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

© 2020 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

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.

I. INTRODUCTION

We consider nonlinear systems of the form

\[ \dot{x} = f(w, x, u), \quad y = h(w, x), \quad e = h_e(w, x) \]  

with state \( x \in \mathbb{R}^n \), control input \( u \in \mathbb{R}^m \), measured outputs \( y \in \mathbb{R}^p \), “regulation error” \( e \in \mathbb{R}^n \), and with \( w \in \mathbb{R}^n \) an exogenous signal generated by the “exosystem”

\[ \dot{w} = s(w). \]  

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

\[ \dot{x}_e = f_e(x_e, y), \quad u = k_e(x_e, y) \]  

achieving the regulation objective \( \lim_{t \to \infty} |e(t)| \leq \epsilon \), 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 \( \epsilon \) can be reduced arbitrarily by opportunistically 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 \( (x^*(t), u^*(t)) \) solution of the so-called regulator equations

\[ \dot{x}^* = f(w, x^*, u^*), \quad 0 = h_e(w, x^*). \]  

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

Regulator structures proposed in the nonlinear context are typically composed by two units, an internal model unit and a stabilizing 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 \( u^*(t) \) required to keep the error at zero in steady state, while the latter is designed to steer the system trajectories to \( x^*(t) \). What makes the design problem particularly challenging is, of course, the fact that \( (x^*, u^*) \) 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 works 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 \( (x^*, u^*) = (\pi(w), c(w)) \) for some \( \pi \) 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 \( u^* \) 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. [7]). This way of proceeding, however, involves a “quantitative” information about how the uncertainties
reflect on the friend that are hard to assume, unless sub-
stantially 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 approxi-
mate 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]. 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], 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,
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 \( x =
\text{col}(x_0, \chi, \zeta) \in \mathbb{R}^{n_x} \) satisfying the following equations

\[ \dot{x}_0 = f_0(w, x) + b(w, x)u \tag{4a} \]

\[ \dot{\chi} = F \chi + H \zeta \tag{4b} \]

\[ \dot{\zeta} = q(w, x) + \Omega(w, x)u \tag{4c} \]

\[ e = C \chi, \quad y = \text{col}(\chi, \zeta), \tag{4d} \]

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_{\zeta}},
\( u \in \mathbb{R}^{n_u} \), with \( n_u \geq n_e \), \( \chi = \text{col}(\chi^1, \ldots, \chi^{n_{\chi}}) \), with
\( \chi^i \in \mathbb{R}^{n_{\chi^i}}, i = 1, \ldots, n_e \), and \( n_{\chi^i} + \ldots + n_{\zeta} = n_{\chi} \).
\( C := \text{blkdiag}(C_1, \ldots, C_{n_e}) \), \( F := \text{blkdiag}(F_1, \ldots, F_{n_e}) \) and
\( H := \text{blkdiag}(H_1, \ldots, H_n) \), with \( C_i := (1 \ 0_{1 \times (n_{\chi^i} - 1)}) \) and

\[ F_i := \begin{pmatrix} 0(n_{\chi^i} - 1) \times 1 \\ n_{\chi^i} \times 1 \end{pmatrix}, \quad H_i := \begin{pmatrix} 0(n_{\chi^i} - 1) \times 1 \\ 1 \end{pmatrix}. \]

The \( \chi \) subsystem, in particular, is described by \( n_e \) chains of integrators with \( \chi \) entering at the bottom and the regulation
error given by the first components \( \chi^1 \) of each chain \( \chi \). Hence, \( \chi \) and \( \zeta \) are linear combinations of the error and its
time derivatives. The functions \( f_0, b, q \) and \( \Omega \) are sufficiently
smooth functions, with \( \Omega(w, x) \in \mathbb{R}^{n_y \times n_{\zeta}} \) 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 re-
spect 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 \( \nu \) 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 \( \chi \) 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. \( x = x_0 \)).

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

\[ \dot{x}_0^* = f_0(w, x^*) + b(w, x^*)u^* \]

\[ 0 = q(w, x^*) + \Omega(w, x^*)u^* \tag{5} \]

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

\[ |x_0(t) - x_0^*(t)| \leq \beta_0(|x_0(0) - x_0^*(0)|, t) + \alpha_0(|\chi, \zeta|[0, t]) \]

for all \( t \geq 0 \).

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

\[ \mathcal{L}^\top \Omega(w, x)\mathcal{L} \geq I_{n_x} \]

