Robust Output Feedback Control of Nonlinear Stochastic Systems Using Neural Networks

Stefano Battilotti, *Senior Member, IEEE,* and Alberto De Santis

***Abstract*—We present an adaptive output feedback controller for a class of uncertain stochastic nonlinear systems. The plant dynamics is represented as a nominal linear system plus nonlinearities. In turn, these nonlinearities are decomposed into a part, obtained as the best approximation given by neural networks, plus a remaining part which is treated as uncertainties, modeling approximation errors, and neglected dynamics. The weights of the neural network are tuned adaptively by a Lyapunov design. The proposed controller is obtained through robust optimal design and combines together parameter projection, control saturation, and high-gain observers. High performances are obtained in terms of large errors tolerance as shown through simulations.**

***Index Terms*—Filtering (93E11, 93E20, 93D15, 93D21, 49K30, 60H10), optimal control, robust stabilization, stochastic nonlinear systems.**

I. Introduction

In modeling of dynamical systems, the stochastic framework is suitable for taking into account either randomly varying system parameters or stochastic exogenous inputs. It is important in many practical situations to require, besides stability in some sense, some optimal and robustness performances, which can be usually described through a suitable cost functional. These performances may include tracking errors and physical constraints, due for example to control actuators or sensors with limited range.

According to the existing literature (see [18]–[20], [22], [23] and the textbooks [15] and [21]), by stability it is usually meant that

- the probability that the trajectory, stemming from x_0, leaves an ϵ-ball around the origin goes to zero as x_0 tends to the origin;
- the trajectory, stemming from x_0, goes asymptotically to zero almost surely.

This stability, usually known as *stability in probability*, is either *local* or *global* according to whether x_0 is in some (small) neighborhood of the origin or, respectively, it is *any point of the state space*. In [15] Lyapunov-based conditions are given for guaranteeing stability in probability and require the solution of partial differential inequalities (PDIs). In [17] and [19], it has been proved that a step-by–step algorithm (*backstepping*) can be successfully implemented for solving globally these PDIs, whenever the state is available for feedback, while in [18], the problem of global output feedback stabilization in probability is solved for a class of systems with output nonlinearities. For deterministic systems, the complexity and the conditions for solving these PDIs can be weakened by relaxing the stability requirements of the closed-loop system. In a deterministic setting, *semiglobal stabilization* was introduced in [6] and requires a local asymptotic stability of the closed-loop system plus a region of attraction containing any *a priori* given compact set of the state space. The basic ingredients for achieving semiglobal stability via output feedback are *control saturations* and *high-gain observers* [10], [11], [14]: Large values of the observer gain guarantee that the error between the state and its estimate, generated by the observer itself, goes to zero "sufficiently fast," while input saturations rule out destabilizing effects such as *peaking* [8], which is a phenomenon occurring when one is trying to force some state variables to zero as fast as possible causing an impulsive-like behavior of some others.

In this paper, we want to study the problem of stabilizing (in some probabilistic sense) the following class of nonlinear stochastic systems:

$$\begin{aligned} dx(t) =& [Ax(t) + B(f(x(t)) + (g(x(t)) + 1)u(t)]\, dt \\ &+ BH(x(t))dw(t) \\ dy(t) =& Cx(t) \end{aligned} \tag{1}$$

where $w(t) \in \mathbb{R}^s$ is a Wiener process, $u(t) \in \mathbb{R}$ is the control, $x(t) \in \mathbb{R}^n$ is the state, $y(t) \in \mathbb{R}$ are the measurements, $f(x), g(x) \in \mathbb{R}$ are model uncertainties and nonlinearities, and

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

When $H = 0$ (i.e., noiseless dynamics), the system (1) is the same as the one considered in [1] with $m = 1$ (i.e., no input derivatives are taken into account).

By calling upon the neural-networks theory [2] for each choice of vector basis functions $f_R(x)$ and $g_R(x)$, it is possible to decompose the functions $f(x)$ and $g(x)$ as follows [29]:

$$\begin{aligned} f(x) =& \vartheta_f^{*T} f_R(x) + \Phi_F(x) \\ g(x) =& \vartheta_g^{*T} g_R(x) + \Phi_G(x) \end{aligned} \tag{2}$$

where

$$\begin{aligned} \vartheta_f^* &= \arg\min_{\vartheta} \left\{ \sup_{x\in\Omega} \left| f(x) - \vartheta_f^T f_R(x) \right| \right\} \\ \vartheta_g^* &= \arg\min_{\vartheta} \left\{ \sup_{x\in\Omega} \left| g(x) - \vartheta_g^T g_R(x) \right| \right\} \end{aligned} \tag{3}$$

Manuscript received April 4, 2001; revised October 31, 2001 and April 12, 2002. This work was supported by MURST under a grant from the Italian Ministry of Scientific and Technological Research.

The authors are with the Dipartimento di Informatica e Sistemistica, 00184 Rome, Italy (e-mail: battilotti@dis.uniroma1.it, desantis@dis.uniroma1.it).

Digital Object Identifier 10.1109/TNN.2002.806609

and $\vartheta_f^* \in \mathbb{R}^{r_f}$ and $\vartheta_g^* \in \mathbb{R}^{r_g}$.

The system

$$\begin{aligned}\dot{x}(t) =& Ax(t) + B\left[\vartheta_f^{*T} f_R(x(t)) + \left(\vartheta_g^{*T} g_R(x(t)) + 1\right) u(t)\right]\\ y(t) =& Cx(t)\end{aligned} \tag{4}$$

(i.e., system under nominal conditions plus best approximation of the nonlinear dynamics) can be understood as the *nominal system* of (1). The model uncertainty is represented by $\Phi(u, x) = \Phi_F(x) + \Phi_G(x)u$.

Since the neural network approximations given in (2) are valid for $x \in \Omega$, one should ensure that the state trajectory $x(t)$ stay inside Ω for all $t \geq 0$ at least with some guaranteed probability. This property is well-covered by the notion of $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$-*SP* stability recently introduced in [4]. Given $0 \leq \alpha$, $\beta \leq 1$ and some desired region of attraction Ω^e and target set $\mathcal{B}^e$, this notion requires that the trajectories of the closed-loop system, resulting from (1), with initial condition in Ω^e remain inside some compact set $\widetilde{\Omega}^e \supseteq \Omega^e$ of the state space, eventually enter any given neighborhood of the target set $\mathcal{B}^e$ in finite time and remain thereinafter with probability at least $(1-\alpha)$ $(1-\beta)$. The numbers α and β are *risk margins*: The first one quantifies the risk of leaving $\widetilde{\Omega}^e$ with initial condition in Ω^e rather than getting close to the target, while the second one gives a risk margin for remaining close to the target. If $\mathcal{B}^e = \{0\}$ and Ω^e can be taken any *a priori* given compact set of the state space and α and β any numbers in $[0, 1)$, our definition extends to a stochastic setting the notion of *semiglobal stabilization* as introduced in [6], and in what follows, we will refer to this property as *semiglobal stabilization in probability*.

The values ϑ_f^* and ϑ_g^* are derived by solving the optimization problem (3). Typically, it is difficult to compute these vaules; however, by some off-line training, one can obtain "good approximations" $\widehat{\vartheta}_{f0}$ and $\widehat{\vartheta}_{g0}$ of the optimal values ϑ_f^* and ϑ_g^*. As a consequence, the nominal system can be written as

$$\begin{aligned}\dot{x}(t) =& Ax(t) + B\left[\widehat{\vartheta}_f^T f_R(x(t)) + \left(\widehat{\vartheta}_g^T g_R(x(t)) + 1\right) u(t)\right]\\ &+ B\left[\left(-\widehat{\vartheta}_f + \vartheta_f^*\right)^T f_R(x(t))\right.\\ &\left.+ \left((-\widehat{\vartheta}_g + \vartheta_g^*)^T g_R(x(t)) + 1\right) u(t)\right]\\ y(t) =& Cx(t)\end{aligned} \tag{5}$$

where $\widehat{\vartheta}_f$ and $\widehat{\vartheta}_g$ are on-line "estimates" of ϑ_f^* and ϑ_g^*, updated through a suitable adaptive law, and $\widehat{\vartheta}_f - \vartheta_f^*$ and $\widehat{\vartheta}_g - \vartheta_g^*$ are the estimation errors. The initial values of $\widehat{\vartheta}_f$ and $\widehat{\vartheta}_g$ may be set to $\widehat{\vartheta}_{f0}^*$ and $\widehat{\vartheta}_{g0}^*$, respectively.

On the other hand, the functions $f(x)$ and $g(x)$ themselves may be known only up to some degree of accuracy so that it is important to take into account the presence of unmodeled dynamics. To cope with these kinds of uncertainties, we define suitable optimality and robustness criteria by means of some admissibility constraints. The optimality criteria are formulated in terms of achieving either a guaranteed value or the minimum of some cost functionals, according to whether multiplicative or additive noise is taken into account. These cost functionals penalize the "distance" from a reference situation for which a "worst-case" controller is designed, and in the linear case, they reduce to a standard quadratic cost (see [20] and [21] for comparisons with other inverse optimal schemes for deterministic and stochastic nonlinear systems).

We show that the problem of finding a stabilizing controller for the class of systems (1) can be split into two lower dimensional problems: One is related to the case in which the *state x is available for feedback* (*state feedback problem*) and the other to the possibility of constructing an observer for estimating the state x (*output injection problem*). This corresponds to a *nonlinear separation* principle. First, we give a semiglobal in probability backstepping design procedure for solving the state feedback problem. While this procedure generalizes the classical semiglobal backstepping design for deterministic systems [14], it stands as a *practical semiglobal* version of the corresponding *global* result proved in [17] and [19]. Finally, we give constructive tools for the observer design. It turns out that *control saturations* and *high-gain observers* are instrumental to the control design exactly as in the case of [1]. In comparison with [1], we do not take into account the presence of a zero dynamics; on the other hand, our control strategy has some important advantages over [1] being robust against model uncertainties and random noise and optimal with respect to given performance indexes. We want to stress that the use of neural network approximation is important to better represent the nominal dynamics of (1) and simplify the structure of what is to be considered as uncertainty. This is even more important in the case that $f(x)$ and $g(x)$ themselves are known only up to some degree of accuracy.

II. Notations and Basic Notions

First, we give some notations extensively used throughout the paper.

- If $\|v\|$ denotes the 2–norm of any given vector v, by $\|A\|$, we denote the induced 2–norm of any given matrix A; by $\|v\|_A$, we denote the A-norm of v, i.e., $\|v\|_A = \sqrt{v^T Av}$; let col $(v_1, \ldots, v_n)$ be the column vector with ith entry equal to v_i.
- By $\mathcal{SP}^n$ (respectively, $\mathcal{SN}^n$), we denote the set of $n \times n$ positive (respectively, negative) definite symmetric matrices; by $\mathcal{SSP}^n$, we denote the set of $n \times n$ positive semidefinite symmetric matrices; $\mathbb{R}^+$ denotes the set of positive real numbers and $\mathbb{R}^{\geq}$ the set of nonnegative real numbers;
- For any vector-valued function $\eta: \mathbb{R}^s \to \mathbb{R}^r$, we denote by η_i (or $[\eta]_i$) its ith component.
- For any given set $\mathcal{S}$, we denote by $\overline{\mathcal{S}}$ its closure and by $\partial\mathcal{S}$ its boundary; moreover, given $\delta > 0$ and a set $\mathcal{S}$, by δ-*neighborhood* of $\mathcal{S}$, we denote the set $\mathcal{S}_\delta = \{z: \inf_{y \in \mathcal{S}} \|z - y\| < \delta\}$;
- For any sequence of sets $\{\mathcal{S}_k\}$, $\liminf_{k\to\infty} \mathcal{S}_k = \cup_{k=1}^{\infty} \cap_{i \geq k} \mathcal{S}_i$, and $\limsup_{k\to\infty} \mathcal{S}_k = \cap_{k=1}^{\infty} \cup_{i\geq k} \mathcal{S}_i$. It is easy to see that if $\liminf_{k\to\infty} \mathcal{S}_k \supset \mathcal{V}$, then there exists k° such that $\mathcal{S}_k \supseteq \mathcal{V}$ for all $k \geq k^\circ$. Similarly, if $\limsup_{k\to\infty} \mathcal{S}_k \subset \mathcal{V}$ then there exists k° such that $\mathcal{S}_k \subseteq \mathcal{V}$ for all $k \geq k^\circ$.

In the remaining part of this section, we shortly recall some notions of stochastic processes, referring the reader for the basic concepts to standard textbooks [25], [26]. We assume that the reader is familiar with the basic notions of probability theory and stochastic processes $\{x(t), t \in \mathbb{R}\}$ on a given probability space $(\Omega, \mathcal{F}, \mathbf{P})$ (we assume that the probability space and all the σ-algebras we consider are completed with all the subsets of sets having null measure). We denote by $\mathbf{E}\{\cdot\}$ the expectation and by $\mathbf{P}\{\cdot|\cdot\}$ $(\mathbf{E}\{\cdot|\cdot\})$ the conditional probability (expectation).

An important definition regards the notion of *Markov time*. Let $\{\mathcal{F}_t, t \in \mathbb{R}\}$ be an increasing family of right continuous σ-algebras contained in $\mathcal{F}$ (*filtration*).

Definition 2.1: A nonnegative random variable τ, $\tau \leq +\infty$, is called an $\mathcal{F}_t$ Markov time if for all $t \geq 0$ $\{\omega{:}\, \tau(\omega) \leq t\} \in \mathcal{F}_t$ (i.e., it is $\mathcal{F}_t$ adapted). If $\mathbf{P}\{\tau < \infty\} = 1$ then τ is called a stopping time.

