Computation of Convergence Bounds for Volterra Series of Linear-Analytic Single-Input Systems

Thomas Hélie and Béatrice Laroche

Abstract—**In this paper, the Volterra series decomposition of a class of single-input time-invariant systems, analytic in state and affine in input, is analyzed. Input-to-state convergence results are obtained for several typical norms ($L^\infty([0,T])$, $L^\infty(\mathbb{R}_+)$ as well as exponentially weighted norms). From the standard recursive construction of Volterra kernels, new estimates of the kernel norms are derived. The singular inversion theorem is then used to obtain the main result of the paper, namely, an easily computable bound of the convergence radius. Guaranteed error bounds for the truncated series are also provided. The relevance of the method is illustrated in several examples.**

Index Terms—**Approximation methods, functional analysis, nonlinear dynamical systems.**

I. Introduction

Volterra series is a functional series expansion of the solution of nonlinear controlled systems, first introduced by the Italian mathematician Volterra [25]. This tool has been extensively used in signal processing and control, electronics and electromagnetic waves, mechanics and acoustics, biomedical engineering, for modeling, identification, and simulation purposes. There exists a vast literature concerning Volterra series. Among others, they were studied in [5], [10], [11], [13], and [19] from the geometric control point of view, and in [9], [23], and [24] from the input–output representation and realization point of view.

However, only a few results on the convergence are available, and most of them require the computation of kernel norms or asymptotic bounds. In [6], the existence of a nonzero convergence radius for complex linear analytic systems with no initial conditions is proved and a convergence criterion is given for invariant real bilinear systems. In [19] and [21], theoretical and local-in-time results are given for control systems, affine in the input, with analytic dynamics and piecewise continuous inputs. Existence results of a convergence radius for continuous inputs are also given in [2] for fading memory systems, and in [14] for results in L^p-spaces. More recently, results in the frequency domain have been developed in [20] and [22], results relying on regular perturbations (which can be related to Volterra series expansion) are given in [7], and results for quadratic systems have been established in [16].

This paper focuses on the computation of guaranteed convergence bounds for the input-to-state Volterra series expansion of a class of single-input time-invariant systems. These systems are assumed to be finite-dimensional, analytic in state, affine in input, with zero initial conditions. The convergence bounds are obtained for several norms for the input and the state: 1) the $L^\infty(\mathbb{R}_+)$ norm is considered to establish bounded signal results for systems with a stable linear part; 2) $L^\infty([0,T])$ norms for $T>0$ makes it possible to relax the convergence condition including for systems with an unstable linear part; 3) an adapted weighted norm is also considered to tackle exponentially fading memory systems.

The paper is organized as follows. Section II defines the notations, the functional setting, and the class of systems under consideration and recalls some general definitions and basic properties of Volterra series. Section III establishes the convergence results and guaranteed truncation error bounds for L^∞ norms. These results are illustrated by several examples in Section IV. Section V extends the results of Section III to weighted norms adapted to exponentially fading memory systems. Finally, conclusions and perspectives are given in Section VI.

II. General Framework

A. Notations and Functional Setting

The following notations are introduced, where $\mathbb{E}, \mathbb{E}_1, \ldots, \mathbb{E}_K$ ($K \geq 2$) and $\mathbb{F}$ are real normed vector spaces.

- $\mathcal{L}(\mathbb{E},\mathbb{F})$ is the vector space of continuous linear functions from an $\mathbb{E}$ to $\mathbb{F}$ with norm $\|f\|_{\mathcal{L}(\mathbb{E},\mathbb{F})} = \sup_{x\in B_\mathbb{E}} \|f(x)\|_\mathbb{F}$, where $B_\mathbb{E}$ is the unit ball in $\mathbb{E}$.
- $\mathcal{ML}(\mathbb{E}_1,\ldots,\mathbb{E}_K,\mathbb{F})$ is the vector space of continuous multilinear functions $f:\mathbb{E}_1\times\cdots\times\mathbb{E}_K\to\mathbb{F}$ with norm

$$\|f\|_{\mathcal{ML}(\mathbb{E}_1,\ldots,\mathbb{E}_K,\mathbb{F})} = \sup_{(x_1,\ldots,x_K)\in B_{\mathbb{E}_1}\times\cdots\times B_{\mathbb{E}_K}} \|f(x_1,\ldots,x_K)\|_\mathbb{F}.$$

- $\mathcal{ML}_{j_1,\ldots,j_K}(\mathbb{E}_1,\ldots,\mathbb{E}_K,\mathbb{F})$ denotes the set $\mathcal{ML}(\underbrace{\mathbb{E}_1,\ldots,\mathbb{E}_1}_{j_1},\underbrace{\mathbb{E}_2,\ldots,\mathbb{E}_2}_{j_2},\ldots,\underbrace{\mathbb{E}_K,\ldots,\mathbb{E}_K}_{j_K},\mathbb{F})$.

The following function spaces are used in the sequel.

- $\mathbb{T}$ denotes the time interval $[0,T]$ with $T>0$ or $\mathbb{R}_+$.
- $L^1(\mathbb{T},\mathbb{E})$ and $L^\infty(\mathbb{T},\mathbb{E})$ are the standard Lebesgue spaces.
- $\mathcal{V}_\mathbb{E}^N$ for $N\in\mathbb{N}^*$ is the set of functions $f:\mathbb{T}\times\mathbb{T}^N\to\mathbb{E}$ such that

$$t\mapsto(\tau\mapsto f(t,\tau))\in L^\infty\big(\mathbb{T},L^1\big(\mathbb{T}^N,\mathbb{E}\big)\big)$$

Manuscript received February 11, 2010; revised May 21, 2010; accepted October 22, 2010. Date of publication November 11, 2010; date of current version September 08, 2011. This work was supported by the National Research Agency and the CONSONNES Project (ANR-05-BLAN-0097–01). A simplified version of this work was presented at the IEEE Conference on Control and Decision, Shanghai, China, December 16–18, 2009. Recommended by Associate Editor L. Marconi.

T. Hélie is with Ircam, Centre Georges Pompidou, CNRS STMS UMR 9912, 75004 Paris, France (e-mail: thomas.helie@ircam.fr).

B. Laroche is with the Laboratoire des Signaux et Systèmes, Université Paris-Sud, CNRS UMR8506, Supélec, 91405 Gif-Sur-Yvette, France (e-mail: beatrice.laroche@lss.supelec.fr).

Digital Object Identifier 10.1109/TAC.2010.2091435

with norm $\|f\|_{\mathcal{V}_{\mathbb{E}}^N} = \sup_{t\in\mathbb{T}} \int_{\mathbb{T}^N} \|f(t,\tau)\|_{\mathbb{E}}\, \mathrm{d}\tau$, where $\mathrm{d}\tau$ denotes the Lebesgue's measure $\mathrm{d}\tau = \mathrm{d}\tau_1 \ldots \mathrm{d}\tau_N$.

B. System Under Consideration

The systems under consideration are analytic systems with zero initial conditions and an affine dependence on the input, that is, for $t \in \mathbb{T}$

$$\dot{x} = f(x,u) = A\,x + B\,u + P(x) + Q(x,u) \quad (1)$$
$$y = g(x,u) \quad (2)$$
$$x(0) = 0 \quad (3)$$

where $u : \mathbb{T} \to \mathbb{R}$, $x : \mathbb{T} \to \mathbb{X} = \mathbb{R}^n$, $y : \mathbb{T} \to \mathbb{R}^p$ with finite dimensions $n \in \mathbb{N}^*$, $p \in \mathbb{N}^*$, where A is an $n \times n$ real matrix, B is a nonzero $n \times 1$ real matrix, g is analytic at $(0,0)$, and where P and Q are analytic functions expressed as a series of homogeneous contributions

$$P(x) = \sum_{k=2}^{\infty} P_k(\underbrace{x,\ldots,x}_{k}) \quad (4)$$

$$Q(x,u) = \sum_{k=2}^{\infty} Q_k(\underbrace{x,\ldots,x}_{k-1},u) \quad (5)$$

with $P_k \in \mathcal{ML}_k(\mathbb{X},\mathbb{X})$ and $Q_k \in \mathcal{ML}_{k-1,1}(\mathbb{X},\mathbb{R},\mathbb{X})$. Note that these systems correspond to the class of so-called "linear analytic systems" or "affine systems" (see [6] and [8])

$$\dot{x} = \widetilde{f}(x) + \widetilde{g}(x)u$$

where we set $\widetilde{f}(x) = Ax + P(x)$ and $\widetilde{g}(x) = Q(x,1)$. As stated in Section VI, some of the previous hypotheses such as zero initial conditions, single input, and "affine in input" will be relaxed in future works.

Definition 1 (Weak and Bounded Solutions): Let $\mathcal{U} = L^\infty(\mathbb{T},\mathbb{R})$ and $\mathcal{X} = L^\infty(\mathbb{T},\mathbb{X})$. Let $u \in \mathcal{U}$. A function $x : \mathbb{T} \to \mathbb{R}$ is a weak solution of (1) and (3) if the following apply.

1) x is absolutely continuous on all bounded intervals of $\mathbb{T}$ and $\mathrm{d}x/\mathrm{d}t \in \mathcal{X}$.
2) Equation (1) is satisfied almost everywhere, and (3) holds.

Moreover, x is said to be a bounded solution if it is a weak solution and it belongs to $\mathcal{X}$.

