# Neural Network-Based Practical/Ideal Integral Sliding Mode Control

Nikolas Sacchi, *Student Member, IEEE*, Gian Paolo Incremona, *Member, IEEE*, and Antonella Ferrara, *Fellow, IEEE*

***Abstract*—This letter deals with the design of a novel neural network based integral sliding mode (NN-ISM) control for nonlinear systems with uncertain drift term and control effectiveness matrix. Specifically, this letter extends the classical integral sliding mode control law to the case of unknown nominal model. The latter is indeed reconstructed by two deep neural networks capable of approximating the unknown terms, which are instrumental to design the so-called integral sliding manifold. In this letter, the ultimate boundedness of the system state is formally proved by using Lyapunov stability arguments, thus providing the conditions to enforce practical integral sliding modes. The possible generation of ideal integral sliding modes is also discussed. Moreover, the effectiveness of the proposed NN-ISM control law is assessed in simulation relying on the classical Duffing oscillator.**

***Index Terms*—Sliding mode control, neural networks, uncertain systems.**

## I. Introduction

SLIDING Mode Control (SMC) has a wide popularity due to its ability to make the controlled system insensitive to matched uncertainty terms whenever the system state lies on a predefined sliding manifold [1]. This property is enabled by the discontinuous nature of the control law, which, on one hand, enables the finite time convergence of the so-called sliding variable to the corresponding sliding manifold, while, on the other hand, it may cause the notorious chattering phenomenon [2]. The amplitude of chattering is considerably affected by the size of the control gain, which, on the other hand, must to be selected in order to dominate the uncertain terms. As such, it is beneficial when non-excessively conservative bounds on the uncertainties are available to the control designer.

Manuscript received March 4, 2022; revised May 15, 2022; accepted June 8, 2022. Date of publication June 14, 2022; date of current version June 24, 2022. This work was supported in part by the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN) under Grant 2017YKXYXJ. Recommended by Senior Editor F. Dabbene. *(Corresponding author: Nikolas Sacchi.)*

Nikolas Sacchi and Antonella Ferrara are with the Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, 27100 Pavia, Italy (e-mail: nikolas.sacchi01@universitadipavia.it; antonella.ferrara@unipv.it).

Gian Paolo Incremona is with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail: gianpaolo.incremona@polimi.it).

Digital Object Identifier 10.1109/LCSYS.2022.3182814

Different solutions have been proposed to alleviate the chattering effect while maintaining the robustness and finite time convergence properties of SMC. Among these solutions, there are for instance higher order sliding mode controllers [3]–[5], adaptive control approaches [6]–[8], or internal model principle based strategies [9]. Instead, as for the improvement of the robustness of SMC, a new paradigm was introduced in [10], based on the concept of Integral Sliding Mode (ISM), which enables to eliminate the so-called reaching phase, during which the controlled system is still sensitive to the uncertainties, thus enabling a sliding mode, with the associated robustness property, from the initial time instant. Extensions of the basic ISM control concept have been proposed in the literature. For instance, the original setting with only matched disturbances has been extended in [11], where also unmatched uncertain terms are taken into account. Moreover, several works have assessed the validity of the ISM control approach for different applications (see e.g., [12]–[14]).

Given a nonlinear system affine in the control input, to design an ISM control the nominal drift term and the matrix multiplying the control input, i.e., the control effectiveness matrix, must be known. This knowledge is not available in many practical implementations, where only conservative bounds can be retrieved from the physics of the process to control or via experimental tests.

In the last twenty years, the so-called deep neural networks (DNNs) and their universal approximation property [15] have become a viable way to provide estimations of the model uncertain terms, even in model based control schemes. For instance, in [15] and [16], neural networks with weight adaptation laws which rely on Lyapunov's stability theory have been introduced. The same concepts have been used in [17]–[19] to directly generate the control laws. Moreover, Lyapunov-based adaptation laws have been used to train DNNs which estimate the unknown dynamics in the case of continuous-time systems in [20], and in that of discrete-time system in [21]. The combination of neural networks with SMC has been also investigated, see e.g., [22]–[25].

In this letter, we propose a novel NN-ISM control approach. Specifically, we extend the algorithm in [10] to the case of unknown nominal dynamics, exploiting the use of two DNNs to design the so-called integral sliding manifold. This allows us to produce a sliding mode control law with a smaller amplitude than the one that would be obtained, to get equal performance, via a conventional SMC approach. This because the NN-ISM control law does not rely on conservative bounds on the uncertain terms, but on the rather accurate approximation of these terms provided by the DNNs included in the control scheme. Moreover, differently from other published solutions which use SMC and neural networks, no knowledge of the bounds of the function reconstruction errors is necessary in our proposal. In the considered setting, the ultimate boundedness of the state to a set depending on the design control parameters (i.e., the enforcement of a practical sliding mode) is formally proved in this letter. The generation of ideal sliding modes is instead proved in some specific conditions.

To the best of our knowledge, this is the first time that ISM and neural networks are combined giving rise to an original solution, which is capable of solving a complex control problem with a minimum amount of available information, and a limited amplitude of the sliding mode control input, thus alleviating all the problems which, in practical implementations, are normally associated with a high control gain.

*Notation:* The used variables and operators are mostly standard. Let $x$ be a vector, then $x^\top$ refers to its transpose. Given a real matrix $A \in \mathbb{R}^{n\times n}$, then $\mathrm{tr}(A)$ is its trace. Given two real matrices $A, B \in \mathbb{R}^{n\times n}$, then $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$, while given $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{m\times n}$, then $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. Given two real column vectors $a, b \in \mathbb{R}^n$, the trace of the outer product is equivalent to the inner product, i.e., $\mathrm{tr}(ba^\top) = a^\top b$.

