ON OPTIMAL ESTIMATION PROBLEMS FOR NONLINEAR SYSTEMS AND THEIR APPROXIMATE SOLUTION

A. Alessandri [*] **C. Cervellera** [*] **A.F. Grassia** [*]
M. Sanguineti [**]

[*] *ISSIA-CNR National Research Council of Italy*
Via De Marini 6, 16149 Genova, Italy
{angelo, cervellera, filippo}@ge.issia.cnr.it
[**] *DIST–University of Genoa*
Via Opera Pia 13, 16145 Genova, Italy
marcello@dist.unige.it

Abstract: An approach based on optimization is described to construct state estimators that provide a stable dynamics of the estimation error and minimize a L_p measure of the estimation error. The state estimator depends on an innovation function made up of two terms: a linear gain and a feedforward neural network. The gain and the weights of the neural network can be chosen in such way to ensure the convergence of the estimation error and minimize the L_p performance index, after a suitable discretization of the state and error space. Simulation results are reported.

Keywords: Optimal state estimation, Lyapunov method, stability, neural networks

1. INTRODUCTION

Many methods to perform state estimation for nonlinear systems are reported in the literature. Two different frameworks are usually considered for both continuous-time and discrete-time systems. The first is related to the case where no noise affects a system, hence uncertainty in the plant is due only to the initial state and to a limited access to the state variables. The solution of such a problem enables one to design an estimator, usually called "observer." The second framework concerns a system setting where disturbances affect both the dynamics and the measures. The solution of the estimation problem in the presence of noise is usually called "filter." The designs of both observers and filters for general nonlinear systems with convergence guarantees are still difficult tasks.

A comprehensive discussion of the various contributions to nonlinear estimation by different approaches is beyond the scope of the paper (the interested reader is referred to (Walcott *et al.*, 1987; Misawa and Hedrick, 1989) for the observer problem and to (Haykin *et al.*, 1997) for a survey of filtering).

The first result on observer design for nonlinear systems was obtained by Thau (Thau, 1973), where sufficient conditions to design an asymptotically stable observer were presented. Further investigations on the so-called high-gain observers were carried out later on (Gauthier *et al.*, 1992). These estimators ensure a fast convergence of the estimation error but require a sufficiently large gain, which can cause instability when used in cascade with a regulator for the purpose of feedback control. Stability conditions on the estimation error for observers were also proposed in (Raghavan and Hedrick, 1994); they are essentially based on the idea of accounting for the nonlinearity by means of the corresponding Lipschitz inequality. An algebraic Riccati equation is introduced to guarantee convergence for the estimator. Further investigations by Rajamani gave new interpretations of these results (Rajamani, 1998), as well as discussions and suggestions to face design issues (Rajamani and Cho, 1998).

As previously pointed out, there exist numerous methods to filter the state of a noisy system. The disturbances that affect the system and/or the measurement equations can be regarded as either unknown deterministic inputs or stochastic random variables within a probabilistic framework. Here we focus on estimators that admit a L_p bounded estimation error, $p \in [1, \infty]$, which can be considered also as a performance index (Alessandri and Sanguineti, 2001*b*).

More specifically, state estimation problems are considered for continuous-time, nonlinear dynamic systems with Lipschitz nonlinearities by means of a full-order observer that is constructed using an innovation function made up of two terms: a linear gain and a parameterized nonlinear structure. Under some regularity assumptions on the system and measurement equations and on the innovation function, a procedure is developed to design such an observer with the parameterized structure implemented by means of a class of approximating networks such as feedforward neural networks.

In order to guarantee the convergence of the estimation error and a satisfying L_p performance index, a quadratic Lyapunov function is sought for. The design parameters (i.e., the linear gain and the weights of the neural network) can be chosen in such a way to suitably constrain the derivative of a quadratic Lyapunov function to be negative on a sampling grid of points on the Cartesian product of the state space and the estimation error space. This is accomplished by minimizing a cost function that penalizes the bad estimation performances and the constraints that are not satisfied in correspondence of the sampling points. It is worth noting that the selection of the design parameters is made completely off line (see also (Alessandri *et al.*, 1999)).