Remark 1: If (x,u) is in the analytic domain of g almost everywhere, then y is also bounded almost everywhere.

Then, in the sequel, we focus on the input-to-state system.

C. Volterra Series: General Definitions and Basic Properties

We restate the standard definition of the Volterra series [3].

Definition 2 (Volterra Series): Let $\mathcal{VS}$ denote the set of sequences of kernels $\{f_m\}_{m\in\mathbb{N}^*}$ such that for all $m \in \mathbb{N}^*$, $f_m \in \mathcal{V}^m$, where $\mathcal{V}^m = \mathcal{V}_{\mathbb{X}}^m$. A causal SI-system can be described by an input-to-state Volterra series if there exist $\{h_m\}_{m\in\mathbb{N}^*} \in \mathcal{VS}$ and $\rho \in \mathbb{R}_+^*$ such that for all input $u \in \mathcal{U}$ satisfying $\|u\|_{\mathcal{U}} < \rho$, the series

$$x(t) = \sum_{m\in\mathbb{N}^*} \int_{[0,t]^m} h_m(t,\tau)\,\left[\Pi_m u\right](\tau)\,\mathrm{d}\tau \quad (6)$$

defined for $t \in \mathbb{T}$, with $\left[\Pi_m u\right](\tau) = \prod_{i=1}^m u(\tau_i)$, is normally convergent for the norm $\|.\|_{\mathcal{X}}$. For $m \in \mathbb{N}^*$, the function h_m is called the kernel of order m.

Remark 2: The input-to-output Volterra series of the system (1)–(5) can be deduced from its input-to-state Volterra series by substituting (6) in (2).

Definition 3 (Gain Bound Function): Let $\{h_m\}_{m\in\mathbb{N}^*} \in \mathcal{VS}$ be such that convergence radius ρ of the formal series $\sum_{m\in\mathbb{N}^*} \|h_m\|_{\mathcal{V}^m} X^m$ belongs to $\mathbb{R}_+^*$. Then, the *gain bound function* φ of $\{h_m\}_{m\in\mathbb{N}^*}$ is defined for all $z \in \mathbb{C}$ such that $|z| < \rho$ by

$$\varphi(z) = \sum_{m\in\mathbb{N}^*} \|h_m\|_{\mathcal{V}^m}\, z^m.$$

Theorem 1 (Bounded-Input Bounded-State Relation): Let $\{h_m\}_{m\in\mathbb{N}^*} \in \mathcal{VS}$ be such that the gain bound function φ has a nonzero radius of convergence $\rho > 0$. Then, the Volterra series is convergent in $\mathcal{X}$ for inputs such that $\|u\|_{\mathcal{U}} < \rho$. In this case, $x \in \mathcal{X}$ satisfies $\|x\|_{\mathcal{X}} \le \varphi\big(\|u\|_{\mathcal{U}}\big) < \infty$.

Proof: Let $u \in \mathcal{U}$ be such that $\|u\|_{\mathcal{U}} < \rho$. Then, $\varphi\big(\|u\|_{\mathcal{U}}\big) < \infty$. Now, for all $m \in \mathbb{N}^*$,

$$\begin{aligned}
&\sup_{t\in\mathbb{T}} \left(\left\| \int_{[0,t]^m} h_m(t,\tau)\,\left[\Pi_m u\right](\tau)\,\mathrm{d}\tau \right\|_{\mathbb{X}} \right) \\
&\le \sup_{t\in\mathbb{T}} \left(\int_{[0,t]^m} \left\|h_m(t,\tau)\right\|_{\mathbb{X}} \left(\|u\|_{\mathcal{U}}\right)^m \mathrm{d}\tau \right) \\
&\le \|h_m\|_{\mathcal{V}^m} \left(\|u\|_{\mathcal{U}}\right)^m.
\end{aligned}$$

Hence, the series $\sum_{m\in\mathbb{N}^*} \int_{[0,t]^m} h_m(t,\tau)\,\left[\Pi_m u\right](\tau)\,\mathrm{d}\tau$ converges normally in the Banach space $\mathcal{X}$ to a limit x such that

$$\begin{aligned}
\|x\|_{\mathcal{X}} &= \sup_{t\in\mathbb{T}} \left(\left\| \sum_{m\in\mathbb{N}^*} \int_{[0,t]^m} h_m(t,\tau)\,\left[\Pi_m u\right](\tau)\,\mathrm{d}\tau \right\|_{\mathbb{X}} \right) \\
&\le \sum_{m\in\mathbb{N}^*} \|h_m\|_{\mathcal{V}^m} \left(\|u\|_{\mathcal{U}}\right)^m = \varphi\big(\|u\|_{\mathcal{U}}\big).
\end{aligned}$$

■

D. Recursive Construction of Kernels

Definition 4 (Index Set and Selection Function): Let $m \in \mathbb{N}^*$ and $K \in \mathbb{N}^*$. The set $\mathbb{M}_m^K$ is defined by

$$\mathbb{M}_m^K = \left\{ p \in (\mathbb{N}^*)^K \mid p_1 + \cdots + p_K = m \right\}.$$

Moreover, for all $p \in \mathbb{M}_m^K$ and for all $k \in [1,K]_{\mathbb{N}}$, the selection function $S_p^k : \mathbb{T}^m \to \mathbb{T}^{p_k}$ is defined by, denoting $\tau = (\tau_1, \tau_2, \ldots, \tau_m)$

$$S_p^k(\tau) = \left(\tau_{p_1+\cdots+p_{k-1}+1}, \tau_{p_1+\cdots+p_{k-1}+2}, \ldots, \tau_{p_1+\cdots+p_k}\right).$$

Note that if $K > m$, then $\mathbb{M}_m^K = \emptyset$.

Following [3], [4], [11], and [23], a recursive construction algorithm for the kernels associated to the system described by (1)–(5) is given.

Proposition 1 (Kernels Recursive Construction): Let the family of kernels $\{h_m\}_{m\in\mathbb{N}^*}$ be defined by $h_1 : \mathbb{T} \times \mathbb{T} \to \mathbb{X}$ with

$$h_1(t,\tau_1) = 1_{\mathbb{R}_+}(t-\tau_1)\,\mathrm{e}^{A(t-\tau_1)}B$$

and for all $m \geq 2$, $h_m : \mathbb{T} \times \mathbb{T}^m \to \mathbb{X}$ with

$$h_m(t,\tau) = 1_{\mathbb{R}_+}(t - \max\tau)\left(\int_{\max\tau}^{t} v_m(t,\theta,\tau)\mathrm{d}\theta + w_m(t,\tau)\right)$$

where $1_{\mathbb{R}_+}$ denotes the Heaviside function and

$$v_m(t,\theta,\tau) = \mathrm{e}^{A(t-\theta)} \sum_{k=2}^{m} \sum_{p\in\mathbb{M}_m^k} P_k\Big(h_{p_1}\big(\theta, S_p^1(\tau)\big), \ldots, h_{p_k}\big(\theta, S_p^k(\tau)\big)\Big) \tag{7}$$

$$w_m(t,\tau) = 1_{\mathbb{R}_+}(\tau_m - \max_{1\leq i<m}\tau_i)\,\mathrm{e}^{A\,(t-\tau_m)} \times \left[\sum_{k=2}^{m} \sum_{\begin{cases} q\in\mathbb{M}_m^k \\ q_k=1\end{cases}} \times Q_k\Big(h_{q_1}\big(\tau_m, S_q^1(\tau)\big), \ldots, h_{q_{k-1}}\big(\tau_m, S_q^{k-1}(\tau)\big), 1\Big)\right]. \tag{8}$$

Then, the series (6) is a formal solution of system (1)–(5).

III. Main Convergence Results

In this section, a computable lower bound $\rho^\star$ of the convergence radius ρ of the Volterra series introduced in Proposition 1 is given (Theorem 2). The resulting algorithm is fed by the norm of the dynamics of the linearized problem and norms of the multilinear operators (P_k, Q_k), but it does not require any iterative kernels norm estimation. The infinite sum is a bounded solution of problem (1)–(5) in the sense of Definition 1 (Theorem 3). Finally, a bound on the remainders $\|x - V_M x\|_{\mathcal{X}}$ is given in Theorem 4, where $V_M x$ denotes the sum of the Volterra series truncated at order M.

Hypothesis 1: The system (1)–(5) is such that $t \mapsto \exp(At)$ belongs to $L^1(\mathbb{T}, \mathbb{R}^{n\times n})$; that is, A is supposed to be Hurwitz if $\mathbb{T} = \mathbb{R}_+$, and there is no assumption on A if $\mathbb{T}$ is a finite interval $[0, T]$.

Then, the following function can be introduced.