for all \((w, x) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \), and the map \( (\Omega(\cdot)\mathcal{L})^{-1}q(\cdot) \) is
Lipschitz.
Finally, the (static) stabiliser is taken as
\[ u = \mathcal{L}(K_\chi \chi + K_\zeta \zeta + K_\eta \eta_1 + K_w \nu(x^*, w)), \] in which the matrices \( K_\chi, K_\zeta \) and \( K_\eta \) are chosen as follows
\[ K_\chi(\ell, \kappa) = \ell K(\kappa), \quad K_\zeta(\ell) = -\ell I_n, \quad K_\eta(\ell, \kappa) = \ell K(\kappa)C^T \]
with \( K(\kappa) = \text{blkdiag}(K^1(\kappa), \ldots, K^{n_\nu}(\kappa)) \), where
\[ K^i(\kappa) = \begin{pmatrix} c^i_1 \kappa^n & c^i_2 \kappa^{n-1} & \ldots & c^i_{n^\nu} \kappa \end{pmatrix} \] for \( i = 1, \ldots, n_\nu \), in which the coefficients \( c^i_j \) do so that the polynomials \( s^{n^\nu} + c^i_1 s^{n^\nu-1} + \cdots + c^i_{n^\nu} s + c^i_1, i = 1, \ldots, n_\nu \), are Hurwitz, and \( \ell, \kappa > 0 \) are design parameters to be fixed. The matrix \( K_w \) and the function \( \nu \) are introduced for sake of generality and are possibly zero. These terms could represent a “feedback” contribution added by the designer by employing possible knowledge of \( w \) and \( x^* \). Likewise, it could represent a term showing up in the normal form \( A^i \) after a preliminary feedback of available measurements that do not vanish in steady state. Similarly to the other matrices in \( B \), the gain matrix \( K_w \) can depend on \( \kappa \) and \( \ell \). The degrees of freedom left to be fixed at this stage are the dimension \( d \) and function \( \psi \) of the internal model unit \( \mathcal{L} \), the data \( (\bar{Z}, n_\eta, \mu, \omega) \) of the identifier \( I \), and the control parameters \( g, \ell \) and \( \kappa \).

C. Design of the internal model as prediction model

A key step in the regulator synthesis is the choice of the internal model \( \mathcal{L} \) and of its adaptation through the design of the identifier \( I \). 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 \((x^*, u^*)\) and the underlying dynamics. With an eye to the last equation of \( B \), we can write
\[ e(t) = \bar{e}(g)(\bar{\eta}_d(t) - \psi(\eta(t), \theta(t))) \] in which \( \bar{e}(g) := (h_d g^d)^{-1} \). Our design strategy to choose \((d, \psi)\) in \( B \) and the identifier \( I \) pivots around the idea that \( \bar{\eta}_d(t) - \psi(\eta(t), \theta(t)) \) can be interpreted as a “prediction error” attained by the “model” \( \psi \) in relating the “next derivative” \( \bar{\eta}_d(t) \) to the “previous derivatives” \( \eta(t) \) and that, by minimising this prediction error, the actual regulation error is also minimised due to \( D \). This clearly suggests to look at the problem of choosing \( d \) and \( \psi \) 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” \( \eta \) to the “output” \( \bar{\eta}_d \), and to set \( M := \{ \psi(\cdot, \theta) : \theta \in \mathbb{R}^{n_\theta} \} \) of all the possible candidate models as the corresponding model set. The choice of \( d \) and of \( \psi \) 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 \( \bar{\eta}_d \) and \( \eta \) are expected to belong. In this context, the steady-state signals \((x^*, u^*)\) resulting from the regulator equations \( B \) are the anchor point on which that knowledge can be drawn. In particular, let
\[ \bar{\eta}_1 := \Upsilon(\ell, \kappa)(w, x^*), \]
in which
\[ \Upsilon(\ell, \kappa)(w, x^*) := -(\Omega(w, x^*)\mathcal{L}K_\eta)^{-1}(q(w, x^*) + \Omega(w, x^*)\mathcal{L}K_w \nu(w, x^*)), \]
and define recursively $\eta_i^\star$, $i = 2, \ldots, d + 1$, as
$$
\eta_i^\star := L_{\kappa}\eta_{i-1}(w, x^\star) + L_{\ell}\eta_i(w, x^\star) + L_f\eta_i(w, x^\star),
$$
Finally let
$$
\hat{\eta}^\star_d := \eta_{d+1}^\star.
$$
In view of $A_2$ and the definition of $K_\eta$, 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 $\eta_i^\star$, $i = 1, \ldots, d + 1$, depend on the design parameters $\kappa$ and $\ell$ yet to be fixed. The dimension $d$ and the function $\psi$ should be then ideally chosen so that, with $\eta^\star = \text{col}(\eta_1^\star, \ldots, \eta_d^\star)$, the following holds
$$
\hat{\eta}^\star_d(t) = \psi(\eta^\star(t), \theta^\star(t)),
$$
for some “ideal” $\theta^\star(t) \in \mathbb{R}^{n_e}$. 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 $(d, \psi)$ so that (11) is fulfilled for all possible steady-state trajectories $(\tilde{\eta}^\star_d, \eta^\star)$, 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\(^2\) in which $(\tilde{\eta}^\star_d, \eta^\star)$ 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\(^2\). 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 $\psi$ is Lipschitz and differentiable with a locally Lipschitz derivative, and the Lipschitz constants do not dependent on $\kappa$ and $\ell$. Moreover, there exists a compact set $\mathcal{H}^\star \subset \mathbb{R}^{n_e} \times \mathbb{R}^{d\kappa\ell}$, independent on $\kappa$ and $\ell$, such that every $(\tilde{\eta}^\star_d, \eta^\star) \in \mathcal{H}^\star$ satisfies $(\tilde{\eta}^\star_d(t), \eta^\star(t)) \in H^\star$ 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 $\mathcal{H}^\star$, and that the inferred prediction model $\psi$ has some strong regularity properties uniform in the control gains $(\kappa, \ell)$. 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 $n_a = n_e$ in [4], the matrix $\Omega(w, x)$ is square and $\kappa$ and $\ell$ do no mix-up with $\Omega(w, x)$, $q(w, x)$ and $\nu(w, x)$ in the definition of $\hat{\eta}_d^\star$ and $\eta^\star$. Therefore $(\hat{\eta}_d^\star, \eta^\star)$ can be always bounded uniformly in $\kappa$ and $\ell$ whereas they are taken larger than 1 and $K_\omega/(\kappa\ell)$ can be bounded uniformly in $\kappa$ and $\ell$.