Under assumptions on the distribution of the sampling points and smoothness of the Lyapunov function, the negative definiteness of the derivative of the Lyapunov function is ensured, thus the resulting observer provides a convergent estimation error. In particular, it is shown that convergence is obtained by using special deterministic sequences that aim at optimizing the dispersion of the sampling points (a measure that quantifies "how uniformly" the points are spread). The use of such sequences in the context of neural network learning has been introduced and analyzed in (Cervellera and Muselli, 2004).

The paper is organized as follows. Section 2 is devoted to the description of the considered class of systems. In Section 3, the focus is on the stability of the estimation error both with and without disturbances. The proposed approach is presented in Section 4 for a class of parameterized estimators that can be constructed to perform state estimation by using the powerful approximation capabilities of neural networks. Final comments are given in Section 5

Before concluding this section, let us briefly introduced the following notations. For $p \in [1, \infty)$ and a positive n, the space L_p^n consists of all Lebesgue-measurable functions $s : [0, \infty) \to \mathbb{R}^n$ such that $\int_0^\infty \|s(t)\|^p \, dt < \infty$. The space L_∞^n is the set of all Lebesgue-measurable functions defined on $\mathbb{R}^n$ that are essentially bounded, i.e., such that $\operatorname{ess.sup}_{t \geq 0} \|s(t)\| < \infty$, where "ess. sup" denotes the essential supremum (i.e., supremum except on sets of measure zero). To deal with possibly unbounded signals, the extension of L_p^n spaces is defined as follows. For $p \in [1, \infty]$, the extended space $L_{p\,e}^n$ is defined as $L_{p\,e}^n \triangleq \left\{ s | s_\tau \in L_p^n, \forall \, \tau \geq 0 \right\}$, where $s_\tau(t) \triangleq \begin{cases} s(t) \ , & \text{if } t \leq \tau \, , \\ 0 \ , & \text{if } t > \tau \, . \end{cases}$ Moreover, for every $s \in L_{p\,e}^n$ and $p \in [1, \infty]$, let $\|s\|_{p,\,\tau} \triangleq \|s_\tau\|_p$.

Given a matrix $M \in \mathbb{R}^{m \times n}$, $\|M\| \triangleq \sup\limits_{x \in \mathbb{R}^n, \, \|x\|=1} \|M\,x\|$ denotes the matrix norm induced by the Euclidean vector norm $\|\cdot\|$. Given a symmetric matrix S, $\lambda_{\min}(S)$ and $\lambda_{\max}(S)$ denote the minimum and maximum eigenvalue of S, respectively. Recall that $\|M\| = \sqrt{\lambda_{\max}(M^{\mathrm{T}}M)}$, i.e., $\|M\|$ equals the square root of the spectral radius of $M^{\mathrm{T}}M$ (as $M^{\mathrm{T}}M$ is positive semidefinite, it has nonnegative eigenvalues). For a symmetric matrix S, $\|S\| = |\lambda_{\max}(S)|$.

2. SYSTEM DESCRIPTION

In the following, we shall consider class of systems characterized by a linear measurement equation and a dynamics with a linear part. The couple of matrices that describe the linear part of the dynamic equation and the measurements is observable. More specifically, let us consider

$$\begin{cases} \dot{x} = A\,x + f(x) \\ y = C\,x \end{cases} , \quad t \geq 0 \ , \tag{1}$$

where $x(t) \in X \subseteq \mathbb{R}^n$ and $y(t) \in Y \subseteq \mathbb{R}^m$. Let the nonlinearity $f : X \to \mathbb{R}^n$ be Lipschitz in X, with the Lipschitz constant k_f, and suppose that (A, C) is observable.