## II. Preliminaries and Problem Statement

The aim of this section is to describe the dynamical system that will be considered in this letter and recall the main features of the ISM control in [10].

Consider the following nonlinear system

$$\dot{x} = f(x(t)) + B(x(t))u(t) + h(x(t), t), \quad x(0) = x_0, \tag{1}$$

where $x \in \mathbb{R}^n$ is the measurable state vector, $x_0 \in \mathbb{R}^n$ is the initial condition, $u \in \mathbb{R}$ is the control variable, $f(x(t)) : \mathbb{R}^n \to \mathbb{R}^n$ represents the drift dynamics, $B(x(t)) : \mathbb{R}^n \to \mathbb{R}^n$ is the control effectiveness term, and $h(x(t), t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is the system perturbation. Moreover, the following assumption, classical in the sliding mode theory [1], [26], holds.

$\mathcal{A}_1$: There exists a known constant $\overline{h} \in \mathbb{R}_{>0}$ such that the perturbation function $h(x(t), t)$ is bounded as

$$\sup_{t\in\mathbb{R}_{\geq 0}} \|h\| \leq \overline{h}. \tag{2}$$

In order to compensate the effect of the external disturbance $h(x(t), t)$ from $t \geq 0$, an ISM controller can be designed. According to [10], the control law is defined as

$$u = u_0 + u_1, \tag{3}$$

where $u_0$ is a control law making the origin be an asymptotically stable equilibrium point for the nominal dynamics, given by (1) when $h = 0$, while $u_1$ is a discontinuous control aimed at rejecting the uncertainties. In particular, it is defined as

$$u_1 = -\rho\,\mathrm{sign}(s), \tag{4}$$

where $\rho \in \mathbb{R}_{>0}$ is the control gain, and $s(x(t)) : \mathbb{R}^n \to \mathbb{R}$ is the so-called *integral sliding variable* given by

$$s(x(t)) = s_0(x(t)) + z(x(t)), \quad s(x_0) = 0. \tag{5}$$

Specifically, in (5) $z$ is the so-called *transient function* selected such that

$$\dot{z} = -\frac{\partial s_0}{\partial x}(f(x) + B(x)u_0), \quad z(0) = -s_0(0). \tag{6}$$

Then, by properly selecting the stabilizing control law $u_0$ and the discontinuous control gain $\rho$, the sliding mode condition and the robustness of the controlled system with respect to the matched perturbation $h$ can be proved (see [10] for further details). Note that, for the sake of simplicity, in the following the dependence of the sliding variable on $x(t)$ is omitted when obvious, leaving only the time dependence.

## III. DNN Based Function Approximation

While in [10] the nominal dynamics of the system is assumed to be fully known, in this letter the functions $f(x)$ and $B(x)$ are assumed to be unknown. Hence, in this section we introduce two DNNs to estimate such terms, relying on the so-called universal approximation property. More precisely, consider the following theorem.

*Theorem 1 (Universal Approximation [15]):* Let $\Omega \subseteq \mathbb{R}^p$ be a compact set and consider a smooth function $g(\alpha) : \mathbb{R}^p \to \mathbb{R}^q$. Then, there exists a two-layer neural network with $L \in \mathbb{N}_{>0}$ neurons in the hidden layer characterized by ideal weights $W \in \mathbb{R}^{L\times q}$ and $\Phi \in \mathbb{R}^{p\times L}$, ideal activation function vector $\sigma(\cdot) : \mathbb{R}^L \to \mathbb{R}^L$ and a constant $\overline{\varepsilon_g} \in \mathbb{R}_{>0}$ such that

$$g(\alpha) = W^\top \sigma(\Phi^\top \alpha) + \varepsilon_g(\alpha), \tag{7}$$

with $\varepsilon_g(\alpha) : \mathbb{R}^q \to \mathbb{R}^q$ being the so-called *function reconstruction error* so that $||\varepsilon_g|| < \overline{\varepsilon_g}$ for all $\alpha \in \Omega$.

The above theorem can be exploited in order to approximate the unknown drift dynamics $f(x)$ and the control effectiveness term $B(x)$ in (1), as

$$f(x) = W_f^\top \sigma_f(\Phi_f^\top x) + \varepsilon_f(x), \tag{8}$$

$$B(x) = W_B^\top \sigma_B(\Phi_B^\top x) + \varepsilon_B(x), \tag{9}$$

for $x \in \Omega \subseteq \mathbb{R}^n$. In particular, $W_f \in \mathbb{R}^{L_f\times n}$, $W_B \in \mathbb{R}^{L_B\times n}$ and $\Phi_f \in \mathbb{R}^{n\times L_f}$, $\Phi_B \in \mathbb{R}^{n\times L_B}$ are the ideal NN weights, $\sigma_f(\cdot) : \mathbb{R}^{L_f} \to \mathbb{R}^{L_f}$ and $\sigma_B(\cdot) : \mathbb{R}^{L_B} \to \mathbb{R}^{L_B}$ are the ideal bounded activation functions vectors, while $\varepsilon_f(x) : \mathbb{R}^n \to \mathbb{R}^n$ and $\varepsilon_B(x) : \mathbb{R}^n \to \mathbb{R}^n$ are the functions reconstruction errors. Since the ideal DNNs are not known, an approximation of them can be used. In particular, the unknown drift dynamics and the unknown control effectiveness term can be estimated by

$$\widehat{f}(x) = \widehat{W}_f^\top \widehat{\sigma}_f(\widehat{\Phi}_f^\top x), \tag{10}$$

$$\widehat{B}(x) = \widehat{W}_B^\top \widehat{\sigma}_B(\widehat{\Phi}_B^\top x), \tag{11}$$

where $\widehat{\sigma}_f(\cdot)$, $\widehat{\sigma}_B(\cdot)$ are the activation functions vectors selected by the designer, which may differ from the ideal ones $\sigma_f(\cdot)$ and $\sigma_B(\cdot)$. As a consequence, the weight estimation errors are expressed as

$$\widetilde{W}_f(t) = W_f - \widehat{W}_f(t), \tag{12a}$$

$$\widetilde{W}_B(t) = W_B - \widehat{W}_B(t), \tag{12b}$$
$$\widetilde{\Phi}_f(t) = \Phi_f - \widehat{\Phi}_f(t), \tag{12c}$$
$$\widetilde{\Phi}_B(t) = \Phi_B - \widehat{\Phi}_B(t). \tag{12d}$$

Now, the following assumption about the bounds of the ideal output layer weights and about the activation functions vectors needs to be introduced.