Definition 5: The function $\mathcal{F}$ is formally defined by

$$\mathcal{F}(X) = \frac{\|h_1\|_{\mathcal{V}^1} + \sum_{k=2}^{\infty} \mathcal{Q}_k X^{k-1}}{1 - \sum_{k=2}^{\infty} \mathcal{P}_k X^{k-1}} \tag{9}$$

where, defining $\kappa = \int_{\mathbb{T}} \|\mathrm{e}^{A\theta}\|_{\mathcal{L}(\mathbb{X},\mathbb{X})}\mathrm{d}\theta$, for all $k \geq 2$

$$\mathcal{P}_k = \kappa\|P_k\|_{\mathcal{ML}_k(\mathbb{X},\mathbb{X})} \quad \mathcal{Q}_k = \kappa\|Q_k\|_{\mathcal{ML}_{k-1,1}(\mathbb{X},\mathbb{R},\mathbb{X})}. \tag{10}$$

Remark 3: This definition of κ is consistent and $h_1 \in \mathcal{V}^1$. This is obvious if $\mathbb{T}$ is a finite interval. If $\mathbb{T} = \mathbb{R}_+$, then A is Hurwitz so that $-a = \max\Big(\Re\mathrm{e}\big(\mathrm{spec}\,A\big)\Big) < 0$ and $\beta \in \mathbb{R}_+$ can be chosen such that for all $\theta \in \mathbb{R}_+$, $\|\mathrm{e}^{A\theta}\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \leq \beta\,\mathrm{e}^{-a\theta}$. Then, $0 < \kappa \leq \int_{\mathbb{R}_+} \|\mathrm{e}^{A\theta}\|_{\mathcal{L}(\mathbb{X},\mathbb{X})}\,\mathrm{d}\theta \leq \beta/a$, and

$$0 < \|h_1\|_{\mathcal{V}^1} \leq \kappa\,\|B\|_{\mathcal{L}(\mathbb{R},\mathbb{X})} \leq \frac{\beta}{a}\|B\|_{\mathcal{L}(\mathbb{R},\mathbb{X})}. \tag{11}$$

Theorem 2 (Lower Bound for the Convergence Radius): The family $\{h_m\}_{m\in\mathbb{N}^*}$ defined in Proposition 1 belongs to $\mathcal{VS}$. Moreover, let $r \in \mathbb{R}_+^* \cup \{+\infty\}$ be the radius of convergence of $\mathcal{F}$ at $x = 0$. Equation $x\,\mathcal{F}'(x) - \mathcal{F}(x) = 0$ has either one solution denoted σ (case 1) or zero solution (case 2), in $]0, r[$. Let $\rho^\star > 0$ be defined by

$$\text{(case 1)} \qquad \rho^\star = \frac{\sigma}{\mathcal{F}(\sigma)} \tag{12}$$

$$\text{(case 2)} \qquad \rho^\star = \lim_{x\to r^-} \frac{x}{\mathcal{F}(x)}. \tag{13}$$

Then, the convergence radius of the gain bound function φ is greater than $\rho^\star$.

The proof of Theorem 2 is based on tools from analytic combinatorics [12]. The key steps are the following.

- Step 1) The recursive kernel construction formula (in Proposition 1) is exploited to obtain a majorizing sequence ψ_m of the kernel norms.
- Step 2) The associated generating function $\Psi(z)$ is proved to satisfy the implicit equation $\Psi(z) = z\,\mathcal{F}\big(\Psi(z)\big)$. The function $\mathcal{F}$ only involves κ, $\|h_1\|_{\mathcal{V}^1}$ and the norms of operators A_k and B_k.
- Step 3) Function Ψ is proved to be analytic at $z = 0$. A lower bound $\rho^\star$ for its convergence radius is derived using the singular inversion theorem (case 1) and the analytic inversion lemma (case 2) (see e.g., [12]). This gives a lower bound for the convergence radius of the gain bound function φ.

Proof:

Step 1: We prove by induction that, for all $m \in \mathbb{N}^*$, h_m belongs to $\mathcal{V}^m$ and satisfies

$$\|h_m\|_{\mathcal{V}^m} \leq \psi_m \tag{14}$$

with $\psi_1 = \|h_1\|_{\mathcal{V}^1}$, and, for all $m \geq 2$

$$\psi_m = \sum_{k=2}^{m}\left(\mathcal{P}_k \sum_{p\in\mathbb{M}_m^k} \prod_{\ell=1}^{k} \psi_{p_\ell} + \mathcal{Q}_k \sum_{\begin{cases} q\in\mathbb{M}_m^k \\ q_k=1\end{cases}} \prod_{\ell=1}^{k-1} \psi_{q_\ell}\right) \tag{15}$$

where $\mathcal{P}_k$ and $\mathcal{Q}_k$ are given in Definition 5.

Indeed, following Remark 3, h_1 belongs to $\mathcal{V}^1$ and (14) is satisfied for $m = 1$. Now, by induction, let $m \geq 2$ and assume that for $1 \leq m' \leq m-1$, $h_{m'} \in \mathcal{V}^{m'}$ and $\|h_{m'}\|_{\mathcal{V}^{m'}} \leq \psi_{m'}$. Let $t \in \mathbb{T}$. Then, from Proposition 1 and denoting $\overline{\tau} = \tau_1 + \tau_2 + \cdots + \tau_N$

$$\int_{[0,t]^m} \|h_m(t,\tau)\|_{\mathbb{X}}\,\mathrm{d}\tau \leq \int_{[0,t]^m} \left\|\int_{\max\tau}^{t} v_m(t,\theta,\tau)\,\mathrm{d}\theta + w_m(t,\tau)\right\|_{\mathbb{X}} \mathrm{d}\tau \leq \sum_{k=2}^{m}\left(\sum_{p\in\mathbb{M}_m^k} \mathcal{A}_p(t) + \sum_{\begin{cases} q\in\mathbb{M}_m^k \\ q_k=1\end{cases}} \mathcal{B}_q(t)\right) \tag{16}$$

with, for all $p \in \mathbb{M}_m^k$ and $q \in \mathbb{M}_m^k$ such that $q_k = 1$

$$\mathcal{A}_p(t) = \int_{[0,t]^m} \int_{\max \tau}^{t} \widetilde{\mathcal{A}}_p(t,\theta,\tau)\,\mathrm{d}\theta\,\mathrm{d}\tau \tag{17}$$

$$\widetilde{\mathcal{A}}_p(t,\theta,\tau) = \Big\| \mathrm{e}^{A(t-\theta)} P_k\Big(h_{p_1}\big(\theta, S_p^1(\tau)\big),\ldots, h_{p_k}\big(\theta, S_p^k(\tau)\big)\Big)\Big\|_{\mathbb{X}} \tag{18}$$

$$\mathcal{B}_q(t) = \int_{[0,t]^m} \widetilde{\mathcal{B}}_q(t,\tau)\,\mathrm{d}\tau \tag{19}$$

$$\widetilde{\mathcal{B}}_q(t,\tau) = \Big\| \mathrm{e}^{A(t-\tau_m)} Q_k\Big(h_{q_1}\big(\tau_m, S_q^1(\tau)\big),\ldots, h_{q_{k-1}}\big(\tau_m, S_q^{k-1}(\tau)\big),1\Big)\Big\|_{\mathbb{X}}. \tag{20}$$

Now, for $2 \le k \le m$, $p \in \mathbb{M}_m^k$, $\theta \in [0,t]$, and $\tau \in [0,t]^m$

$$\begin{aligned}\widetilde{\mathcal{A}}_p(t,\theta,\tau) &\le \big\|\mathrm{e}^{A(t-\theta)}\big\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \Big\| P_k\Big(h_{p_1}\big(\theta, S_p^1(\tau)\big),\ldots, h_{p_k}\big(\theta, S_p^k(\tau)\big)\Big)\Big\|_{\mathbb{X}} \\ &\le \big\|\mathrm{e}^{A(t-\theta)}\big\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \big\|P_k\big\|_{\mathcal{ML}_k(\mathbb{X},\mathbb{X})} \times \prod_{\ell=1}^{k} \Big\| h_{p_\ell}\big(\theta, S_p^\ell(\tau)\big)\Big\|_{\mathbb{X}}.\end{aligned} \tag{21}$$

Moreover, for $1 \le \ell \le k$ and $\theta \in [0,t]$

$$\begin{aligned}\int_{[0,t]^{p_\ell}} \big\|h_{p_\ell}(\theta,\eta)\big\|_{\mathbb{X}}\mathrm{d}\eta &\le \int_{\mathbb{T}^{p_\ell}} \|h_{p_\ell}(\theta,\eta)\|_{\mathbb{X}}\mathrm{d}\eta \\ &\le \sup_{\theta\in\mathbb{T}} \int_{\mathbb{T}^{p_\ell}} \|h_{p_\ell}(\theta,\eta)\|_{\mathbb{X}}\,\mathrm{d}\eta \\ &\le \|h_{p_\ell}\|_{\mathcal{V}^{p_\ell}}\,.\end{aligned} \tag{22}$$

Hence, from (17)–(22) and since

$$\sup_{t\in\mathbb{T}} \int_0^t \big\|\mathrm{e}^{A(t-\theta)}\big\|_{\mathcal{L}(\mathbb{X},\mathbb{X})}\mathrm{d}\theta = \kappa$$

$$\mathcal{A}_p(t) \le \mathcal{P}_k \prod_{\ell=1}^{k} \|h_{p_\ell}\|_{\mathcal{V}^{p_\ell}} \le \mathcal{P}_k \prod_{\ell=1}^{k} \psi_{p_\ell} \tag{23}$$

is finite.