A stochastic process $\{x(t), t \in \mathbb{R}\}$ is a **Wiener process** (with respect to $\{\mathcal{F}_t, t \in \mathbb{R}\}$) if $\mathbf{E}\{x(t)|\mathcal{F}_s\} = x(s)$ and $\mathbf{E}\{(x(t) - x(s))^2|\mathcal{F}_s\} = t - s$ for $t \geq s$. A stochastic process $\{x(t), t \in \mathbb{R}\}$ is a Markov process if for any collections $t_1 < \ldots < t_N$ and $r_1, \ldots, r_N$

$$\mathbf{P}\{x_{t_N} < r_N | x_{t_1} = r_1, \ldots, x_{t_{N-1}} = r_{N-1}\} = \mathbf{P}\{x_{t_N} < r_N | x_{t_{N-1}} = r_{N-1}\}. \quad (6)$$

For the corresponding definitions in the multidimensional case, we refer to [27].

By a *stochastic differential equation* , we mean the following:

$$dx(t) = f(x(t), t)dt + g(x(t), t)dw(t) \quad (7)$$

with initial condition $x(t_0) = \overline{x}$, where $\{w(t), t \in \mathbb{R}\}$ is a Wiener process (with respect to $\{\mathcal{F}_t, t \in \mathbb{R}\}$). The solution $x(t, t_0, \overline{x})$ of (7), whenever it exists, is a Markov process satisfying

$$x(t, t_0, \overline{x}) = \overline{x} + \int_{t_0}^{t} f(x(s, t_0, \overline{x}), s)ds + \int_{t_0}^{t} g(x(s, t_0, \overline{x}), s)dw(s) \quad (8)$$

almost surely (a.s.). The last integral is called *Itô integral*. It is well known ([15]) that if

$$\begin{aligned} &\|f(t, x_1) - f(x_2, t)\| + \|g(t, x_1) - g(t, x_2)\| \\ &\quad \leq K\|x_1 - x_2\| \\ &\|f(t, x)\| + \|g(t, x)\| \\ &\quad \leq H(1 + \|x\|) \end{aligned} \quad (9)$$

for all (x_1, t), (x_2, t), and (x, t) in $\mathcal{Z} \times [t_0, T]$, with $\mathcal{Z}$ a compact set containing $\overline{x}$, then there exists an a.s. unique stochastic process $x(t)$, sample continuous and satisfying (8) on $[t_0, \tau_{\mathcal{Z},T}(t)]$, where $\tau_{\mathcal{Z},T}(t) = \min(t, \tau_{\mathcal{Z}}, T)$, and $\tau_{\mathcal{Z}}$ is the Markov time (relatively to the σ-algebra generated by $\{x(s), s \leq t\}$) defined as the first time at which $x(t)$ reaches the boundary of $\mathcal{Z}$ [15].

An important property of solutions of stochastic differential equations is *regularity*. Consider a sequence of increasing bounded domains $\{\mathcal{Z}(n)\}$, containing the origin, such that the distance of the boundary from the origin goes to infinity as n tends to infinity, and let $\{\tau_{\mathcal{Z}(n)}\}$ be the corresponding sequence of Markov times. Since $\{\tau_{\mathcal{Z}(n)}\}$ is nondecreasing, its limit exists. We will say that the solution is *regular* if $\lim_{n\to\infty} \tau_{\mathcal{Z}(n)} = \infty$ a.s. Any regular solution can be uniquely (a.s.) extended for all $t \geq t_0$.

Any solution $x(t)$ of (7) satisfies the following *strong Markov property* [26]:

$$\mathbf{P}\{x(t+\tau, t_0, \overline{x}) \in A\} = \int \mathbf{P}\{\tau \in ds;\ x(\tau, t_0, \overline{x}) \in dz\}\mathbf{P}\{x(t+\tau, s, z) \in A\} \quad (10)$$

where τ is any given Markov time (relatively to the σ-algebra generated by $\{x(s), s \leq t\}$). In (10), we can substitute $\mathbf{P}\{\cdot\}$ with its conditioned version $\mathbf{P}\{\cdot|\cdot\}$ as long $\mathbf{P}\{\cdot|\cdot\}$ is regular, i.e., it is a function $p(\omega, A)$, measurable for each fixed A and a probability for each fixed ω.

From now on, we will denote $x(t, t_0, \overline{x})$, if not otherwise stated, simply by $x(t)$. Given a C^2 (measurable) function $V{:}\, \mathbb{R}^n \to \mathbb{R}$, define

$$\mathcal{L}V(x) = \frac{\partial V}{\partial x}(x)f(x,t) + \frac{1}{2}\mathbf{Tr}\left\{g^T(x,t)\frac{\partial^2 V}{\partial x^2}(x)g(x,t)\right\}. \quad (11)$$

Proposition 2.1 (Dynkin's Formula): Let $\overline{x} \in \mathcal{Z}$ a.s. The solution $x(t)$ of (7) satisfies on $[t_0, \tau_{\mathcal{Z},T}(t)]$ the following equation:

$$\mathbf{E}\{V(x(\tau_{\mathcal{Z},T}(t)))\} - V(\overline{x}) = \mathbf{E}\left\{\int_{t_0}^{\tau_{\mathcal{B},T}(t)} \mathcal{L}V(x(s))ds\right\}. \quad (12)$$

The integral appearing in the right-hand side of (12) is meant in the sense that

$$\int_{t_0}^{\tau_{\mathcal{Z},T}(t)} \mathcal{L}V(x(s))ds = \int_{t_0}^{t} \xi_{\tau_{\mathcal{Z},T}>t}\mathcal{L}V(x(s))ds$$

where $\xi_{\tau_{\mathcal{Z},T}>t}$ is the indicator function corresponding to the event $\{\tau_{\mathcal{Z},T} > t\}$.

Also, we will use extensively the following (generalized) Chebyshev inequality

$$\mathbf{P}\{\eta \notin \mathcal{S}\} \leq \frac{\mathbf{E}\{V(\eta)\}}{\inf_{s\in\mathbb{R}^n\setminus\mathcal{S}}\{V(s)\}} \quad (13)$$

where $\mathcal{S} \subset \mathbb{R}^n$, $V(\cdot)$ is real nonnegative, and η is a given random variable such that $\mathbf{E}\{V(\eta)\}$ exists. Finally, we recall the following fundamental formula of the differential calculus.

Proposition 2.2 (Itô Rule): Given a C^2 function $\varphi{:}\, \mathbb{R}^n \to \mathbb{R}$ and if $x(t)$ is a solution of (7), then

$$d\varphi(x(t)) = \frac{\partial \varphi}{\partial x}(x(t))dx(t) + \frac{1}{2}\mathbf{Tr}\left\{g^T(x(t),t)\frac{\partial^2 \varphi}{\partial x^2}(x(t))g(x(t),t)\right\}dt. \quad (14)$$

III. Basic Assumptions and Problem Formulation

The optimal ϑ_f^* and ϑ_g^* are contained in some known compact sets. Off-line training allows to obtain good estimates $\widehat{\vartheta}_{f0}$ and $\widehat{\vartheta}_{g0}$ of ϑ_f^* and ϑ_g^*, respectively. Correspondingly, one can select some compact sets Ω_f and Ω_g in such a way that $f(x) - (\vartheta_f^*)^T f_R(x)$ and $g(x) - (\vartheta_g^*)^T g_R(x)$ are comparable to $f(x) - \widehat{\vartheta}_{f0}^T f_R(x)$ and $g(x) - \widehat{\vartheta}_{g0}^T g_R(x)$, respectively.

Assume that Ω_f and Ω_g are convex hypercubes, i.e.,

$$\begin{aligned}\Omega_f =& \{\vartheta \in \mathbb{R}^{r_f} | a_i \le \vartheta_{fi} \le b_i, 1 \le i \le r_f\}\\ \Omega_g =& \{\vartheta \in \mathbb{R}^{r_g} | a_i \le \vartheta_{gi} \le b_i\\ & r_f + 1 \le i \le r_f + r_g\}\end{aligned} \tag{15}$$

and let $\widehat{\Omega} = \Omega_f \times \Omega_g$. Moreover, let

$$\Omega_{\delta f} = \{\vartheta_f \in \mathbb{R}^{r_f} | a_i - \delta \le \vartheta_{fi} \le b_i + \delta, 1 \le i \le r_f\}. \tag{16}$$

Similarly, define $\Omega_{\delta g}$, and let $\widehat{\Omega}_\delta = \widehat{\Omega}_{\delta f} \times \widehat{\Omega}_{\delta g}$. Let $\widehat{g}(x) = \widehat{\vartheta}_g^T g_R(x) + 1$.

Assumption 1: $\widehat{g}(x) \ge k_1 > 0$ for all x and $\widehat{\vartheta}_g \in \Omega_{\delta g}$.

Let $\sigma \in \mathbb{R}^n$ and $\widehat{\vartheta} = \mathrm{col}(\widehat{\vartheta}_f, \widehat{\vartheta}_g) = \mathrm{col}(\widehat{\vartheta}_{f1}, \ldots, \widehat{\vartheta}_{fr_f}, \widehat{\vartheta}_{g1}, \ldots, \widehat{\vartheta}_{gr_g})$. We consider nonlinear stochastic systems Σ of the form (1), where $x(t) \in \mathbb{R}^n$, $u(t)$, $y(t) \in \mathbb{R}$, $w(t)$ is an s-dimensional Wiener process and $\Phi = \Phi_F + \Phi_G u \in \mathbb{R}^r$ represents model uncertainties and nonlinearities. Moreover, for each sequence of real positive extended numbers $\{\Delta(k)\}$, with $k \in \mathbb{R}^+$, we consider the following class of controllers $\{\mathcal{C}(k)\}$:

$$\begin{aligned}u =& \eta_1\left(\frac{F(k)\sigma - \widehat{\vartheta}_f^T f_R(\sigma)}{\widehat{\vartheta}_g^T g_R(\sigma) + 1}\right)\\ d\sigma =& L(k)\sigma + B\left(\widehat{\vartheta}_f^T f_R(\sigma) + \widehat{\vartheta}_g^T (g_R(\sigma)+1)u\right)dt\\ &+ G(k)dy\\ d\widehat{\vartheta} =& \eta_2(\mathrm{Proj}(\widehat{\vartheta}, \phi(\sigma, u)))dt\\ \widehat{\vartheta}(0) =& \mathrm{col}\left(\widehat{\vartheta}_{f0}, \widehat{\vartheta}_{g0}\right)\end{aligned} \tag{17}$$

with (18), shown at the bottom of the page, where $\phi(\sigma, u) = \mathrm{col}(\phi_f(\sigma, u), \phi_g(\sigma, u))$, $\phi_f: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{r_f}$, $\phi_g: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{r_g}$, $\{\gamma(k)\}$, $\{P_{SF}(k)\}$ and $\{G(k)\}$ are suitable sequences, with $\gamma(k) > 0$ and $P_{SF}(k) \in \mathcal{SP}^n$, and η_j, $j = 1, 2$, are suitable C^0 functions such that

$$\|\eta_j(s)\| \le \Delta(k), \quad \forall\, s \tag{19}$$

$$\eta_j(s) = s, \quad \|s\| \le s_0 \tag{20}$$

for some $s_0 > 0$ and for $j = 1, 2$. While η_j, $j = 1, 2$, are designed in such a way to counteract the destabilizing effects due to large values of $G(k)$ (peaking), $\Delta(k)$ accounts for possible limitations on the control u (as an example, saturations of the control actuators).

For the analysis of stability of the closed-loop system (1)–(17), we also define a class of *candidate Lyapunov functions* $\{V_k^e\}$

$$\begin{aligned}V_{SF,k}(x, \widetilde{\vartheta}) =& \|x\|^2_{P_{SF}(k)} + \frac{1}{2}\widetilde{\vartheta}_f^T \Gamma_f^{-1}(k)\widetilde{\vartheta}_f + \frac{1}{2}\widetilde{\vartheta}_g^T \Gamma_g^{-1}(k)\widetilde{\vartheta}_g\\ V_k^e(x, \sigma, \widetilde{\vartheta}) =& V_{SF,k}(x, \widetilde{\vartheta}) + \varphi\left(\|x - \sigma\|^2_{P_m(k)}\right)\end{aligned} \tag{21}$$

where $\widetilde{\vartheta}_f = \widehat{\vartheta}_f - \vartheta_f^*$, $\widetilde{\vartheta}_g = \widehat{\vartheta}_g - \vartheta_g^*$ and $\{P_m(k)\}$ is a sequence in $\mathcal{SP}^n$, $\Gamma_f^{-1}(k)$ and $\Gamma_g^{-1}(k)$ are sequences of diagonal matrices in $\mathcal{SP}^n$ and $\varphi: \mathbb{R}^{\ge} \to \mathbb{R}$ is a suitable (at least) C^2, positive definite and proper function such that

$$\frac{\partial^2 \varphi}{\partial s^2}(s) \le 0 < \frac{\partial \varphi}{\partial s}(s) \le 1 \tag{22}$$

for all $s \ge 0$. Conditions (22) imply that over any compact set containing the origin any *candidate Lyapunov function is bounded from below and above by a quadratic function* and are is instrumental in enlarging the region of attraction of the closed-loop system.