$\mathcal{A}_2$: By virtue of the universal approximation property, there exist known constants $\overline{W}_F$, $\overline{W}_B$, $\overline{\sigma}_f$, $\overline{\sigma}_B$, $\overline{\widehat{\sigma}}_f$, $\overline{\widehat{\sigma}}_B \in \mathbb{R}_{>0}$ such that the unknown ideal output layer weights $W_f$, $W_B$, the unknown ideal activation functions vectors $\sigma_f(\cdot)$, $\sigma_B(\cdot)$ and the designer-selected activation functions vectors $\widehat{\sigma}_f(\cdot)$, $\widehat{\sigma}_B(\cdot)$ are bounded as

$$\sup_{x(t)\in\Omega} \|W_f\| \leq \overline{W}_f, \quad \sup_{x(t)\in\Omega} \|W_B\| \leq \overline{W}_B,$$
$$\sup_{x(t)\in\Omega} \|\sigma_f\| \leq \overline{\sigma}_f, \quad \sup_{x(t)\in\Omega} \|\sigma_B\| \leq \overline{\sigma}_B,$$
$$\sup_{x(t)\in\Omega} \|\widehat{\sigma}_f\| \leq \overline{\widehat{\sigma}}_f, \quad \sup_{x(t)\in\Omega} \|\widehat{\sigma}_B\| \leq \overline{\widehat{\sigma}}_B.$$

As for the functions reconstruction errors, the following assumption instead holds.

$\mathcal{A}_3$: There exist unknown constants $\overline{\varepsilon}_f, \overline{\varepsilon}_B \in \mathbb{R}_{>0}$ such that the function reconstruction errors $\varepsilon_f$ and $\varepsilon_B$ are bounded, i.e.,

$$\sup_{x(t)\in\Omega} \|\varepsilon_f\| \leq \overline{\varepsilon}_f, \quad \sup_{x(t)\in\Omega} \|\varepsilon_B\| \leq \overline{\varepsilon}_B.$$

## IV. NN-ISM Controller Design

We are now in a position to introduce the proposed NN-ISM control algorithm.

First, select the component $s_0$ of the sliding variable in (5) as the linear combination of the system states, i.e., $s_0 = \lambda^\top x$, with $\lambda \in \mathbb{R}^n$ being a design vector such that $\lambda^\top \mathbb{1} > 0$, and $\mathbb{1} \in \mathbb{R}^n$ being a column vector of all ones. Then, by using the approximations defined in (10) and (11), it is possible to approximate the dynamics (6) as

$$\dot{z} = -\lambda^\top \big(\widehat{f}(x) + \widehat{B}(x)u_0\big), \quad z(0) = -s_0(0),$$

which, exploiting (10) and (11), can be rewritten as

$$\dot{z} = -\lambda^\top \Big(\widehat{W}_f^\top \widehat{\sigma}_f(\widehat{\Phi}_f^\top x) + \widehat{W}_B^\top \widehat{\sigma}_B(\widehat{\Phi}_B^\top x)u_0\Big), \tag{13}$$

with $\widehat{\sigma}_B(\cdot)$ and $\widehat{\sigma}_f(\cdot)$ being selected as vectors of *logistic sigmoid* functions. For the sake of simplicity, in the rest of this letter the quantities $\sigma_f(\widehat{\Phi}_f^\top x)$, $\sigma_B(\widehat{\Phi}_B^\top x)$, $\widehat{\sigma}_f(\widehat{\Phi}_f^\top x)$ and $\widehat{\sigma}_B(\widehat{\Phi}_B^\top x)$ are indicated as $\sigma_f$, $\sigma_B$, $\widehat{\sigma}_f$ and $\widehat{\sigma}_B$, respectively.

Now, in analogy with the classical ISM approach, in order to make the origin be an asymptotically stable equilibrium point for the approximated nominal dynamics

$$\dot{x} = \widehat{f}(x) + \widehat{B}(x)u_0 = \widehat{W}_f^\top \widehat{\sigma}_f + \widehat{W}_B^\top \widehat{\sigma}_B u_0, \tag{14}$$

the control law $u_0$ can be selected as

$$u_0 = -kx, \tag{15}$$

with $k \in \mathbb{R}^{1\times n}$ being a row vector of positive gains $k_i$, $\forall i = 1, 2, \ldots, n$. Then, the overall control input is obtained as the sum of (4) and (15), i.e.,

$$u = -kx - \rho\,\text{sign}\,(s). \tag{16}$$

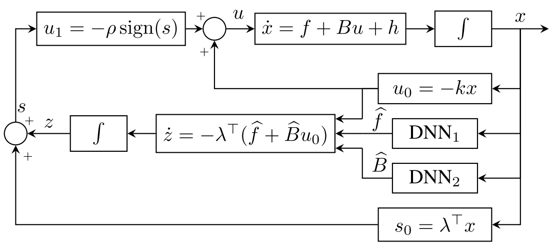



Fig. 1. Block diagram of the proposed NN-ISM control scheme.