D. The design of the identifier

With $d$ and $\psi$ 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
$$
\varepsilon^\star(t, \theta) := \hat{\eta}_d^\star(t) - \psi(\eta^\star(t), \theta),
$$
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 $\eta^\star(t)$ 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 $C^0(\mathbb{R}^{n_e}, \mathbb{R}_{\geq 0})$ the space of continuous functions $\mathbb{R}^{n_e} \to \mathbb{R}_{\geq 0}$, with each pair $(\hat{\eta}_d^\star, \eta^\star) \in \mathcal{H}^\star$ we associate the map $J_{\hat{\eta}_d^\star, \eta^\star}(t)(\theta) : \mathbb{R}_{\geq 0} \to C^0(\mathbb{R}^{n_e}, \mathbb{R}_{\geq 0})$ given by
$$
J_{\hat{\eta}_d^\star, \eta^\star}(t)(\theta) := \int_0^t c_\varepsilon(s, t) \left\| \varepsilon^\star(s, \theta) \right\| ds + c_r(\theta),
$$
with $c_\varepsilon : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $c_r : \mathbb{R}^{n_e} \to \mathbb{R}_{\geq 0}$ 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 $\theta$ fits the historical data, while $c_r(\theta)$ plays the role of a regularisation factor. With $J_{\hat{\eta}_d^\star, \eta^\star}$ we associate the set-valued map $\partial^o_{\hat{\eta}_d^\star, \eta^\star}(t) : \mathbb{R}^{n_e} \to \mathbb{R}^{n_e}$ defined as
$$
\partial^o_{\hat{\eta}_d^\star, \eta^\star}(t)(\theta) := \text{argmin}_{\theta \in \mathbb{R}^{n_e}} J_{\hat{\eta}_d^\star, \eta^\star}(t)(\theta).
$$
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 $J_{\hat{\eta}_d^\star, \eta^\star}(t)$, i.e. satisfies $\theta^\star(t) \in \partial^o_{\hat{\eta}_d^\star, \eta^\star}(t)$ for all $t \geq 0$. In particular, the identifier (7) is chosen as a system with state
$$
z = \text{col}(\xi, c), \quad \xi \in \mathbb{R}^{n_e}, \quad c \in \mathbb{Z}_c,
$$
in which $\mathbb{Z}_c$ is a finite-dimensional normed vector space, $\mathcal{Z} = \mathbb{R}^{2n_e} \times \mathbb{Z}_c$, and the pair $(\mu, \omega)$ is chosen so that, with $\xi_1, \xi_2 \in \mathbb{R}^{n_e}$ such that $\xi = \text{col}(\xi_1, \xi_2)$, the equations (7) read as
$$
\begin{align*}
\dot{\xi}_1 &= \xi_2 - m_1 \rho (\xi_1 - \eta_d) \\
\dot{\xi}_2 &= \psi(\xi_2, \eta, c) - m_2 \rho^2 (\xi_1 - \eta_d) \\
\dot{c} &= \varphi(s, \xi_2, \eta) \\
\theta &= \gamma(c)
\end{align*}
$$
\(^\text{1}\)We denote by $L_g f$ the Lie derivative of $f$ along $g$.
\(^\text{2}\)Formally, $\mathcal{H}^\star$ is a subset of the space of functions $\mathbb{R}_{\geq 0} \to \mathbb{R}^{n_e} \times \mathbb{R}^{d\kappa\ell}$. 
where \(m_1, m_2 > 0\) are arbitrary, \(\rho > 0\) is a design parameter, \(\psi : \mathbb{R}^{n_x} \times \mathbb{R}^{dn_x} \times \mathbb{Z}_c \to \mathbb{R}^{n_x}\) is a function fixed below, and \((\varphi, \gamma)\) is chosen to satisfy the following requirement.

**Requirement 1 (Identifier Requirement).** The pair \((\varphi, \gamma)\) is said to satisfy the identifier requirement relative to a class \(\mathcal{H}^*\) and a cost functional \((13)\), if \(\varphi\) is locally Lipschitz, \(\gamma\) is Lipschitz and differentiable with locally Lipschitz derivative, and there exist \(\beta \in K\mathcal{L}\), a compact set \(S^* \subset \mathbb{Z}_c\), \(\alpha_c > 0\) and, for each \((\hat{\eta}_d, \eta^*) \in \mathcal{H}^*\), a unique \(\varsigma^* : \mathbb{R}_0^+ \to S^*\), such that:

- for every locally integrable \(\delta : \mathbb{R} \to \mathbb{R}^{n_x}\) and \(\delta^* : \mathbb{R} \to \mathbb{R}^{dn_x}\), all the maximal solutions to the system
  \[
  \dot{\varsigma}(t) = \varphi\left(\varsigma, \hat{\eta}_d^* + \delta^*, \eta^* + \delta^*\right)\]
  complete and satisfy
  \[
  |\varsigma(t) - \varsigma^*(t)| \leq \beta(t, |\varsigma(0) - \varsigma^*(0)|, t) + \alpha_c (|\delta^*|, |\delta^*|)(0, t)
  \]
  for all \(t \in \mathbb{R}_0^+\);
- the signal \(\theta^* : \mathbb{R}_0^+ \to \mathbb{R}\) satisfies \(\theta^*(t) \in \mathbb{D}_c\) for all \(t \in \mathbb{R}_0^+\).

