid, by asserting that the internal model is in general required to incorporate a mixture of the residual plant's and exosystem's dynamics, in this way making meaningless the distinction between the plant and the exosystem from a design viewpoint (and thus between robust and adaptive output regulation).

"Friend-centric" internal models. Many of the existing regulators are strongly "friend-centric", namely the design of the internal model unit is definitely tailored around the specific $u^\star$ resulting from the regulator equations. This, in turn, leads to fragile designs in which unexpected variations of the system/exosystem easily lead to ineffective regulators with unpredictable asymptotic properties. Uncertainties in the system/exosystem are typically handled by parametrising the internal model in terms of uncertain parameters and by looking for "adaptive" mechanisms according to the actual regulation error (see e.g. [5,7]). This way of proceeding, however, involves a "quantitative" information about how the uncertainties reflect on the friend that are hard to assume, unless substantially limiting the topology describing system/exosystem variations. These difficulties pushed the authors of [1] to conjecture that asymptotic regulation in a general nonlinear and uncertain context is unachievable with finite dimensional regulators and to promote approaches looking for approximate regulators, which possibly become asymptotic if certain fortunate conditions happen. In general, how a "qualitative" information about the friend can be transferred into the design of an internal model that behaves "well" for a "wide" range of system/exosystem variations is still an open point in literature.

Pre- versus post-processing schemes. A taxonomy recently introduced in the literature regards the distinction between *pre-processing* and *post-processing* internal models [8,9]. In the latter, the internal model unit directly processes the regulation error, while the stabilising unit stabilises the cascade of the system driving the internal model

unit. In the former, conversely, the two units are somehow "swapped", with the internal model directly generating the feedforward input and the stabiliser stabilising the cascade of the internal model unit driving the system. The regulator structures proposed so far are definitely biased on pre-processing solutions and, as such, limited to deal with single input-single error systems (i.e. $n_u = n_e = 1$) or some "square" extensions with $n_u = n_e$ (see, e.g., [10]). As observed in [8,9], post-processing solutions seem more suited to handling general multivariable contexts with possibly $n_u > n_e$. The latter, in turn, are also more promising to handle contexts in which, besides the regulation errors, also extra measurements are available that do not necessarily vanish at the steady state. Not surprisingly, the general regulator structure for linear systems is post-processing [11]. The drift towards post-processing solutions for nonlinear systems, however, substantially complicates the design of the nonlinear regulator by raising an intertwining in the design of the internal model and stabiliser (referred to as *chicken-egg dilemma* in [12]) not present in pre-processing approaches. To the best knowledge of the authors, a general post-processing nonlinear framework is still unavailable in literature with just some attempts done in [13] and [14] for simplified exosystems.

In this paper we propose a design technique based on the aforementioned non-equilibrium context, in which the effects of the system and exosystem dynamics on the steady state are jointly considered in the design of the internal model. The proposed regulator embeds a "post-processing" internal model that applies to multivariable systems not necessarily square, and whose construction is not "friend-centric" but rather it is based on a "qualitative" information on the ideal error-zeroing steady state.

II. Main Result

A. The class of systems

We consider a subclass of systems (1) with state $x = \mathrm{col}(x_0, \chi, \zeta) \in \mathbb{R}^{n_x}$ satisfying the following equations

$$\dot{x}_0 = f_0(w, x) + b(w, x)u \tag{4a}$$
$$\dot{\chi} = F\chi + H\zeta \tag{4b}$$
$$\dot{\zeta} = q(w, x) + \Omega(w, x)u \tag{4c}$$
$$e = C\chi\,, \qquad y = \mathrm{col}(\chi, \zeta), \tag{4d}$$

in which $x_0 \in \mathbb{R}^{n_0}$, $y \in \mathbb{R}^{n_y}$, $e \in \mathbb{R}^{n_e}$, $\zeta \in \mathbb{R}^{n_e}$, $u \in \mathbb{R}^{n_u}$, with $n_u \geq n_e$, $\chi = \mathrm{col}(\chi^1, \dots, \chi^{n_e})$, with $\chi^i \in \mathbb{R}^{n_\chi^i}$, $i = 1, \dots, n_e$, and $n_\chi^1 + \dots + n_\chi^{n_e} =: n_\chi$, $C := \mathrm{blkdiag}(C_1, \dots, C_{n_e})$, $F := \mathrm{blkdiag}(F_1, \dots, F_{n_e})$ and $H := \mathrm{blkdiag}(H_1, \dots, H_{n_e})$, with $C_i := \begin{pmatrix} 1 & 0_{1\times(n_\chi^i-1)} \end{pmatrix}$ and

$$F_i := \begin{pmatrix} 0_{(n_\chi^i-1)\times 1} & I_{n_\chi^i-1} \\ 0 & 0_{1\times(n_\chi^i-1)} \end{pmatrix}, \quad H_i := \begin{pmatrix} 0_{(n_\chi^i-1)\times 1} \\ 1 \end{pmatrix}.$$

The χ subsystem, in particular, is described by n_e chains of integrators with ζ entering at the bottom and the regulation error given by the first components χ_1^i of each chain χ^i. Hence, χ and ζ are linear combinations of the error and its time derivatives. The functions f_0, b, q and Ω are sufficiently smooth functions, with $\Omega(w, x) \in \mathbb{R}^{n_e \times n_u}$ denoting the so-called "high-frequency matrix". The form (4) is representative of different frameworks addressed in literature. For instance, systems having a well-defined vector relative degree with respect to the input-output pair (u, e) and admitting a *canonical normal form* fit in the proposed framework. In this case the x_0 dynamics in (4) does not depend on u and it represents the *zero dynamics* of the system relative to the indicated input-output pair. On the other hand (4), with a slightly different structure of χ and of the matrices F and H, is also representative of systems that are "just" (globally) strongly invertible in the sense of [15,16] and feedback linearisable with respect to the input-output pair (u, e) and, as such, can be transformed in *partial normal form*, see [17]. In this case the dynamics (4b)-(4c) are the partial normal form of the system and the subsystem (4a) is indeed the whole plant (i.e. $x = x_0$).

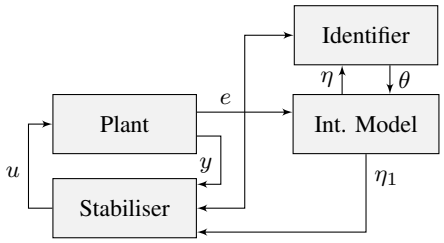


Fig. 1. Block-diagram of the regulator.

We observe that the measurable outputs y are assumed to be linear combinations of the error and its time derivatives, namely we look for a *partial state feedback* solution. A pure error feedback regulator only processing e can be obtained by replacing the time derivatives with appropriate estimates via standard high-gain techniques (see [18]) whose details are not presented here.

In the rest of the paper we assume the following.

A1) *There exist $\beta_0 \in \mathcal{KL}$, $\alpha_0 > 0$ and, for each solution w of (2), each input u, and each solution x of (4) corresponding to (w, u), there exist $x_0^\star : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_0}$ and $u^\star : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_u}$ fulfilling*

$$\begin{aligned} \dot{x}_0^\star &= f_0(w, x^\star) + b(w, x^\star)u^\star \\ 0 &= q(w, x^\star) + \Omega(w, x^\star)u^\star \end{aligned} \tag{5}$$

in which $x^\star := (x_0^\star, 0, 0)$, and

$$|x_0(t) - x_0^\star(t)| \leq \beta_0\big(|x_0(0) - x_0^\star(0)|, t\big) + \alpha_0|(\chi, \zeta)|_{[0,t)}$$

for all $t \geq 0$.

A2) *There exists a full-rank matrix $\mathcal{L} \in \mathbb{R}^{n_u \times n_e}$ such that the (square) matrix $\Omega(w, x)\mathcal{L}$ is bounded, it satisfies*

$$\mathcal{L}^\top \Omega(w, x)^\top + \Omega(w, x)\mathcal{L} \geq I_{n_e}$$

for all $(w, x) \in \mathbb{R}^{n_w} \times \mathbb{R}^{n_x}$, and the map $(\Omega(\cdot)\mathcal{L})^{-1}q(\cdot)$ is Lipschitz.