For $2 \le k \le m$, $q \in \mathbb{M}_m^k$ and $\tau \in [0,t]^m$

$$\begin{aligned}\widetilde{\mathcal{B}}_q(t,\tau) &\le \big\|\mathrm{e}^{A(t-\tau_m)}\big\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \Big\| Q_k\Big(h_{q_1}\big(\tau_m, S_q^1(\tau)\big),\ldots, h_{q_{k-1}}\big(\tau_m, S_q^{k-1}(\tau)\big),1\Big)\Big\|_{\mathbb{X}} \\ &\le \big\|\mathrm{e}^{A(t-\tau_m)}\big\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \big\|Q_k\big\|_{\mathcal{ML}_{k-1,1}(\mathbb{X},\mathbb{R},\mathbb{X})} \times \prod_{\ell=1}^{k-1} \Big\| h_{q_\ell}\big(\tau_m, S_q^\ell(\tau)\big)\Big\|_{\mathbb{X}}.\end{aligned}$$

Hence, from (19)–(20) and using (22) with $\theta = \tau_m \in [0,t]$ and q_ℓ instead of p_ℓ

$$\mathcal{B}_q(t) \le \mathcal{Q}_k \prod_{\ell=1}^{k-1} \|h_{q_\ell}\|_{\mathcal{V}^{q_\ell}} \le \mathcal{Q}_k \prod_{\ell=1}^{k-1} \psi_{q_\ell} \tag{24}$$

is finite. From (16), (23), and (24), h_m is in $\mathcal{V}^m$ and (14) holds.

Step 2: Consider the formal series $\Psi(X) = \sum_{m\in\mathbb{N}^*} \psi_m X^m$ and $\mathcal{R}(X) = \mathcal{P}\big(\Psi(X)\big) + X\mathcal{Q}\big(\Psi(X)\big)$, with $\mathcal{P}(X) = \sum_{k=2}^{\infty} \mathcal{P}_k X^k$ and $\mathcal{Q}(X) = \sum_{k=2}^{\infty} \mathcal{Q}_k X^{k-1}$. Then, from (15)

$$\begin{aligned}\mathcal{R}(X) &= \sum_{k=2}^{\infty}\left(\mathcal{P}_k\Big(\sum_{m\in\mathbb{N}^*}\psi_m X^m\Big)^k + X\mathcal{Q}_k\Big(\sum_{m\in\mathbb{N}^*}\psi_m X^m\Big)^{k-1}\right) \\ &= \sum_{m=2}^{\infty}\sum_{k=2}^{m}\left(\mathcal{P}_k \sum_{p\in\mathbb{M}_m^k}\prod_{\ell=1}^{k}\psi_{p_\ell} + \mathcal{Q}_k \sum_{\left\{\substack{q\in\mathbb{M}_m^k\\ q_k=1}\right.}\prod_{\ell=1}^{k-1}\psi_{q_\ell}\right)X^m \\ &= \sum_{m=2}^{\infty}\psi_m X^m = \Psi(X) - \psi_1 X.\end{aligned}$$

Therefore, $\Psi(X) - \mathcal{P}\big(\Psi(X)\big) = X\left(\mathcal{Q}\big(\Psi(X)\big) + \psi_1\right)$. Since $\psi_1 = \|h_1\|_{\mathcal{V}^1}$, it follows that $\Psi(X) = X\,\mathcal{F}\big(\Psi(X)\big)$, where $\mathcal{F}$ is defined by (9).

Step 3: From (14), Ψ is a majorizing series of the gain bound function φ. Moreover, from Step 2 and Lemma A (in Appendix A), Ψ is analytic at 0, and its convergence radius is given by (12) and (13), which concludes the proof. ■

The following algorithm for the computation of $\rho^\star$ is deduced from Theorem 2.

Algorithm 1 (Computation of $\rho^\star$): The following computation steps can be performed either numerically or analytically.

1) Compute $\|h_1\|_{\mathcal{V}^1}$, $\mathcal{P}_k$, $\mathcal{Q}_k$ and $\mathcal{F}$ (see Definition 5).
2) Compute the positive solution $\sigma \in]0,r[$ of equation $\sigma\,\mathcal{F}'(\sigma) - \mathcal{F}(\sigma) = 0$ if any.
3) Compute $\rho^\star$ using (12) and (13).

From Theorems 1 and 2, the convergence of the Volterra series expansion with the kernels given in Proposition 1 is guaranteed if $u \in \mathcal{D}^\star$, where

$$\mathcal{D}^\star = \{u \in \mathcal{U} \mid \|u\|_{\mathcal{U}} < \rho^\star\}. \tag{25}$$

Remark 4: In practice, most systems fall into case 1. The class of systems with entire nonlinearities P, Q falls into case 1 except if $P=0$ and $Q=Q_1$ ($\mathcal{F}$ is affine), which falls into case 2. Another example for case 2 is obtained for system $\dot{x} = -x/2 + u + (1/2)\big(-x+2x^2-(x-1)^2\ln(1-x)\big)u$. In this case, for $\mathbb{T} = \mathbb{R}_+$, $\|h_1\|_{\mathcal{V}^1} = \kappa = 2$, $\mathcal{F}(x) = 1 + 2x - (x-1)^2\ln(1-x)/x$ with $r=1$, and $\rho^\star = \lim_{x->1^-} x/\mathcal{F}(x) = 1/3$. Simulations show that this is the stability limit of the system, so $\rho^\star$ is a tight bound.

Remark 5: The bound $\rho^\star$ is not guaranteed to be optimal because of (21), (23), and (24), however optimality is reached in some cases (see examples in Section IV).

Remark 6: When both P and Q are collinear to B, κ can be replaced by $\|h_1\|_{\mathcal{V}^1}/\|B\|_{\mathbb{X}}$ in the proof of Theorem 2 and in the algorithm that provides more accurate bounds.

Remark 7: In the case where exact computations of $\|h_1\|_{\mathcal{V}^1}$, $\mathcal{P}_k$, or $\mathcal{Q}_k$ are not possible, upper bounds of these coefficients can be used in Algorithm 1. Indeed, the induction in Step 1 remains valid when using such upper bounds. However, obviously, this leads to underestimated values for $\rho^\star$.

Remark 8: If $Q=0$, the following simplifications occur.

- $\mathcal{F}$ has the form

$$\mathcal{F}(X) = \frac{\|h_1\|_{\mathcal{V}^1}}{1 - \kappa\widetilde{\mathcal{P}}(X)} \tag{26}$$

with $\widetilde{\mathcal{P}}(X) = \mathcal{P}(X)/\kappa = \sum_{k=2}^{\infty} \big\|P_k\big\|_{\mathcal{ML}_k(\mathbb{X},\mathbb{X})} X^{k-1}$,

- the equation satisfied by σ (case 1) becomes

$$\widetilde{\mathcal{P}}(X) + X\,\widetilde{\mathcal{P}}'(X) - \frac{1}{\kappa} = 0. \tag{27}$$

- $\rho^\star$ is given by

$$\rho^\star = \frac{\sigma\big(1 - \kappa\widetilde{\mathcal{P}}(\sigma)\big)}{\|h_1\|_{\mathcal{V}^1}}. \tag{28}$$

Moreover, since $\|h_1\|_{\mathcal{V}^1} \leq \kappa\|B\|_{\mathcal{L}(\mathbb{X},\mathbb{R})}$, an underestimated bound of $\rho^\star$ is given by $\rho^\star_- = \sigma\big(1/\kappa - \widetilde{\mathcal{P}}(\sigma)\big)/\|B\|_{\mathcal{L}(\mathbb{X},\mathbb{R})}$.

The Volterra series is a bounded solution of (1)–(5), as stated in the following theorem.

Theorem 3 (Bounded Solution): For all $u \in \mathcal{D}^\star$, the Volterra series with kernels $\{h_m\}_{m\in\mathbb{N}^*}$ converges to a bounded solution of (1)–(5) in the sense of definition 1.

The proof is detailed in Appendix B.

Finally, a $\mathcal{X}$-norm bound of the truncation error of the series can be computed.

Theorem 4 (Truncation Error Bound): For all $M \in \mathbb{N}^*$, denote $V_M x(t) = \sum_{m=1}^{M} \int_{[0,t]^m} h_m(t,\tau)\,[\Pi_m u](\tau)\,\mathrm{d}\tau$ and $R_M\Psi(X) = \sum_{m=M+1}^{+\infty} \psi_m X^m$. Then, for all $u \in \mathcal{D}^\star$, the remainder of the truncated Volterra series satisfies

$$\big\|x - V_M x\big\|_{\mathcal{X}} \leq R_M\Psi\big(\|u\|_{\mathcal{U}}\big).$$

Moreover, in case 1 of Theorem 2

$$R_M\Psi\big(\|u\|_{\mathcal{U}}\big) \leq \sigma\,\frac{\left(\frac{\|u\|_{\mathcal{U}}}{\rho^\star}\right)^{M+1}}{1 - \frac{\|u\|_{\mathcal{U}}}{\rho^\star}}. \tag{29}$$

Proof: The gain bound function φ is dominated by the generating function Ψ (see the proof of Theorem 2, Steps 2–3). Then, for all $u \in \mathcal{D}^\star$

$$\|x - V_M x\|_{\mathcal{X}} \leq \sum_{m=M+1}^{+\infty} \|h_m\|_{\mathcal{V}^m}\big(\|u\|_{\mathcal{U}}\big)^m \leq R_M\Psi(\|u\|_{\mathcal{U}}) < +\infty.$$

Moreover, in case 1, the restriction of Ψ to $[0,\rho^\star[$ is a positive strictly increasing bijection from $[0,\rho^\star[$ to $[0,\sigma[$ with positive Taylor coefficients (see Lemma A in Appendix A). Function Ψ is normally convergent on any closed disk $D(\rho) \subset \mathbb{C}$ with radius $\rho < \rho^\star$. Hence, Cauchy estimates on $D(\rho)$ yield $\psi_m = \Psi^{(m)}(0)/m! \leq \sup_{z\in D(\rho)} |\Psi(z)|/\rho^m \leq \sigma/\rho^m, \forall m \in \mathbb{N}^*$. For $\rho \to \rho^\star$, the limit leads to $\psi_m \leq \sigma/(\rho^\star)^m$. Finally, Cauchy estimates for Ψ yield (29), which concludes the proof. ■

IV. Examples

In this section, several examples are presented, for which analytic computations are possible.