With \(H^*\) and \(S^*\) the compact sets introduced, respectively, in \(A3\) and in the identifier requirement, and with \(D\psi\) and \(D\gamma\) denoting the Jacobian of \(\psi\) and \(\gamma\) respectively, we define \(\psi\) as any function satisfying

\[
\psi(\xi, \eta, \varsigma) = D\psi(\eta, \theta) \text{col}\left(\Phi(\eta, \gamma(\varsigma)), D\gamma(\varsigma)\varphi(\xi, \eta, \varsigma)\right)
\]

for all \((\xi, \eta, \varsigma) \in H^* \times S^*\). With this construction, since under \(A3\) and the identifier requirement, \(D\psi\), \(D\varphi\), \(\gamma\) and \(D\gamma\) are locally Lipschitz, and \(\psi\) is bounded, there exists \(L_\psi > 0\) such that

\[
|\psi(\xi, \eta, \varsigma) - \psi(\hat{\eta}_d, \eta^*, \varsigma^*)| \leq L_\psi (|\xi - \hat{\eta}_d, \eta - \eta^*, \varsigma - \varsigma^*|)
\]

for all \((\xi, \eta, \varsigma) \in \mathbb{R}^{n_x} \times \mathbb{R}^{dn_x} \times \mathbb{Z}_c\) and \((\hat{\eta}_d, \eta^*, \varsigma^*) \in H^* \times S^*\).

The identifier \((14)\) is thus composed of the two subsystems \(\xi\) and \(\varsigma\). The dynamics and output maps \((\varphi, \gamma)\) of \(\varsigma\) are designed to fulfill the identifier requirement. When driven by the “ideal” input pair \((\hat{\eta}_d, \eta^*)\), the subsystem \(\varsigma\) is supposed to have an attractive steady-state solution \(\varsigma^*\) along which its output \(\theta^*\) leads to the best model in the model set \(M\) according to \((13)\).

In addition, a robustness property, given in terms of input-to-state stability with respect to the additive inputs \((\delta, \delta^*)\), is required. This additional property is needed since \((\hat{\eta}_d, \eta^*)\) is not available by the input \((\xi, \eta, \varsigma)\), the latter playing the role of a “proxy” for \((\hat{\eta}_d, \eta^*)\). While it is clear that \(\eta\) carries some information on \(\eta^*\), the fact that \(\xi\) acts as a proxy of \(\hat{\eta}_d\) follows by the definition of \(\xi\), which is indeed designed as a derivative observer of the derivative \(\eta_d\) of \(\eta_d\), providing the missing information on \(\eta_d\).

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 \(\hat{\eta}_d^*\) and \(\varsigma^*\), as evident for instance in the definition of \(S^*\) and \(H^*\). We remark, however, that this information concerns high-level properties of the class \(\mathcal{H}^*\), such as a uniform bound on its elements, and not the precise knowledge of the actual \((\hat{\eta}_d, \eta^*)\). In Section \(III\) pair \((\varphi, \gamma)\) 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

\[
\begin{align*}
\dot{\eta} &= \Phi(\eta, \gamma(\varsigma)) + Ge \\
\dot{\varsigma} &= \varphi(\varsigma, \xi_2, \eta) \\
\xi_1 &= \xi_2 - m_1 \rho(\xi_1 - \eta_d) \\
\xi_2 &= \psi(\xi_2, \eta, \varsigma) - m_2 \rho^2(\xi_1 - \eta_d) \\
u &= \mathcal{L}(K\varsigma + K\varsigma + K\eta_d + K_w \nu(x^*, w))
\end{align*}
\]

We finally show that the design parameters \((g, \ell, \kappa, \rho)\) 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 \(\mathcal{H}^*\) and \(\psi\) satisfying \(A3\) and \((\varphi, \gamma)\) fulfilling the identifier requirement relative to \(\mathcal{H}^*\) and a cost functional \((13)\). Suppose moreover that \((\hat{\eta}_d, \eta^*) \in \mathcal{H}^*\) for all \(\kappa > 1\) and \(\ell > 1\). Then there exist \(c, \rho^*, g^*(\rho), \kappa^*(\rho), \ell^*(\kappa, \rho)\) such that, for all \(\rho \geq \rho^*, g \geq g^*(\rho), \kappa \geq \kappa^*(\rho), \ell \geq \ell^*(\kappa, \rho)\), every solution of the closed-loop system \((4)\), \((17)\) satisfies

\[
\limsup_{t \to \infty} |e(t)| \leq \frac{c}{g^d} \sup_{t \to \infty} |e^*(t, \theta^*(t))|
\]

with \(c\) not dependent on the control parameters.

Theorem \(II\) is proved in the Appendix. Its claim is an approximate regulation result, which becomes asymptotic whenever \(e^*(t, \theta^*(t)) = 0\). This, in turn, happens when a “real” model exists and belongs to the chosen model set \(M\). As Assumption \(A3\) and the identifier requirement imply that \(e^*\) 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 \(\psi\) is implemented in the internal model unit \((3)\) in place of \(\psi\) (for instance by saturating \(\psi\) on \(H^* \times \gamma(S^*)\) in the same way as it is done in \((13)\) for \(\psi\)), and if \((\hat{\eta}_d, \eta^*)\) is bounded uniformly in the control parameters (which is always true in the square case as remarked in Section \(II\)), then a practical regulation result is still preserved also in the case in which \((\hat{\eta}_d, \eta^*) \notin \mathcal{H}^*\), thus paraphrasing the “canonical” pre-processing results (see e.g. \((12)\)). In this case, however, the asymptotic bound on \(e(t)\) cannot be related to \(e^*\) any more.