Equations (5) are the specialisation of the regulator equations (3) in this non-equilibrium context. The steady state $(x^\star, u^\star)$ might be dependent on the initial conditions of the system coherently with [2]. Condition A1 asks for *uniform (in u) detectability* of the ideal steady state $x^\star$ (see [19]). In case of systems with canonical normal form in which (4a) does not depend on u, this assumption boils down to a conventional *minimum-phase* requirement, far to be necessary although typically assumed in the pertinent literature. A2, instead, is a robust stabilisability requirement and it implies that $\Omega(w, x)$ is everywhere full rank. As a consequence, $u^\star$ in (5) is given by

$$u^\star = -\Omega(w, x^\star)^\top \big(\Omega(w, x^\star)\Omega(w, x^\star)^\top\big)^{-1} q(w, x^\star)\,.$$

As clear from A1 and A2, we deal with a simplified case in which a global result is sought under quite restrictive global Lipschitz and boundedness conditions. Nevertheless, we remark that the proposed result can be extended to a *semiglobal* setting by asking w to evolve in a compact space and A1 and A2 to hold only locally (namely on each compact subset of $\mathbb{R}^{n_x}$). For reason of space and since the extension follows by well-known arguments (see e.g. [18,20]) we omit this extension and we focus on the new adaptive framework.

B. The regulator structure

The proposed regulator structure is depicted in Figure 1. The post-processing internal model unit has the form

$$\dot\eta = \Phi(\eta,\theta) + Ge, \qquad \eta \in \mathbb{R}^{dn_e} \tag{6}$$

with $d \in \mathbb{N}$, $\eta = (\eta_1, \dots, \eta_d)$, $\eta_i \in \mathbb{R}^{n_e}$, and

$$\Phi(\eta,\theta) = \begin{pmatrix} \eta_2 \\ \cdots \\ \eta_d \\ \psi(\eta,\theta) \end{pmatrix}, \; G = \begin{pmatrix} gh_1 I_{n_e} \\ g^2 h_2 I_{n_e} \\ \cdots \\ g^d h_d I_{n_e} \end{pmatrix},$$

in which h_i, $i = 1, \dots, d$, are fixed so that the polynomial $s^d + h_1 s^{d-1} + \cdots + h_{d-1}s + h_d$ is Hurwitz, $g > 0$ is a parameter to be designed, $\psi : \mathbb{R}^{dn_e} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_e}$ is a function to be fixed, and $\theta \in \mathbb{R}^{n_\theta}$, $n_\theta \in \mathbb{N}$, is an "adaptive" parameter generated by the *identifier* subsystem, whose dynamics is described by

$$\dot z = \mu(z,\eta,e), \qquad \theta = \omega(z), \tag{7}$$

in which $\mu : \mathcal{Z} \times \mathbb{R}^{dn_e} \times \mathbb{R}^{n_e} \to \mathcal{Z}$ and $\omega : \mathcal{Z} \to \mathbb{R}^{n_\theta}$, with $\mathcal{Z}$ a normed vector space of finite dimension, have to be fixed. Finally, the (static) stabiliser is taken as

$$u \;=\; \mathcal{L}\big(K_\chi \chi + K_\zeta \zeta + K_\eta \eta_1 + K_w \nu(x^\star, w)\big), \tag{8}$$

in which the matrices K_χ, K_ζ and K_η are chosen as follows

$$K_\chi(\ell,\kappa) = \ell K(\kappa), \;\; K_\zeta(\ell) = -\ell I_{n_e}, \;\; K_\eta(\ell,\kappa) = \ell K(\kappa) C^\top$$

with $K(\kappa) = \mathrm{blkdiag}(K^1(\kappa), \dots, K^{n_e}(\kappa))$, where

$$K^i(\kappa) = -\begin{pmatrix} c_1^i \kappa^{n_\chi^i} & c_2^i \kappa^{n_\chi^i - 1} & \dots & c_{n_\chi^i}^i \kappa \end{pmatrix} \tag{9}$$

for $i = 1, \dots, n_e$, in which the coefficients c_j^i are chosen so that the polynomials $s^{n_\chi^i} + c_{n_\chi^i}^i s^{n_\chi^i - 1} + \cdots + c_2^i s + c_1^i$, $i = 1, \dots, n_e$, are Hurwitz, and ℓ, $\kappa > 0$ are design parameters to be fixed. The matrix K_w and the function ν are introduced for sake of generality and are possibly zero. These terms could represent a "feedforward" contribution added by the designer by employing possible knowledge of w and $x^\star$. Likewise, it could represent a term showing up in the normal form (4) after a preliminary feedback of available measurements that do not vanish in steady state. Similarly to the other matrices in (8), the gain matrix K_w can depend on κ and ℓ. The degrees of freedom left to be fixed at this stage are the dimension d and function ψ of the internal model unit (6), the data $(\mathcal{Z}, n_\theta, \mu, \omega)$ of the identifier (7), and the control parameters g, ℓ and κ.

C. Design of the internal model as prediction model

A key step in the regulator synthesis is the choice of the internal model (6) and of its adaptation through the design of the identifier (7). Consistently with the discussion in Section I, this must be done to achieve a small, possibly zero, asymptotic regulation error in spite of uncertainties involving $(x^\star, u^\star)$ and the underlying dynamics. With an eye to the last equation of (6), we can write

$$e(t) = \bar c(g)\,\big(\dot\eta_d(t) - \psi(\eta(t), \theta(t))\big) \tag{10}$$

in which $\bar c(g) := (h_d g^d)^{-1}$. Our design strategy to choose (d, ψ) in (6) and the identifier (7) pivots around the idea that $\dot\eta_d(t) - \psi(\eta(t), \theta(t))$ can be interpreted as a "prediction error" attained by the "model" ψ in relating the "next derivative" $\dot\eta_d(t)$ to the "previous derivatives" $\eta(t)$, and that, by minimising this prediction error, the actual regulation error is also minimised due to (10). This clearly suggests to look at the problem of choosing d and ψ as an *identification problem* and, by borrowing the notation typically adopted in that literature [21], to refer to the map $\psi(\cdot,\theta)$ as the *prediction model* relating the "input data" η to the "output" $\dot\eta_d$, and to the set $\mathcal{M} := \{\psi(\cdot,\theta) : \theta \in \mathbb{R}^{n_\theta}\}$ of all the possible candidate models as the corresponding *model set*. The choice of d and of ψ thus must be done in such a way that the attainable prediction error is minimised. Unless relying on "universal" infinite-dimensional models, however, this selection must be grounded on some preliminary knowledge about the *class of signals* to which $\dot\eta_d$ and η are expected to belong. In this context, the steady-state signals $(x^\star, u^\star)$ resulting from the regulator equations (5) are the anchor point on which that knowledge can be drawn. In particular, let

$$\eta_1^\star := \Upsilon_{(\ell,\kappa)}(w, x^\star),$$

in which

$$\Upsilon_{(\ell,\kappa)}(w,x^\star) := \\ -\big(\Omega(w,x^\star)\mathcal{L}K_\eta\big)^{-1}\big(q(w,x^\star) + \Omega(w,x^\star)\mathcal{L}K_\omega \nu(w,x^\star)\big),$$

and define recursively $\eta_i^\star$, $i = 2, \dots, d+1$, as[1]

$$\eta_i^\star := L_{s(w)}^{i-1}\Upsilon_{(\ell,\kappa)}(w,x^\star) + L_{f_0(w,x)+b(w,x)u^\star}^{i-1}\Upsilon_{(\ell,\kappa)}(w,x^\star).$$

Finally let

$$\dot\eta_d^\star := \eta_{d+1}^\star.$$

In view of A2 and the definition of K_η, the matrix $\Omega(w,x)\mathcal{L}K_\eta$ is everywhere invertible and, thus, all the previous quantities are well-defined. Moreover, we observe that the quantities $\eta_i^\star$, $i = 1, \dots, d+1$, depend on the design parameters κ and ℓ yet to be fixed. The dimension d and the function ψ should be then ideally chosen so that, with $\eta^\star = \mathrm{col}(\eta_1^\star, \dots, \eta_d^\star)$, the following holds

$$\dot\eta_d^\star(t) = \psi(\eta^\star(t), \theta^\star(t)), \tag{11}$$

for some "ideal" $\theta^\star(t) \in \mathbb{R}^{n_\theta}$. This, in fact, would make $(x^\star, \eta^\star)$ a trajectory of the closed-loop system in which the associated regulation error is identically zero. The design of the pair (d, ψ) so that (11) is fulfilled for all possible steady-state trajectories $(\dot\eta_d^\star, \eta^\star)$, however, is not realistic unless limiting even further the class of treatable nonlinear systems and of manageable uncertainties on the solution of (5). Furthermore, even in the fortunate case in which the ideal relation (11) could be fulfilled with a perfect parametrisation (maybe playing with large values of d), this might require an unacceptable complexity of the internal model, and an approximated model with a possibly lower d would be preferable. Along this direction, we rather assume that the designer has a *qualitative* knowledge about a "class" $\mathcal{H}^\star$ of signals[2] to which $(\dot\eta_d^\star, \eta^\star)$ belongs in order to fix a model set $\mathcal{M}$ necessarily approximated but optimised for the specific class. This is the "modelling part", in which the "touch" of the designer and the knowledge on the steady-state trajectories come into play. The class $\mathcal{H}^\star$, in turn, is fixed on the basis of the knowledge on the nominal solution $(x^\star, u^\star)$ to (5), and after considering all the expected system/exosystem uncertainties that may affect it. The problem of handling the overall uncertainty on $(x^\star, u^\star)$ is thus transferred to the adaptation side, and the idea of relying on system identification techniques for it is further motivated by the fact that, typically, identification methods structurally manage large classes of signals [21]. From now on we suppose that the designer has fixed a class $\mathcal{H}^\star$ and, accordingly, a model set $\mathcal{M}$, so that the following assumption holds.