Equations (1) refer to a general class of nonlinear, essentially observable systems and to all systems that are diffeomorphic to (1). For example, in (Shim *et al.*, 2001) necessary and sufficient conditions are given to ensure the existence of a diffeomorphism that transforms a quite general nonlinear system in the form (1).

An admissible observer for (1) can be expressed as:

$$\dot{\hat{x}} = A\,\hat{x} + f(\hat{x}) + g\,(y - C\,\hat{x}) \ ,$$

where $z \triangleq y - C\hat{x} \in Z \subseteq \mathbb{R}^m$ and the innovation function $g : Z \to \mathbb{R}^n$ is Lipschitz in Z. Let the innovation function g be chosen as the summation of two contributions:

$$g\,(y - C\,\hat{x}) = L\,(y - C\,\hat{x}) + \gamma\,(y - C\,\hat{x}) \ ,$$

where L is an $n \times m$ matrix and $\gamma : Z \to \mathbb{R}^n$ is Lipschitz in Z, with a Lipschitz constant k_γ, and $\gamma(0) = 0$. Hence the estimator dynamics is given by

$$\dot{\hat{x}} = A\,\hat{x} + f(\hat{x}) + L\,(y - C\,\hat{x}) + \gamma\,(y - C\,\hat{x}) \ . \quad (2)$$

In the following, we shall consider the behavior of the estimation error $e(t) \triangleq x(t) - \hat{x}(t)$ in different settings.

3. STABILITY OF OBSERVERS WITH AND WITHOUT THE PRESENCE OF DISTURBANCES

Consider the dynamics of the estimation error for system (1) with the observer (2):

$$\dot{e} = (A - L\,C\,)\;e + f(x) - f(\hat{x}) - \gamma\,(y - C\,e) \, . \quad (3)$$

Consider the Lyapunov function $V \triangleq e^{\mathrm{T}} P\,e$, where the matrix P is positive definite and symmetric; the derivative of V is given by

$$\begin{aligned}\dot{V} &= e^{\mathrm{T}} \left[(A - LC)^{\mathrm{T}} P + P\,(A - LC) \right] e \\ &\quad + 2\,[f(x) - f(\hat{x})]^{\mathrm{T}} P\,e - 2\,\gamma^{\mathrm{T}}\,(C\,e)\;P\,e \quad (4)\end{aligned}$$

As is usually done in the design of observers for dynamic systems with Lipschitz nonlinearities (Rajamani, 1998; Alessandri, 2002), by a simple algebra we can compute upper bounds to the last two terms of (4):

$$\begin{aligned}&2\,[f(x) - f(\hat{x})]^{\mathrm{T}}\,P\,(x - \hat{x}) \\ &\leq 2\,k_f\,\|x - \hat{x}\|\;\|P\,(x - \hat{x})\| \\ &\leq k_f^2\,(x - \hat{x})^{\mathrm{T}}\;P\,P\,(x - \hat{x}) + (x - \hat{x})^{\mathrm{T}}\,(x - \hat{x})\end{aligned}$$

and

$$\begin{aligned}&-2\gamma^{\mathrm{T}}\,[C\,(x - \hat{x})]\;P\,(x - \hat{x}) \\ &\leq 2\,k_\gamma\,\|C\,(x - \hat{x})\,\|\;\|P\,(x - \hat{x})\| \\ &\leq k_\gamma^2\,\|C\|^2\,(x - \hat{x})^{\mathrm{T}}\;P\,P\,(x - \hat{x}) \\ &+ (x - \hat{x})^{\mathrm{T}}\,(x - \hat{x})\,.\end{aligned}$$

From (4), using the above-written inequalities we obtain

$$\begin{aligned}\dot{V} &\leq e^{\mathrm{T}} \Big[(A - LC)^{\mathrm{T}} P \\ &\quad + P\,(A - LC) + \left(k_f^2 + k_\gamma^2\,\|C\|^2\right)\,P\,P + 2\,I \Big] e\,.\end{aligned}$$