Next, we define some *admissibility constraints* for the noise coefficient H and for the uncertainty term Φ. For, define the following compact sets:

$$\begin{aligned}\Omega(k) =& \left\{x \in \mathbb{R}^n, \widetilde{\vartheta} \in \mathbb{R}^{r_f + r_g}: V_{SF,k}(x, \widetilde{\vartheta}) \le k\right\}\\ \mathcal{U}_{\Delta(k)} =& \{u \in \mathbb{R}^m: \|u\| \le \Delta(k)\}.\end{aligned} \tag{23}$$

Let $\{E(k)\}$, $\{R_1(k)\}$ and $\{c_1(k)\}$ be sequences in $\mathcal{SSP}^n$, $\mathcal{SP}^m$ and $\mathbb{R}^{\ge}$, respectively. Define (24), shown at the bottom of the page, where e is the estimation error, and

$$v = \widehat{\vartheta}_f^T f_R(x) + \left(\widehat{\vartheta}_g^T g_R(x) + 1\right)u. \tag{25}$$

Let $\mathcal{F}(k)$ be the class of C^0 functions $\Phi: \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^r$ such that $\widetilde{\mathcal{P}}_1(u, x, e, k) \ge 0$ for all for all e, $(x, \widetilde{\vartheta}) \in \Omega(k)$, $\vartheta^* \in \widehat{\Omega}$ and $u \in \mathcal{U}_{\Delta(k)}$. Note that $\mathcal{P}_1$ penalizes the distance of Φ from being the worst-case uncertainty

$$\Phi_*^e = \frac{1}{\gamma^2(k)}\left[B^T P_{SF}(k)x + \left.\frac{\partial \varphi}{\partial s}\right|_{s = \|e\|^2_{P_m(k)}} B^T P_m(k)e\right]$$

and, at the same time, from being a linear (parameterized by k) function of x and u, representing a worst-case situation with

$$L(k) = A + \frac{1}{\gamma^2(k)}BB^T P_{SF}(k) - G(k)C$$

$$\mathrm{Proj}(\widehat{\vartheta}, \phi(\sigma, u)) =: \begin{cases}\phi_i(\sigma, u)\left(1 + \frac{b_i - \widehat{\vartheta}_i}{\delta}\right), & \text{if } \widehat{\vartheta}_i > b_i \text{ and } \phi_i(\sigma, u) > 0\\ \phi_i(\sigma, u)\left(1 + \frac{\widehat{\vartheta}_i - a_i}{\delta}\right), & \text{if } \widehat{\vartheta}_i < a_i \text{ and } \phi_i(\sigma, u) < 0\\ \varphi_i, & \text{otherwise}\end{cases} \tag{18}$$

respect to which the matrices $F(k)$ and $G(k)$ are designed. Note also that $c_1(k)$ accounts for additive uncertainty (i.e., $f(0) \neq 0$).

Let $\{\widehat{H}_j(k)\}$, $j = 1, \ldots, s$, be a sequence in $\mathbb{R}^{n\times 1}$ and $\{c_2(k)\}$ a sequence in $\mathbb{R}^{\geq}$. Define

$$\begin{aligned}\mathcal{P}_2(u,x,e,\widehat{\vartheta},k) = &-\mathbf{Tr}\left\{H^T(x)B^T P_{SF}(k)BH(x))\right\} \\ &+\sum_{j=1}^{s} x^T \widehat{H}_j^T(k)B^T P_{SF}(k)B\widehat{H}_j(k)x + c_2(k).\end{aligned} \tag{26}$$

Note that $\mathcal{P}_2$ penalizes the distance of $\mathbf{Tr}\{H^T(x)B^T P_{SF}(k)BH(x))\}$ from a sum of quadratic functions. Let $\mathcal{H}(k)$ be the class of C^0 functions $H\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^{n\times s}$ such that $\mathcal{P}_2(u,x,e,\widehat{\vartheta},k) \geq 0$, for all e, $(x,\widetilde{\vartheta}) \in \Omega(k)$, $\vartheta^* \in \widehat{\Omega}$ and $u \in \mathcal{U}_{\Delta(k)}$. Since $\Omega(k)$ and $\mathcal{U}_{\Delta(k)}$ are compact sets, the admissibility constraints on $\widetilde{\mathcal{P}}_1$ and $\mathcal{P}_2$ can be always met whenever Φ and H are *locally Lipschitz*. Note also that $c_2(k)$ accounts for additive noise (i.e., $H(0) \neq 0$).

Let $\{Q_m(k)\}$ and $\{c_3(k)\}$ be sequences in $\mathcal{SP}^n$ and $\mathbb{R}^{\geq}$ respectively. Define $\mathcal{D}(k)$ as the class of triples of C^0 functions $\eta_1\colon \mathbb{R}\to\mathbb{R}$, $\eta_2\colon \mathbb{R}^{r_f+r_g} \to \mathbb{R}^{r_f+r_g}$ and $\varphi\colon \mathbb{R}^{\geq} \to \mathbb{R}$ satisfying (19), (20), and (22) and such that

$$\begin{aligned}\mathcal{P}_3(u,x,e,k) = &-\|v-F(k)x\|^2_{R_1(k)} + \|x\|^2_{Q_{SF}(k)} \\ &+ c_3(k) + \left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|^2_{P_m(k)}} \\ &\times\Big[\|F(k)e\|^2_{R_1(k)} + \|e\|^2_{Q_m(k)} \\ &\quad - 2e^T P_m(k) \\ &\quad \times B\Big((\widehat{\vartheta}_f)^T(f_R(x)-f_R(x-e)) \\ &\quad +(\widehat{\vartheta}_g)^T(g_R(x)-g_R(x-e))\Big) \\ &\quad - \mathbf{Tr}\{H^T(x)P_m(k)H(x)\}\Big] \\ &- \widetilde{\vartheta}^T \Gamma^{-1}(k) \\ &\times[\eta_2(\mathrm{Proj}(\widehat{\vartheta},\phi(x-e,u))) \\ &\quad - \phi(x,u)] > 0\end{aligned} \tag{27}$$

with $\Gamma(k) = \mathrm{diag} = \{\Gamma_f(k), \Gamma_g(k)\}$, $\widetilde{\vartheta} = \mathrm{diag}\{\widetilde{\vartheta}_f, \widetilde{\vartheta}_g\}$ and u is as in (17), for all $(x,\widetilde{\vartheta},e) \in \Omega(k)\times\mathbb{R}^n$, $(x,e)\neq 0$ and $\vartheta^* \in \widehat{\Omega}$. The functions η_1, η_2 are typically designed to avoid the peaking phenomenon, while φ is instrumental in enlarging the region of attraction of the closed-loop system. Moreover, similar remarks can be repeated for the constants $c_3(k)$.

Finally, let

$$\begin{aligned}\mathcal{P}_4(u,x,e,\widehat{\vartheta},k) = &\frac{1}{\gamma^2(k)}\left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|^2_{P_m(k)}} \\ &\times\left[\left(1-\left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|^2_{P_m(k)}}\right)e^T P_m(k)\right. \\ &\quad\times BB^T P_m(k)e + e^T P_m(k)G(k)R_2 \\ &\quad\times(k)G^T(k)P_m(k)e\Big] \\ &-\left.\frac{\partial^2\varphi}{\partial s^2}\right|_{s=\|e\|^2_{P_m(k)}} \\ &\times e^T P_m(k)H(x)H^T(x)P_m(k)e\end{aligned} \tag{28}$$

for some sequence $\{R_2(k)\}$ in $\mathcal{SP}^p$. Note that, by (22), (28) is nonnegative for all $e \in \mathbb{R}^n$. Note also that $\mathcal{P}_4$ penalizes the distance from the situation for which φ is linear (i.e., quadratic Lyapunov functions).

In what follows, we will refer to $\Phi \in \mathcal{F}(k)$, $H \in \mathcal{H}(k)$ and $(\eta_1,\eta_2,\varphi) \in \mathcal{D}(k)$ as *admissible functions*. Moreover, any choice of $\{P_{SF}(k)\}$, $\{P_m(k)\}$, $\{Q_{SF}(k)\}$, $\{Q_m(k)\}$, $\{R_1(k)\}$, $\{R_2(k)\}$, $\{\gamma(k)\}$, $\{E(k)\}$, $\{c_j(k)\}$, $j = 1,2,3$, $\{\widehat{H}_j(k)\}$, $j = 1,\ldots,s$, for which $\Phi \in \mathcal{F}(k)$ and $H \in \mathcal{H}(k)$ will be referred to as *admissible parameterization*.

Denote by $x_k^e(t,t_0,x_0^e) = \mathrm{col}(x_k(t,t_0,x_0^e), \sigma_k(t,t_0,x_0^e), \widetilde{\vartheta}_k(t,t_0,x_0^e))$ the trajectory of the closed-loop system $\Sigma\circ\mathcal{C}(k)$ at time $t \geq t_0$ stemming from $x_0^e = \mathrm{col}(x_0,\sigma_0,\widetilde{\vartheta}_0)$. With some abuse of notation, wherever there is no ambiguity, we will use $x_k^e(t)$ instead of $x_k^e(t,t_0,x_0^e)$. Moreover, let $e_k(t) = x_k(t) - \sigma_k(t)$. We introduce a sequence of cost functionals $\{J(k)\}$ as follows:

$$\begin{aligned}J(k) = &\lim_{T\to\infty}\frac{1}{T-t_0} \\ &\times\int_{t_0}^{T}\mathbf{E}\left\{\sum_{j=1}^{4}\mathcal{P}_j(u(t),x_k(t),e_k(t),\widehat{\vartheta}(t),k)\right\}dt.\end{aligned} \tag{29}$$

Note that $J(k) \geq 0$ for any $\Phi \in \mathcal{F}(k)$, $H \in \mathcal{H}(k)$ and $(\eta_1,\eta_2,\varphi) \in \mathcal{D}(k)$. Dividing by $T - t_0$ allows to cope with the case of both *additive and multiplicative noise*, since otherwise $c_j(k) \neq 0$ for at least one j would make the cost diverge.

The aim of this paper is to study under which conditions it is possible to modify the behavior of (1) in such a way that $J(k)$ achieves a guaranteed value and to obtain stability in some "stochastic" sense. To make the last point precise, let us give the following definition.

$$\begin{aligned}\mathcal{P}_1(u,x,e,\widehat{\vartheta},k) = &\widetilde{\mathcal{P}}_1(u,x,e,\widehat{\vartheta},k) + \gamma^2\left\|\Phi(u,x) - \frac{1}{\gamma^2(k)}\left[B^T P_{SF}(k)x + \left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|^2_{P_m(k)}} B^T P_m(k)e)\right]\right\|^2 \\ \widetilde{\mathcal{P}}_1(u,x,e,\widehat{\vartheta},k) = &-\gamma^2(k)\|\Phi(u,x)\|^2 + \|x\|^2_{E(k)} + \|v\|^2_{R_1(k)} + c_1(k)\end{aligned} \tag{24}$$

Definition 3.1: Let $\alpha, \beta \in [0,1)$ and $\Omega^e, \mathcal{B}^e \subset \mathbb{R}^{2n+r_f+r_g}$ be compact sets. The system (1) is said to be $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$-stabilizable in probability (or $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$-SP) if there exist a sequence of admissible control laws $\{\mathcal{C}(k)\}$, a sequence of compact sets $\{\Omega^e(k)\}$ and open sets $\{\mathcal{B}^e(k)\}$ of $\mathbb{R}^{2n+r_f+r_g}$ such that

1) there exists k^o such that $\Omega^e(k) \supset \Omega^e \supset \mathcal{B}^e(k) \forall k \geq k^o$;
2) for each $\delta > 0$ and $\Phi \in \mathcal{F}(k)$

$$\liminf_{k\to\infty} \inf_{x_0^e \in \overline{\mathcal{B}}^e(k)} \mathbf{P}\left\{x_k^e(t) \in \overline{\mathcal{B}}_\delta^e \;\forall\, t \geq t_0\right\} \geq 1-\beta; \tag{30}$$

3) for each $\delta > 0$ and $\Phi \in \mathcal{F}(k)$

$$\begin{aligned}
&\liminf_{k\to\infty} \inf_{x_0^e \in \Omega^e \setminus \mathcal{B}^e(k)} \\
&\quad \mathbf{P}\big\{x_k^e(t) \in \Omega^e(k) \;\forall\, t \geq t_0 \text{ and} \\
&\qquad\qquad x_k^e(t + \tau_{\mathbb{R}^{2n+r_f+r_g}\setminus \mathcal{B}^e(k)}) \in \mathcal{B}_\delta^e \;\forall\, t \geq 0 \\
&\qquad \text{and } \tau_{\mathbb{R}^{2n+r_f+r_g}\setminus \mathcal{B}^e(k)} < \infty\big\} \\
&\quad \geq (1-\alpha)(1-\beta).
\end{aligned} \tag{31}$$

Note that the events in (30) and (31) are measurable by separability and measurability (on the product σ-algebra $\mathcal{F} \times \mathcal{I}$, where $\mathcal{I}$ is the Borel σ-algebra on the line) of the process $x_k^e(t)$ and being $\{\tau_{\mathbb{R}^{2n+r_f+r_g}\setminus \mathcal{B}^e(k)} \leq t\}$ adapted to the σ-algebra generated by $\{x_k^e(s), s \leq t\}$.

The set Ω^e gives the *guaranteed region of attraction* of the closed-loop system $\Sigma \circ \mathcal{C}(k)$, while $\mathcal{B}^e$ represents the *target set*. From 1), it follows that there exists k° such that $\Omega^e(k) \supset \Omega^e \supset \mathcal{B}^e \supset \mathcal{B}^e(k) \subset \mathcal{B}_\delta^e$ for all $k \geq k^\circ$. Property 2) is a *local* property with respect to $\mathcal{B}^e$: For each δ-neighborhood of $\mathcal{B}^e$, there exists sufficiently large k° for which the probability that the trajectories $x_k^e(t)$ of the closed-loop system $\Sigma \circ \mathcal{C}(k)$, starting from $\overline{\mathcal{B}}^e(k)$, stay forever in $\overline{\mathcal{B}}_\delta^e$ is at least $1-\beta$ for all $k \geq k^\circ$. Property 3) is a property *in the large* with respect to Ω^e: There exists sufficiently large k° for which the trajectories of $\Sigma \circ \mathcal{C}(k)$ starting inside Ω^e remain inside $\Omega^e(k)$, eventually enter any given δ-neighborhood of the target set $\mathcal{B}^e$ in finite time, and remain thereinafter with probability at least $(1-\alpha)(1-\beta)$ for all $k \geq k^\circ$. The numbers α and β are given *risk margins: the first one quantifies the risk of leaving the compact set* $\Omega^e(k)$ *with initial condition in* Ω^e *rather than getting close to the target, while the second one gives a risk margin for remaining close to the target*. Note also that 3) requires that $\tau_{\mathbb{R}^{2n+r_f+r_g}\setminus \mathcal{B}^e(k)} < \infty$. As it will be clear in Section IV, under the standard assumptions of local existence and uniqueness a.s. of trajectories, each Markov time $\tau_{\mathbb{R}^{2n+r_f+r_g}\setminus \mathcal{B}^e(k)}$, conditioned to $x_k(t) \in \Omega^e(k)$ for all $t \geq t_0$, is always finite and $\tau_{\mathbb{R}^{2n+r_f+r_g}\setminus \mathcal{B}^e(k)} \to \infty$ as $k \to \infty$ as long as $\limsup_{k\to\infty} \mathcal{B}^e(k) = \{0\}$. In particular, this implies that, if $\mathcal{B}^e = \{0\}$, the trajectory approaches the origin as $t \to \infty$.