**III. CONTINUOUS-TIME LEAST SQUARES IDENTIFIERS**

We develop here an example of a pair \((\varphi, \gamma)\) that fulfils the identifier requirement when the model \(\psi(\cdot, \theta)\) is a finite linear combination of known functions of the form

\[
\psi(\cdot, \theta) = \sum_{i=1}^{n_\theta} \theta_i \sigma_i(\cdot),
\]

\[(18)\]

3This can be deduced by the proof of Theorem \(II\) by neglecting the identifier’s dynamics and by noticing that, in \((21)\), \(\psi(\eta, \varsigma, \eta^*, \varsigma^*) - e^* = (\eta(\gamma(\varsigma)) - \hat{\eta}_d\) can be bounded uniformly in \(c\).

4For ease of exposition we present here the case in which \(n_c\) is not in the case in which \(n_c > 1\) with the remark that an identifier of the same kind for \(n_c > 1\) can always be obtained as the composition of \(n_c\) single-variable identifiers.
Proof. Relative to Lipschitz, smooth in \((20)\) omitted for reason of space.

 relative to (19) as follows is invertible and the singular values of can be simply obtained by saturating the expression in (20) on the compact “norm” of the prediction errors associated with all the past \(\epsilon\). With \(\mathcal{SP}_{\eta} \subseteq \mathbb{R}^{n_x \times n_x}\), we let \(Z_{\varsigma} := \mathcal{SP}_{\eta} \times \mathbb{R}^{n_x}\) and, by partitioning the template as \(\varsigma = (\varsigma_1, \varsigma_2)\), with \(\varsigma_1 \in \mathcal{SP}_{\eta}\) and \(\varsigma_2 \in \mathbb{R}^{n_x}\), we equip \(Z_{\varsigma}\) with the norm \(|\varsigma| := |\varsigma_1| + |\varsigma_2|\). We thus construct a pair \((\varphi, \gamma)\) satisfying the identifier requirement relative to (19) as follows

\[
\varsigma_1 = -\lambda \varsigma_1 + \lambda \sigma(\eta) \sigma(\eta)^T, \quad \gamma(\varsigma, \theta) = (\varsigma_1 + \Gamma)^{-1} \varsigma_2, \quad \varsigma \in Z_{\varsigma} \tag{20}
\]

The claim is formalized by the following proposition.

**Proposition 1.** With \(c > 0\) arbitrary, let \(\mathcal{H}^{*}\) be a class of locally integrable functions \((\hat{\theta}_d, \eta^*) : \mathbb{R}_\geq 0 \to \mathbb{R} \times \mathbb{R}^d\) satisfying \(|(\hat{\theta}_d, \eta^*)|_\infty \leq c\). Then, if \(\Gamma > 0\), the pair \((\varphi, \gamma)\) constructed in (20) satisfies the identifier requirement locally relative to \(\mathcal{H}^{*}\) and the least-squares functional (18) with \(\beta(s, t) = s \exp(-\lambda t)\) and with \(\alpha = 2\sigma\).

**Proof.** As \(\sigma\) is Lipschitz and bounded, then \(\varphi(\varsigma_1, \xi_2, \eta) := (-\lambda \varsigma_1 + \lambda \sigma(\eta) \sigma(\eta)^T, -\lambda \varsigma_2 + \lambda \sigma(\eta) \xi_2)\) is locally Lipschitz. Pick an eigenvalue \(\epsilon(t)\) of \(\varsigma_1(t) + \Gamma\), and let \(v(t) \neq 0\) be a corresponding eigenvector. Then \(v(t)^T (\varsigma_1(t) + \Gamma) v(t) = \epsilon(t) |v(t)|^2\), and since \(\Gamma > 0\) and \(\varsigma_1(t) \in \mathcal{SP}_{\eta}\), this implies \(\epsilon(t) \geq p\), with \(p > 0\) the smallest eigenvalue of \(\Gamma\). 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 locally Lipschitz, smooth in \(\varsigma\) and, as a consequence, its derivative is locally Lipschitz.

Pick now \(\xi_2 = \hat{\eta}_d^2 + \delta^t\) and \(\eta = \eta^* + \delta^T\), with \((\hat{\eta}_d^2, \eta^*) \in \mathcal{H}^{*}\) and \((\delta^t, \delta^T)\) locally integrable. Forward completeness follows by noticing that (20) is a stable linear system driven by the locally integrable input \((\sigma(\eta) \sigma(\eta)^T, \sigma(\eta)\xi_2)\) and that, as \(\sigma(\eta) \sigma(\eta)^T \in \mathcal{SP}_{\eta}\), then \(\mathcal{SP}_{\eta}\) is forward invariant for \(\varsigma_1\). With \(\Sigma(\eta^*, \delta^T) := \sigma(\eta^* + \delta^T \sigma(\eta^* + \delta^T)^T\) and \(\pi(\eta^*, \hat{\eta}_d^2, \delta^T, \delta^T) := \sigma(\eta^* + \delta^T) \hat{\eta}_d^2 + \delta^T\), define

\[
\varsigma_1^T(t) := \int_0^t e^{-\lambda(t-s)} \Sigma(\eta^*(s), 0)ds
\]