A3) *The map ψ is Lipschitz and differentiable with a locally Lipschitz derivative, and the Lipschitz constants do not dependent on κ and ℓ.*

[1] We denote by $L_g f$ the Lie derivative of f along g.

[2] Formally, $\mathcal{H}^\star$ is a subset of the space of functions $\mathbb{R}_{\geq 0} \to \mathbb{R}^{n_e} \times \mathbb{R}^{dn_e}$.

Moreover, there exists a compact set $H^\star \subset \mathbb{R}^{n_e} \times \mathbb{R}^{dn_e}$, independent on κ and ℓ, such that every $(\dot\eta_d^\star, \eta^\star) \in \mathcal{H}^\star$ satisfies $(\dot\eta_d^\star(t), \eta^\star(t)) \in H^\star$ for all $t \in \mathbb{R}_{\geq 0}$.

The previous assumption formalizes the "quantitative" properties required to the members of the class $\mathcal{H}^\star$ on which the design of the internal model and the identifier is grounded. In particular, it is asked that the elements of $\mathcal{H}^\star$ stay in a known compact set $H^\star$, and that the inferred prediction model ψ has some strong regularity properties uniform in the control gains (κ, ℓ). These requirements, in principle not needed in the design of the identifier and internal model, are rather needed for the successive embedding of the two units in the overall regulator, as they permit to break the "chicken-egg dilemma" and sequence the design of the remaining degrees of freedom.

We remark, moreover, that in the "square" case, namely when $n_u = n_e$ in (4), the matrix $\Omega(w,x)$ is square and κ and ℓ do no mix-up with $\Omega(w,x)$, $q(w,x)$ and $\nu(w,x)$ in the definition of $\dot\eta_d^\star$ and $\eta^\star$. Therefore $(\dot\eta_d^\star, \eta^\star)$ can be always bounded uniformly in κ and ℓ whenever they are taken larger than 1 and $K_\omega/(\kappa\ell)$ can be bounded uniformly in κ and ℓ.

D. The design of the identifier

With d and ψ fixed, we shift our attention to the design of the identifier. The fact that (11) is not attainable exactly suggests to define a *steady state prediction error* as

$$\varepsilon^\star(t,\theta) := \dot\eta_d^\star(t) - \psi(\eta^\star(t), \theta) \tag{12}$$

and to look for a dynamical system which is able to select the best parameter, say $\theta^\star$, whose corresponding model $\psi(\cdot, \theta^\star(t))$ is, at each t, the "best" model in $\mathcal{M}$ relating $\dot\eta^\star(t)$ and $\eta^\star(t)$, minimising in some sense $\varepsilon^\star$. As customary in system identification, the meaning of "best" in the model selection is based on the definition of a *fitness criteria* assigning to each model $\psi(\cdot,\theta) \in \mathcal{M}$ a suitable and comparable value. In particular, with $C^0(\mathbb{R}^{n_\theta}, \mathbb{R}_{\geq 0})$ the space of continuous functions $\mathbb{R}^{n_\theta} \to \mathbb{R}_{\geq 0}$, with each pair $(\dot\eta_d^\star, \eta^\star) \in \mathcal{H}^\star$ we associate the map $\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)} : \mathbb{R}_{\geq 0} \to C^0(\mathbb{R}^{n_\theta}, \mathbb{R}_{\geq 0})$ given by

$$\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}(t)(\theta) := \int_0^t c_\varepsilon\Big(t, s, |\varepsilon^\star(s,\theta)|\Big) ds + c_r(\theta), \tag{13}$$

with $c_\varepsilon : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $c_r : \mathbb{R}^{n_\theta} \to \mathbb{R}_{\geq 0}$ some user-defined positive functions characterising the particular underlying identification problem. More precisely, the integral term of (13) measures how well a given choice of θ fits the historical data, while $c_r(\theta)$ plays the role of a regularisation factor. With $\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}$ we associate the set-valued map $\vartheta^\circ_{(\dot\eta_d^\star, \eta^\star)} : \mathbb{R}_{\geq 0} \rightrightarrows \mathbb{R}^{n_\theta}$ defined as

$$\vartheta^\circ_{(\dot\eta_d^\star, \eta^\star)}(t) := \operatorname*{argmin}_{\theta \in \mathbb{R}^{n_\theta}} \mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}(t)(\theta)\,.$$

Once a cost functional of the form (13) is defined, the identifier subsystem (7) is constructed to guarantee the existence of an "optimal" steady state $z^\star$, which is robustly asymptotically stable for (7), and whose corresponding output $\theta^\star = \omega(z^\star)$ is a pointwise minimiser of $\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}(t)$, i.e. satisfies $\theta^\star(t) \in \vartheta^\circ_{(\dot\eta_d^\star, \eta^\star)}(t)$ for all $t \geq 0$. In particular, the identifier (7) is chosen as a system with state

$$z = \mathrm{col}(\xi, \varsigma), \qquad \xi \in \mathbb{R}^{2n_e},\ \varsigma \in \mathcal{Z}_\varsigma,$$

in which $\mathcal{Z}_\varsigma$ is a finite-dimensional normed vector space, $\mathcal{Z} = \mathbb{R}^{2n_e} \times \mathcal{Z}_\varsigma$, and the pair (μ, ω) is chosen so that, with $\xi_1,\ \xi_2 \in \mathbb{R}^{n_e}$ such that $\xi = \mathrm{col}(\xi_1, \xi_2)$, the equations (7) read as

$$\begin{array}{rcl} \dot\xi_1 &=& \xi_2 - m_1\,\rho\,(\xi_1 - \eta_d) \\ \dot\xi_2 &=& \dot\psi(\xi_2, \eta, \varsigma) - m_2\,\rho^2\,(\xi_1 - \eta_d) \\ \dot\varsigma &=& \varphi(\varsigma, \xi_2, \eta) \\ \theta &=& \gamma(\varsigma) \end{array} \tag{14}$$

where $m_1,\ m_2 > 0$ are arbitrary, $\rho > 0$ is a design parameter, $\dot\psi : \mathbb{R}^{n_e} \times \mathbb{R}^{dn_e} \times \mathcal{Z}_\varsigma \to \mathbb{R}^{n_e}$ is a function fixed below, and (φ, γ) is chosen to satisfy the following requirement.