As shown in (Rajamani, 1998), if there exist a gain matrix L and a symmetric positive definite matrix Q such that $A - LC$ is stable and the algebraic Riccati equation

$$\begin{aligned}&(A - LC)^{\mathrm{T}}\;P + P\,(A - LC) \\ &+ \left(k_f^2 + k_\gamma^2\,\|C\|^2\right) P\,P + 2\,I = -Q \quad (5)\end{aligned}$$

has a symmetric, positive definite matrix P as solution, then the estimator (2) admits a quadratic Lyapunov function such that $c_1\,\|e\|^2 \leq V \leq c_2\,\|e\|^2$ and $\dot{V} \leq -c_3\,\|e\|^2$, with $c_1 \triangleq \lambda_{\min}(P)$, $c_2 \triangleq \lambda_{\max}(P)$, and $c_3 \triangleq \lambda_{\min}(Q)$. Hence we conclude that (2) is a global exponential estimator for the system (1). Note that the observability hypothesis about (A, C) is a necessary condition for the existence of a solution of (5).

Suppose that additive L_p disturbances affect the dynamics and the measurement equations, that is,

$$\begin{cases} \dot{x} = A\,x + f(x) + w \\ y = C\,x + v \end{cases} \quad (6)$$

where $w \in L^n_{p\,e}$ and $v \in L^m_{p\,e}$. The estimation error dynamics is

$$\begin{aligned}\dot{e} &= (A - L\,C\,)\;e + f(x) \\ &\quad - f(\hat{x}) + w - L\,v - \gamma\,(C\,e + v) \ . \quad (7)\end{aligned}$$

It is easy to verify that the fulfilment of the Riccati equation (5), we obtain

$$\|e\|_{p,\tau} \leq \eta\,\|w\|_{p,\tau} + \lambda\,\|v\|_{p,\tau} + \beta \quad (8)$$

for all $\tau \in [0, \infty)$ with

$$\eta = \frac{2\lambda^2_{\max}(P)\,(\|A\| + k_f)}{\lambda_{\min}(P)\,\lambda_{\min}(Q)}$$

$$\lambda = \frac{2\lambda^2_{\max}(P)\;(\|L\| + k_\gamma)}{\lambda_{\min}(P)\,\lambda_{\min}(Q)}$$

$$\beta = \sqrt{\frac{\lambda_{\min}(P)}{\lambda_{\max}(P)}}\;\|e(0)\|\;\rho$$

where

$$\rho = \begin{cases} 1 & , \quad p = \infty \\ \left(\dfrac{2\,\lambda_{\max}(P)}{\lambda_{\min}(Q)\,p} \right)^{1/p} & , \quad p \in [1, \infty)\,. \end{cases}$$

So far the problem of constructing observers has been reduced to the solution of the Riccati equation (5). Unfortunately, it is difficult to find such solutions for large values of the Lipschitz constants k_f and k_γ. Thus, we shall address the construction of state estimators according to a different approach presented in the following.

4. DESIGN OF A PARAMETERIZED ESTIMATOR

The design of the an estimator can be stated as an optimization problem that aims both at obtaining asymptotic stability and satisfying performances of the estimation error. A particular class of approximating networks can be used as part of the innovation function of the estimator. More specifically, we consider estimators described by the following equation

$$\dot{\hat{x}} = A\,\hat{x} + f(\hat{x}) + L\,(y - C\hat{x}) + \gamma_\nu\,(y - C\hat{x}, w_\nu)\ ,(9)$$

where $L \in \mathbb{R}^{n\times m}$ is a gain matrix and, for a positive integer ν, γ_ν is a function belonging to the class A_ν, defined as follows.