The adaptation laws for the estimated weights are instead chosen as

$$\dot{\widehat{W}}_f = \Gamma_f \widehat{\sigma}_f s \lambda^\top, \tag{17a}$$
$$\dot{\widehat{W}}_B = -\Gamma_B \widehat{\sigma}_B k x s \lambda^\top, \tag{17b}$$
$$\dot{\widehat{\Phi}}_f = \Theta_f x \Big(\dot{\widehat{\sigma}}_f^\top \widehat{W}_f s \lambda\Big)^\top, \tag{17c}$$
$$\dot{\widehat{\Phi}}_B = -\Theta_B x \Big(\dot{\widehat{\sigma}}_B^\top \widehat{W}_B k x s \lambda\Big)^\top, \tag{17d}$$

where $\Gamma_f \in \mathbb{R}^{L_f\times L_f}$, $\Gamma_B \in \mathbb{R}^{L_B\times L_B}$, $\Theta_f \in \mathbb{R}^{n\times n}$ and $\Theta_B \in \mathbb{R}^{n\times n}$ are diagonal gain matrices, and

$$\dot{\widehat{\sigma}}_f = \text{diag}\,\{\widehat{\sigma}_f\}(I_{L_f\times L_f} - \text{diag}\,\{\widehat{\sigma}_f\}), \tag{18a}$$
$$\dot{\widehat{\sigma}}_B = \text{diag}\,\{\widehat{\sigma}_B\}(I_{L_B\times L_B} - \text{diag}\,\{\widehat{\sigma}_B\}), \tag{18b}$$

with $I_{L_f\times L_f}$ and $I_{L_B\times L_B}$ being identity matrices.

The dynamics of the sliding variable, that is, $\dot{s} = \lambda^\top \dot{x} + \dot{z}$, can be computed. In particular, using (1), (8), (9), (13), (15), (16) and exploiting relations (12a) and (12b) to express $\widehat{W}_f = W_f - \widetilde{W}_f$ and $\widehat{W}_B = W_B - \widetilde{W}_B$, one obtains

$$\begin{aligned}\dot{s} = \lambda^\top \Big(& h + \varepsilon_f - \varepsilon_B(\rho\,\text{sign}\,(s) + kx) \\ & - W_B^\top \sigma_B \rho\,\text{sign}\,(s) + W_f^\top(\sigma_f - \widehat{\sigma}_f) \\ & - W_B^\top(\sigma_B - \widehat{\sigma}_B)kx + \widetilde{W}_f^\top \widehat{\sigma}_f - \widetilde{W}_B^\top \widehat{\sigma}_B kx\Big).\end{aligned} \tag{19}$$

The proposed control scheme is illustrated in Fig. 1.

## V. Stability Analysis

In this section, the main theoretical results relevant to the proposed control approach are presented. In particular, the following theorem proves the ultimate boundedness of the system state, providing a bound on the convergence set depending on the control parameters. Moreover, conditions for the enforcement of both practical and ideal integral sliding modes are indicated.

*Theorem 2 (Ultimate Boundedness):* Consider the nonlinear system (1) controlled by (16), with $x_0 \in \Omega$, sliding variable as in (5) and (13), and output layer weights adaptation laws (17a) and (17b). If $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ hold, and

$$\rho > \frac{\overline{h} + \overline{W}_f(\overline{\sigma}_f + \overline{\widehat{\sigma}}_f)}{\overline{W}_B \overline{\sigma}_B}, \tag{20}$$

then $\forall t \geq \tilde{t}$, with $\tilde{t} \geq 0$, the state of the controlled system $x(t)$ is ultimately bounded in the set $\mathcal{X} \subseteq \Omega$ given by

$$\mathcal{X} := \left\{ x \in \Omega \mid \|x\| \leq \frac{\overline{\varepsilon}_f}{\|k\|\big(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B)\big)} \right\}. \tag{21}$$

*Proof:* Select the following Lyapunov-like candidate function $V$,

$$V = \frac{1}{2}s^2 + \frac{1}{2}\mathrm{tr}(\widetilde{W}_f^\top \Gamma_f^{-1}\widetilde{W}_f) + \frac{1}{2}\mathrm{tr}(\widetilde{W}_B^\top \Gamma_B^{-1}\widetilde{W}_B) \quad (22)$$

where $s$ is the sliding variable defined in (5). By differentiating with respect to time, one has

$$\dot{V} = s\dot{s} + \mathrm{tr}(\widetilde{W}_f^\top \Gamma_f^{-1}\dot{\widetilde{W}}_f) + \mathrm{tr}(\widetilde{W}_B^\top \Gamma_B^{-1}\dot{\widetilde{W}}_B). \quad (23)$$

Exploiting (19) and, since by computing the derivatives of (12a) and (12b) one has $\dot{\widetilde{W}}_f = -\dot{\widehat{W}}_f$ and $\dot{\widetilde{W}}_B = -\dot{\widehat{W}}_B$, then the derivative becomes

$$\begin{aligned}\dot{V} = s\lambda^\top\Big(&h + \varepsilon_f - \varepsilon_B(\rho\,\mathrm{sign}\,(s) + kx)\\ &- W_B^\top\sigma_B\rho\,\mathrm{sign}\,(s) + W_f^\top(\sigma_f - \widehat{\sigma}_f)\\ &- W_B^\top(\sigma_B - \widehat{\sigma}_B)kx + \widetilde{W}_f^\top\widehat{\sigma}_f - \widetilde{W}_B^\top\widehat{\sigma}_B kx\Big)\\ &- \mathrm{tr}(\widetilde{W}_f^\top\Gamma_f^{-1}\dot{\widehat{W}}_f) - \mathrm{tr}(\widetilde{W}_B^\top\Gamma_B^{-1}\dot{\widehat{W}}_B). \quad (24)\end{aligned}$$