The role of the risk margins, region of attraction, and target set is peculiar of our setup and become unessential in the classical definitions given in [15]. If $\mathcal{B}^e = \{0\}$, $\alpha = \beta = 0$ and $\{\mathcal{C}(k)\} = \mathcal{C}$ for all k, Definition 3.1 recovers the classical definition of *asymptotic stability in probability* [15]. If in addition $\Omega^e = \mathbb{R}^{2n+r_f+r_g}$, Definition 3.1 gives the notion of *asymptotic stability in probability in the large* [15]. On the other hand, if $\mathcal{B}^e = \{0\}$ and Ω^e can be taken any *a priori* given compact set of $\mathbb{R}^{2n+r_f+r_g}$ and α and β any *a priori* given numbers in $[0,1)$, our definition gives a stochastic analog of the concept of *semiglobal stabilization*, as introduced in [6]. If $\Omega^e = \mathbb{R}^{2n+r_f+r_g}$ and $\mathcal{B}^e$ can be taken any *a priori* given compact set of $\mathbb{R}^{2n+r_f+r_g}$ and α and β any *a priori* given number in $[0,1)$, Definition 3.1 extends to a stochastic setting the concept of *practical stabilization*.

All the above remarks can be straightforwardly extended to the definition of stability in quadratic mean.

Remark 3.1: Definition 3.1 requires convergence of parameter estimates to the optimal values in some probabilistic sense. However, if one is not interested in the parameter convergence this definition should be slightly modified in 3) with regard to the parameter estimation errors $\tilde{\vartheta}$ by only requiring their boundedness in probability. For sake of simplicity, in the rest of the paper by $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$-stability in probability, we will refer to this situation, being only interested in the asymptotic stability of the state trajectories.

We are ready to formulate our problems.

Nonlinear Stabilization in Probability with Guaranteed Cost (NSPG): Let $\Phi \in \mathcal{F}(k)$, $H \in \mathcal{H}(k)$, $\mathcal{B}^e \subset \Omega^e$ be compact sets of $\mathbb{R}^{2n+r_f+r_g}$, $\alpha, \beta \in [0,1)$, $x_0^e \in \Omega^e$ and $\{\Delta(k)\}$ and $\{\overline{\omega}(k)\}$ given sequences in $(0,\infty]$ and $\mathbb{R}^{\geq}$, respectively. Find an admissible parameterization and $(\eta_1, \eta_2, \varphi) \in \mathcal{D}(k)$ such that

- *(Guaranteed cost)* along the trajectories $x_k^e(t)$ of the closed-loop systems $\Sigma \circ \mathcal{C}(k)$
$$\liminf_{k\to\infty} \mathbf{Pr}\{J(k) \leq \overline{\omega}(k)\} \geq (1-\alpha). \tag{32}$$
- *(Stability)* $\Sigma \circ \mathcal{C}(k)$ is $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$-stable in probability.

IV. Main Results

Let

$$H(x) = (\,H_1(x) \quad \ldots \quad H_s(x)\,). \tag{33}$$

Theorem 4.1: Assume that there exist an admissible parameterization and $(\eta_1, \eta_2, \varphi) \in \mathcal{D}(k)$ such that

- *state feedback (SF)*
$$\begin{aligned}
&A^T P_{SF}(k) + P_{SF}(k)A + \frac{1}{\gamma^2(k)} P_{SF}(k)BB^T P_{SF}(k) \\
&\quad + E(k) - F^T(k)R_1(k)F(k) \\
&\quad + \sum_{j=1}^{s} \widehat{H}_j^T(k)P_{SF}(k)\widehat{H}_j(k) \leq -Q_{SF}(k)
\end{aligned} \tag{34}$$
where
$$F(k) = -R_1^{-1}(k)B^T P_{SF}(k). \tag{35}$$
- *output injection (OI)* for all $e \in \mathbb{R}^n$ and $x \in \Omega(k)$
$$\begin{aligned}
&P_m(k)\left(A + \frac{1}{\gamma^2(k)}BB^T P_{SF}(k)\right) \\
&\quad + \left(A + \frac{1}{\gamma^2(k)}BB^T P_{SF}(k)\right)^T P_m(k) \\
&\quad + \frac{1}{\gamma^2(k)} P_m(k)BB^T P_m(k) - \gamma^2(k)C^T R_2^{-1}(k)C \\
&\quad + F^T(k)R_1(k)F(k) \leq -Q_m(k)
\end{aligned} \tag{36}$$

- *risk margins (RM)* if

$$\Omega^e(k) = \left\{(x,\sigma,\widetilde{\vartheta}) \in \mathbb{R}^{2n+r_f+r_g} : V_k^e(x,\sigma,\widetilde{\vartheta}) \le k\right\} \quad (37)$$

and $\{\mathcal{B}^e(k)\}$, a sequence of open sets of $\mathbb{R}^{2n+r_f+r_g}$, are such that

$$\limsup_{k\to\infty} \mathcal{B}^e(k) \subseteq \mathcal{B}^e \subset \Omega^e \subset \liminf_{k\to\infty} \Omega^e(k) \quad (38)$$

then for each $\delta > 0$

$$\limsup_{k\to\infty} \sup_{(x,\sigma,\widetilde{\vartheta})\in\Omega^e\setminus\mathcal{B}^e(k)} \limsup_{k\to\infty} \frac{V_k^e(x,\sigma\widetilde{\vartheta})}{k} \le \alpha \quad (39)$$

$$\limsup_{k\to\infty} \sup_{(x,\sigma,\widetilde{\vartheta})\in\partial\mathcal{B}^e(k)} \frac{V_k^e(x,\sigma,\widetilde{\vartheta})}{\inf_{(s_1,s_2,s_3)\in\mathbf{R}^{2n+r_f+r_g}\setminus\overline{\mathcal{B}}_\delta^e} V_k^e(s_1,s_2,s_3)} \le \beta \quad (40)$$

for all t and $(x,\sigma,\widetilde{\vartheta}) \in \Omega^e(k)\setminus\mathcal{B}^e(k)$.

Under the above assumptions, the controller (17) with $F(k)$ as in (35), and

$$G(k) = \gamma^2(k) P_m^{-1}(k) C^T R_2^{-1}(k) \quad (41)$$

solves NSPG with $\overline{\omega}(k) = \sum_{j=1}^3 c_j(k)$.

Proof: Throughout the proof, unless otherwise stated, we will omit k and the arguments of Φ and H. Moreover, we can assume $k \ge k^*$, where k^* is such that $\Omega^e(k) \supseteq \Omega^e \supset \mathcal{B}^e(k)$ for all $k \ge k^*$ (this is always possible by (38)).

Let V_{SF}, V_k^e as in Section III and $e = x - \sigma$. The closed-loop system is

$$\begin{aligned} dx =& \left(\left(A + \frac{1}{\gamma^2}BB^TP_{SF}\right)x \right.\\ &\left. + B\left(v + \widetilde{\vartheta}_f^T f_R(x) - \widetilde{\vartheta}_g^T(g_R(x)+1)u\right) + B\widetilde{\Phi}\right)dt \\ &+ H dw \\ de =& \left(Le + B\left(\widehat{\vartheta}_f^T f_R(x) - \widehat{\vartheta}_f^T f_R(\sigma) + \widehat{\vartheta}_g^T(g_R(x)+1)u \right.\right.\\ &\left.\left. - \widehat{\vartheta}_g^T(g_R(\sigma)+1)u\right) + B\widetilde{\Phi}\right)dt + H dw \\ d\widehat{\vartheta} =& \eta_2\left(\mathrm{Proj}(\widehat{\vartheta},\phi(\sigma,u))\right)dt \end{aligned} \quad (42)$$

where v and u are as in (17) and (25) and $\widetilde{\Phi} = \Phi - (1/\gamma^2)B^TP_{SF}x$.

By (18) and (11)

$$\begin{aligned} \mathcal{L}\varphi =& \frac{\partial\varphi}{\partial s}_{s=\|e\|_{P_m}^2} 2e^TP_m \\ &\times\left[Le + B\left(\widehat{\vartheta}_f^T f_R(x) - \widehat{\vartheta}_f^T f_R(\sigma) \right.\right.\\ &\left.\left. + \widehat{\vartheta}_g^T(g_R(x)+1)u - \widehat{\vartheta}_g^T(g_R(\sigma)+1)u\right) + B\widetilde{\Phi}\right] \\ &+ \frac{1}{2}\mathbf{Tr}\left\{H^T\frac{\partial^2\varphi}{\partial e^2}H\right\} \\ \mathcal{L}V_{SF,k} \le& 2x^TP_{SF}\left[\left(A + \frac{1}{\gamma^2}BB^TP_{SF}\right)x + Bv + B\widetilde{\Phi}\right] \\ &+ \mathbf{Tr}\left\{H^TP_{SF}H\right\} \\ &+ \widetilde{\vartheta}^T\Gamma^{-1}\left[\eta_2\left(\mathrm{Proj}(\widehat{\vartheta},\phi(\sigma,u))\right) - \phi(x,u)\right] \end{aligned} \quad (43)$$

with $\phi_f(x,u) = \Gamma_f(2x^TP_{SF}Bf_R(x))^T$ and $\phi_g(x,u) = \Gamma_g(2x^TP_{SF}B(g_R(x)+1)u)^T$.

Moreover

$$\frac{\partial^2\varphi}{\partial e^2} = 2\left.\frac{\partial^2\varphi}{\partial s^2}\right|_{s=\|e\|_{P_m}^2} P_m ee^T P_m + 2\left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|_{P_m}^2} P_m. \quad (44)$$

Since $\mathbf{Tr}(AB) = \mathbf{Tr}(BA)$, by (44) one has

$$\begin{aligned} \frac{1}{2}\mathbf{Tr}\left\{H^T\frac{\partial^2\varphi}{\partial e^2}H\right\} =& \left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|_{P_m}^2} \mathbf{Tr}\left\{H^TP_mH\right\} \\ &+ \left.\frac{\partial^2\varphi}{\partial s^2}\right|_{s=\|e\|_{P_m}^2} e^TP_mHH^TP_me. \end{aligned} \quad (45a)$$

Using (36) and (45a), we have

$$\begin{aligned} &\mathcal{L}\varphi - \gamma^2\|\widetilde{\Phi}\|^2 + \left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|_{P_m}^2} \|F(k)e\|_{R_1(k)}^2 \\ &= -\|v - Fx\|_{R_1}^2 + \|x\|_{Q_{SF}}^2 - \widetilde{\vartheta}^T\Gamma^{-1} \\ &\quad\times\left[\eta_2\left(\mathrm{Proj}\left(\widehat{\vartheta},\phi(\sigma,u)\right)\right) - \phi(x,u)\right] \\ &\quad - \mathcal{P}_1 + \widetilde{\mathcal{P}}_1 - \mathcal{P}_3 - \mathcal{P}_4 + c_3 + \left.\frac{\partial\varphi}{\partial s}\right|_{s=\|e\|_{P_m}^2} \\ &\quad\times\|F(k)e\|_{R_1(k)}^2. \end{aligned} \quad (45b)$$

Moreover, for all u by completing the square and using (34)

$$\begin{aligned} \mathcal{L}V_{SF,k} \le& \|v - Fx\|_{R_1}^2 - \gamma^2\|\widetilde{\Phi}\|^2 - \|x\|_{Q_{SF}}^2 \\ &- \widetilde{\mathcal{P}}_1 - \mathcal{P}_2 + c_1 + c_2 + \widetilde{\vartheta}^T\Gamma^{-1} \\ &\times\left[\eta_2\left(\mathrm{Proj}\left(\widehat{\vartheta},\phi(\sigma,u)\right)\right) - \phi(x,u)\right]. \end{aligned} \quad (46)$$

Summing up together (45b) and (46), we conclude that

$$\mathcal{L}V_k^e + \sum_{j=1}^4 \mathcal{P}_j \le \overline{\omega}. \quad (47)$$

We are left with proving the following facts:

— $J(k) \le \overline{\omega}(k)$, conditionally to the event $x_k^e(t) \in \Omega^e(k)$ for all $t \ge t_0$;

— $\liminf_{k\to\infty} \mathbf{Pr}\{(x_k^e(t) \in \Omega^e(k)\} \ge 1 - \alpha$ for all $t \ge t_0$;

— $\Sigma \circ \mathcal{C}(k)$ is $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$ stable in probability.

First, we prove that $J(k) \le \overline{\omega}(k)$, conditionally to the event $x_k^e(t) \in \Omega^e(k)$ for all $t \ge t_0$. By (47) and Dynkin's formula for each $T > t_0$

$$\begin{aligned} &\frac{1}{T - t_0}\int_{t_0}^T \mathbf{E}\left\{\sum_{j=1}^4 \mathcal{P}_j(u, x_k^e(s), k)\right\} ds \\ &\le \frac{1}{T - t_0}\left(V_k^e\left(x_0^e\right) - \mathbf{E}\left\{V_k^e\left(x_k^e(T)\right)\right\}\right) + \overline{\omega}(k) \end{aligned} \quad (48)$$

for all $x_0^e \in \Omega^e(k)$. From (48), letting $T \to \infty$ and since $V_k^e \ge 0$, we obtain $0 \le J(k) \le \overline{\omega}(k)$.