A. 1-D System With Third-Order Nonlinearity

We start with a very simple 1-D toy example with a polynomial nonlinearity. Let $a \in]0,1[$, $\varepsilon \in \mathbb{R}$, and consider the following system:

$$\forall t > 0, \qquad \dot{x} = -a\,x + \varepsilon x^3 + u \tag{30}$$

with zero initial conditions $x(0) = 0$ and scalar bounded signal input u. It has the form (1)–(5) with $\mathbb{X} = \mathbb{R}$, $A = -a$, $B = 1$, $P(x) = \varepsilon x^3$, $Q(x,u) = 0$. Following the steps of Algorithm 1, we compute the convergence radius of the Volterra series for an infinite time horizon ($\mathbb{T} = \mathbb{R}_+$).

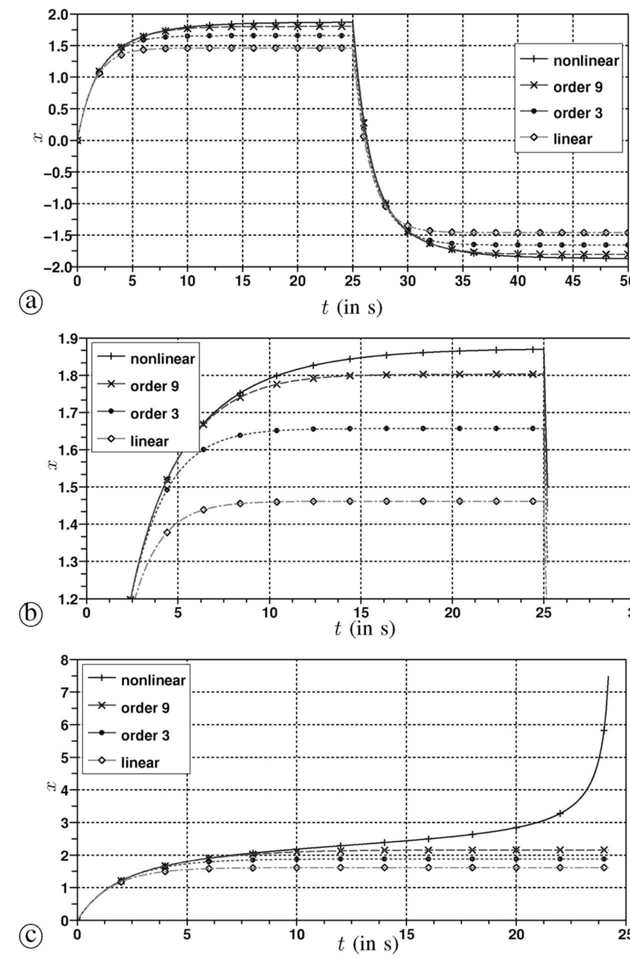

Fig. 1. (Example IV-A, case $\varepsilon > 0$) Numerical computation of x for $a = 0.65$, $\varepsilon = 0.04068$ with $e = 0.95 < \rho^\star = 1$ in (a) and (b) and $e = 1.05 > \rho^\star$ in (c). (b) is a zoom of (a).

Step 1) $\|h_1\|_{\mathcal{V}^1} = 1/a$, $\kappa = 1/a$, $\mathcal{P}_k = 0$ for all $k \in \mathbb{N}^*$, except for $k = 3$, where $\mathcal{P}_3 = \kappa|\varepsilon|$, and $\mathcal{Q}_k = 0$ for all $k \in \mathbb{N}^*$, so $\mathcal{F}$ is given by (26) with

$$\widetilde{\mathcal{P}}(X) = |\varepsilon|\,X^2.$$

Step 2) σ satisfies (27), that is, $3|\varepsilon|\,X^2 - a = 0$. The unique positive solution is

$$\sigma = \sqrt{\frac{a}{3|\varepsilon|}}.$$

Step 3) We compute $\rho^\star$, e.g., using (28), which leads to

$$\rho^\star = a\sigma\Big(1 - \frac{|\varepsilon|\sigma^2}{a}\Big) = \frac{2}{3}\sqrt{\frac{a^3}{3|\varepsilon|}}.$$

Numerical simulations are performed with $a = 0.65$ and $\varepsilon = \pm 0.04068$ so that $\rho^\star \approx 1$. The input $u(t)$ is constant and equal to e on [0, 25] and jumps to $-e$ on [25, 50].

1) Case $\varepsilon > 0$ (Fig. 1): Simulations show that $\rho^\star$ is indeed the radius of convergence of the Volterra series: For $e < 1$, it converges to the trajectory of the nonlinear system [see

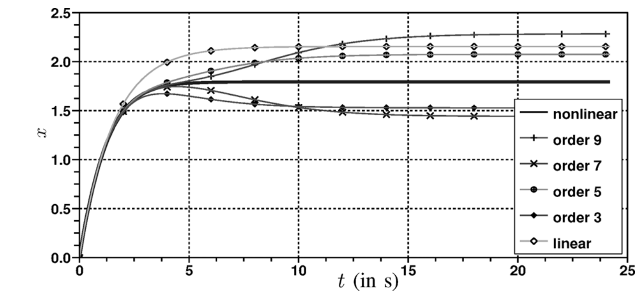

Fig. 2. (Example IV-A, case $\varepsilon < 0$) Numerical computation of x for $a = 0.65$, $\varepsilon = -0.04068$ with $e = 1.4 > \rho^\star$. A high value of e is chosen in order to see the otherwise slow divergent behavior of the Volterra series expansion.

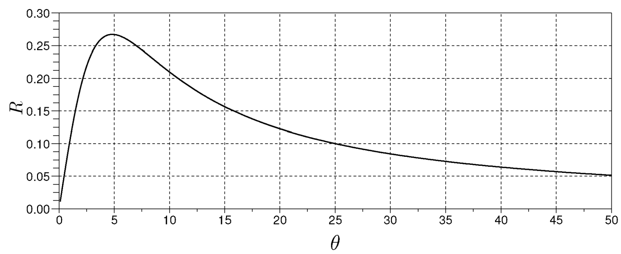

Fig. 3. (Example IV-A) Convergence radius proposed in [7] is a function of a parameter θ (denoted ρ in [7]). The best estimate is obtained here at $\theta \approx 4.8$.

Fig. 1(a) and (b)], whereas for $e > 1$, it becomes divergent [see Fig. 1(c)]. In Fig. 1(c), it can also be seen that $\rho^\star$ is the BIBO stability limit of the system at 0 since the trajectory of the nonlinear system is not bounded any more.

2) Case $\varepsilon < 0$ (Fig. 2) : As pointed out in [6], considering (30) for complex parameters, ε, a (with $\Re\mathrm{e}(a) > 0$) or complex signals u, the convergence radius is also a tight bound. Numerical simulations show that the Volterra series expansion diverges very slowly from the nonlinear system for $e > \rho^\star$ (see Fig. 2), which shows once again that the bound obtained for the radius of convergence is tight. However, the nonlinear system restricted to real valued signals is BIBO stable at 0 for any bounded input. Hence, for this type of system, Volterra series expansion convergence radius does not coincide with the BIBO stability limit of the system at 0.

3) Comparison to Other Results: Most convergence results on Volterra series require an explicit estimation of the asymptotic behavior of the kernel norms. This is not easy except for particular systems (exact computations are available for bilinear and quadratic systems [6], [7], [16], and our algorithm provides similar results). For systems (1)–(5) such that $Q = 0$, a convergence condition has been proposed in [7, Theorem 3.1], which is given by (using our notations) $\exists\theta > 0$, $\|u\|_{\mathcal{U}} < R(\theta) = \frac{\theta}{n\kappa} + 2\nu(\theta)\Big(1 - \sqrt{1 + \frac{\theta}{n\,\kappa\,\nu(\theta)}}\Big)$, where $n = \dim \mathbb{X}$ and $\nu(\theta) = \sup_{\substack{z\in\mathbb{C}\\ |z|<\theta}} \|Ax + P(x)\|_{\mathbb{X}}$. Here, $\nu(\theta) = a\theta + |\varepsilon|\theta^3$ and

$$R(\theta) = a\theta\left[1 + 2\Big(1 + \frac{|\varepsilon|}{a}\theta^2\Big)\Big(1 - \sqrt{1 - \Big(1 + \frac{|\varepsilon|}{a}\theta^2\Big)^{-1}}\Big)\right]$$

reaches its maximal value $R^\star \approx 0.267$ at $\theta \approx 4.8$ (see Fig. 3). Although comparisons between $\theta \mapsto R(\theta)$ and $\rho^\star$ are not presented for the following examples, it has been checked that $R^\star$ (here, 0.267) is lower than $\rho^\star$ (here, 1).