Using the adaptive laws (17a) and (17b) one has

$$\begin{aligned}\dot{V} = s\lambda^\top\Big(&h + \varepsilon_f - \varepsilon_B(\rho\,\mathrm{sign}\,(s) + kx)\\ &- W_B^\top\sigma_B\rho\,\mathrm{sign}\,(s) + W_f^\top(\sigma_f - \widehat{\sigma}_f)\\ &- W_B^\top(\sigma_B - \widehat{\sigma}_B)kx + \widetilde{W}_f^\top\widehat{\sigma}_f - \widetilde{W}_B^\top\widehat{\sigma}_B kx\Big)\\ &- \mathrm{tr}(\widetilde{W}_f^\top\widehat{\sigma}_f s\lambda^\top) + \mathrm{tr}(\widetilde{W}_B^\top\widehat{\sigma}_B kxs\lambda^\top)\\ = s\lambda^\top\Big(&h + \varepsilon_f - \varepsilon_B(\rho\,\mathrm{sign}\,(s) + kx)\\ &- W_B^\top\sigma_B\rho\,\mathrm{sign}\,(s) + W_f^\top(\sigma_f - \widehat{\sigma}_f)\\ &- W_B^\top(\sigma_B - \widehat{\sigma}_B)kx + \widetilde{W}_f^\top\widehat{\sigma}_f - \widetilde{W}_B^\top\widehat{\sigma}_B kx\Big)\\ &- s\lambda^\top\widetilde{W}_f^\top\widehat{\sigma}_f + s\lambda^\top\widetilde{W}_B^\top\widehat{\sigma}_B kx. \quad (25)\end{aligned}$$

Rearranging the previous expression, $\dot{V}$ can be written as

$$\begin{aligned}\dot{V} = &-s\lambda^\top\Big(-h - W_f^\top(\sigma_f - \widehat{\sigma}_f)\Big) - s\lambda^\top\varepsilon_B\rho\,\mathrm{sign}\,(s)\\ &- s\lambda^\top\varepsilon_B kx - s\lambda^\top W_B^\top\sigma_B\rho\,\mathrm{sign}\,(s)\\ &- s\lambda^\top\Big(-\varepsilon_f + W_B^\top(\sigma_B - \widehat{\sigma}_B)kx\Big). \quad (26)\end{aligned}$$

Since $\boldsymbol{\mathcal{A}_1}$, $\boldsymbol{\mathcal{A}_2}$ and $\boldsymbol{\mathcal{A}_3}$ hold, one can write

$$\begin{aligned}\dot{V} \le &-\lambda^\top\mathbb{1}\big(-\overline{h} - \overline{W}_f(\overline{\sigma}_f + \overline{\widehat{\sigma}}_f)\big)|s| - \lambda^\top\mathbb{1}\overline{\varepsilon}_B\rho|s|\\ &- \lambda^\top\mathbb{1}\overline{\varepsilon}_B\|k\|\|x\||s| - \lambda^\top\mathbb{1}\overline{W}_B\overline{\sigma}_B\rho|s|\\ &- \lambda^\top\mathbb{1}\big(-\overline{\varepsilon}_f + \overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B)\|k\|\|x\|\big)|s|\\ \le &-\lambda^\top\mathbb{1}\big(\overline{W}_B\overline{\sigma}_B\rho - \big(\overline{h} + \overline{W}_f(\overline{\sigma}_f + \overline{\widehat{\sigma}}_f)\big)\big)|s|\\ &- \lambda^\top\mathbb{1}\big(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B)\|k\|\|x\| - \overline{\varepsilon}_f\big)|s|. \quad (27)\end{aligned}$$

Using $\rho$ as in (20), the first term of inequality (27) is negative, while for the second term two cases can occur. If

$$\|x\| \le \frac{\overline{\varepsilon}_f}{\|k\|\big(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B)\big)},$$

then $\dot{V} < 0$ or $\dot{V} \ge 0$. In this second subcase, one has that, inside the ball of radius $\frac{\overline{\varepsilon}_f}{\|k\|(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B))}$, $V$ is an increasing function and $\|x\|$ increases until $\|x\| = \frac{\overline{\varepsilon}_f}{\|k\|(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B))}$. On the other hand, if

$$\|x\| > \frac{\overline{\varepsilon}_f}{\|k\|\big(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B)\big)},$$

then $\dot{V} < 0$, and, outside the ball, $\|x\|$ decreases until $\|x\| \le \frac{\overline{\varepsilon}_f}{\|k\|(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B))}$. Therefore, $\forall\, t \ge \tilde{t}$, with $\tilde{t} \ge 0$ the state of the controlled system $x(t)$ is ultimately bounded in the set $\mathcal{X}$ in (21). Moreover, by using (22) and (27) one may observe that $V$ is bounded, which in turn implies that the weights $\widehat{W}_f$ and $\widehat{W}_B$ are bounded as well. ■

*Remark 1 (Size of the Convergence Set):* Note that, according to Theorem 2, a degree of freedom for the designer to reduce the radius of the set $\mathcal{X}$ is provided by the possibility of sizing the gain vector norm $\|k\|$.

*Remark 2 (Practical and Ideal Integral Sliding Mode):* Note that, if $x(t)$ is ultimately bounded in the set $\mathcal{X}$, as it happens if the assumptions of Theorem 2 hold, then according to the definition of $s_0$, also $|s_0|$ is ultimately bounded. This implies that a practical integral sliding mode is enforced in a boundary layer around $s = 0$, the amplitude of which can be reduced by acting on $\|k\|$. This will be illustrated in § VI.