and let \(\varsigma^* := (\varsigma_1^*, \varsigma_2^*). If \(|(\hat{\eta}_d^2, \eta^*)| \leq c\) for some \(c > 0\), then clearly there exists \(\epsilon' > 0\) such that \(\varsigma^*(t) \in S^\varsigma := \{\varsigma \in Z_{\varsigma} : |\varsigma| \leq c\}'\). Furthermore, since \(\sigma\) is Lipschitz and bounded, there exists \(l_{\sigma} > 0\) (possibly depending on \(c\)) such that \(|\Sigma(\eta^*, \delta^T) - \Sigma(\eta^*, 0)| \leq l_{\sigma} |\delta^T|\) and \(|\pi(\eta^*, \hat{\eta}_d^2, \delta^T, \delta^T) - \pi(\eta^*, \hat{\eta}_d^2, 0, 0)| \leq l_{\sigma}(|\delta^T, \delta^T)|\) for all \((\delta^T, \delta^T)\) \(\in \mathbb{R}^d \times \mathbb{R}^d\). Hence, by integration of (20), and using \(\varsigma_1^T(0) = 0\), we obtain

\[
|\varsigma_1(t) - \varsigma_1^T(t)| \leq e^{-\lambda_{1,2}}|\varsigma(0) - \varsigma_1^T(0)| + l_{\sigma} |(\delta^T, \delta^T)|_{0, t},
\]

and a similar bound holds for \(|\varsigma_2(t) - \varsigma_2^T(t)|\), thus implying the first item of the identifier requirement with \(\beta(s, t) = s \exp(-\lambda t)\) and with \(\alpha = 2\sigma\).

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 \(\mathcal{A}\). As evident in (19), indeed, having \(\Gamma \geq 0\) means that, even if \(\theta\) annihilates the prediction error \(\epsilon^*\), and thus the integral term of (19), it also produces a positive addend \(\theta^T \Gamma \theta\), thus possibly making such \(\theta\) a non-stationary point of \(\mathcal{J}(\hat{\theta}_d, \eta^*)\). In this case, \(\theta\) approaches a neighbourhood of \(\theta^*\) of a size that depends on the maximum eigenvalue of \(\Gamma\) that, however, can be taken as small as desired. Nevertheless, \(\Gamma\) 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 \(\varsigma_1 + \Gamma\) 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 \(\gamma\) 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 \(\rho, g, \kappa\) and \(\ell\), applicable only to the solutions carrying sufficient excitation.

**IV. EXAMPLE: CONTROL OF THE VTOL**

Consider the lateral \((p_1, p_2)\) and angular \((p_3, p_4)\) dynamics of a VTOL aircraft described by (23)

\[
p_1 = p_2, \quad p_2 = d(w) - \rho \tan p_3 + v, \quad p_3 = p_4, \quad p_4 = Bu
\]
with \( q > 0 \) 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 \( d(w) := M^{-1}d_0(w) \), with \( d_0(w) \) the lateral wind force disturbance, and \( M > 0 \) the VTOL mass. The control goal is to eliminate the wind action from the lateral position dynamics, i.e. the regulation error is defined as \( e(t) = \eta_1(t) \). We also suppose to have available for feedback the entire state, namely \( y = p \). Let \( w \) be generated by an exosystem of the form (2) and change variables as \( p \mapsto x := (\chi, \zeta) \), with \( \chi := (p_1, p_2, -\vartheta \tan p_3 + d(w)) \) and \( \zeta := L_\vartheta d(w) - gp_4/(\cos p_3)^2 \).

In the new coordinates the following equations hold

\[
\begin{align*}
\dot{\chi}_1 &= \chi_2 \\
\dot{\chi}_2 &= \chi_3 = \zeta \\
\dot{\chi}_3 &= q(w, x) + \Omega(x, w, u),
\end{align*}
\]

in which \( \Omega(x, w, u) := -gB/(\cos(\tan^{-1}(d(w) - \chi_3/g)))^2 \) and \( q(w, x) \) properly defined. This system is in the form (4), with \( A_1 \) trivially fulfilled (\( \Omega \) being absent) by \( \vartheta = 0 \) and \( \nu(u, x, w) = g = 1/\vartheta \). We also let \( \eta_1 \) be the first state of the internal model fixed later. In the new coordinates \( \chi, \zeta \), this control law is of the form (8), with \( K_w = \ell (c_3 \kappa^2) \) and \( \nu(x^*, w) = \text{col}(d(w), L_\vartheta d(w)) \).

Regarding the design of the internal model unit, we observe that, by following Section 4, \( Y(\ell, x, w) = Q(\ell, \kappa)D(w) \), in which \( Q(\ell, \kappa) := (c_3/\ell c_1^2) 1/(c_1^2 \ell C_0) \) and \( D(w) = \text{col}(d(w), L_\vartheta d(w), -Ω(w, 0), q(w, 0)) \). Thus, \( \eta^*_\vartheta = Q(\ell, \kappa) L_\vartheta^{-1} D(w) \), \( \ell = 1, \ldots, d \), and \( \eta^*_\vartheta = Q(\ell, \kappa) L_\vartheta^{-1} D(w) \). The form of \( Q \) and the fact \( \kappa \) and \( \ell \) have large values show that the dominant elements in \( \eta^*_\vartheta \) and \( \eta^*_\vartheta \) are \( L_\vartheta^{-1} d(w) \) and \( L_\vartheta^{-1} d(w) \), regardless the value of the dimension \( d \) of \( \eta \). Now, by choosing \( d(w) \) consists of a single harmonic at an unknown frequency. The design \( (d, \psi) \) and the identifier to reject \( d(w) \) is then performed by considering a single oscillator as the model set, obtained with \( d = 2, \nu_\theta = 2 \), and \( \psi(\vartheta, \eta, \theta) := \theta^\vartheta \). The adaptation phase, in turn, can be set up by using the least-squares identifier presented in Section 2 with \( \nu_\theta = 2 \) and \( \sigma \) any bounded function satisfying \( \sigma(\eta) = \eta \) in the region where \( \text{col}(d(w), L_\vartheta d(w)) \cdot c_3/\ell c_1^2 \) is supposed to range.