Definition 4.1.
$A_\nu \triangleq \{\ \gamma_\nu : K \times \mathbb{R}^l \to \mathbb{R}^n\ ,\ K$ compact , such that

(i) $\gamma_{\nu\, j}(\xi, \omega_{\nu_j}) = \sum_{i=1}^{\nu} c_{ij}\,\varphi_i(\xi, \kappa_i)$, $\varphi_i : K \times \mathbb{R}^l \to \mathbb{R}$, $|c_{ij}| \le C$, $C \in \mathbb{R}^+$, $\kappa_i \in \mathbb{R}^l$, $i = 1, \ldots, \nu$, $j = 1, \ldots, n$, $\omega_{\nu\, j} \triangleq \mathrm{col}(c_{ij}, \kappa_i : i = 1, \ldots, \nu)$;
(ii) the functions $\varphi_i(\cdot, \kappa_i)$ are bounded in aggregate, i.e., $\exists\, M \in \mathbb{R}^+$ such that $\forall i = 1, \ldots, \nu$, $\forall \kappa_i \in \mathbb{R}^l$, $\sup_{\xi\in K}|\varphi_i(\xi, \kappa_i)| \le M$;
(iii) the functions $\varphi_i(\cdot, \kappa_i)$ are equicontinuous,i.e., $\forall\, \epsilon > 0\ \exists\, \delta_\epsilon > 0$ such that $\forall\, i = 1, \ldots, \nu$, $\forall\, \kappa_i, \kappa_i' \in \mathbb{R}^l$ if $\|\xi - \xi'\| < \delta_\epsilon$ then $|\varphi_i(\xi, \kappa_i) - \varphi_i(\xi', \kappa_i')| \le \epsilon$;
(iv) the functions $\varphi_i(\cdot, \kappa_i)$ are Lipschitz, i.e., $\forall\, i = 1, \ldots, \nu\ \exists\, L_i \in \mathbb{R}^+$ such that $\forall \kappa_i \in \mathbb{R}^l$, $|\varphi_i(\xi, \kappa_i) - \varphi_i(\xi', \kappa_i)| \le L_i |\xi - \xi'|$;
(v) $\bigcup_{\nu=1}^{\infty} A_\nu$ is dense in $\mathcal{G}$ with respect to the supremum $\}$.

The class A_ν defines a set of admissible innovation functions. Once a type of approximating networks, i.e., a mother function $\varphi_i(\cdot, \cdot)$ is chosen, the Lyapunov function for the estimator (9) depends on the values of L and w_ν. As to the function φ_i, it is suitable to make a choice generating sets A_ν whose closure is as large as possible: loosely speaking, the larger such a closure, the wider the choice at our disposal for a Lyapunov function. So we shall use approximating networks that are dense in the space of continuous functions on compact sets, i.e., in the neural-network parlance, that enjoy the "universal approximation property." Well-known examples of approximating networks are feedforward neural networks of the perceptron type, with at most ν hidden units and bounded parameters and radial-basis-functions with at most ν hidden units and bounded input weights and variances. The proofs of the fact that such functions are provided with the density property in the space $\mathcal{C}(K, \mathbb{R}^n)$, where $K \subset \mathbb{R}^m$ is compact, can be found, for example, in (Leshno *et al.*, 1993) and (Kůrková, 1995).

To guarantee the possibility of finding an estimator implemented with a "small" number ν of basis functions also for vectors x to be estimated with a large number of components, we shall employ so-called "polynomially-complex approximating networks". Such networks have the desirable property that the number ν of basis functions required to guarantee a fixed approximation accuracy has to grow at most polynomially with the number of variables (in the case of the estimator (9), the dimension m of the measurement vector); see (Kurková and Sanguineti, 2001; Kůrková and Sanguineti, 2002; Zoppoli *et al.*, 2002) for details.

Let us consider the following problem.