Requirement 1 (Identifier Requirement) *The pair (φ, γ) is said to satisfy the identifier requirement relative to a class $\mathcal{H}^\star$ and a cost functional* (13)*, if φ is locally Lipschitz, γ is Lipschitz and differentiable with locally Lipschitz derivative, and there exist $\beta_\varsigma \in \mathcal{KL}$, a compact set $S^\star \subset \mathcal{Z}_\varsigma$, $\alpha_\varsigma > 0$ and, for each $(\dot\eta_d^\star, \eta^\star) \in \mathcal{H}^\star$, a unique $\varsigma^\star : \mathbb{R}_{\geq 0} \to S^\star$, such that:*

- *for every locally integrable $\delta' : \mathbb{R} \to \mathbb{R}^{n_e}$ and $\delta'' : \mathbb{R} \to \mathbb{R}^{dn_e}$, all the maximal solutions to the system $\dot\varsigma = \varphi(\varsigma, \dot\eta_d^\star + \delta', \eta^\star + \delta'')$ are complete and satisfy*

$$|\varsigma(t) - \varsigma^\star(t)| \leq \beta_\varsigma(|\varsigma(0) - \varsigma^\star(0)|, t) + \alpha_\varsigma\,|(\delta', \delta'')|_{[0,t)}$$

  for all $t \in \mathbb{R}_{\geq 0}$;
- *the signal $\theta^\star(t) := \gamma(\varsigma^\star(t))$ satisfies $\theta^\star(t) \in \vartheta^\circ_{(\dot\eta_d^\star, \eta^\star)}(t)$ for all $t \in \mathbb{R}_{\geq 0}$.*

With $H^\star$ and $S^\star$ the compact sets introduced, respectively, in A3 and in the identifier requirement, and with $\mathrm{D}\psi$ and $\mathrm{D}\gamma$ denoting the Jacobian of ψ and γ respectively, we define $\dot\psi$ as any bounded function satisfying

$$\dot\psi(\xi_2, \eta, \varsigma) = \mathrm{D}\psi(\eta, \theta)\,\mathrm{col}\left(\Phi(\eta, \gamma(\varsigma)),\ \mathrm{D}\gamma(\varsigma)\varphi(\varsigma, \xi_2, \eta)\right) \tag{15}$$

for all $(\xi_2, \eta, \varsigma) \in H^\star \times S^\star$. With this construction, since under A3 and the identifier requirement, ψ, $\mathrm{D}\psi$, φ, γ and $\mathrm{D}\gamma$ are locally Lipschitz, there exists $l_\psi > 0$ such that

$$|\dot\psi(\xi_2, \eta, \varsigma) - \dot\psi(\dot\eta_d^\star, \eta^\star, \varsigma^\star)| \leq l_\psi |(\xi_2 - \dot\eta_d^\star, \eta - \eta^\star, \varsigma - \varsigma^\star)| \tag{16}$$

for all $(\xi_2, \eta, \varsigma) \in \mathbb{R}^{n_e} \times \mathbb{R}^{dn_e} \times \mathcal{Z}_\varsigma$ and $(\dot\eta_d^\star, \eta^\star, \varsigma^\star) \in H^\star \times S^\star$.

The identifier (14) is thus composed of the two subsystems ξ and ς. The dynamics and output maps (φ, γ) of ς are designed to fulfil the identifier requirement. When driven by the "ideal" input pair $(\dot\eta_d^\star, \eta^\star)$, the subsystem ς is supposed to have an attractive steady-state solution $\varsigma^\star$ along which its output $\theta^\star$ leads to the best model in the model set $\mathcal{M}$ according to (13). In addition, a robustness property, given in terms of input-to-state stability with respect to the additive inputs (δ', δ''), is required. This additional property is needed since $(\dot\eta_d^\star, \eta^\star)$ is not available for feedback and in (14) the system ς is instead driven by the input (ξ_2, η), the latter playing the role of a "proxy" for $(\dot\eta_d^\star, \eta^\star)$. While it is clear that η carries some information on $\eta^\star$, the fact that ξ_2 acts as a proxy of $\dot\eta_d^\star$ follows by the definition of ξ, which is indeed designed as a *derivative observer* of the derivative $\dot\eta_d$ of η_d, providing the missing information on $\dot\eta_d^\star$.

We stress that the ability to construct an identifier satisfying the requirement as indicated above hides the need of qualitative and quantitative knowledge on the ideal steady-state signals $\dot\eta_d^\star$, $\eta^\star$ and $\varsigma^\star$, as evident for instance in the definition of $S^\star$ and $H^\star$. We remark, however, that this information concerns high-level properties of the class $\mathcal{H}^\star$, such as a uniform bound on its elements, and not the precise knowledge of the actual $(\dot\eta_d^\star, \eta^\star)$. In Section III, pair (φ, γ) fulfilling the identifier requirement when the model $\psi(\cdot,\theta)$ is linearly parametrised and (13) is a least square functional is presented.

E. The asymptotic stability result

The overall regulator reads as follows

$$\begin{array}{rcl} \dot\eta &=& \Phi(\eta, \gamma(\varsigma)) + Ge \\ \dot\varsigma &=& \varphi(\varsigma, \xi_2, \eta) \\ \dot\xi_1 &=& \xi_2 - m_1\rho(\xi_1 - \eta_d) \\ \dot\xi_2 &=& \dot\psi(\xi_2, \eta, \varsigma) - m_2\rho^2(\xi_1 - \eta_d) \\ u &=& \mathcal{L}\big(K_\chi\chi + K_\varsigma\zeta + K_\eta\eta_1 + K_w\nu(x^\star, w)\big) \end{array} \tag{17}$$

We finally show that the design parameters (g, ℓ, κ, ρ) can be chosen so that the closed-loop system has an asymptotic regulation error that is bounded by a function of the best attainable prediction error. The result is precisely formulated in the following theorem.

Theorem 1 *Suppose that A1 and A2 hold, and consider the regulator* (17) *constructed in the previous sections with $\mathcal{H}^\star$ and ψ satisfying A3 and (φ, γ) fulfilling the identifier requirement relative to $\mathcal{H}^\star$ and a cost functional* (13)*. Suppose moreover that $(\dot\eta_d^\star, \eta^\star) \in \mathcal{H}^\star$ for all $\kappa > 1$ and $\ell > 1$. Then there exist c, $\rho^\star$, $g^\star(\rho)$, $\kappa^\star(g)$, $\ell^\star(g,\kappa) > 0$ such that, for all $\rho \geq \rho^\star$, $g \geq g^\star(\rho)$, $\kappa \geq \kappa^\star(g)$ and $\ell \geq \ell^\star(g,\kappa)$, every solution of the closed-loop system* (4)*,* (17) *satisfies*

$$\limsup_{t\to\infty} |e(t)| \leq \frac{c}{g^d} \limsup_{t\to\infty} |\varepsilon^\star(t, \theta^\star(t))|,$$

with c not dependent on the control parameters.

Theorem 1 is proved in the Appendix. Its claim is an *approximate* regulation result, which becomes *asymptotic* whenever $\varepsilon^\star(t, \theta^\star(t)) = 0$. This, in turn, happens when a "real" model exists and belongs to the chosen model set $\mathcal{M}$. As Assumption A3 and the identifier requirement imply that $\varepsilon^\star$ can be bounded uniformly in the control parameters, the claim of the theorem is also a *practical* regulation result, with the bound on the regulation error that can be reduced arbitrarily by increasing g. Finally, we remark that, if a "saturated version" of ψ is implemented in the internal model unit (6) in place of ψ (for instance by saturating ψ on $H^\star \times \gamma(S^\star)$ in the same way as it is done in (15) for $\dot\psi$), and if $(\dot\eta_d^\star, \eta^\star)$ is bounded uniformly in the control parameters (which is always true in the square case as remarked in Section II-C), then a *practical* regulation result is still preserved[3] also in the case in which $(\dot\eta_d^\star, \eta^\star) \notin \mathcal{H}^\star$, thus paralleling the "canonical" pre-processing results (see e.g. [4,22]). In this case, however, the asymptotic bound on $e(t)$ cannot be related to $\varepsilon^\star$ any more.

III. Continuous-Time Least Squares Identifiers

We develop here an example of a pair (φ, γ) that fulfils the identifier requirement when the model $\psi(\cdot, \theta)$ is a finite *linear combination of known functions* of the form[4]

$$\psi(\cdot, \theta) = \sum_{i=1}^{n_\theta} \theta_i \sigma_i(\cdot)\,, \tag{18}$$

in which $n_\theta \in \mathbb{N}$ is arbitrary and $\sigma_i : \mathbb{R}^d \to \mathbb{R}$ are known Lipschitz and bounded functions. In this case the model set $\mathcal{M}$ is the family of functions of the form $\sigma(\cdot)^\top\theta$, having defined $\sigma(\cdot) := \mathrm{col}(\sigma_1(\cdot), \ldots, \sigma_{n_\theta}(\cdot))$ and $\theta := \mathrm{col}(\theta_1, \ldots, \theta_{n_\theta})$. We associate with $\mathcal{M}$ the following cost functional, obtained by letting in (13) $c_\varepsilon(t, s, \cdot) := \exp(-\lambda(t-s))|\cdot|^2$ and $c_r(\theta) := \theta^\top\Gamma\theta$, with $\lambda > 0$ and $\Gamma \in \mathbb{R}^{n_\theta \times n_\theta}$ symmetric and positive semi-definite

$$\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}(t)(\theta) = \lambda \int_0^t e^{-\lambda(t-s)} \left|\varepsilon^\star(s, \theta)\right|^2 ds + \theta^\top\Gamma\theta \tag{19}$$

in which the prediction error (12) at time s reads as

$$\varepsilon^\star(s, \theta) := \dot\eta_d^\star(s) - \sigma(\eta^\star(s))^\top\theta\,.$$

The optimisation problem associated with (19) is recognised to be a (weighted) least squares problem with regularisation, in which λ and Γ play the role of the forgetting factor and the regulariser respectively. Namely, except for the regularisation term, minimising (19) means minimising a weighted squared "norm" of the prediction errors associated with all the past data.