In case $\dot{V}$ in (23) would result in being negative-definite both outside and inside $\mathcal{X}$, then, with $\rho$ as in (20), the *reaching condition* [1, Ch. 1] would be verified. Therefore, $s$ would become zero in a finite time $\bar{t}$, and remain zero $\forall\, t \ge \bar{t}$. Since, according to (13), $s(x_0) = 0$, then $\bar{t} = 0$, which means that an ideal integral sliding mode on $s(t) = 0$ would be enforced since the initial time instant. Notice that, when $s = 0$, the controlled system, which is in sliding mode, becomes equal to system (14) (i.e., to the approximated nominal dynamics) with $u_0$ as in (15). By design, $u_0$ makes the origin be an asymptotically stable equilibrium point for the approximated nominal dynamics. Hence, when an ideal integral sliding mode on $s = 0$ is produced, it follows that the origin results in being an asymptotically stable equilibrium point also for system (1) controlled by means of the proposed control $u$ indicated in (16).

*Remark 3 (Conservativeness of the Control Gain):* Note that Theorem 2 holds under the assumption $\boldsymbol{\mathcal{A}_3}$, which implies the control gain $\rho$ in (20), and the set $\mathcal{X}$ in (21). However, if we could relax $\boldsymbol{\mathcal{A}_3}$ by assuming to know the bound of the input function reconstruction error $\varepsilon_B$, that is $\overline{\varepsilon}_B \in \mathbb{R}_{>0}$, the sliding mode control gain would be given by $\overline{\rho} > \frac{\overline{h} + \overline{W}_f(\overline{\sigma}_f + \overline{\widehat{\sigma}}_f)}{\overline{W}_B\overline{\sigma}_B + \overline{\varepsilon}_B}$, with $\overline{\rho} < \rho$. As a consequence, the new convergence set $\overline{\mathcal{X}} \subset \mathcal{X}$ would be

$$\overline{\mathcal{X}} := \left\{x \in \Omega \mid \|x\| \le \frac{\overline{\varepsilon}_f}{\|k\|\big(\overline{W}_B(\overline{\sigma}_B + \overline{\widehat{\sigma}}_B + \overline{\varepsilon}_B)\big)}\right\}. \quad (28)$$

## VI. Numerical Example

In this section, the proposed control scheme is assessed in simulation, considering as process to control the Duffing oscillator (see [1, Ch. 1]) described by

$$f(x) = \begin{bmatrix} x_2 \\ \frac{g}{l}\sin(x_1) - \frac{\beta}{ml^2}x_1 - \frac{\gamma}{ml^2}x_2 \end{bmatrix}, \quad B(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$h(x,t) = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.15\sin\left(\frac{1}{2}t\right) + 0.1\cos(t) \end{bmatrix},$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top$, with $x_1$ being the position and $x_2$ the velocity, $m$ the mass of the oscillator, and $l$ its length. If the

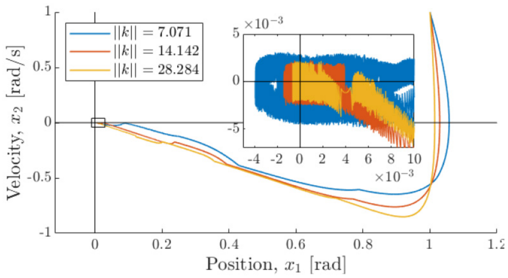



Fig. 2. State-space trajectories in the case of different values of $\|k\|$.

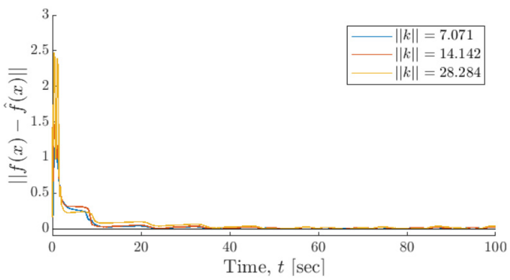



Fig. 3. Time evolution of the norm of the drift term estimation error when the system is controlled with different values of $\|k\|$.

model parameters $\beta$ and $\gamma$ are chosen such that $\frac{g}{l} - \frac{\beta}{ml^2} = 1$, $\frac{g}{6l} = 1$ and $\frac{\gamma}{ml^2} = 1$, then the dynamics can be rewritten as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_1 = x_1 - x_1^3 - x_2 + u + h_2. \end{cases} \tag{29}$$

Note that the knowledge of this model is assumed not to be available for the controller design.

### A. Settings

To asses the validity of the NN-ISM control algorithm introduced in § IV, three different simulations has been carried out, analyzing the behavior of system (29) for different values of the control gain vector $k$: $k = \begin{bmatrix} 5 & 5 \end{bmatrix}$, $k = \begin{bmatrix} 10 & 10 \end{bmatrix}$ and $k = \begin{bmatrix} 20 & 20 \end{bmatrix}$, with $\|k\|$ equal to 7.071, 14.142 and 28.284, respectively. The drift dynamics and the control effectiveness term used for updating the transient function $z$ have been estimated by using 2 DNNs with one output layer with 2 neurons and one hidden layer with $L_f = L_B = 50$ neurons. Such DNNs are characterized by output layer weights $\widehat{W}_f$, $\widehat{W}_B$ and hidden layer weights $\widehat{\Phi}_f$, $\widehat{\Phi}_B$, whose values are initialized as small random numbers (between 0 and 0.01) and then adapted using (17), with $\Gamma_f = \Gamma_B = 0.7 \cdot I_{50\times 50}$ and $\Theta_f = \Theta_B = 0.9 \cdot I_{2\times 2}$. All the simulations have the same initial conditions $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^\top$ and duration equal to 100 seconds. Moreover, the sliding variable parameter vector has been chosen as $\lambda = \begin{bmatrix} 1 & 1 \end{bmatrix}^\top$, while the discontinuous control gain has been selected as $\rho = 0.3$, which, assuming upper bounds $\overline{h} = 0.25$, $\overline{W}_f = 0.2$, $\overline{W}_B = 2.2$ and $\overline{\tilde{\sigma}}_f = \overline{\sigma}_f = \overline{\sigma}_B = 1$, satisfies (20).