Problem OEP$_\nu$. *Given* $p \in [1, \infty]$ *and* $T > 0$, *solve*

$$\inf_{\gamma_\nu \in A_\nu} J(\gamma_\nu), \tag{10}$$

where $J(\gamma_\nu) = \|x - \hat{x}\|_{p,T}$, x , $\hat{x} \in L^n_{p\,e}$, $w \in L^n_{p\,e}$, $v \in L^m_{p\,e}$, *and*

$$\begin{cases} \dot{x} = f\,(x, w) \\ y = h\,(x, v) \\ \dot{\hat{x}}_\nu = f(\hat{x}_\nu, 0) + \gamma_\nu\,(\omega_\nu, y - h\,(\hat{x}_\nu, 0))\,. \end{cases} \tag{11}$$

□

As each A_ν is a set of parameterized functions with a fixed structure, the minimization has to be performed with respect to the finite-dimensional vector of parameters $\omega_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}$, whereas Problem OEP entails an infinite-dimensional minimization. This turns out to be evident by substituting γ_ν into the differential equation of the estimator and then into J . The cost functional is a function of the parameter vector ω_ν . With a little notational abuse, we denote such a function by $J_\nu(\omega_\nu)$. Thus, for each positive integer

ν the minimization with respect to the infinite-dimensional set A_ν is replaced by the minimization with respect to the finite-dimensional vector $\omega_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}$.

Thus, we can define the following.

Problem OEP$'_\nu$. *Given $T > 0$, find*

$$\inf_{\omega_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}} J_\nu(\omega_\nu)\,. \tag{12}$$

where $J_\nu(\omega_\nu) = \|x-\hat{x}\|_{p,T}$, x, $\hat{x} \in L^n_{pe}$, $w \in L^n_{pe}$, $v \in L^m_{pe}$, and

$$\begin{cases} \dot{x} = f\,(x, w) \\ y = h\,(x, v) \\ \dot{\hat{x}}_\nu = f(\hat{x}_\nu, 0) + \gamma_\nu\,(\omega_\nu, y - h\,(\hat{x}_\nu, 0))\,. \end{cases} \tag{13}$$

□

It is worth noting that the solution of this last problem may not be unique, even if there exists a unique minimum of Problem OEP$_\nu$, as it might happen that there is no one-to-one correspondence between a vector $\omega_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}$ and an element $\gamma_\nu \in A_\nu$.

The solution of Problem OEP$'_\nu$ may provide satisfying estimation performances, which can be preferably associated with stability requirements for the estimation error. To this end, suitable constraints must be added to the minimization problem, as we will see in the following. Note that, once the output y is known and the parameters vector ω_ν are chosen, the evolution of the estimated state vector $\hat{x}$ is completely determined. Of course, the selection of L and γ_ν must ensure the stability of the estimation error, whose dynamics is given by

$$\dot{e} = (A - LC)\,e + f(x) - f(\hat{x}) - \gamma_\nu\,(C\,e, \omega_\nu)\,. \tag{14}$$

If a quadratic Lyapunov function $V = e^{\mathrm{T}} P\, e$ is considered with P symmetric positive definite matrix, we obtain

$$\begin{aligned} \dot{V} = e^{\mathrm{T}}\, &\left[(A - LC)^{\mathrm{T}} P + P\,(A - LC)\right] e \\ &+ 2\,[f(x) - f(\hat{x}) - \gamma_\nu\,(C\,e, w_\nu)]\; P\,e\,. \end{aligned} \tag{15}$$

Since the pair (A, C) is observable, there exist a gain matrix L and a unique symmetric positive definite matrix P as solution of the Lyapunov equation

$$(A - LC)^{\mathrm{T}} P + P\,(A - LC) = -Q \tag{16}$$

where Q is a given symmetric positive definite matrix.