Next, we show that $\Sigma \circ \mathcal{C}(k)$ is $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$ stable in probability. Using (47) and (RM), the $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$ stability of $\Sigma \circ \mathcal{C}(k)$ is a consequence of the following lemma, which can be proved by suitably modifying the proof of lemma 5.1 in [5]. □

Lemma 4.1: The system (1) is $(\Omega^e, \mathcal{B}^e, \alpha, \beta)$-SP if there exist a sequence of admissible control laws $\{\mathcal{C}(k)\}$, a sequence of (at least) C^2, positive definite and proper functions $\{V_k^e(x^e)\}$, a sequence of C^0, positive definite functions $\{Q_k^e(x,\sigma)\}$ and open sets $\{\mathcal{B}^e(k)\}$, $\mathcal{B}^e(k) \subset \mathbb{R}^{2n+r_f+r_g}$, containing the origin, such that

4) there exists k^o such that $\Omega^e(k) \supset \Omega^e \supset \mathcal{B}^e(k) \forall k \geq k^o$;, where

$$\Omega^e(k) = \left\{z \in \mathbb{R}^{2n+r_f+r_g} : V_k^e(z) \leq k\right\}$$

5) $\mathcal{L}V_k^e(x^e) \leq -Q_k^e(x,\sigma)$ for all k, t, $\Phi \in \mathcal{F}(k)$ and $x^e \in \Omega^e(k) \backslash \mathcal{B}^e(k)$;

6) $\limsup_{k\to\infty} \sup_{x^e \in \Omega^e \backslash \mathcal{B}^e(k)} (V_k^e(x^e)/k) \leq \alpha$ and $\limsup_{k\to\infty} \sup_{x^e \in \partial\mathcal{B}^e(k)} (V_k^e(x^e)/\inf_{z\in\mathbb{R}^{2n+r_f+r_g}\backslash \overline{\mathcal{B}}_\delta^e} V_k^e(z)) \leq \beta$ for each $\delta > 0$.

Remark 4.1: We note that, as a consequence of (39), if

$$\limsup_{k\to\infty} \frac{V_k^e(x^e)}{k} = 0 \tag{49}$$

for each x^e then the risk margin α can be taken *any* number in $[0, 1)$, and any *a priori* given compact set can be included in Ω^e. Thus, (49), together with 4)–6) of Lemma 4.1, guarantee *semiglobal stabilization in probability*.

On the other hand, if $\limsup_{k\to\infty} \mathcal{B}^e(k) = \{0\}$, then for each $\delta > 0$

$$\limsup_{k\to\infty} \sup_{x^e \in \partial\mathcal{B}^e(k)} \frac{V_k^e(x^e)}{\inf_{z\in\mathbb{R}^{2n+r_f+r_g}\backslash \overline{\mathcal{B}}_\delta^e} V_k^e(z)} = 0 \tag{50}$$

and the risk margin β can be taken *any* number in $[0, 1)$. Moreover, any *a priori* given compact set can be chosen as target set and (50), together with 4)–6) of Lemma 4.1, guarantee *practical stabilization in probability*. □

Remark 4.2: The proof of Lemma 4.1 is based on a *probabilistic invariance property* which extends to a stochastic setup the following well-known property: If there exists a C^1 proper and positive definite function $V_k^e : \mathbb{R}^{2n+r_f+r_g} \to \mathbb{R}$ such that, along the trajectories $x_k^e(t, t_0, x_0^e)$ of $\Sigma \circ \mathcal{C}(k)$, $\dot{V}_k^e$ is definite negative on $\Omega^e(k) \backslash \mathcal{B}^e(k)$ and 4) and 6) hold, then any trajectory $x_k^e(t, t_0, x_0^e)$ starting from $\Omega^e \subseteq \Omega^e(k)$ stays forever in $\Omega^e(k)$, eventually enters any given δ-neighborhood of $\mathcal{B}^e$ in finite time and remains thereinafter. In our setting, this invariance property corresponds to an event which occurs with probability at least $(1-\alpha)(1-\beta)$. For the above reasons, α and β can be thought of as *risk margins*. In the deterministic case, 6) corresponds to a precise geometric property of the level sets of V_k^e for sufficiently large k: one is that Ω^e is contained in $\Omega^e(k)$ and the other is that $\overline{\mathcal{B}}^e(k)$ is contained in some level set of V_k^e which is, on turn, contained in $\overline{\mathcal{B}}_\delta^e$. □

V. Stochastic Stabilization With Guaranteed Cost for Feedback Linearizable Systems

The conditions of Theorem 4.1 do not provide any constructive procedure to find an admissible parameterization and the functions η_1, η_2 and φ. In Sections V-A and B, we want to outline algorithms for accomplishing this task for the class of nonlinear stochastic systems (1), with $g(x) = 0$ and (A, B, C) invertible with no invariant zeroes, $\Phi(u, x)$ and $H(x)$ normbounded from above by a locally Lipschitz function of x and u. Moreover, we will assume that $\Phi(0, 0) = 0$ and $H(0) = 0$. The case $\Phi(0, 0) \neq 0$ or $H(0) \neq 0$ as well as $g(x) \neq 0$ can be treated in a similar way but with more complicate calculations.

First, we give a *semiglobal in probability backstepping* design procedure for solving the state feedback problem (SF); then a recursive procedure to solve the filtering problem (OI) and (RM). Note that by assuming (A1), we can define as new control input $\widetilde{u} = (1 + (\vartheta_g^*)^T g_R(x))u + (\vartheta_f^*)^T f_R(x)$, and for simplicity, we denote $\widetilde{u}$ directly by u. We also remark that Theorem 4.1 still holds if one replaces (34) with

$$\begin{aligned}\mathcal{L}V_{SF,k} =& \|v - Fx\|_{R_1}^2 - \gamma^2\|\widetilde{\Phi}\|^2 - \|x\|_{Q_{SF}}^2 \\ &- \widetilde{\mathcal{P}}_1 - \mathcal{P}_2 + c_1 + c_2 + \widetilde{\vartheta}^T \Gamma^{-1} \\ &\times \left[\eta_2 \left(\mathrm{Proj}\left(\widehat{\vartheta}, \phi(\sigma, u)\right)\right) - \phi(x, u)\right]\end{aligned} \tag{51}$$

and

$$\begin{aligned}\widetilde{\mathcal{P}}_1(u, x, e, k) = & -\gamma^2(k)\|\Phi(u, x)\|^2 + \|x\|_{E(k)}^2 \\ & + \|v - \Xi(k)x\|_{R_1(k)}^2 + c_1(k)\end{aligned} \tag{52}$$

for some sequence of matrices $\Xi(k)$ and with $F(k) = -R_1^{-1}(k)B^T P_{SF}(k) + \Lambda(k)$. In order to keep the backstepping algorithm as simple as possible, it is convenient to satisfy (51) rather than (34). Moreover, the choice of $\Gamma(k)$ gives an additional flexibility in the optimal control design.

Preliminarly, by [28], there exists a change of coordinates $z = Zx$ such that (1) reads out in the new coordinates

$$\begin{aligned}dz =& \left[\widehat{A}z + \widehat{B}\left(u + \left(\vartheta_f^*\right)^T f_R(Z^{-1}z)\right.\right. \\ &\left.\left. + \Phi_F\left(t, u, Z^{-1}z\right)\right)\right] dt + H(t, Z^{-1}z)dw \\ y =& \widehat{C}z\end{aligned} \tag{53}$$

where

$$\widehat{A} = \begin{pmatrix} a_{11} & 1 & 0 & \dots & 0 & 0 \\ a_{12} & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ a_{n-1,1} & 0 & 0 & \dots & 0 & 1 \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-1} & a_{nn} \end{pmatrix}$$

$$\widehat{B} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad \widehat{C} = (1 \quad 0 \quad 0 \quad \dots \quad 0 \quad 0).$$

To simplify notations, we will omit the hats and denote z by x.

A. Backstepping Design

The main result of this section is the following.

Theorem 5.1: The system (53) is semiglobally stabilizable in probability with guaranteed cost through a state feedback controller. □

As a first step toward the proof of Theorem 5.1, rewrite (53) as

$$d\pi_0 = (A_0\pi_0 + B_0 x_n)dt \tag{54}$$

$$dx_n = \left(v + \widetilde{\vartheta}_f^T f_R(x) + \widetilde{f}_n(x)\right) dt + \widetilde{h}_n(x) dw \quad (55)$$

with $\pi_0 = \mathrm{col}(x_1, \ldots, x_{n-1})$ and $v = u + \widehat{\vartheta}_f^T f_R(x)$.

Definition 5.1: We will say that

$$\begin{aligned} d\pi =& (A(k)\pi + B(k)\Phi(u,\pi) + B(k)v)\, dt + H(\pi) dw \\ d\widetilde{\vartheta} =& v_a \end{aligned} \quad (56)$$

satisfies the property DI if there exist C^0 functions $P_{SF}, Q_{SF}: \mathbb{R}^+ \to \mathcal{SP}^n$, $R_1, \gamma: \mathbb{R}^+ \to \mathbb{R}^+$, $E: \mathbb{R}^+ \to \mathcal{SSP}^n$ and $\Xi: \mathbb{R}^n \times \mathbb{R}$ such that $\Phi \in \mathcal{F}(k)$, $H \in \mathcal{H}(k)$ and for all $(\pi, \widetilde{\vartheta}) \in \Omega(k) = \{\pi \in \mathbb{R}^n, \widetilde{\vartheta} \in \mathbb{R}^{r_f + r_g}: V_{SF,k}(\pi, \widetilde{\vartheta}) \le k\}$ and u one has

$$\|\Phi(u,\pi)\|^2 \le \frac{\|\pi\|^2_{E(k)} + \|v - \Xi(k)\pi\|^2_{R_1(k)}}{\gamma^2(k)} \quad (57)$$

and

$$\begin{aligned} &\mathcal{L}V_{SF,k} + \|\pi\|^2_{E(k)} + \|v - \Xi(k)\pi\|^2_{R_1(k)} - \gamma^2(k)\|\Phi(u,\pi)\|^2 \\ &\quad - \mathbf{Tr}\left\{H^T(\pi) P_{SF} H(\pi)\right\} + \sum_{j=1}^r \pi^T \widehat{H}_j(k) P_{SF} \widehat{H}_j(k)\pi \\ &\quad = -\|\pi\|^2_{Q_{SF}(k)} + \|v - F(k)\pi\|^2_{R_1(k)} - \gamma^2(k) \\ &\qquad \times \left\|\Phi(u,\pi) - \frac{1}{\gamma^2(k)} B^T(k) P_{SF}(k)\pi\right\|^2 \\ &\qquad + \widetilde{\vartheta}^T \Gamma^{-1}\left[v_a - \phi(\pi, u)\right] \end{aligned} \quad (58)$$

where

$$\begin{aligned} V_{SF,k}(\pi, \widetilde{\vartheta}) =& \|\pi\|^2_{P_{SF}(k)} + \frac{1}{2}\widetilde{\vartheta}_f^T \Gamma_f^{-1}(k)\widetilde{\vartheta}_f \\ &+ \frac{1}{2}\widetilde{\vartheta}_g^T \Gamma_g^{-1}(k)\widetilde{\vartheta}_g \end{aligned} \quad (59)$$

$$F(k) = -R_1^{-1}(k) B_2^T(k) P_{SF}(k) + \Xi(k). \quad (60)$$

We have the following result, which roughly states that (55) satisfies the DI property in some new coordinates π.

Lemma 5.1: There exists a C^0 function $\lambda: \mathbb{R}^+ \to (0,1)$ such that (56) satisfies DI, with

$$\begin{aligned} \pi =& \begin{pmatrix} \pi_0 \\ \zeta \end{pmatrix} \\ \zeta =& \lambda(k)(x_n - F_0\pi_0) \\ F_0 =& -R_0^{-1} B_0^T P_0 \\ A(k) =& \begin{pmatrix} A_0 + B_0 F_0 & \frac{B_0}{\lambda(k)} \\ -\lambda(k) F_0 (A_0 + B_0 F_0) & -F_0 B_0 \end{pmatrix} \\ B(k) =& \lambda(k) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ H(\pi) =& \lambda(k) \begin{pmatrix} 0 \\ \widetilde{h}_n\left(t, \frac{\zeta}{\lambda(k)} + F_0\pi_0\right) \end{pmatrix} \\ \Phi(u,\pi) =& \lambda(k) \begin{pmatrix} 0 \\ \widetilde{f}_n\left(\frac{\zeta}{\lambda(k)} + F_0\pi_0\right) \end{pmatrix} \end{aligned} \quad (61)$$

and $R_0 > 0$ and $P_0 \in \mathcal{SP}^{n-1}$ such that

$$A_0 P_0 + P_0 A_0^T - P_0 B_0 R_0^{-1} B_0^T P_0 + I = 0. \quad (62)$$

Proof: For simplicity, throughout the proof and whenever there is no ambiguity, we omit the arguments of the functions involved.

We have by our assumptions on $\widetilde{h}$

$$\begin{aligned} \|\widetilde{h}_n\|^2 \le& 2\left[\left\|\widetilde{h}_n - \widetilde{h}_n\Big|_{\zeta=0}\right\|^2 + \left\|\widetilde{h}_n\Big|_{\zeta=0}\right\|^2\right] \\ \le& \|\pi_0\|^2_{\widetilde{M}(k)} + \widetilde{N}(k)\zeta^2 \end{aligned} \quad (63)$$

for all π and u such that $\pi \in \Omega(k)$, where $\widetilde{h}_n|_{\zeta=0}$ denotes $\widetilde{h}_n$ evaluated for $\zeta = 0$ and for some C^0 functions $\widetilde{M}: \mathbb{R}^+ \to \mathcal{SSP}^{n-1}$ and $\widetilde{N}: \mathbb{R}^+ \to \mathbb{R}^{\ge}$. We remark that $\widetilde{M}$ can be chosen as function of P_0 only.