With $\mathcal{SP}_{n_\theta}$ the space of symmetric positive semi-definite matrices in $\mathbb{R}^{n_\theta \times n_\theta}$, we let $\mathcal{Z}_\varsigma := \mathcal{SP}_{n_\theta} \times \mathbb{R}^{n_\theta}$ and, by partitioning the state as $\varsigma = (\varsigma_1, \varsigma_2)$, with $\varsigma_1 \in \mathcal{SP}_{n_\theta}$ and $\varsigma_2 \in \mathbb{R}^{n_\theta}$, we equip $\mathcal{Z}_\varsigma$ with the norm $|\varsigma| := |\varsigma_1| + |\varsigma_2|$. We thus construct a pair (φ, γ) satisfying the identifier requirement relative to (19) as follows

$$\begin{array}{rcll} \dot\varsigma_1 & = & -\lambda\varsigma_1 + \lambda\sigma(\eta)\sigma(\eta)^\top & \\ \dot\varsigma_2 & = & -\lambda\varsigma_2 + \lambda\sigma(\eta)\xi_2 & \varsigma \in \mathcal{Z}_\varsigma \\ \theta & = & (\varsigma_1 + \Gamma)^{-1}\varsigma_2, & \end{array} \tag{20}$$

The claim is formalized by the following proposition.

Proposition 1 *With $c > 0$ arbitrary, let $\mathcal{H}^\star$ be a class of locally integrable functions $(\dot\eta_d^\star, \eta^\star) : \mathbb{R}_{\geq 0} \to \mathbb{R} \times \mathbb{R}^d$ satisfying $|(\dot\eta_d^\star, \eta^\star)|_\infty \leq c$. Then, if $\Gamma > 0$, the pair (φ, γ) constructed in* (20) *satisfies the identifier requirement relative to $\mathcal{H}^\star$ and the least-squares functional (18) with $\beta_\varsigma(s,t) = s\exp(-\lambda t)$.*

Proof: As σ is Lipschitz and bounded, then $\varphi(\varsigma, \xi_2, \eta) := (-\lambda\varsigma_1 + \lambda\sigma(\eta)\sigma(\eta)^\top, -\lambda\varsigma_2 + \lambda\sigma(\eta)\xi_2)$ is locally Lipschitz. Pick an eigenvalue $\epsilon(t)$ of $\varsigma_1(t) + \Gamma$, and let $v(t) \neq 0$ be a corresponding eigenvector. Then $v(t)^\top(\varsigma_1(t) + \Gamma)v(t) = \epsilon(t)|v(t)|^2$, and since $\Gamma > 0$ and $\varsigma_1(t) \in \mathcal{SP}_{n_\theta}$, this implies $\epsilon(t) \geq p$, with $p > 0$ the smallest eigenvalue of Γ. Thus $\varsigma_1 + \Gamma$ is invertible and the singular values of $(\varsigma_1 + \Gamma)^{-1}$ are bounded by p^{-1}, which implies that $\gamma(\varsigma) := (\varsigma_1 + \Gamma)^{-1}\varsigma_2$ is Lipschitz, smooth in ς and, as a consequence, its derivative is locally Lipschitz.

Pick now $\xi_2 = \dot\eta_d^\star + \delta'$ and $\eta = \eta^\star + \delta''$, with $(\dot\eta_d^\star, \eta^\star) \in \mathcal{H}^\star$ and (δ', δ'') locally integrable. Forward completeness follows by noticing that (20) is a stable linear system driven by the locally integrable input $(\sigma(\eta)\sigma(\eta)^\top, \sigma(\eta)\xi_2)$ and that, as $\sigma(\eta)\sigma(\eta)^\top \in \mathcal{SP}_{n_\theta}$, then $\mathcal{SP}_{n_\theta}$ is forward invariant for ς_1. With $\Sigma(\eta^\star, \delta'') := \sigma(\eta^\star + \delta'')\sigma(\eta^\star + \delta'')^\top$ and $\pi(\eta^\star, \dot\eta_d^\star, \delta', \delta'') := \sigma(\eta^\star + \delta'')(\dot\eta_d^\star + \delta')$, define

$$\begin{aligned} \varsigma_1^\star(t) &:= \lambda \int_0^t e^{-\lambda(t-s)}\Sigma(\eta^\star(s), 0)ds \\ \varsigma_2^\star(t) &:= \lambda \int_0^t e^{-\lambda(t-s)}\pi(\eta^\star(s), \dot\eta_d^\star(s), 0, 0)ds, \end{aligned}$$

and let $\varsigma^\star = (\varsigma_1^\star, \varsigma_2^\star)$. If $|(\dot\eta_d^\star, \eta^\star)| \leq c$ for some $c > 0$, then clearly there exists $c' > 0$ such that $\varsigma^\star(t) \in S^\star := \{\varsigma \in \mathcal{Z}_\varsigma : |\varsigma| \leq c'\}$. Furthermore, since σ is Lipschitz and bounded, there exists $l_\sigma > 0$ (possibly depending on c) such that $|\Sigma(\eta^\star, \delta'') - \Sigma(\eta^\star, 0)| \leq l_\sigma|\delta''|$ and $|\pi(\eta^\star, \dot\eta_d^\star, \delta', \delta'') - \pi(\eta^\star, \dot\eta_d^\star, 0, 0)| \leq l_\sigma|(\delta', \delta'')|$ for all $(\delta', \delta'') \in \mathbb{R} \times \mathbb{R}^d$. Hence, by integration of (20), and using $\varsigma_1^\star(0) = 0$, we obtain

$$|\varsigma_1(t) - \varsigma_1^\star(t)| \leq e^{-\lambda t}|\varsigma_1(0) - \varsigma_1^\star(0)| + l_\sigma|(\delta', \delta'')|_{[0,t)},$$

and a similar bound holds for $|\varsigma_2(t) - \varsigma_2^\star(t)|$, thus implying the first item of the identifier requirement with $\beta_\varsigma(s,t) = s\exp(-\lambda t)$ and with $\alpha_\varsigma = 2l_\sigma$. For fixed $t \in \mathbb{R}_{\geq 0}$, differentiating (19) with respect to θ yields $\mathrm{D}_\theta\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}(t)(\theta) = 2((\varsigma_1^\star(t) + \Gamma)\theta - \varsigma_2^\star(t))$. Since, $\vartheta^\circ_{(\dot\eta_d^\star, \eta^\star)}(t) := \{\theta \in \mathbb{R}^{n_\theta} : \mathrm{D}_\theta\mathcal{J}_{(\dot\eta_d^\star, \eta^\star)}(t)(\theta) = 0\}$, then $\theta^\star(t) = (\varsigma_1^\star(t) + \Gamma)^{-1}\varsigma_2^\star(t) \in \vartheta^\circ_{(\dot\eta_d^\star, \eta^\star)}(t)$, which is the second item of the requirement, thus concluding the proof. ∎

We observe that the regularisation matrix $\Gamma > 0$ plays a fundamental role in Proposition 1, as it ensures that $\varsigma_1 + \Gamma$ is uniformly nonsingular. However, its presence frustrates the possibility of having asymptotic regulation also when the "right" internal model belongs to the model set (18). As evident in (19), indeed, having $\Gamma \geq 0$

[3]This can be deduced by the proof of Theorem 1 by neglecting the identifier's dynamics and by noticing that, in (21), $\tilde\psi(\tilde\eta, \tilde\varsigma, \eta^\star, \varsigma^\star) - \varepsilon^\star = \psi(\eta, \gamma(\varsigma)) - \dot\eta_d^\star$ can be bounded uniformly in ς.