We consider a compact set $\bar{E}$ for the estimation error. Given the compact set $S = X \times \bar{E}$, let us denote by S_M a set of M sample points $s_i = \mathrm{col}(x_i, e_i) \triangleq (x_i^{\mathrm{T}}, e_i^{\mathrm{T}})^{\mathrm{T}}$, $e_i \neq 0$, $i = 1, 2, \ldots, M$, that belong to S. Let us define the dispersion of S_M as

$$\theta(S_M) \triangleq \sup_{s \in S} \min_{1 \le i \le M} \|s - s_i\|\,.$$

Therefore, in order to guarantee the asymptotic stability of the estimation error, using Proposition 4 in (Alessandri and Sanguineti, 2001*a*) we impose that for some ω_ν,

$$\begin{aligned} &2\,[f(x_i) - f(x_i - e_i) - \gamma_\nu\,(C\,e_i, w_\nu)]\; P\,e_i \\ &-e_i^{\mathrm{T}}\, Q\, e_i \le -c\|e_i\|^2 \end{aligned} \tag{17}$$

where $c > 0$, $l > 1$, and $\theta(S_M) < \dfrac{\varepsilon_M}{L_F}$, $s_i = \mathrm{col}(x_i, e_i) \in S_M$, as (i) is trivially satisfied.

Summing up, in order to construct an estimator that solves Problem OEP$_\nu$ and has a stable (in some sense) estimation error, we have to perform the following steps.

1) Choose a time horizon T, an L_p measure for the estimation error, and compact sets $X, \bar{E} \subset \mathbb{R}^n$.
2) Choose a composite model A_ν with ν basis functions φ; the admissible innovation functions have to belong to A_ν.
3) Choose a set $S_M \subset X \times \bar{E}$ of M sample points $s_i \triangleq \mathrm{col}(x_i, e_i)$, e_i, $i = 1, \ldots, M$ with good dispersion properties (e.g., belonging to a low-discrepancy sequence).
4) Given $c > 0$, a gain matrix L, and two symmetric positive definite matrices P and Q such that (16) is satisfied, find $\omega^*_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}$ such that
   4.1) $\omega^*_\nu = \operatorname*{argmin}_{\omega_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}} \|e(\omega_\nu)\|_{p,T}$;
   4.2) $\dot{V}\,(s_i, \bar{\omega}) < -c\,\|e_i\|^2$ for $s_i \in S_M$, $i = 1, 2, \ldots, M$, where $c > 0$;
   4.3) $\theta(S_M) < \dfrac{\varepsilon_M}{L_F}$;
   where L_F be the Lipschitz constant of $F(s, \bar{\omega}) \triangleq \dot{V}\,(s, \bar{\omega}) + c\,\|e\|^l$ with respect to s and $\varepsilon_M \triangleq -\max_{1 \le i \le M} F(s_i, w) > 0$.

In order to satisfy the constraints in Step 4.2), we employ a sum of suitable penalty functions

$$J_{\mathrm{stab}}(\omega_\nu) = \sum_{i=1}^{M} \left(\max\left\{0, \dot{V}\,(s_i, \omega_\nu) + c\|e_i\|^2\right\}\right)^2. \tag{18}$$

This leads to a new problem that contains both the performance and stability requirements, and can be defined as

Problem S-OEP$'_\nu$. *Given $T > 0$, find*

$$\inf_{\omega_\nu \in \mathbb{R}^{\mathcal{N}(\nu)}} J_\nu(\omega_\nu) + \alpha J_{\rm stab}(\omega_\nu)\,, \tag{19}$$

where $\alpha > 0$, $J_\nu(\omega_\nu) = \|x - \hat{x}\|_{p,T}$, x, $\hat{x} \in L^n_{p\,e}$, $w \in L^n_{p\,e}$, $v \in L^m_{p\,e}$, $J_{\rm stab}$ *is given by (18), and*

$$\begin{cases} \dot{x} = f\,(x, w) \\ y = h\,(x, v) \\ \dot{\hat{x}}_\nu = f(\hat{x}_\nu, 0) + \gamma_\nu\,(\omega_\nu, y - h\,(\hat{x}_\nu, 0))\,. \end{cases} \tag{20}$$

□

Results on exact penalization can be found in (Bertsekas, 1999) for a convenient choice of the penalty parameter α.