### B. Results and Discussion

By virtue of Theorem 2, and according to Remark 1 the size of the convergence set $\mathcal{X}$ decreases as $\|k\|$ increases. Fig. 2 shows indeed that, larger $\|k\|$ is, smaller the radius of set $\mathcal{X}$,

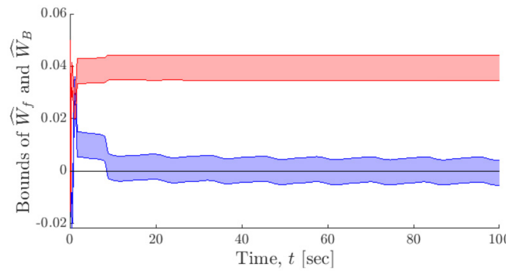
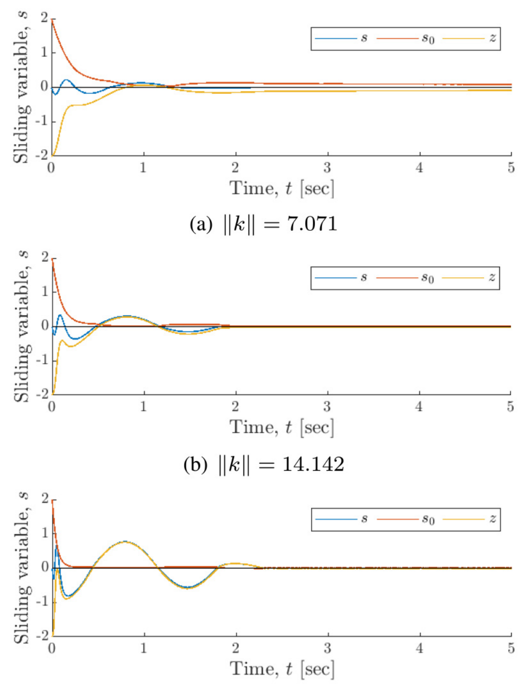



Fig. 4. Lower and upper bounds of $\widehat{W}_f$ (in blue) and $\widehat{W}_B$ (in red) when $\|k\| = 28.284$.




(c) $\|k\| = 28.284$

Fig. 5. Time evolution of $s = s_0 + z$ when $\|k\| = 7.071$ (a), $\|k\| = 14.142$ (b) and $\|k\| = 28.284$ (c).

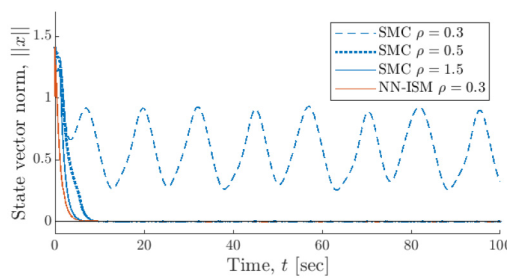



Fig. 6. Time evolution of $\|x\|$ when the system is controlled via a conventional SMC, with different control amplitudes, and via the proposed NN-ISM control.

namely $r_{\mathcal{X}}$, becomes. This result is also evident from Table I, where $r_{\mathcal{X}}$ is estimated relying on (21). Evidence of the DNN capability of estimating, for instance, the uncertain drift term is provided in Fig. 3, where the norm of the drift term estimation error is reported versus time for different values of $\|k\|$.

In Fig. 4 the evolution of the output layer weights $\widehat{W}_f$ and $\widehat{W}_B$ is shown. As expected from Theorem 2, they are also

TABLE I
SIZE OF THE CONVERGENCE SET $\mathcal{X}$ WITH RESPECT TO $\|k\|$

| $\|k\|$ | $r_{\mathcal{X}}$ |
|---|---|
| 7.071 | $4.2 \cdot 10^{-3}$ |
| 14.142 | $1.3 \cdot 10^{-3}$ |
| 28.284 | $7 \cdot 10^{-4}$ |

bounded. In Fig. 5 the time evolution of the sliding variable $s$ is illustrated for different value of $\|k\|$. Starting from $s = 0$, then for a short transient of 2 seconds the sliding mode is temporarily lost until the DNN weights $\widehat{W}_f$, $\widehat{W}_B$, $\widehat{\Phi}_f$, $\widehat{\Phi}_B$ are properly adjusted, and then it is again steered to zero. Finally, a comparison between the proposed NN-ISM control approach and a classical SMC [26], with amplitude of the control input based on the knowledge of the upper bounds of the uncertain terms, is reported in Fig. 6. Notice that to obtain a performance comparable with that of the NN-ISM control, the amplitude of the classical SMC must be about 5 times higher.

## VII. CONCLUSION

In this letter, a novel NN-ISM control algorithm is proposed for a class of uncertain nonlinear systems. In order to use the classical ISM control approach, the nominal model of the system should be known, which is not true in the considered case. Thus, in this letter, the integral sliding manifold is designed relying on two DNNs. The theoretical analysis reported in this letter provides conditions for the enforcement of practical integral sliding modes, as well as results on the boundedness of the DNNs weights. Moreover, the possible generation of ideal integral sliding modes is discussed. The proposed control approach provides satisfactory performance, as assessed in simulation.

## REFERENCES


[1] A. Ferrara, G. P. Incremona, and M. Cucuzzella, *Advanced and Optimization Based Sliding Mode Control: Theory and Applications*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 2019.

[2] A. Levant, “Chattering analysis,” *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1380–1389, Jun. 2010.