[4]For ease of exposition we present here the case in which $n_e = 1$, with the remark that an identifier of the same kind for $n_e > 1$ can be always obtained as the composition of n_e single-variable identifiers.

means that, even if θ annihilates the prediction error $\varepsilon^\star$, and thus the integral term of (19), it also produces a positive addend $\theta^\top\Gamma\theta$, thus possibly making such θ a non-stationary point of $\mathcal{J}_{(\dot\eta_d^\star,\eta^\star)}(t)$. In this case, θ approaches a neighbourhood of $\theta^\star(t)$ of a size that depends on the maximum eigenvalue of Γ that, however, can be taken as small as desired. Nevertheless, Γ can be chosen positive semi-definite (and possibly zero). In this case, (7) can still be used by substituting the inverse operator with a pseudo-inverse (indeed $\sigma_1+\Gamma$ needs not be invertible in this case), and the claim of Proposition 1 applies only if the minimum non-zero singular value of $\varsigma_1+\Gamma$ is bounded away from zero uniformly in t, which can be seen as a *persistence of excitation condition*. We also remark that, in this case, the Lipschitz constant of γ and its derivative becomes dependent on how large is the minimum non-zero singular value of $\varsigma_1+\Gamma$, thus making the result of Theorem 1 obtained for a certain value of the gains ρ, g, κ and ℓ, applicable only to the solutions carrying sufficient excitation.

IV. Example: Control of the VTOL

Consider the lateral $(\mathrm{p}_1,\mathrm{p}_2)$ and angular $(\mathrm{p}_3,\mathrm{p}_4)$ dynamics of a VTOL aircraft described by [23]

$$\begin{array}{rclcrcl}\dot{\mathrm{p}}_1 &=& \mathrm{p}_2 & & \dot{\mathrm{p}}_3 &=& \mathrm{p}_4\\ \dot{\mathrm{p}}_2 &=& \mathrm{d}(w)-\varrho\tan\mathrm{p}_3+v & & \dot{\mathrm{p}}_4 &=& Bu\end{array}$$

with $M>0$ the VTOL mass, $\varrho>0$ the gravitational constant and $B=2LJ^{-1}>0$, with $L>0$ the length of the wings and J the moment of inertia (typically uncertain). The input u is the force on the wingtips, v is a vanishing input taking into account the (controlled) vertical dynamics (not considered here) and $\mathrm{d}(w):=M^{-1}\mathrm{d}_0(w)$, with $\mathrm{d}_0(w)$ that is the lateral wind force disturbance. The control goal is to eliminate the wind action from the lateral position dynamics, i.e. the regulation error is defined as $e(t)=\mathrm{p}_1(t)$. We also suppose to have available for feedback the entire state, namely $y=\mathrm{p}$. Let w be generated by an exosystem of the form (2) and change variables as $\mathrm{p}\mapsto x:=(\chi,\zeta)$, with $\chi:=(\mathrm{p}_1,\,\mathrm{p}_2,\,-\varrho\tan\mathrm{p}_3+\mathrm{d}(w))$ and $\zeta:=L_s\mathrm{d}(w)-\varrho\mathrm{p}_4/(\cos\mathrm{p}_3)^2$. In the new coordinates the following equations hold

$$\begin{array}{rclcrcl}\dot\chi_1 &=& \chi_2 & & \dot\chi_3 &=& \zeta\\ \dot\chi_2 &=& \chi_3 & & \dot\zeta &=& q(w,x)+\Omega(w,x)u,\end{array}$$

in which[5] $\Omega(w,x):=-\varrho B/(\cos(\tan^{-1}(\mathrm{d}(w)-\chi_3)/\varrho)))^2$ and $q(w,x)$ properly defined. This system is in the form (4), with A1 trivially fulfilled (x_0 being absent) by $x^\star=0$ and $u^\star=\varrho B^{-1}(L_s^2\mathrm{d}(w)-2\mathrm{d(w)}^2L_s\mathrm{d(w)})/(\mathrm{d(w)}^2+\varrho^2)$, and A2 fulfilled on each compact set with $\mathcal{L}$ a negative number[6]. With (c_1,c_2,c_3) the coefficients of a Hurwitz polynomial and $\kappa,\ell>0$ design parameters, we fix the control law as

$$\begin{aligned}u=-\mathcal{L}\big(&c_1\ell\kappa^3(\mathrm{p}_1+\eta_1)+c_2\ell\kappa^2\mathrm{p}_2+c_3\ell\kappa(-\varrho\tan\mathrm{p}_3)\\&+\ell(-\varrho\mathrm{p}_4/\cos^2\mathrm{p}_3)\big),\end{aligned}$$

with η_1 the first state of the internal model fixed later. In the new coordinates (χ,ζ), this control law is of the form (8), with $\mathcal{K}_w=\ell(c_3\kappa\;\;1)$ and $\nu(x^\star,w)=\mathrm{col}(\mathrm{d}(w),\,L_s\mathrm{d}(w))$.

Regarding the design of the internal model unit, we observe that, by following Section II-C, $\Upsilon_{(\ell,\kappa)}(w)=Q(\ell,\kappa)D(w)$, in which $Q(\ell,\kappa):=(c_3/(c_1\kappa^2)\;\;1/(c_1\kappa^3)\;\;1/(c_1\ell\mathcal{L}\kappa^3))$ and $D(w)=\mathrm{col}(\mathrm{d}(w),\,L_s\mathrm{d}(w),\,-\Omega(w,0)^{-1}q(w,0))$. Thus, $\eta_i^\star=Q(\ell,\kappa)L_s^{i-1}D(w)$, $i=1,\dots,d$, and $\dot\eta_d^\star=Q(\ell,\kappa)L_s^dD(w)$. The form of Q and the fact κ and ℓ have large values show that the dominant elements in $\eta_i^\star$ and $\dot\eta_d^\star$ are $L_s^{i-1}\mathrm{d}(w)$ and $L_s^d\mathrm{d}(w)$, regardless the value of the dimension d of η. Now, suppose that $\mathrm{d}(w)$ consists of a single harmonic at an unknown frequency. The design of (d,ψ) and the identifier to reject $\mathrm{d}(w)$ is then carried out by considering a single oscillator as the model set, obtained with $d=2$, $n_\theta=2$, and $\psi(\eta,\theta):=\theta^\top\eta$. The adaptation phase, in turn, can be set up by using the least-squares identifier presented in Section III with $n_\theta=2$ and σ any bounded function satisfying $\sigma(\eta)=\eta$ in the region where $\mathrm{col}(\mathrm{d}(w),L_s\mathrm{d}(w))\cdot c_3/c_1\kappa^2$ is supposed to range.

[5]Recall that $\cos(\tan^{-1}(s))=1/\sqrt{s^2+1}$.

[6]In this respect, we observe that the ideal steady-state value of the measurements $(\mathrm{p}_3,\mathrm{p}_4)$ is given by $(\mathrm{p}_3^\star,\mathrm{p}_4^\star):=(\tan^{-1}(\mathrm{d}(w)/\varrho),\varrho L_s\mathrm{d}(w)/(\mathrm{d}(w)^2+1))$, and thus y is not in general vanishing at the steady state.

V. Conclusions

The paper presented a post-processing design procedure for a class of multivariable nonlinear systems stabilisable by high-gain feedback hinging on a "non-equilibrium" framework. The internal model is adaptive with the adaptation mechanisms cast as an identification problem and with the asymptotic regulation error that is directly related to the identification error. The framework does not rely on an exact knowledge of the steady state friend, nor on an exact parametrisation of it. Rather, it assumes the knowledge of some qualitative\quantitative information about the class of steady state signals used to choose the model set of the underlying identification problem. The paper fits in the research direction of [1,14] in which approximate, rather than asymptotic, regulation is envisioned as the right perspective in presence of general uncertainties.

Appendix I
Proof of Theorem 1

With $\kappa>1$ and $\ell>1$, pick a solution $(x,\chi,\zeta,\eta,\varsigma,\xi)$ to the closed-loop system (4), (17) and let $(x^\star,u^\star,\eta^\star,\dot\eta_d^\star)$ be given by A1 and Section II-C. Assume that $(\dot\eta_d^\star,\eta^\star)\in\mathcal{H}^\star$, and let $(\varsigma^\star,\theta^\star)$ be produced by the identifier requirement. Consider the following change of variables

$$\begin{array}{rclcrcl}\eta &\mapsto& \tilde\eta:=\eta-\eta^\star & & \varsigma &\mapsto& \tilde\varsigma:=\varsigma-\varsigma^\star\\ \chi &\mapsto& \tilde\chi:=\chi+C^\top\tilde\eta_1 & & \zeta &\mapsto& \tilde\zeta:=\zeta-K(\kappa)\tilde\chi\\ \xi &\mapsto& \tilde\xi:=\xi-\begin{pmatrix}\eta_d^\star\\ \psi(\eta^\star,\theta^\star)\end{pmatrix} & & e &\mapsto& \tilde e:=e+\tilde\eta_1,\end{array}$$