5. CONCLUDING REMARKS

In Section 3, we pointed out the issues in designing observers and filters for the considered class of systems, which are related to the problem of solving the Riccati equation (5). Unfortunately, the solution of such equation is impossible for large values of the Lipschitz constants k_f and k_γ. This is a serious difficulty, which has suggested to address the construction of state estimators according to the approach formalized in Section 4. The main drawbacks of such an approach is the necessity of suitably sampling the domain of both state and estimation error spaces. Thus, in practice we can apply it only to systems that admits an invariant set and construct estimators with a regional kind of convergence for the estimation error, i.e., in other words assuming that $e(t)$ belongs to a bounded set.

REFERENCES

Alessandri, A. (2002). Sliding-mode estimators for a class of nonlinear systems affected by bounded disturbances. *Int. J. of Control* **76**(3), 226–236.

Alessandri, A. and M. Sanguineti (2001*a*). $\mathcal{L}_p$–stable and asymptotic estimators for nonlinear dynamic systems. In: *Proc. American Control Conference*. Arlington, VA. pp. 1991–1996.

Alessandri, A. and M. Sanguineti (2001*b*). On estimators for nonlinear systems in $\mathcal{L}_p$ spaces. In: *Proc. 40th Conference on Decision and Control*. Orlando, FL. pp. 3372–3377.

Alessandri, A., M. Baglietto, T. Parisini and R. Zoppoli (1999). A neural state estimator with bounded errors for nonlinear systems. *IEEE Trans. on Automatic Control* **44**(11), 2028–2042.

Bertsekas, D. P. (1999). *Nonlinear Programming*. Athena Scientific, Belmont, MA.

Cervellera, C. and M. Muselli (2004). Deterministic design for neural network learning: an approach based on discrepancy. *IEEE Trans. on Neural Networks* **15**(3), 533–544.

Gauthier, J. P., H. Hammouri and S. Othman (1992). A simple observer for nonlinear systems applications to bioreactors. *IEEE Trans. on Automatic Control* **37**(6), 875–880.

Haykin, S., P. Yee and E. Derbez (1997). Optimum nonlinear filtering. *IEEE Trans. on Signal Processing* **45**(11), 2774–2786.

Kůrková, V. (1995). Approximation of functions by perceptron networks with bounded number of hidden units. *Neural Networks* **8**, 745–750.

Kůrková, V. and M. Sanguineti (2002). Comparison of worst-case errors in linear and neural network approximation. *IEEE Trans. on Information Theory* **48**(1), 265–275.

Kurková, V. and M. Sanguineti (2001). Bounds on rates of variable-basis and neural-network approximation. *IEEE Trans. on Information Theory* **47**, 2659–2665.

Leshno, M., V. Ya, A. Pinkus and S. Schocken (1993). Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks* **6**, 861–867.

Misawa, E. A. and J. K. Hedrick (1989). Nonlinear observers – A state-of-the-art survey. *J. of Dynamic Systems, Measurement, and Control* **111**, 344–352.

Raghavan, S. and J. K. Hedrick (1994). Observer design for a class of nonlinear systems. *Int. J. of Control* **59**(2), 515–528.

Rajamani, R. (1998). Observers for Lipschitz nonlinear systems. *IEEE Trans. on Automatic Control* **43**(3), 397–401.

Rajamani, R. and Y. M. Cho (1998). Existence and design of observers for nonlinear systems: relation to distance to unobservability. *Int. J. of Control* **69**(5), 717–731.

Shim, H., Y. I. Son and J. H. Seo (2001). Semi-global observer for multi-output nonlinear systems. *System & Control Letters* **42**(3), 233–244.

Thau, F. E. (1973). Observing the state of nonlinear dynamic systems. *Int. J. of Control* **17**(3), 471–479.

Walcott, B. L., M. J. Corless and S. H. Zak (1987). Comparative study of nonlinear state-observation techniques. *Int. J. of Control* **45**(6), 2109–2132.

Zoppoli, R., M. Sanguineti and T. Parisini (2002). Approximating networks and the extended Ritz method for the solution of functional optimization problems. *J. of Optimization Theory and Applications* **112**(2), 403–439.