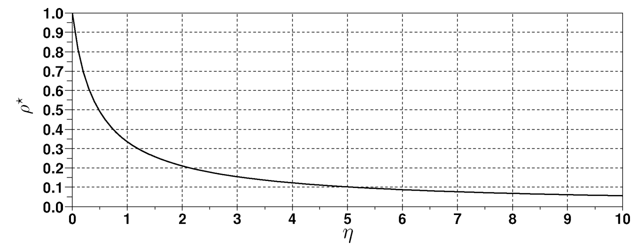

Fig. 4. (Example IV-B) Radius of convergence $\rho^\star$ as a function of η for $\eta \in [0, 1]$ and $a = 0.65$ and $|\varepsilon| = 0.04068$.

B. Similar Example With a Nonzero Contribution Q

Consider the slightly complexified version of the previous example given by

$$\forall t > 0, \qquad \dot{x} = -a\,x + \varepsilon x^3 + \eta x\,u + u \tag{31}$$

which has a nonzero Q, given by $Q(x, u) = \eta x u$ with $\eta \in \mathbb{R}$. Computations in Algorithm 1 are modified as follows.

Step 1) $\mathcal{Q}_2 = \kappa|\eta|$ so $\mathcal{Q}(X) = \kappa|\eta|X$ and $\mathcal{F}$ is given by (9) with

$$\mathcal{F}(X) = \frac{\|h_1\|_{\mathcal{V}^1} + \kappa|\eta|\,X}{1 - \kappa|\varepsilon|\,X^2} = \frac{1 + |\eta|\,X}{a - |\varepsilon|\,X^2}.$$

Step 2) σ (computed numerically) is the unique positive solution of $\sigma\,\mathcal{F}'(\sigma) - \mathcal{F}(\sigma) = 0$, that is, of $a - 3|\varepsilon|X^2 - 2|\varepsilon\eta|X^3 = 0$.

Step 3) We numerically compute $\rho^\star = \sigma/\mathcal{F}(\sigma)$.

The radius of convergence is plotted in Fig. 4 for $\eta \in [0, 1]$ (values for η and $-\eta$ are equal) and the same parameters as in Section IV-A ($a = 0.65$ and $|\varepsilon| = 0.04068$). As expected, the convergence radius is 1 for $\eta = 0$ and decreases for $\eta > 0$. Moreover, for $\eta \to +\infty$, $\sigma \sim \sqrt[3]{a/|2\varepsilon\eta|}$ and $\rho^\star \sim a/\eta$.

C. Saturated Stabilization

Consider here the following 1-D system:

$$\dot{x} = ax - \tanh(\gamma x) + u \tag{32}$$

where $\gamma > a > 0$ and the initial state $x(0) = 0$. It is representative of the frequently encountered situation where an unstable linear system ($\dot{x} = ax + u$) is stabilized by a linear state feedback (γx), but the stabilizing signal is saturated by a $\tanh$ static function.

For constant causal inputs $u(t) = U \in \mathbb{R}$, the stability analysis shows that the system is stable if $|U| < U_{\lim}$ with

$$U_{\lim} = \sqrt{1-\xi} - \xi\,\mathrm{atanh}\sqrt{1-\xi}, \qquad \text{with } \xi = \frac{a}{\gamma}$$

and is unstable if $|U| > U_{\lim}$.

The Taylor expansion of the saturating function around $x = 0$ is given by $\tanh y = \sum_{n\in\mathbb{N}^*} c_{2k-1} y^{2k-1}$ with $c_1 = 1$, $c_3 = -1/3$, $c_5 = 2/15$, and more generally, $c_{2k-1} = (-1)^{k+1} 2^{2k}(2^{2k} - 1)B_{2k}/(2k)!$, where the B_{2k}'s denote the Bernoulli numbers [1, (4.5.64)]. For an infinite time horizon ($\mathbb{T} = \mathbb{R}_+$), we have the following.

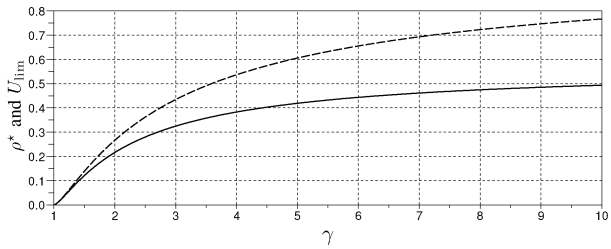

Fig. 5. (Example IV-C) $\rho^\star$ (solid line) and U_{lim} (dashed line) as functions of γ.

Step 1) $\kappa = \|h_1\|_{\mathcal{V}^1} = 1/(\gamma - a)$, $\mathcal{Q}_k = 0$ for all $k \in \mathbb{N}^*$ and $\mathcal{P}_k = \kappa |c_k| \gamma^k$. Since $\sum_{n\in\mathbb{N}^*} |c_{2k-1}| y^{2k-1} = \tan y$, it follows that $\widetilde{\mathcal{P}}(X) = \big(\tan(\gamma X) - \gamma X\big)/X$ and that the gain bound function is given by

$$\mathcal{F}(X) = \frac{X}{(2\gamma - a)X - \tan(\gamma X)}.$$

Step 2) σ satisfies (27), that is, $\gamma - a - \tan^2(\gamma X) = 0$. The unique positive root of this equation is

$$\sigma = \frac{1}{\gamma} \arctan\left(\sqrt{\frac{\gamma - a}{\gamma}}\right).$$

Step 3) Using (28), we obtain

$$\rho^\star = \Big[(2-\xi)\arctan\sqrt{1-\xi}\Big] - \sqrt{1-\xi}, \qquad \text{with } \xi = \frac{a}{\gamma}.$$

Fig. 5 displays the values of U_{lim} and $\rho^\star$ for $a = 1$, as a function of the feedback gain γ. It shows that for moderate values of γ, $\rho^\star$, and U_{lim} are close. For constant inputs greater than $\rho^\star$, we checked numerically that the Volterra series is slowly divergent. Nevertheless, interesting approximations are given by the series truncated at order N such that $\psi_N U^N$ is the smallest term of the Taylor expansion of the dominant function $\Psi(U)$. Further investigations on divergent series will be performed in a future work.

Fig. 6 shows a simulation obtained with a constant input of $0.8\rho^\star$ during 10 s, followed by a jump to a sinusoidal input centered on $0.5\rho^\star$, with an amplitude of $0.5\rho^\star$ and a frequency of 0.8 Hz until $t = 20$ s. From $t = 20$ s to $t = 30$ s, the input is a decreasing ramp to zero. From 0 to 10 s, the input is a low frequency (constant) signal, close to the convergence bound of the series. The linear approximation is very poor, the third-order one is much better, and the series truncated at order 5 give extremely good results. Between 10 and 20 s, the signal is set on a higher frequency than the bandwidth of the underlying linear system, so that, due to the low pass effect, the linear approximation is acceptable and the third-order one is very good, although the input reaches the convergence bound in infinite norm. The end of the simulation corresponds to an input with decreasing amplitude, so that the linear approximation becomes better at the end of the simulation, and the third-order one is quite accurate.

D. Two Second-Order Systems

1) System With a Nonlinear Damping: Let $a > 0$, $b \in \mathbb{R}$, $\omega > 0$, and consider the following system:

$$\forall t > 0, \qquad \ddot{x}_1 + 2a\,\dot{x}_1 + b\,\dot{x}_1^3 + \omega^2 x_1 = u \qquad (33)$$

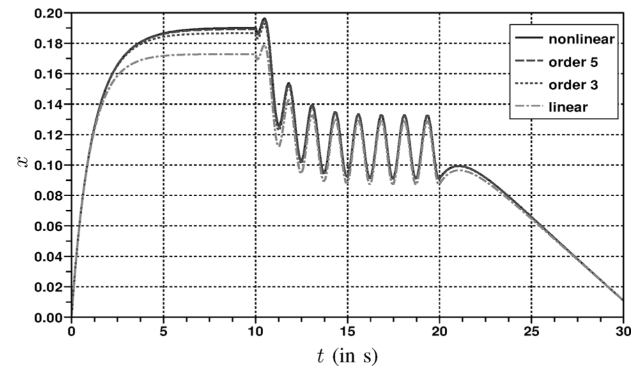

Fig. 6. (Example IV-C) Simulation for a dynamic input.

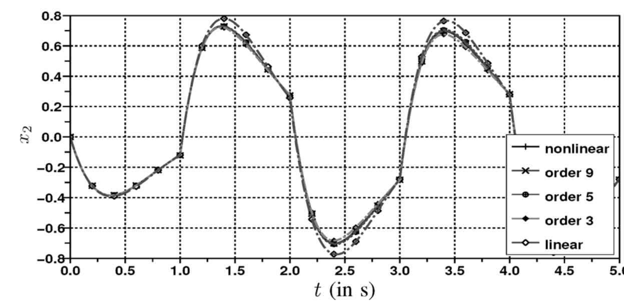

Fig. 7. Example IV-D1 : Numerical computation of the speed x_2 for $a = 2$, $b = 1$, with $e = \rho^\star$.

with zero initial conditions and scalar bounded signal input u. The nonlinearity corresponds to a damping if $b > 0$. It takes the form (1)–(5), where $\mathbb{X} = \mathbb{R}^2$ is associated with the euclidean norm. The state is $x = (x_1, x_2)^T$ and $A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2a \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $P(x) = \begin{pmatrix} 0 \\ -b x_2^3 \end{pmatrix} = -b x_2^3 B$, $Q(x,u) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Note that $\max\Big(\Re\mathrm{e}\big(\operatorname{spec} A\big)\Big) = -a < 0$.