where we recall that $K(\kappa)$ is defined in (8) and $\tilde\eta_1\in\mathbb{R}^{n_e}$ represents the first n_e components of $\tilde\eta$. By definition of $\eta^\star$, $\dot\eta_i^\star=\eta_{i+1}^\star$, and in the new coordinates we obtain

$$\begin{array}{rcll}\dot{\tilde\eta}_i &=& \tilde\eta_{i+1}-h_ig^i\tilde\eta_1+g^ih_i\tilde e, & i=1,\dots,d-1\\ \dot{\tilde\eta}_d &=& -h_dg^d\tilde\eta_1+\tilde\psi(\tilde\eta,\tilde\varsigma,\eta^\star,\varsigma^\star)+h_dg^d\tilde e-\varepsilon^\star.\end{array}\tag{21}$$

with $\varepsilon^\star=\varepsilon^\star(t,\theta^\star)$ given by (12) and with $\tilde\psi(\tilde\eta,\tilde\varsigma,\eta^\star,\varsigma^\star):=\psi(\tilde\eta+\eta^\star,\gamma(\tilde\varsigma+\varsigma^\star))-\psi(\eta^\star,\gamma(\varsigma^\star))$ that, since A3 and the identifier requirement imply that ψ and γ are Lipschitz, fulfils $|\tilde\psi(\tilde\eta,\tilde\varsigma,\eta^\star,\varsigma^\star)|\le c_{\psi,\gamma}|(\tilde\eta,\tilde\varsigma)|$ for some $c_{\psi,\gamma}>0$ independent on the control parameters. Thus, standard high-gain arguments (see e.g. [24]) show that there exist $a_0,a_1,a_2,a_3>0$ and $g_0^\star>0$ such that for all $g\ge g_0^\star$ the following bound holds

$$\begin{aligned}|\tilde\eta_i(t)|\le{}& a_0g^{i-1}|\tilde\eta(0)|\mathrm{e}^{-a_1gt}+a_2g^{i-d-1}|(\tilde\varsigma,\varepsilon^\star)|_{[0,t)}\\&+a_3g^{i-1}|\tilde e|_{[0,t)}\end{aligned}\tag{22}$$

for all $t\in\mathbb{R}_{\ge0}$ and each $i=1,\dots,d$. Moreover, $\tilde\xi$ satisfies

$$\begin{array}{rcl}\dot{\tilde\xi}_1 &=& \tilde\xi_2-m_1\rho\tilde\xi_1+m_1\rho\tilde\eta_d-\varepsilon^\star\\ \dot{\tilde\xi}_2 &=& -m_2\rho^2\tilde\xi_1+\tilde\mu(\tilde\eta,\tilde\varsigma,\tilde\xi_2,\eta^\star,\varsigma^\star)+m_2\rho^2\tilde\eta_d,\end{array}$$

in which, since by A3 $(\dot\eta_d^\star,\eta^\star)\in\mathcal{H}^\star$ implies $(\dot\eta_d^\star(t),\eta^\star(t))\in H^\star$, and by the identifier requirement we have $\varsigma^\star(t)\in S^\star$, in view of (15) $\tilde\mu$ reads as $\tilde\mu(\tilde\eta,\tilde\varsigma,\tilde\xi_2,\eta^\star,\varsigma^\star):=\dot\psi(\tilde\xi_2+\psi(\eta^\star,\theta^\star),\tilde\eta+\eta^\star,\tilde\varsigma+\varsigma^\star)-$

$\dot{\psi}(\dot{\eta}_d^\star, \eta^\star, \varsigma^\star)$. Moreover, in view of (16), there exists $l_\psi > 0$ such that $|\tilde{\mu}(\tilde{\eta}, \tilde{\varsigma}, \tilde{\xi}_2, \eta^\star, \varsigma^\star)| \leq l_\psi\big(|(\tilde{\eta}, \tilde{\varsigma}, \tilde{\xi}_2)| + |\varepsilon^\star|\big)$. Hence, customary high-gain arguments show that there exist $\rho_0^\star > 1$ and $a_4, a_5, a_6 > 0$ such that, for all $\rho \geq \rho_0^\star$, the following holds

$$|\tilde{\xi}(t)| \leq a_4\rho|\tilde{\xi}(0)|\mathrm{e}^{-a_5\rho t} + a_6\left(\rho|\tilde{\eta}|_{[0,t)} + \rho^{-1}|\tilde{\varsigma}|_{[0,t)} + |\varepsilon^\star|_{[0,t)}\right). \tag{23}$$

We can write $\xi_2 = \dot{\eta}_d^\star + \delta'$ and $\eta = \eta^\star + \delta''$, with $\delta' := \tilde{\xi}_2 - \varepsilon^\star$ and $\delta'' := \tilde{\eta}$, so that the identifier requirement yields

$$|\tilde{\varsigma}(t)| \leq \beta_\varsigma(|\tilde{\varsigma}(0)|, t) + \alpha_\varsigma|(\tilde{\eta}, \tilde{\xi}, \varepsilon^\star)|_{[0,t)}. \tag{24}$$

In view of standard small-gain arguments (see e.g. [25]), the bounds (22), (23), (24) yield the existence of $\beta_1 \in \mathcal{KL}$, $a_7 > 0$, $\rho^\star \geq \rho_0^\star$ and $g^\star(\rho) \geq g_0^\star$ such that, for all $\rho > \rho^\star$ and $g \geq g^\star(\rho)$, we have

$$\begin{aligned}|(\tilde{\eta}(t), \tilde{\varsigma}(t), \tilde{\xi}(t))| &\leq \beta_1(|(\tilde{\eta}(0), \tilde{\varsigma}(0), \tilde{\xi}(0))|, t)\\ &\quad + a_7\big(g^{d-1}|\tilde{e}|_{[0,t)} + |\varepsilon^\star|_{[0,t)}\big)\\ |\tilde{\eta}_i(t)| &\leq \beta_1(|(\tilde{\eta}(0), \tilde{\varsigma}(0), \tilde{\xi}(0))|, t)\\ &\quad + a_7\big(g^{i-1}|\tilde{e}|_{[0,t)} + g^{i-1-d}|\varepsilon^\star|_{[0,t)}\big).\end{aligned} \tag{25}$$

By noticing that $\tilde{e} = C\tilde{\chi}$, differentiating $\tilde{\chi}$ yields

$$\dot{\tilde{\chi}} = (F + HK(\kappa) + gh_1C^\top C)\tilde{\chi} + H\tilde{\zeta} + C^\top(\tilde{\eta}_2 - gh_1\tilde{\eta}_1),$$

so that, in view of (9), quite standard high-gain arguments (see e.g. [20]) show that there exists $\kappa_0^\star(g) > 1$ such that, for all $\kappa > \kappa_0^\star(g)$ the following hold

$$\begin{aligned}|\tilde{\chi}(t)| &\leq a_9(\kappa)|\tilde{\chi}(0)|\mathrm{e}^{-a_{10}\kappa t} + \frac{a_{11}}{\kappa}|\tilde{\zeta}|_{[0,t)}\\ &\quad + a_{12}(\kappa)\big(g|\tilde{\eta}_1|_{[0,t)} + |\tilde{\eta}_2|_{[0,t)}\big)\\ |\tilde{e}(t)| &\leq a_9(\kappa)|\tilde{\chi}(0)|\mathrm{e}^{-a_{10}\kappa t} + \frac{a_{11}}{\kappa}|\tilde{\zeta}|_{[0,t)}\\ &\quad + \frac{a_{13}}{\kappa}\big(g|\tilde{\eta}_1|_{[0,t)} + |\tilde{\eta}_2|_{[0,t)}\big)\end{aligned} \tag{26}$$

for some $a_9(\kappa), a_{10}, a_{11}, a_{12}(\kappa), a_{13} > 0$. Furthermore, in the new coordinates, the control law (8) becomes $u = -\ell\mathcal{L}\tilde{\zeta} - \mathcal{L}\big(\Omega(w, x^\star)\mathcal{L}\big)^{-1}q(w, x^\star)$, and differentiating $\tilde{\zeta}$ yields

$$\dot{\tilde{\zeta}} = \delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) + \tilde{\phi}(w, x, x^\star) - \ell\Omega(w, x)\mathcal{L}\tilde{\zeta} \tag{27}$$