Pick $\lambda: \mathbb{R}^+ \to (0,1)$ such that

$$\lambda^2(k)\widetilde{M}(k) \le \frac{I}{4}. \quad (64)$$

By the Îto rule

$$d\zeta = \lambda(k)((v - \Xi(k)\pi + \widetilde{\vartheta}^T f_R(x) + \widetilde{f}_n)dt + \widetilde{h} d\widetilde{w}) \quad (65)$$

where

$$\Xi(k) = F_0 \left(A_0 + B_0 F_0 \quad \frac{B_0}{\lambda(k)} \right).$$

Pick $\widetilde{P}(k)$, $\Gamma_f^{-1}(k)$ and $\Gamma_g^{-1}(k)$ such that $\lim_{k\to\infty} \widetilde{P}(k) = 0$, $\lim_{k\to\infty} \Gamma_f^{-1}(k) = 0$ and $\lim_{k\to\infty} \Gamma_g^{-1}(k) = 0$, and define

$$\begin{aligned} \Omega(k) = \Big\{&(v_1, v_2, v_3, v_4) \in \mathbb{R}^{n + r_f + r_g}: \|v_1\|^2_{P_0} + \widetilde{P}(k) v_2^2 \\ &+ \frac{1}{2} v_3^T \Gamma_f^{-1}(k) v_3 + \frac{1}{2} v_4^T \Gamma_g^{-1}(k) v_4 \le k\Big\}. \end{aligned} \quad (66)$$

Find C^0 functions $\widetilde{Q}: \mathbb{R}^+ \to \mathbb{R}^+$, $\gamma: \mathbb{R}^+ \to \mathbb{R}^+$, $\widetilde{E}_1: \mathbb{R}^+ \to \mathbb{R}^{\ge}$ and $\widetilde{E}_2: \mathbb{R}^+ \to \mathcal{SSP}^{n-1}$ such that

- For all π, u such that $\pi \in \Omega(k)$, one has

$$\lambda^2(k)\widetilde{f}_n^2 \le \frac{\zeta^2 \widetilde{E}_1(k) + \|\pi_0\|^2_{\widetilde{E}_2(k)} + |u - \Xi(k)\pi|^2 \widetilde{R}(k)}{\gamma^2(k)} \quad (67)$$

  with any C^0 function $\widetilde{R}: \mathbb{R}^+ \to \mathbb{R}^+$, and

$$\widetilde{E}_2(k) \le \frac{I}{4} \quad (68)$$

- the following equality holds

$$\frac{\widetilde{P}^2(k)}{\gamma^2(k)} - \frac{\lambda^2(k)\widetilde{P}^2(k)}{\widetilde{R}(k)} + \widetilde{E}_1(k) + \lambda^2(k)\widetilde{P}(k)\widetilde{N}(k) = -\widetilde{Q}(k) \quad (69)$$

  with

$$\frac{\widetilde{Q}(k)}{2} \ge \frac{R_0}{\lambda^2(k)}. \quad (70)$$

Let $\widetilde{V}_k(\zeta,\widetilde{\vartheta}) = \widetilde{P}(k)\zeta^2 + (1/2)\widetilde{\vartheta}_f^T\Gamma_f^{-1}(k)\widetilde{\vartheta}_f + (1/2)\widetilde{\vartheta}_g^T\Gamma_g^{-1}(k)\widetilde{\vartheta}_g$, and $\widetilde{F}(k) = -\lambda(k)\widetilde{R}^{-1}(k)\widetilde{P}(k)$. From (69)

$$\begin{aligned}\mathcal{L}\widetilde{V}_k &+ \|\pi_0\|^2_{\widetilde{E}_2(k)} + \zeta^2\widetilde{E}_1(k) + |v - \Xi(k)\pi|^2\widetilde{R}(k)\\ &- \gamma^2(k)\lambda^2(k)\widetilde{f}_n^2 + \lambda^2\widetilde{P}\\ &\times\left(-\left\|\widetilde{h}_n|_{x_n=(\zeta/\lambda(k))+F_0(k)\pi_0}\right\|^2\right.\\ &\quad\left. + \|\pi_0\|^2_{\widetilde{M}}(k) + \zeta^2\widetilde{N}(k)\right)\\ &= -\zeta^2\widetilde{Q}(k) + \|\pi_0\|^2_{\widetilde{E}_2(k)+\lambda^2(k)\widetilde{P}(k)\widetilde{M}(k)}\\ &\quad + \left|v - \widetilde{F}(k)\zeta - \Xi(k)\pi\right|^2_{\widetilde{R}(k)}\\ &\quad - \gamma^2(k)\left|\lambda(k)\widetilde{f}_n - \frac{1}{\gamma^2(k)}\widetilde{P}(k)\zeta\right|^2\\ &\quad + \widetilde{\vartheta}^T|\ \Gamma^{-1}\left[v_a - \phi(\pi,u)\right].\end{aligned}\tag{71}$$

Define

$$P_{SF}(k) = \begin{pmatrix} P_0 & 0\\ 0 & \widetilde{P}(k)\end{pmatrix}.\tag{72}$$

With our definitions

$$\mathbf{Tr}\left\{H^T P_{SF}(k)H\right\} = \lambda^2(k)\widetilde{P}(k)\left\|\widetilde{h}_n\Big|_{x_n=(\zeta/\lambda(k))+F_0(k)\pi_0}\right\|^2\tag{73}$$

From (62), (64), (68), (70), (71) and (73), with $V_0(\pi_0) = \|\pi_0\|^2_{P_0}$ and $V_{SF,k}(\pi) = V_0(\pi_0) + \widetilde{V}_k(\zeta,\widetilde{\vartheta})$ it follows that

$$\begin{aligned}\mathcal{L}V_0 &+ x_n^2R_0 + \mathcal{L}\widetilde{V}_k + \|\pi_0\|^2_{\widetilde{E}_2(k)} + \zeta^2\widetilde{E}_1(k)\\ &+ |v - \Xi(k)\pi|^2\widetilde{R}(k) - \lambda^2(k)\gamma^2(k)\widetilde{f}_n^2 + \lambda^2\widetilde{P}\\ &\times\left(-\left\|\widetilde{h}_n\Big|_{x_n=(\zeta/\lambda(k))+F_0(k)\pi_0}\right\|^2\right.\\ &\quad\left. + \|\pi_0\|^2_{\widetilde{M}(k)} + \zeta^2\widetilde{N}(k)\right)\\ &= \mathcal{L}V_{SF,k} + \|\pi\|^2_{E(k)} + |v - \Xi(k)\pi|^2R_1(k)\\ &\quad - \gamma^2(k)\|\Phi\|^2 - \mathbf{Tr}\left\{H^T(\pi)P_{SF}H(\pi)\right\}\\ &\quad + \sum_{j=1}^{s}\pi^T\widehat{H}_j(k)P_{SF}\widehat{H}_j(k)\pi\\ &\quad + \widetilde{\vartheta}^T\Gamma^{-1}[v_a - \phi(\pi,u)]\\ &= -\|\pi\|^2_{Q_{SF}(k)} + |v - F(k)\pi|^2R_1(k) - \gamma^2(k)\\ &\quad\times\left\|\Phi - \frac{1}{\gamma^2(k)}B^T(k)P_{SF}(k)\pi\right\|^2\\ &\quad + \widetilde{\vartheta}^T\Gamma^{-1}[v_a - \phi(\pi,u)]\end{aligned}\tag{74}$$

for a suitably defined $E(k)$, with $R_1(k) = \widetilde{R}(k)$ and shown in (75) at the bottom of the page. This proves (58).

Finally, from (63), (67), and (73), it follows that $\widetilde{\mathcal{P}}_1 \geq 0$ and $\mathcal{P}_2 \geq 0$ for all $(\pi,\widetilde{\vartheta}) \in \Omega(k)$. This proves the admissibility of Φ and H. □

We are ready to obtain Theorem 5.1. Note that each $\widetilde{P}(k)$, $\Gamma_f^{-1}(k)$ and $\Gamma_g^{-1}(k)$ can be chosen in such a way that

$$\begin{aligned}\lim_{k\to\infty}\frac{\widetilde{P}(k)}{k} &= 0\\ \lim_{k\to\infty}\frac{\Gamma_f^{-1}(k)}{k} &= 0\\ \lim_{k\to\infty}\frac{\Gamma_g^{-1}(k)}{k} &= 0\end{aligned}\tag{76}$$

which implies

$$\lim_{k\to\infty}\frac{V_{SF,k}(\pi,\widetilde{\vartheta})}{k} = 0\tag{77}$$

for each $(\pi,\widetilde{\vartheta}) \in \mathbb{R}^{n+r_f+r_g}$. From (74), (77), and Theorem 4.1 with (34) replaced by (51), since $c_j(k) = 0$ for all j so that the sequence $\{\mathcal{B}^e(k)\}$ can be chosen arbitrarily, we conclude that (56) is *semiglobally stabilizable in probability with arbitrary guaranteed cost* through the state feedback $v = F(k)\pi$ and $v_a = \phi(\pi,u)$.

Since F_0 is independent of k, (76) is sufficient to conclude that also the original system (53) is *semiglobally stabilizable in probability with arbitrary guaranteed cost* through state-feedback, which proves Theorem 5.1.

B. Filter Design

In this section, we want to show that for some admissible parameterization and $(\eta,\varphi) \in \mathcal{D}(k)$ also (OI) and (RM) can be met for (55). First, let us prove (OI). Define

$$\begin{aligned}Q_m(k) &= 2\epsilon^2(k)P_m(k) + m_0(k)I\\ P_m(k) &= \widetilde{P}_m(\epsilon(k))\\ &= \mathrm{diag}\left\{\epsilon^{2(n-1)}(k),\epsilon^{2(n-2)}(k),\ldots,1\right\}P_1(\epsilon(k))\\ &\quad\times\mathrm{diag}\left\{\epsilon^{2(n-1)}(k),\epsilon^{2(n-2)}(k),\ldots,1\right\}\end{aligned}\tag{78}$$

$$\begin{aligned}\widehat{H}_j(k) &= \lambda(k)\begin{pmatrix} P_0^{-1/2}\sqrt{\widetilde{M}(k)\widetilde{P}(k)} & 0\\ 0 & \sqrt{\widetilde{N}(k)}\end{pmatrix}\\ Q_{SF}(k) &= \begin{pmatrix} I - \widetilde{E}_2(k) - \lambda^2(k)\widetilde{M}(k)\widetilde{P}(k) & 0\\ 0 & \widetilde{Q}(k) - \frac{R_0}{\lambda^2(k)}\end{pmatrix}.\end{aligned}\tag{75}$$

where $\epsilon: \mathbb{R}^+ \to \mathbb{R}^+$ is a C^0 function (to be specified later) such that $\lim_{k\to\infty} \epsilon(k) = \infty$.

In what follows, for sake of simplicity, we will omit the argument k when there is no ambiguity. We claim that there exists a C^0 function $\epsilon_1^*: \mathbb{R}^+ \to \mathbb{R}^+$ such that Q_m and P_m, defined in (78), solve (OI) with $R_2 = (\gamma^2/\epsilon^{2(2n-1)})$, for all C^0 functions $\epsilon \geq \epsilon_1^*$. Indeed, substituting in (OI), left and right-multiplying by $\mathrm{diag}\{\epsilon^{-2(n-1)}, \epsilon^{-2(n-2)}, \ldots, 1\}$, and dividing both members by ϵ^2, we find out that solving (OI) amounts to satisfying

$$P_1(\epsilon)(J+I+S_1(\epsilon)) + (J+I+S_1(\epsilon))^T P_1(\epsilon) - C^T C + \frac{1}{\gamma^2\epsilon^2} P_1(\epsilon) B B^T P_1(\epsilon) + S_2(\epsilon) = 0 \quad (79)$$

where $S_1, S_2: \mathbb{R}^+ \to \mathbb{R}^{n\times n}$ are C^0 functions such that $\lim_{\epsilon\to\infty} S_j(\epsilon) = 0$, $j = 1, 2$, $S_2(\epsilon)$ is symmetric and positive semidefinite, and

$$J = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & \frac{1}{\lambda} \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

It can be shown that there exists $\epsilon_1^*: \mathbb{R}^+ \to \mathbb{R}^+$ such that (79), and thus, (OI) holds for each $\epsilon \geq \epsilon_1^*$ and for some $P_1(\epsilon) \in \mathcal{SP}^n$. Indeed, by continuity and since $(J+I+S_1(\epsilon), B)$ is controllable and $(C, J+I+S_1(\epsilon))$ is observable for sufficiently large ϵ, for each sufficiently large ϵ and for some $P_1^0(\epsilon) \in \mathcal{SP}^n$

$$P_1^0(\epsilon)(J+I+S_1(\epsilon)) + (J+I+S_1(\epsilon))^T P_1^0(\epsilon) - C^T C + \frac{1}{\gamma^2\epsilon^2} P_1^0(\epsilon) B B^T P_1^0(\epsilon) = 0. \quad (80)$$

Moreover, if $\epsilon \to \infty$, then $P_1^0(\epsilon) \to P_1^0(\infty)$. Finally, by standard arguments, for ϵ large enough, it can be shown the existence of some $P_1(\epsilon) \in \mathcal{SP}^n$ satisfying (79) and such that $P_1(\epsilon) > P_1^0(\epsilon)$ and $P_1(\epsilon) \to P_1^0(\infty)$ as $\epsilon \to \infty$. This proves (OI).