For a given time interval $\mathbb{T} = [0, T]$, the corresponding convergence radius $\rho_T^\star$ is computed as follows.

Step 1) For all $k \in \mathbb{N}^*$, $\mathcal{Q}_k = 0$, $\mathcal{P}_k = 0$ except $\mathcal{P}_3 = \kappa\,|b|$, so $\mathcal{F}$ is given by (26) with $\widetilde{\mathcal{P}}(X) = |b| X^2$.

Step 2) σ satisfies (27), that is, $3\,|b|\,X^2 - 1/\kappa = 0$. The unique positive solution is $\sigma = 1/\sqrt{3\kappa\,|b|}$.

Step 3) Computing $\rho^\star$, e.g., using (28), leads to

$$\rho^\star = \frac{1}{\left(\sqrt{3\kappa\,|b|}\;\|h_1\|_{\mathcal{V}^1}\right)}.$$

Moreover, in this example and following Remark 6, κ can be replaced by $\|h_1\|_{\mathcal{V}^1}/1$ in the algorithm. For each $\mathbb{T}$, the value of $\|h_1\|_{\mathcal{V}^1}$ is computed numerically.

Fig. 7 displays a numerical simulation with $a = 2$, $b = 1$, $\omega = 3$, and $T \to +\infty$ so that $\rho^\star \approx 2.5$. The input $u(t)$ switches every time unit from e to $-e$, with $e = \rho^\star$. The system is adequately approximated by the series truncated at order 5. For higher values of e, the Voterra series expansion becomes slowly divergent.

2) Damped Pendulum: Let $a \in]0, 1[$, $\omega > 0$, and consider the system zero initial conditions and input u governed by

$$\forall t > 0, \qquad \ddot{x}_1 + 2a\dot{x}_1 + \omega^2 \sin(x_1) = u. \qquad (34)$$

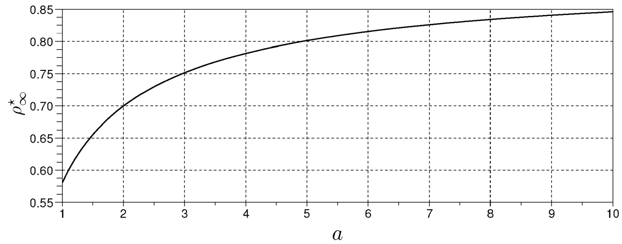

Fig. 8. (Example IV-D2) $\rho^\star_\infty$ as a function of a.

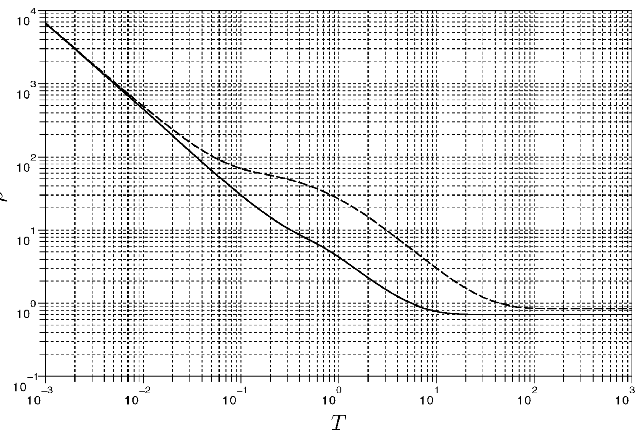

Fig. 9. (Example IV-D2) Function $T \mapsto \rho^\star$ for a damping $a = 2$ (solid line) and $a = 10$ (dashed line). The convergence radius for $T \to +\infty$ ($\mathbb{T} = \mathbb{R}_+$) is $\rho^\star \approx 0.7$ for $a = 2$ and $\rho^\star \approx 0.85$ for $a = 10$.

Like in the previous example, the state is $x = (x_1, x_2)^T$ with $A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2a \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $P(x) = \begin{pmatrix} 0 \\ -\omega^2(\sin x_1 - x_1) \end{pmatrix} = -\omega^2(\sin x_1 - x_1)B$ and $Q(x,u) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Note that $\max\left(\Re\mathrm{e}\big(\mathrm{spec}\, A\big)\right) = -a < 0$.

In this case, for $\mathbb{T} = [0, T]$, $\rho^\star_T$ is computed as the following.

Step 1) For all $k \in \mathbb{N}^*$, $\mathcal{Q}_k = 0$, $\mathcal{P}_{2k} = 0$, and $\mathcal{P}_{2k+1} = \kappa\omega^2/(2k+1)!$, so $\mathcal{F}$ is given by (26) with

$$\widetilde{\mathcal{P}}(X) = \frac{\omega^2(\sinh X - X)}{X}.$$

Step 2) σ satisfies (27), that is, $\omega^2 \cosh X - \omega^2 - 1/\kappa = 0$. The unique positive solution is

$$\sigma = \ln\left(1 + \frac{1 + \sqrt{1 + 2\kappa\omega^2}}{\kappa\omega^2}\right).$$

Step 3) Computing $\rho^\star$, e.g., using (28), leads to

$$\rho^\star = \frac{\sigma - \kappa\omega^2\big(\sinh\sigma - \sigma\big)}{\|h_1\|_{\mathcal{V}^1}}.$$

Again (see Remark 6), κ can be replaced by $\|h_1\|_{\mathcal{V}^1}/1$ in the algorithm. The stability limit for constant inputs of the nonlinear system is $U_{\lim} = 1$. Fig. 8 displays the values of $\rho^\star_\infty$ as a function of the damping a, and Fig. 9 the values of $\rho^\star_T$ as a function of T. As it could be expected, $\rho^\star$ is an increasing function of a and a decreasing function of T.

In Fig. 10, a simulation is performed for $a = 2$ and a constant input $u = \rho^\star_T$ with $T = 1.5$ ($\rho^\star_3 \approx 3$). The angular position

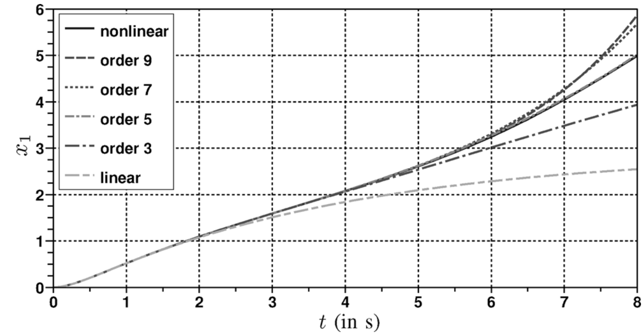

Fig. 10. (Example IV-D2) Simulation for $a = 2, T = 1.5$, and a constant input $e = \rho^\star_T \approx 3$. The angular position of the nonlinear system (solid line) and those predicted by the successive Volterra series expansions are plotted.

$x_1(t)$ is a slowly increasing function of time since $\rho^\star_T > U_{\lim}$. It is plotted for the system and for its Volterra series expansions up to order 9. As guaranteed by Theorem 2, the convergence is achieved if $t < T = 1.5$ (here, a first-order expansion proves accurate on $\mathbb{T}$). However, the convergence is not guaranteed for $t > T$, so the series is not representative of the behavior of the system in this case. As an indicator of divergence, the smallest term of the series expansion for orders $m \in \{1, 3, 5, 7, 9\}$ corresponds to $m = 9$ for $t < 9.23$, $m = 7$ for $9.23 < t < 11.4$, and $m = 1$ for $t > 11.4$.

V. Generalization to Exponentially Damped Input–Output Results

In this section, we concentrate on results on the infinite interval $\mathbb{T} = \mathbb{R}_+$.

A. Definitions and Convergence Results

Definition 6 (Spaces $\mathcal{U}_\lambda$, $\mathcal{X}_\lambda$, $\mathcal{V}^m_{\lambda,\mu}$, and $\mathcal{VS}_{\lambda,\mu}$): For all λ, μ in $\mathbb{R}_+$ and $m \in \mathbb{N}^*$, we introduce the following sets.

- $\mathcal{X}_\lambda$ is the set of functions x such that $t \mapsto \mathrm{e}^{\lambda t} x(t) \in L^\infty\big(\mathbb{R}_+, \mathbb{X}\big)$ endowed with the norm

$$\|x\|_{\mathcal{X}_\lambda} = \sup_{t\in\mathbb{R}_+}\left(\mathrm{e}^{\lambda t}\,\|x(t)\|_{\mathbb{X}}\right).$$

  Note that if $\lambda_2 > \lambda_1 \geq 0$, then $\mathcal{X}_{\lambda_2} \subset \mathcal{X}_{\lambda_1} \subseteq \mathcal{X}_0$.
- $\mathcal{U}_\lambda$ is defined in the same way as $\mathcal{X}_\lambda$, replacing $\mathbb{X}$ by $\mathbb{R}$ and $\|\cdot\|_{\mathbb{X}}$ by $|\cdot|$.
- $\mathcal{V}^m_{\lambda,\mu}$ is the set of functions $f : \mathbb{R}_+ \times \mathbb{R}^m_+ \to \mathbb{X}$ such that

$$t \mapsto \big(\tau \mapsto \mathrm{e}^{\mu t - \lambda\overline{\tau}} f(t,\tau)\big) \in L^\infty\big(\mathbb{R}_+, L^1\big(\mathbb{R}^m_+, \mathbb{X}\big)\big)$$

  where $\forall \tau = (\tau_1, \ldots, \tau_m) \in \mathbb{R}^m_+$, $\overline{\tau} = \tau_1 + \tau_2 + \cdots + \tau_m$. This set is endowed with the norm defined $\|f\|_{\mathcal{V}^m_{\lambda,\mu}} = \sup_{t\in\mathbb{R}_+}\left(\mathrm{e}^{\mu t}\int_{\mathbb{R}^m_+}\|f(t,\tau)\|_{\mathbb{X}}\,\mathrm{e}^{-\lambda\overline{\tau}}\mathrm{d}\tau\right)$.
- $\mathcal{VS}_{\lambda,\mu}$ is the set of the series $(f_m)_{m\in\mathbb{N}^*}$ such that for all $m \in \mathbb{N}^*$, $f_m \in \mathcal{V}^m_{\lambda,\mu}$.