with $\delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) := -K(\kappa)((F + HK(\kappa) + gh_1C^\top C)\tilde{\chi} + H\tilde{\zeta} + C^\top(\tilde{\eta}_2 - h_1g\tilde{\eta}_1))$ that satisfies $|\delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta})| \leq a_{14}(\kappa, g)|(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta})|$, for some $a_{14}(g, \kappa) > 0$, and with $\tilde{\phi}(w, x, x^\star) := \Omega(w, x)\mathcal{L}((\Omega(w, x)\mathcal{L})^{-1}q(w, x) - (\Omega(w, x^\star)\mathcal{L})^{-1}q(w, x^\star))$ that, in view of A2 and since $|\chi| \leq |\tilde{\chi}| + |\tilde{\eta}_1|$ and $|\zeta| \leq |\tilde{\zeta}| + |K(\kappa)\tilde{\chi}|$, satisfies $|\tilde{\phi}(w, x, x^\star)| \leq a_{15}(\kappa)|(\tilde{x}_0, \tilde{\chi}, \tilde{\zeta}, \tilde{\eta}_1)|$, for some $a_{15}(\kappa) > 0$ and with $\tilde{x}_0 := x_0 - x_0^\star$. Hence, usual high-gain arguments show that, under A2, there exists an $\ell_0^\star(\kappa, g) > 0$ such that, for all $\ell > \ell_0^\star(\kappa, g)$ the following bound holds

$$\begin{aligned}|\tilde{\zeta}(t)| \leq& a_{16}|\tilde{\zeta}(0)|\mathrm{e}^{-a_{17}\ell t} + \frac{a_{18}(\kappa)}{\ell}|\tilde{\chi}|_{[0,t)}\\ &+ \frac{a_{19}}{\ell}|\tilde{x}_0|_{[0,t)} + \frac{a_{20}(\kappa)}{\ell}\big(g|\tilde{\eta}_1|_{[0,t)} + |\tilde{\eta}_2|_{[0,t)}\big)\end{aligned} \tag{28}$$

for some $a_{16}, a_{17}, a_{18}(\kappa), a_{19}, a_{20}(\kappa) > 0$. Furthermore, by noticing that $|\chi| \leq |\tilde{\chi}| + |\tilde{\eta}_1|$ and $|\zeta| \leq |\tilde{\zeta}| + |K(\kappa)\tilde{\chi}|$, A1 yields the existence of $b_2, b_3(\kappa) > 0$ such that

$$|\tilde{x}_0(t)| \leq \beta_0(|\tilde{x}_0(0)|, t) + b_2|(\tilde{\eta}_1, \tilde{\zeta})|_{[0,t)} + b_3(\kappa)|\tilde{\chi}|_{[0,t)}. \tag{29}$$

Therefore, in view of (25), (26), (28) and (29), repeating the small-gain arguments of [25] yields the existence of a $\kappa^\star(g) \geq \kappa_0^\star(g)$ and of an $\ell^\star(\kappa, g) \geq \ell_0^\star(\kappa, g)$ such that, for each $\rho > \rho^\star$, $g \geq g^\star(\rho)$, $\kappa \geq \kappa^\star(g)$, and $\ell \geq \ell^\star(g, \kappa)$, it holds that

$$|(\tilde{x}(t), \tilde{\eta}(t), \tilde{\varsigma}(t), \tilde{\xi}(t))| \leq \beta(|(\tilde{x}(0), \tilde{\eta}(0), \tilde{\varsigma}(0), \tilde{\xi}(0))|, t)$$
$$+ p_1(\kappa, g)|\varepsilon^\star|_{[0,t)}$$

and

$$\limsup_{t\to\infty}|\tilde{e}(t)| \leq p_2\kappa^{-1}g^{1-d}\limsup_{t\to\infty}|\varepsilon^\star(t)|$$
$$\limsup_{t\to\infty}|\tilde{\eta}_1(t)| \leq p_3g^{-d}\limsup_{t\to\infty}|\varepsilon^\star(t)|$$

for some $\beta \in \mathcal{KL}$ and $p_1(\kappa, g), p_2, p_3 > 0$, and the result follows with $c = p_2 + p_3$ by noticing that $|e| \leq |\tilde{e}| + |\tilde{\eta}_1|$ and that, since $\kappa^\star(g)$ can be taken to be larger than g, then $\kappa^{-1} \leq g^{-1}$.

REFERENCES

[1] M. Bin, D. Astolfi, L. Marconi, and L. Praly, "About robustness of internal model-based control for linear and nonlinear systems," in *2018 IEEE 57th Conf. Decision and Control*, 2018.

[2] C. I. Byrnes and A. Isidori, "Limit sets, zero dynamics and internal models in the problem of nonlinear output regulation," *IEEE Trans. Autom. Control*, vol. 48, pp. 1712–1723, 2003.

[3] J. Huang and C. F. Lin, "On a robust nonlinear servomechanism problem," *IEEE Trans. Autom. Control*, vol. 39, no. 7, pp. 1510–1513, 1994.

[4] C. I. Byrnes and A. Isidori, "Nonlinear internal models for output regulation," *IEEE Trans. Autom. Control*, vol. 49, pp. 2244–2247, 2004.

[5] F. D. Priscoli, L. Marconi, and A. Isidori, "A new approach to adaptive nonlinear regulation," *SIAM J. Control Optim.*, vol. 45, no. 3, pp. 829–855, 2006.

[6] L. Marconi, L. Praly, and A. Isidori, "Output stabilization via nonlinear Luenberger observers," *SIAM J. Control Optim.*, vol. 45, pp. 2277–2298, 2007.

[7] A. Serrani, A. Isidori, and L. Marconi, "Semiglobal nonlinear output regulation with adaptive internal model," *IEEE Trans. Autom. Control*, vol. 46, no. 8, pp. 1178–1194, 2001.

[8] A. Isidori and L. Marconi, "Shifting the internal model from control input to controlled output in nonlinear output regulation," in *2012 IEEE 51st Conference on Decision and Control*, 2012, pp. 4900–4905.

[9] M. Bin and L. Marconi, "About a post-processing design of regression-like nonlinear internal models," in *20th IFAC World Congress*, Toulouse, France, Jul. 2017.

[10] L. Wang, A. Isidori, H. Su, and L. Marconi, "Nonlinear output regulation for invertible nonlinear MIMO systems," *Int. J. Robust Nonlinear Control*, vol. 26, pp. 2401–2417, 2016.

[11] E. J. Davison, "The robust control of a servomechanism problem for linear time-invariant multivariable systems," *IEEE Trans. Autom. Control*, vol. 21, pp. 25–34, Feb. 1976.

[12] M. Bin and L. Marconi, "The chicken-egg dilemma and the robustness issue in nonlinear output regulation with a look towards adaptation and universal approximators," in *2018 IEEE 57th Conference on Decision and Control*, 2018.

[13] D. Astolfi and L. Praly, "Integral action in output feedback for multi-input multi-output nonlinear systems," *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 1559–1574, 2017.

[14] D. Astolfi, L. Praly, and L. Marconi, "Approximate regulation for nonlinear systems in presence of periodic disturbances," in *Proc. 54th IEEE Conference on Decision and Control*, 2015, pp. 7665–7670.

[15] M. R. Hirschorn, "Invertibility of multivariable nonlinear control systems," *IEEE Trans. Autom. Control*, vol. 24, no. 6, pp. 855–865, 1979.

[16] S. N. Singh, "A modified algorithm for invertibility of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 26, no. 2, pp. 595–598, 1981.

[17] L. Wang, A. Isidori, and H. Su, "Global stabilization of a class of invertible MIMO nonlinear systems," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 616–631, 2015.

[18] A. R. Teel and L. Praly, "Tools for semiglobal stabilization by partial state and output feedback," *SIAM J. Control Optim.*, vol. 33, pp. 1443–1488, 1995.

[19] D. Liberzon, A. S. Morse, and E. D. Sontag, "Output-input stability and minimum-phase nonlinear systems," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 422–436, 2002.

[20] A. Isidori, *Nonlinear Control Systems II*. London, Great Britain: Springer-Verlag, 1999.

[21] L. Ljung, *System Identification. Theory for the user*. Prentice Hall, 1999.

[22] A. Isidori, L. Marconi, and L. Praly, "Robust design of nonlinear internal models without adaptation," *Automatica*, vol. 48, pp. 2409–2419, 2012.

[23] A. Isidori, L. Marconi, and A. Serrani, *Robust Autonomous Guidance. An internal Model Approach*. Springer-Verlag, 2003.

[24] H. K. Khalil, *Nonlinear Systems. Third Edition*. Prentice Hall, 2002.
[25] Z.-P. Jiang, A. R. Teel, and L. Praly, "Small-gain theorem for ISS systems and applications," *Math. Control Signals Systems*, vol. 7, pp. 95–120, 1994.