Next, we define an admissible pair (φ, η). Choose $\varphi(s) = (1/\epsilon)\ln(1+s)$ if $s \geq 0, \eta_2(s) = \mathrm{col}(\eta_{21}(s_1), \ldots, \eta_{2,r_f+r_g}(s_{r_f+r_g}))$

$$\eta_1(s) = \begin{cases} s, & \text{if } |s| \leq M \\ \frac{s}{|s|}M, & \text{otherwise} \end{cases} \quad (81)$$

$$\eta_{2i}(s_i) = \begin{cases} s_i, & \text{if } |s_i| \leq U_i \\ \frac{s_i}{|s_i|}U_i, & \text{otherwise} \end{cases} \quad (82)$$

where

$$M = \max_{x\in\Omega} \left| \frac{Fx - \widehat{\vartheta}_f^T f_R(x)}{\widehat{\vartheta}_g^T g_R(x) + 1} \right| \quad (83)$$

$$U_i = \max_{x\in\Omega;\, \widehat{\vartheta}\in\widehat{\Omega}} \left| [\mathrm{Proj}(\widehat{\vartheta}, \phi(x, u(x)))]_i \right| \quad (84)$$

where $[\cdot]_i$ denotes the ith component, $u(x) = (Fx - \widehat{\vartheta}_f^T f_R(x)/1 + g_R(x))$, and F and Ω are as in (66). Note that $(\partial^2\varphi/\partial s^2) \leq 0 < (\partial\varphi/\partial s) \leq 1$ for all $s \geq 0$ and the functions η_j, $j = 1, 2$, are bounded and linear near the origin. Moreover, the triple $(\varphi, \eta_1, \eta_2)$ is admissible if there exists a C^0 function $\epsilon_2^*: \mathbb{R}^+ \to \mathbb{R}^+$ such that $\mathcal{P}_3 > 0$ for all C^0 functions $\epsilon \geq \epsilon_2^*$ and $(\pi, e) \in (\Omega \times \mathbb{R}^n)\backslash(0,0)$, with P_{nn} being the (n,n) entry (independent of k) of P_m.

In order to prove this, find a covering $\cup_{j=1}^3 \mathcal{M}_j$ of $\{(\pi, e) \in \Omega \times \mathbb{R}^n\}$, with

$$\begin{aligned} \mathcal{M}_1 &= \left\{(\pi, e) \in \Omega \times \mathbb{R}^n : \|\pi - e\| \leq \vartheta_1;\ \|e\| \leq \frac{\vartheta_1}{2}\right\} \\ \vartheta_1 &> 0 \\ \mathcal{M}_2 &= \{(\pi, e) \in \Omega \times \mathbb{R}^n : \|\pi - e\| \geq \vartheta_1;\ \|e\| \leq \vartheta_2\} \\ \vartheta_2 &\leq \frac{\vartheta_1}{2} \\ \mathcal{M}_3 &= \{(\pi, e) \in \Omega \times \mathbb{R}^n : \|e\| \geq \vartheta_2\}. \end{aligned} \quad (85)$$

Pick $\vartheta_1 > 0$ such that $\eta(F(\pi - e)) = F(\pi - e)$ for all $\|\pi - e\| \leq \vartheta_1$.

Note that since $P_1(\epsilon) \to P_1^0(\infty)$ as $\epsilon \to \infty$, the term $-\left(\|\widetilde{h}_n\|^2 P_{nn}/\epsilon\left(1 + \|e\|^2_{\widetilde{P}_m(\epsilon)}\right)\right) + \|\pi\|^2_{Q_{SF}}$ can be rendered positive over the set Ω by choosing ϵ large enough.

First of all, it is easy to see that there exists a C^0 function $\epsilon_3^*: \mathbb{R}^+ \to \mathbb{R}^+$ such that $\mathcal{P}_3 > 0$ holds for all C^0 functions $\epsilon \geq \epsilon_3^*$ and $(\pi, e) \in \mathcal{M}_1$.

Moreover, $\mathcal{P}_3 > 0$ holds for all $(\pi, e) \in \mathcal{M}_2$ for some $\vartheta_2 \leq (\vartheta_1/2)$ and for all $k > 0$. Indeed, $(0, e) \notin \mathcal{M}_2$ since $\vartheta_1 > \vartheta_2$. Thus, we have $\|x\|^2_{Q_{SF}} > 0$ on $\mathcal{M}_2$. It follows that for any such x by continuity, there exists $e_\pi > 0$ such that $\mathcal{P}_3 > 0$ holds for all $\|e\| \leq e_\pi$ and for all C^0 functions $\epsilon \geq \epsilon_3^*$. Since $\mathcal{O} = \{\pi \in \Omega: \|\pi\| \geq (\vartheta_1/2)\}$ is compact and $\vartheta_1 > 0$, one can take $\vartheta_2 = \min\{(\vartheta_1/2), \min_{\pi\in\mathcal{O}} e_\pi\}$.

We are left with proving that there exists $\epsilon_4^*: \mathbb{R}^+ \to \mathbb{R}^+$ such that $\mathcal{P}_3 > 0$ holds for all $(\pi, e) \in \mathcal{M}_3$ and for all C^0 functions $\epsilon \geq \epsilon_4^*$. On the other hand, this readily follows by the boundedness of η and since

$$\lim_{\epsilon\to\infty} \inf_{\|e\|\geq\vartheta_2} \frac{\epsilon\|e\|^2_{\widetilde{P}_m(\epsilon)}}{1 + \|e\|^2_{\widetilde{P}_m(\epsilon)}} = \infty. \quad (86)$$

Pick $\epsilon_2^* \geq \max\{\epsilon_3^*, \epsilon_4^*\}$.

Finally, we will show how to satisfy (RM). Since

$$\lim_{s\to 0} s \ln s^{-r} = 0, \quad \forall\, r \geq 0$$

by (77) and the definition of $P_m(k)$, there exists a C^0 function $\epsilon_5^*: \mathbb{R}^+ \to \mathbb{R}^+$ such that for all C^0 functions $\epsilon \geq \epsilon_5^*$

$$\lim_{k\to\infty} \frac{\|\pi\|^2_{P_{SF}(k)} + \ln\left(1 + \|e\|^2_{\widetilde{P}_m(\epsilon(k))}\right)}{k} = 0 \quad (87)$$

for each $(\pi, e) \in \mathbb{R}^{2n+r_f+r_g}$, which proves (39). On the other hand, by choosing properly the sequence $\{\mathcal{B}^e(k)\}$, we can also satisfy (40). We conclude that (OI)–(RM) hold as long as $\epsilon: \mathbb{R}^+ \to \mathbb{R}^+$ is any C^0 function such that $\epsilon \geq \max\{\epsilon_1^*, \epsilon_2^*, \epsilon_5^*\}$. By using Theorem 5.1, with (34) replaced by (51), and the above results, we obtain the following theorem.

Theorem 5.2: Any system (54) and (55) is semiglobally stabilizable in probability with arbitrary guaranteed cost through output feedback. $\square$

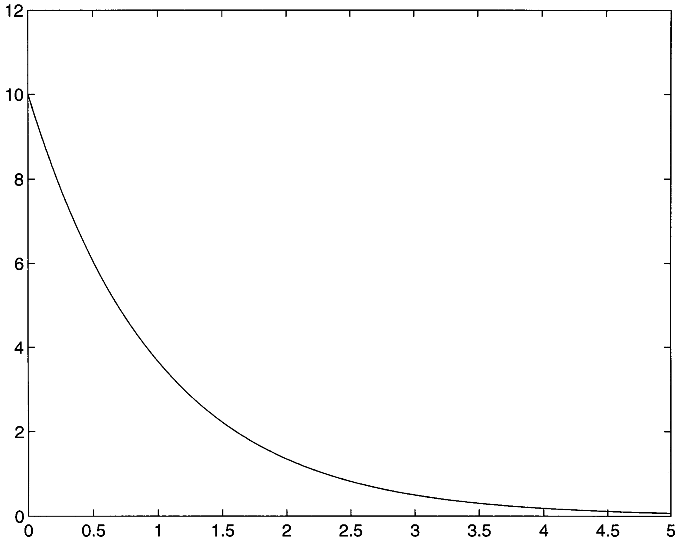


Fig. 1. State x_1 versus time in seconds.

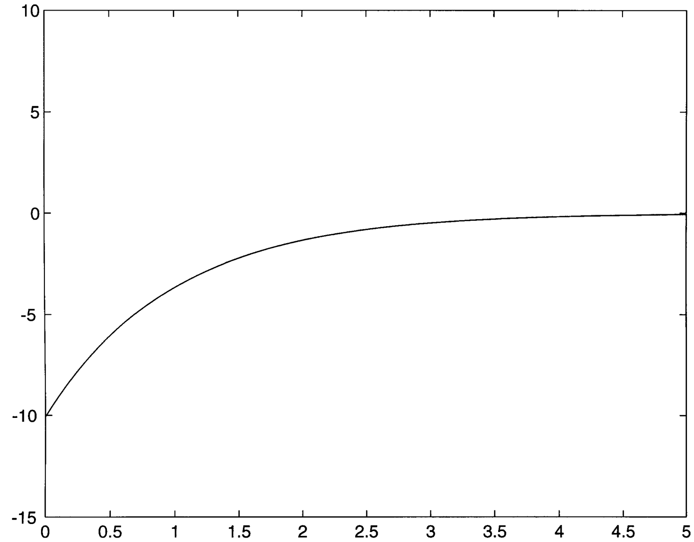


Fig. 2. State x_2 versus time in seconds.

We conclude with an example. Let us consider

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= \left(u + x_2^3\right) dt \\ y &= x_1. \end{aligned} \tag{88}$$

It is known from [13] that (88) is not globally stabilizable by any continuous dynamic output feedback controller. However, we select a guaranteed region of attraction $\Omega = \{(x_1, x_2): |x_j| \leq 5\}$, and we approximate on the compact interval $[-10, 10]$ the function $f(x) = x_2^3$ through a standard two-layer neural network

$$f_N(x) = \sum_{j=1}^{10} \theta_j \sigma\left(x_2 w^{(j)} + w_0^{(j)}\right) + \theta_{11} = \widehat{\vartheta}_{f0}^T f_R(x) \tag{89}$$

where the first layer consists of ten standard neurons (tanh) with input weights

$$\begin{aligned} \theta_1 &= -67.6129, \quad \theta_2 = -6.2793, \quad \theta_3 = -36.1681 \\ \theta_4 &= -145.2503, \quad \theta_5 = -39.7344 \\ \theta_6 &= -1.5481, \quad \theta_7 = -89.1238, \quad \theta_8 = 10.9402 \\ \theta_9 &= 71.7582, \quad \theta_{10} = 64.1357 \end{aligned} \tag{90}$$

and $\theta_{11} = 94.2859$, output weights

$$\begin{aligned} w^{(1)} &= -5.5332, \quad w^{(2)} = 9.5610, \quad w^{(3)} = -7.0830 \\ w^{(4)} &= -3.7208, \quad w^{(5)} = 6.5889 \\ w^{(6)} &= -0.4971, \quad w^{(7)} = -4.5200, \quad w^{(8)} = -1.7649 \\ w^{(9)} &= 5.7562, \quad w^{(10)} = 6.0560 \end{aligned} \tag{91}$$

and biases

$$\begin{aligned} w_0^{(1)} &= -29.1293, \quad w_0^{(2)} = -24.0213, \quad w_0^{(3)} = 21.2418 \\ w_0^{(4)} &= 26.1952, \quad w_0^{(5)} = 17.7946 \\ w_0^{(6)} &= 13.1256, \quad w_0^{(7)} = 22.7060, \quad w_0^{(8)} = 8.2112 \\ w_0^{(9)} &= 4.2590, \quad w_0^{(10)} = 6.8861. \end{aligned} \tag{92}$$

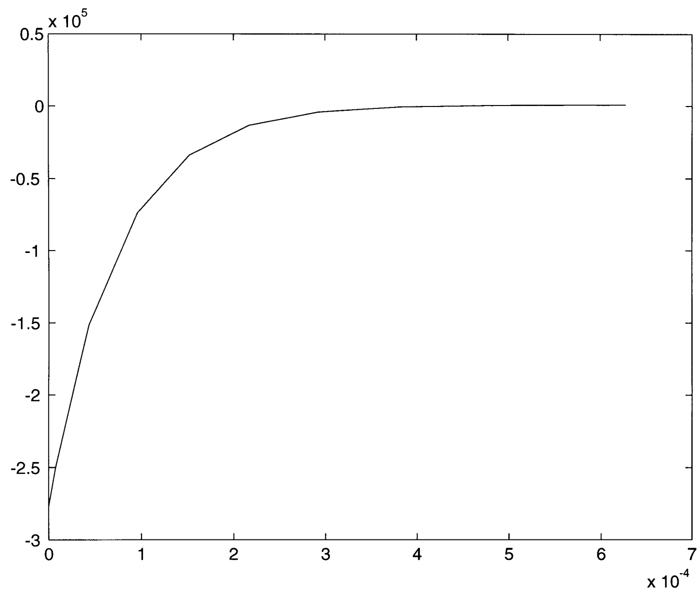


Fig. 3. Control u versus time in seconds.

Proceeding as in [1], one can show that (88) is stabilizable through a dynamic output feedback controller, based on control saturation and high gain observer, and with region of attraction containing Ω. Assume that a Wiener process affects the second equation of (88) as follows:

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= \left(u + x_2^3\right) dt + x_1^2 dw \\ y &= x_1. \end{aligned} \tag{93}$$

Using (2), we rewrite (93) as

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= \left(u + \vartheta_f^* f_R(x) + \Phi_F(x)\right) dt + x_1^2 dw \\ y &= x_1. \end{aligned} \tag{94}$$

Applying Theorem 5.2, we conclude that (93) is stabilizable in probability with region of attraction containing Ω. In Figs. 1–3, the evolution of the states x_1 and x_2 and the control u of the closed-loop system versus t are plotted, starting from initial conditions $x_1(0) = x_2(0) = 10$, $\sigma_1(0) = \sigma_2(0) = 0$ and with a Gaussian random noise.

APPENDIX

Proof of Lemma 4.1

Throughout the proof, $\tau_{\mathcal{S}}(t) = \min\{t, \tau_{\mathcal{S}}\}$, where $\tau_{\mathcal{S}}$ is the Markov time (relatively to the σ-algebra generated by $\{x_k^e(s), s \leq t\}$) defined as the first time at which the trajectory of (1) reaches the boundary of $\mathcal{S}$. By 4), we can assume $k \geq k^*$, with k^* such that $\Omega^e(k) \supseteq \Omega^e$ for all $k \geq k^*$, and we fix any $\Phi \in \mathcal{D}_{\Xi}$.