V. CONCLUSIONS

The paper presented a post-processing design procedure for a class of multivariable nonlinear systems stabilizable 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 any 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 [11][12] in which approximate, rather than asymptotic, regulation is envisioned as the right perspective in presence of general uncertainties.

APPENDIX A

PROOF OF THEOREM A

With \( \kappa > 1 \) and \( \ell > 1 \), pick a solution \( (x, \chi, \zeta, \eta, \xi, \zeta) \) to the closed-loop system (4), (17) and let \( (x^*, u^*, \eta^*, \xi^*) \) be given by \( A_1 \) and Section 4-C. Assume that \( (\eta^*_\vartheta, \eta^*_\vartheta) \in H^* \), and let \( (\xi^*, \theta^*) \) be produced by the identifier requirement. Consider the following change of variables

\[
\begin{align*}
\eta &\mapsto \tilde{\eta} := \eta - \eta^*_\vartheta \\
\chi &\mapsto \tilde{\chi} := \chi + C^\tau \tilde{\eta} \\
\zeta &\mapsto \tilde{\zeta} := \zeta - K(k) \tilde{\chi} \\
\xi &\mapsto \tilde{\xi} := \xi - \left( \psi(\eta^*(\theta^*), \theta^*) \right) e \mapsto \tilde{e} := e + \tilde{\eta},
\end{align*}
\]

where we recall that \( K(k) \) is defined in (5) and \( \tilde{\eta} \in R^n \) represents the first \( n_e \) components of \( \tilde{\eta} \). By definition of \( \tilde{\eta} \), \( \tilde{\eta} = \eta^*_\vartheta + 1 \), and in the new coordinates we obtain

\[
\begin{align*}
\dot{\tilde{\eta}}_1 &= \tilde{\eta}_{i+1} - h_0 g_\vartheta \tilde{\eta}_{i+1} + g h_\vartheta e, \quad i = 1, \ldots, d - 1 \\
\dot{\tilde{\eta}}_d &= -h_0 g^6 \tilde{\eta}_d + \psi(\tilde{\eta}, \zeta, \xi^*, \eta^*) + h_0 g^6 e - \vartheta. \tag{21}
\end{align*}
\]

with \( \vartheta = \vartheta^*(t, \theta^*) \) given by (12) and with \( \psi(\tilde{\eta}, \zeta, \xi^*, \eta^*) := \psi(\tilde{\eta} + \eta^*, \gamma(\xi + \sigma^*)) - \psi(\tilde{\eta} + \gamma(\xi^*)) \) that, since \( A_3 \) and the identifier requirement imply that \( \psi \) and \( \gamma \) are Lipschitz, fulfills

\[
|\psi(\tilde{\eta}, \zeta, \xi^*, \eta^*)| \leq c_{\psi, \gamma}(\tilde{\eta}, \zeta) \tag{22}
\]

for all \( t \in R_0 > 0 \) and each \( i = 1, \ldots, d \). Moreover, \( \tilde{\xi} \) satisfies

\[
\begin{align*}
\dot{\tilde{\xi}}_1 &= \tilde{\xi}_2 - m_1 \nu \tilde{\xi}_1 + m_1 \nu \tilde{\eta}_d - \vartheta \\
\dot{\tilde{\xi}}_2 &= -m_2 \nu^2 \tilde{\xi}_1 + \tilde{\mu}(\tilde{\eta}, \tilde{\zeta}, \tilde{\xi}_2, \eta^*, \xi^*) + m_2 \nu \tilde{\eta}_d, \\
\end{align*}
\]

in which, since by \( A_3(\tilde{\eta}^*_\vartheta, \eta^*) \in H^* \) implies \( (\tilde{\eta}^*_\vartheta(t), \eta^*(t)) \in H^* \), and by the identifier requirement we have \( \xi^*(t) \in S^* \), in view of (15) \( \tilde{\mu} \) reads as \( \tilde{\mu}(\tilde{\eta}, \tilde{\zeta}, \tilde{\xi}_2, \eta^*, \xi^*) := \psi(\tilde{\xi}_2 + \psi(\eta^*(\theta^*), \tilde{\eta} + \eta^*, \tilde{\xi}^*), \tilde{\eta} + \eta^*, \tilde{\xi}^*, \xi^*) - \psi(\tilde{\eta}^*_\vartheta(t), \xi^*) \). Moreover, in view of (16), there exists \( L_\psi > 0 \) such that

\[
|\psi(\tilde{\eta}, \tilde{\zeta}, \tilde{\xi}_2, \eta^*, \xi^*)| \leq L_\psi (|\tilde{\eta}| + |\tilde{\zeta}| + |\tilde{\xi}_2| + |\xi^*|). \tag{23}
\]