[3] G. Bartolini, A. Ferrara, A. Levant, and E. Usai, “On second order sliding mode controllers,” in *Variable Structure Systems, Sliding Mode and Nonlinear Control* (Lecture Notes in Control and Information), K. D. Young and Ü. Özgüner, Eds. London, U.K.: Springer-Verlag, 1999, pp. 329–350.

[4] A. Levant, “Higher-order sliding modes, differentiation and output-feedback control,” *Int. J. Control*, vol. 76, nos. 9–10, pp. 924–941, 2003.

[5] L. Fridman, J. A. Moreno, B. Bandyopadhyay, S. Kamal, and A. Chalanga, “Continuous nested algorithms : The fifth generation of sliding mode controllers,” in *Recent Advances in Sliding Modes: From Control to Intelligent Mechatronics*, X. Yu and M. Ö. Efe, Eds. Cham, Switzerland: Springer Int., 2015, pp. 5–35.

[6] G. Bartolini, A. Ferrara, L. Giacomini, and E. Usai, “A combined backstepping/second order sliding mode approach to control a class of nonlinear systems,” in *Proc. IEEE Int. Workshop Variable Struct. Syst.*, Tokyo, Japan, Dec. 1996, pp. 205–210.

[7] G. Bartolini, A. Ferrara, A. Pisano, and E. Usai, “Adaptive reduction of the control effort in chattering-free sliding-mode control of uncertain nonlinear systems,” *Appl. Math. Comput. Sci.*, vol. 8, no. 1, pp. 51–71, 1998.

[8] A. Pisano, M. Tanelli, and A. Ferrara, “Switched/time-based adaptation for second-order sliding mode control,” *Automatica*, vol. 64, pp. 126–132, Feb. 2016.

[9] G. P. Incremona, L. Mirkin, and P. Colaneri, “Integral sliding-mode control with internal model: A separation,” *IEEE Control Syst. Lett.*, vol. 6, pp. 446–451, 2022.

[10] V. Utkin and J. Shi, “Integral sliding mode in systems operating under uncertainty conditions,” in *Proc. 35th IEEE Conf. Decis. Control*, vol. 4. Kobe, Japan, Dec. 1996, pp. 4591–4596.

[11] M. Rubagotti, A. Estrada, F. Castaños, A. Ferrara, and L. Fridman, “Integral sliding mode control for nonlinear systems with matched and unmatched perturbations,” *IEEE Trans. Autom. Control*, vol. 56, no. 11, pp. 2699–2704, Nov. 2011.

[12] A. Ferrara, G. P. Incremona, and B. Sangiovanni, “Tracking control via switched integral sliding mode with application to robot manipulators,” *Control Eng. Pract.*, vol. 90, pp. 257–266, Sep. 2019.

[13] A. Ferrara and G. P. Incremona, “Design of an integral suboptimal second-order sliding mode controller for the robust motion control of robot manipulators,” *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 6, pp. 2316–2325, Nov. 2015.

[14] G. P. Incremona, A. Ferrara, and L. Magni, “MPC for robot manipulators with integral sliding modes generation,” *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 3, pp. 1299–1307, Jun. 2017.

[15] F. Lewis, S. Jagannathan, and A. Yesildirak, *Neural Network Control of Robot Manipulators and Non-Linear Systems*. London, U.K.: Taylor Francis, 1999.

[16] F. Lewis, “Nonlinear network structures for feedback control,” *Asian J. Control*, vol. 1, no. 4, pp. 205–228, 1999.

[17] K. Esfandiari, F. Abdollahi, and H. A. Talebi, *Neural Network-Based Adaptive Control of Uncertain Nonlinear Systems*. Cham, Switzerland: Springer, 2022.

[18] R. Safaric and K. Jezernik, “Trajectory tracking neural network controller for a robot mechanism and Lyapunov theory of stability,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 1. Munich, Germany, Sep. 1994, pp. 626–633.

[19] L. Cheng, Z. Wang, F. Jiang, and J. Li, “Adaptive neural network control of nonlinear systems with unknown dynamics,” *Adv. Space Res.*, vol. 67, no. 3, pp. 1114–1123, 2021.

[20] R. Sun, M. L. Greene, D. M. Le, Z. I. Bell, G. Chowdhary, and W. E. Dixon, “Lyapunov-based real-time and iterative adjustment of deep neural networks,” *IEEE Contr. Syst. Lett.*, vol. 6, pp. 193–198, 2022.

[21] A. Guarneros-Sandoval, M. Ballesteros, I. Salgado, J. Rodríguez-Santillán, and I. Chairez, “Lyapunov stable learning laws for multilayer recurrent neural networks,” *Neurocomputing*, vol. 491, pp. 644–657, Jun. 2022.

[22] L. Wang, T. Chai, and L. Zhai, “Neural-network-based terminal sliding-mode control of robotic manipulators including actuator dynamics,” *IEEE Trans. Ind. Electron.*, vol. 56, no. 9, pp. 3296–3304, Sep. 2009.

[23] N. T. Tai and K. K. Ahn, “A RBF neural network sliding mode controller for SMA actuator,” *Int. J. Control Autom. Syst.*, vol. 8, no. 6, pp. 1296–1305, 2010.

[24] J. Fei and H. Ding, “Adaptive sliding mode control of dynamic system using RBF neural network,” *Nonlinear Dyn.*, vol. 70, no. 2, pp. 1563–1573, 2012.

[25] J. Fei and C. Lu, “Adaptive sliding mode control of dynamic systems using double loop recurrent neural network structure,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1275–1286, Apr. 2018.

[26] V. I. Utkin, *Sliding Modes in Control and Optimization* (Communication and Control Engineering). Berlin, Germany: Springer, 1992.


Open Access funding provided by ‘Università degli Studi di Pavia’ within the CRUI-CARE Agreement