We have to show only 2) and 3) of Definition 3.1. As a consequence of the Dynkin's formula (with $\mathcal{Z} = \Omega^e(k) \backslash \mathcal{B}^e(k)$ and $T = \infty$), since $\mathcal{L}V_k^e$ is negative definite on $\Omega^e(k) \backslash \mathcal{B}^e(k)$

$$\mathbf{E}\{V_k^e(x_k^e(\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}(t), t_0, x_0^e))\} \leq V_k^e(x_0^e) \tag{95}$$

for all $x_0^e \in \partial\mathcal{B}^e(k)$. By 4), for each $\delta > 0$, there exists k° such that $\mathcal{B}_\delta^e \supset \mathcal{B}^e(k)$. By the Chebyshev inequality (with $\mathcal{S} = \overline{\mathcal{B}}_\delta^e$, $\eta = x_k^e(\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}(r), t_0, x_0^e)$, $r \geq t_0$ ranging over the rationals, and $V = V_k^e$), (95) and 5), we have

$$\mathbf{P}\left\{x_k^e(r, t_0, x_0^e) \notin \overline{\mathcal{B}}_\delta^e \text{ for some rational } r \geq t_0\right\} \leq \frac{V_k^e(x_0^e)}{\inf_{z \in \mathbb{R}^{2n+r_f+r_g}\backslash\overline{\mathcal{B}}_\delta^e} V_k^e(z)} \tag{96}$$

for all $k \geq k^\circ$ and $x_0^e \in \partial\mathcal{B}^e(k)$. Since, by $\mathcal{B}^e(k) \subset \mathcal{B}_\delta^e$ for all $k \geq k^\circ$

$$\mathbf{P}\{x_k^e(r, t_0, x_0^e) \notin \overline{\mathcal{B}}_\delta^e \text{ for some rational } r \in [t_0, \tau_{\mathcal{B}^e(k)})\} = 0$$

for all $k \geq k^\circ$ and $x_0^e \in \mathcal{B}^e(k)$, by 6) and continuity of $x_k^e(t, \cdot, \cdot)$ we conclude 2) of Definition 3.1.

To prove 3), we implicitly assume that $x_0^e \in \Omega^e \backslash \mathcal{B}^e(k)$. First of all, we prove

$$\mathbf{P}\{\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)} < \infty\} = 1 \tag{97}$$

in the case that $x_k^e(t, t_0, x_0^e)$ is regular, since otherwise, it is trivially true. Since $Q_k^e(x^e)$ is continuous on its domain, we have $\mathcal{L}V_k^e \leq -\nu(k) < 0$ for all $x^e \in \Omega^e(k)\backslash\mathcal{B}^e(k)$ and for some $\nu(k) > 0$. Directly from the Dynkin's formula (with $\mathcal{Z} = \Omega^e(k)\backslash\mathcal{B}^e(k)$ and $T = \infty$), we obtain

$$\nu(k)\mathbf{E}\{\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}(t) - t_0\} \leq V_k^e(x_0^e). \tag{98}$$

Thus, by the Chebyshev inequality (with $\mathcal{S} = \mathbb{R}^{2n}\backslash(\Omega^e(k)\backslash\mathcal{B}^e(k))$, $\eta = x_k^e(\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}(r), t_0, x_0^e)$, with $r \geq t_0$ ranging over the rationals, and $V(\eta) = \tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)} - t_0$)

$$\mathbf{P}\{\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)} \geq r\} \leq \frac{V_k^e(x_0^e)}{\nu(k)(r - t_0)}. \tag{99}$$

Since $V_k^e(x_0^e)/\nu(k)(r - t_0) \to 0$ as $r \to \infty$, from (99) and (sequential) continuity of $\mathbf{Pr}\{\cdot\}$, we obtain (97).

Next, we show that

$$\liminf_{k\to\infty} \inf_{x_0^e \in \Omega^e\backslash\mathcal{B}^e(k)} \mathbf{P}\{\tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)} < \tau_{\Omega^e(k)}\} \geq 1 - \alpha. \tag{100}$$

From the Dinkyn's formula (with $\mathcal{Z}$ and T as above) and since $\mathcal{L}V_k^e$ is negative definite on $\Omega^e(k)\backslash\mathcal{B}^e(k)$, it follows that

$$\mathbf{E}\{V_k^e(x_k^e(\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}(t), t_0, x_0^e))\} \leq V_k^e(x_0^e). \tag{101}$$

By (101)

$$\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}(t) = \tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}, \quad \text{a.s.} \tag{102}$$

From (13) and (103)

$$\mathbf{P}\left\{\tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)} > \tau_{\Omega^e(k)}\right\} \leq \mathbf{P}\left\{\frac{V_k^e(x_k^e(\tau_{\Omega^e(k)\backslash\mathcal{B}^e(k)}, t_0, x_0^e))}{k} \geq 1\right\} \leq \frac{V_k^e(x_0^e)}{k}. \tag{103}$$

By (103) and since $\limsup_{k\to\infty} \sup_{x_0^e \in \Omega^e\backslash\mathcal{B}^e(k)} V_k^e(x_0^e)/k \leq \alpha$ by 6), we get

$$\limsup_{k\to\infty} \sup_{x_0^e \in \Omega^e\backslash\mathcal{B}^e(k)} \mathbf{P}\{\tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)} > \tau_{\Omega^e(k)}\} \leq \alpha. \tag{104}$$

By continuity of $x_k^e(t, \cdot, \cdot)$ and since $\mathcal{B}^e(k) \subset \Omega^e(k)$, $\mathbf{P}\{\tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)} = \tau_{\Omega^e(k)}\} = 0$, which along with (38) implies (39).

By (96) and (vi) and using the continuity of $x_k^e(t, \cdot, \cdot)$, for each $\epsilon, \delta > 0$, there exists k° such that $\mathcal{B}_\delta^e \supset \mathcal{B}^e(k)$ and

$$\mathbf{P}\left\{x_k^e(t, s, z) \notin \overline{\mathcal{B}}_\delta^e \text{ for some } t \geq s\right\} < \beta + \epsilon \tag{105}$$

for all $k \geq k^\circ$ and $z \in \partial\mathcal{B}^e(k)$.

Finally, let $\mathcal{F}(\Omega^e(k), \mathcal{B}^e(k))$ denote the σ-algebra generated by the events $\left\{x_k^e(r, t_0, x_0^e) \in \Omega^e(k), r \leq \tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)}\right\}$. Since $\mathcal{F}(\Omega^e(k), \mathcal{B}^e(k))$ is a sub–σ-algebra of the one generated by $\left\{x_k^e(r, t_0, x_0^e), r \leq \tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)}\right\}$ and by the strong Markov property, for each $\epsilon, \delta > 0$ and for all $k \geq k^\circ$

$$\begin{aligned}
&\mathbf{P}\Big\{x_k^e(t + \tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)}, t_0, x_0^e) \notin {}_\delta^e \text{ for some } t \\
&\quad \geq 0 |\, \mathcal{F}(\Omega^e(k), \mathcal{B}^e(k))\} \\
&\quad = \int_{t_0}^{\infty} \int_z \in \partial\mathcal{B}^e(k) \\
&\qquad \times \Big(\mathbf{P}\{\tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)} \in ds x_k^e \\
&\qquad\quad \times (\tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)}, t_0, x_0^e) \\
&\qquad\quad \times \in dz \Big| \mathcal{F}(\Omega^e(k), \mathcal{B}^e(k))\} \\
&\qquad\quad \cdot \mathbf{P}\{x_k^e(t, s, z) \notin \overline{\mathcal{B}}_{\triangleq lta}^e \text{ for some } t \\
&\qquad\quad \in [s, \infty) \Big| \mathcal{F}(\Omega^e(k), \mathcal{B}^e(k))\}\Big) < \beta + \epsilon
\end{aligned} \tag{106}$$

which implies for each $\delta > 0$

$$\liminf_{k\to\infty} \inf_{x_0^e \in \Omega^e\backslash\mathcal{B}^e(k)} \mathbf{P}\left\{x_k^e(t + \tau_{\mathbb{R}^{2n+r_f+r_g}\backslash\mathcal{B}^e(k)}, t_0, x_0^e) \in \overline{\mathcal{B}}_\delta^e \; \forall\, t \geq 0 |\, \mathcal{F}(\Omega^e(k), \mathcal{B}^e(k))\right\} \geq 1 - \beta \tag{107}$$

REFERENCES

[1] S. Seshagiri and H. K. Khalil, "Output feedback control of nonlinear systems using RBF neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 69–79, Jan. 2000.

[2] S. S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York: MacMillan, 1994.

[3] S. Battilotti, "A unifying framework for the semiglobal stabilization of nonlinear uncertain systems via measurement feedback," *IEEE Trans. Automat. Contr.*, vol. 46, pp. 3–16, Jan. 2001.

[4] S. Battilotti and A. De Santis, "A new notion of stabilization in probability for nonlinear stochastic systems," in *Mathematical Theory on Networks and Systems*. Perpignan, France: Perpignan, 2000.

[5] ——, "Stabilization in probability of nonlinear stochastic systems with guaranteed cost," SIAM J. Contr. Optimization, 2002, to be published.

[6] A. Bacciotti, "Further remarks on potentially global stabilizability," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 637–638, June 1989.

[7] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 831–847, May 1989.

[8] H. J. Sussmann and P. V. Kokotovic, "The peaking phenomenon and the global stabilization of nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 424–439, Apr. 1991.

[9] P. Kokotovic and R. Marino, "On vanishing stability regions in nonlinear systems with high gain feedback," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 967–970, July 1986.

[10] F. Esfandiari and H. K. Khalil, "Output feedback stabilization of fully linearizable systems," *Int. J. Contr.*, vol. 56, pp. 1007–1037, 1992.

[11] H. K. Khalil and F. Esfandiari, "Semiglobal stabilization of a class of nonlinear system using output feedback," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 1412–1415, Sept. 1993.

[12] A. Saberi, Z. Lin, and A. Teel, "Control of linear systems with saturating actuators," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 368–378, Mar. 1996.

[13] F. Mazenc, L. Praly, and W. P. Dayawansa, "Global stabilization by output feedback: Examples and counterexamples," *IEEE Trans. Automat. Contr.*, vol. 22, 1994.

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

[15] R. Z. Khas'minskii, *Stochastic Stability of Differential Equations*. Rockville, MD: Sjithoff and Noordhoff, 1980.

[16] D. Hinrichsen and A. J. Pritchard, "Stochastic $\mathcal{H}_\infty$," *SIAM J. Contr. Optim.*, vol. 36, pp. 1504–1538, 1998.

[17] Z. Pan and T. Basar, "Backstepping controller design for nonlinear stochastic systems under a risk sensitive cost criterion," *SIAM J. Contr. Optim.*, vol. 37, pp. 957–995, 1999.

[18] H. Deng and M. Krstic, "Output feedback stochastic nonlinear stabilization," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 328–333, Feb. 1999.

[19] ——, "Stochastic nonlinear stabilization—Part I: A backstepping design," *Syst. Contr. Lett.*, vol. 32, pp. 143–150, 1997.

[20] ——, "Stochastic nonlinear stabilization—Part II: Inverse optimality," *Syst. Contr. Lett.*, vol. 32, pp. 151–159, 1997.

[21] M. Krstic and H. Deng, *Stabilization of Nonlinear Uncertain Systems*. London, U.K.: Springer-Verlag, 1998.

[22] P. Florchinger, "Lyapunov techniques for stochastic stability," *SIAM J. Contr. Optim.*, vol. 33, pp. 1151–1169, 1995.

[23] ——, "A universal formula for the stabilization of control stochastic differential equations," *Stoch. Anal. Applicat.*, vol. 11, pp. 155–162, 1993.

[24] M. Zakai, "On the optimal filtering of diffusion processes," *Z. Wahrscheinlich. Verw. Geb.*, vol. 11, pp. 230–243, 1969.

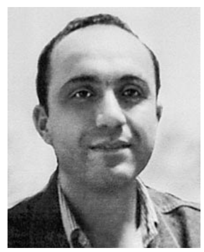
[25] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*. New York: Springer-Verlag, 1984.

[26] I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*. New York: Springer-Verlag, 1972.

[27] B. Øksendal, *Stochastic Differential Equations: An Introduction With Applications*. New York: Springer Verlag, 1985.

[28] A. Saberi, Z. Lin, and A. Teel, "Control of linear systems with saturating actuators," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 368–378, Mar. 1996.

[29] R. M. Sanner and J.-J. Slotine, "Gaussian networks for direct adaptive control," *IEEE Trans. Neural Networks*, vol. 3, pp. 837–863, June 1992.



Stefano Battilotti (S'89–M'95–SM'00) was born in 1962 in Rome, Italy. He received the Laurea degree in 1987 and the Ph.D. degree in 1992 from the University of Rome "La Sapienza."

From 1994 to 1998, he was Assistant Professor in the Department of Informatica e Sistemistica "Antonio Ruberti," and since 1998, he has been an Associate Professor of Automatic Control with the same department. His research interests are focused on the analysis and control of nonlinear deterministic and stochastic systems. He has authored *Noninteracting control with stability for nonlinear systems* (New York: Springer-Verlag, 1994).

Alberto De Santis was born in Rome, Italy, in 1958. He received the Laurea degree in electrical engineering in 1984 and the postgraduate Diploma degree in system and control theory on 1986, both from the University of Rome "La Sapienza."

From 1987 to 1992, he was with the IASI-CNR of Rome. He then joined the "Antonio Ruberti" Department of Informatica e Sistemistica, Faculty of Engineering, University of Rome "La Sapienza," where he is currently Associate Professor. In 1991, he spent one year at the Henry Samueli School of Engineering and Applied Science, University of California, Los Angeles. His research interests are in system identification and signal processing, control of distributed parameter mechanical systems, and nonlinear stochastic control.