Hence, customary high-gain arguments show that there exist \( a_0, a_1, a_2, a_3 > 0 \) such that, for all \( \rho \geq \rho_0^* \), the following holds

\[
|\tilde{\xi}(t)| \leq a_0 \rho^i(t)|\tilde{\xi}(0)|e^{-a_0 \rho^i} + a_2 g^4 - d - 1(|\xi^*|)|0, t| + a_3 g^4 - 1|0, t|. \tag{24}
\]

Recall that \( \cos(\tan^{-1}(\kappa)) = 1/\sqrt{\kappa^2 + 1} \).
In view of standard small-gain arguments (see e.g. [23]), the bounds (22), (23), (24) yield the existence of $\beta_1 \in K\mathcal{L}$, $a_7 > 0$, $p^* \geq p^0$ and $g^*(\rho) \geq g^*_0$ such that, for all $\rho > p^*$ and $g \geq g^*(\rho)$, we have

$$
|\tilde{\eta}(t)| \leq \beta_1((\tilde{\eta}(0), \zeta(0), \tilde{\xi}(0)), t) + \sigma(g^{*} - 1\|\tilde{e}(t)\|_{0,t}) + \varepsilon^*|_{0,t}
$$

for some $\beta > 0$, $\sigma > 0$, and $\varepsilon^* > 0$. Thus, for $g \geq g^*(\rho)$, we have

$$
|\tilde{\eta}(t)| \leq \beta_1((\tilde{\eta}(0), \zeta(0), \tilde{\xi}(0)), t) + \sigma(g^{*} - 1\|\tilde{e}(t)\|_{0,t}) + \varepsilon^*|_{0,t}.
$$

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

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

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

$$
|\tilde{\chi}(t)| \leq a_9(g)|\tilde{\chi}(0)|e^{-a_{10} \kappa t} + \frac{a_{11}(\kappa)}{\kappa} |\tilde{\eta}_1(0)|_{0,t} + \frac{a_{12}(\kappa)}{\kappa} |\tilde{\eta}_2(0)|_{0,t}
$$

and

$$
|\tilde{e}(t)| \leq a_9(g)|\tilde{\chi}(0)|e^{-a_{10} \kappa t} + \frac{a_{11}(\kappa)}{\kappa} |\tilde{\eta}_1(0)|_{0,t} + \frac{a_{12}(\kappa)}{\kappa} |\tilde{\eta}_2(0)|_{0,t}
$$

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

$$
\dot{\tilde{\chi}} = \delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta}) + \phi(w, x, x^*) - \varepsilon^* L(\Omega(w, x) L^{-1}q(w, x^*))
$$

with

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

satisfying $|\delta(\tilde{\eta}, \tilde{\chi}, \tilde{\zeta})| \leq a_{14}(\kappa, g) |\tilde{\eta}_1, \tilde{\chi}, \tilde{\zeta}|$, for some $a_{14}(\kappa, g) > 0$, and with $\phi(w, x, x^*) := \Omega(w, x) L(\Omega(w, x) L^{-1}q(w, x) - (\Omega(w, x^*) L^{-1}q(w, x^*)$ that, in view of A2, and since $|\chi| \leq |\tilde{\chi}| + |\tilde{\eta}_1|$ and $|\zeta| \leq |\tilde{\xi}| + |K(\kappa)\tilde{\chi}|$, satisfies $|\phi(w, x, x^*)| \leq a_{15}(\kappa, g, \tilde{\eta}_1, \tilde{\chi}, \tilde{\eta}_2)$, for some $a_{15}(\kappa, g) > 0$ and by $\tilde{x}_0 := \tilde{x}_0 - \xi_0^0$. Hence, usual high-gain arguments show that, under A2, there exists $\ell^*_0(\kappa, g) > 0$ such that, for all $\ell > \ell^*_0(\kappa, g)$, the following holds

$$
|\tilde{\chi}(t)| \leq a_{10}(\kappa, g) e^{-a_{17} \kappa t} + \frac{a_{11}(\kappa)}{\kappa} |\tilde{\eta}_1(0)|_{0,t} + \frac{a_{12}(\kappa)}{\kappa} |\tilde{\eta}_2(0)|_{0,t}
$$

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

$$
|\tilde{x}_0(t)| \leq b_2(\tilde{x}_0(0), t) + b_3(\tilde{\eta}_1, \tilde{\chi}, \tilde{\eta}_2) \leq b_2(\tilde{x}_0(0), t) + a_{13}(\kappa, g)|\tilde{\eta}_1(0)|_{0,t} + |\tilde{\eta}_2(0)|_{0,t}.
$$

Therefore, in view of (22), (26), (28) and (29), repeating the small-gain arguments of (23) yields the existence of $\kappa^*(g) \geq \kappa_0^*(g)$ and of an $\ell^*(\kappa, g) \geq \ell^*_0(\kappa, g)$ such that, for each $\rho > p^*$, $g \geq g^*(\rho)$, $\kappa > \kappa^*(g)$, and $\ell > \ell^*(\kappa, g)$, it holds that

$$
|\tilde{\eta}(t), \tilde{\chi}(t), \tilde{\zeta}(t), \tilde{\xi}(t))| \leq \beta((\tilde{\eta}(0), \tilde{\eta}(0), \tilde{\xi}(0), \tilde{\xi}(0)), t) + p_1(\kappa, g)|\varepsilon^*|_{0,t}
$$

and

$$
\limsup_{t \to \infty} |\tilde{\eta}(t)| \leq p_2 \kappa^{-1} \gamma^{-1-d} \limsup_{t \to \infty} |\varepsilon^*(t)|
$$

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