Proposition 2 (Coefficients $\kappa_{k,\mu}$ and Norm of h_1): Let the system (1)–(5) and $\beta > 0$ be such that $-a = \max\big(\Re\mathrm{e}\big(\mathrm{spec}\, A\big)\big) < 0$ and $\forall t \in \mathbb{R}_+$, $\left\|\mathrm{e}^{At}\right\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \leq \beta\,\mathrm{e}^{-at}$.

Then, for all $\mu \in [0, a)$, $\lambda \geq \mu$, the coefficients defined by, for all $k \in \mathbb{N}^*$

$$\kappa_{k,\mu} = \sup_{t\in\mathbb{R}_+}\left(\mathrm{e}^{\mu t}\int_0^t \left\|\mathrm{e}^{A(t-\theta)}\right\|_{\mathcal{L}(\mathbb{X},\mathbb{X})}\mathrm{e}^{-k\mu\theta}\,\mathrm{d}\theta\right) \quad (35)$$

are finite and such that

$$\kappa_{k,\mu} \le \beta / \Big((a-\mu)\, \xi\big((k-1)\mu/(a-\mu)\big) \Big),$$

where $\xi : \mathbb{R}_+ \to [1, +\infty)$ is the continuous bijective increasing function defined for all $x \in \mathbb{R}_+$ by

$$\xi(0) = 1 \quad \xi(1) = \mathrm{e}^1$$

and

$$\xi(x) = x^{x/(x-1)} \qquad \text{otherwise.}$$

Moreover, $h_1 \in \mathcal{V}^1_{\lambda,\mu}$ and $\|h_1\|_{\mathcal{V}^1_{\lambda,\mu}} \le \beta \|B\|_{\mathcal{L}(\mathbb{R},\mathbb{X})} / \Big((a-\mu)\, \xi\big((\lambda-\mu)/(a-\mu)\big) \Big)$.

The proof is straightforward. The following function can then be introduced.

Definition 7 (Function $\mathcal{F}_{\lambda,\mu}$): For all $\mu \in [0, a)$ and $\lambda \ge \mu$, the function $\mathcal{F}_{\lambda,\mu}$ is formally defined by

$$\mathcal{F}_{\lambda,\mu}(X) = \frac{\|h_1\|_{\mathcal{V}^1_{\lambda,\mu}} + \sum_{k=2}^{\infty} \mathcal{Q}_{k,\mu}\, X^{k-1}}{1 - \sum_{k=2}^{\infty} \mathcal{P}_{k,\mu}\, X^{k-1}}$$

where, for all $k \in \mathbb{N}^*$, $\mathcal{P}_{k,\mu} = \kappa_{k,\mu} \big\|P_k\big\|_{\mathcal{ML}_k(\mathbb{X},\mathbb{X})}$ and $\mathcal{Q}_{k,\mu} = \kappa_{k,\mu} \big\|Q_k\big\|_{\mathcal{ML}_{k-1,1}(\mathbb{X},\mathbb{R},\mathbb{X})}$.

Remark 9 (Links With Section III): All the definitions of this section correspond to those of Section III for $\lambda = \mu = 0$. More precisely, $\mathcal{U}_0 = \mathcal{U}$, $\mathcal{X}_0 = \mathcal{X}$, $\mathcal{V}^m_{0,0} = \mathcal{V}^m$, $\mathcal{VS}_{0,0} = \mathcal{VS}$, $\kappa_k^{0,0} = \kappa$ (for all k in $\mathbb{N}^*$), and $\mathcal{F}_{0,0} = \mathcal{F}$.

Theorem 5 (Generalization of Theorems 1–4): Let $\mu \in [0, a)$ and $\lambda \ge \mu$. Then, results given in Theorems 1–4 and Algorithm 1 are still available when replacing $\mathcal{U}$, $\mathcal{X}$, $\mathcal{V}^m$, $\mathcal{VS}$, and $\mathcal{F}$ by $\mathcal{U}_\lambda$, $\mathcal{X}_\mu$, $\mathcal{V}^m_{\lambda,\mu}$, $\mathcal{VS}_{\lambda,\mu}$, and $\mathcal{F}_{\lambda,\mu}$, respectively.

The proof is given in Appendix C.

These results indicate that if $0 \le \mu < a$, we can compute a (nonempty) convergence domain $\mathcal{D}_{\lambda,\mu} \subset \mathcal{U}_\lambda \subset \mathcal{U}$ of inputs decreasing at least like $\mathrm{e}^{-\lambda t}$ with $\lambda \ge \mu$ for which the Volterra series converges to a state that is guaranteed to decrease at least like $\mathrm{e}^{-\mu t}$.

Remark 10: Following Remark 7, replacing $\|h_1\|_{\mathcal{V}^1_{\lambda,\mu}}$ and $\kappa_{k,\mu}$ in Algorithm 1 by the overestimated bounds given in Proposition 2 still yields valid results.

B. Example of a 1-D System

Consider a 1-D damped system described by (1)–(5) with $\mathbb{X} = \mathbb{R}$, $A = -a < 0$, $B = 1$, $P(x) = \sum_{k=2}^{\infty} a_k x^k$, $Q = 0$, zero initial conditions. Let $\mu \in [0, a)$, $\lambda \ge \mu$. Moreover, denote $\alpha = a - \mu > 0$, $\xi_{k,\mu} = \xi\big((k-1)\mu/(a-\mu)\big) \ge 1$ ($k \in \mathbb{N}^*$) and $\zeta_{\lambda,\mu} = \xi\big((\lambda-\mu)/(a-\mu)\big) \ge 1$.

For such this system, the bounds given in Proposition 2 with $\beta = 1$ yield the following exact quantities $\|h_1\|_{\mathcal{V}^1_{\lambda,\mu}} = 1/(\alpha\, \zeta_{\lambda,\mu})$, $\kappa_{k,\mu} = 1/(\alpha\, \xi_k)$. Then, $\mathcal{F}_{\lambda,\mu}(X) = \Big(\zeta_{\lambda,\mu}\big(\alpha - \widetilde{\mathcal{P}}_\mu(X)\big)\Big)^{-1}$ with $\widetilde{\mathcal{P}}_\mu(X) = \sum_{k=2}^{+\infty} (|a_k|/\xi_{k,\mu})\, X^{k-1}$, σ_μ is the positive solution of $\widetilde{\mathcal{P}}_\mu(X) + X\, \widetilde{\mathcal{P}}'_\mu(X) - \alpha = 0$ and

$$\rho^\star_{\lambda,\mu} = \zeta_{\lambda,\mu}\, \sigma_\mu \big(\alpha - \widetilde{\mathcal{P}}_\mu(\sigma_\mu)\big).$$

For Example 1 (see Section IV-A), this yields $\widetilde{\mathcal{P}}_\mu(X) = (\varepsilon/\xi_{3,\mu})\, X^2$, so that $\sigma_\mu = \sqrt{\alpha\, \xi_{3,\mu}/(3\,\varepsilon)}$ and

$$\rho^\star_{\lambda,\mu} = \frac{2\,\zeta_{\lambda,\mu}}{3} \sqrt{\frac{\alpha^3\, \xi_{3,\mu}}{3\,\varepsilon}}.$$

The case ($\mu = 0, \lambda \ge 0$) leads to $\alpha = a$, $\xi_{3,0} = 1$, $\zeta_{\lambda,0} = \xi(\lambda/a)$, $\sigma_0 = \sqrt{a/(3\varepsilon)}$ and

$$\rho^\star_{\lambda,0} = \rho^\star_{0,0}\, \xi\Big(\frac{\lambda}{a}\Big) > \rho^\star_{0,0}$$

is an increasing function of λ/a (note also that $\xi(\lambda/a) \underset{\lambda\to+\infty}{\sim} \lambda/a$). In particular, this illustrates that the radius of convergence that ensures bounded states ($\mu = 0$) is known to be greater for damped inputs ($\lambda > 0$) than for bounded inputs ($\lambda = 0$). Hence, the initial value of a damped input can be chosen greater than the convergence radius computed for the BIBO case. For instance, for $\lambda = 2a$, the initial value can be chosen $1/\xi(2) = 4$ times greater than for $\lambda = 0$.

The case ($\mu = a/2, \lambda \ge \mu$) leads to $\alpha = a/2$, $\xi_{3,a/2} = \xi(2) = 4$, $\sigma_{a/2} = \sqrt{2a/(3\varepsilon)} = \sqrt{2}\, \sigma_0$ and

$$\rho^\star_{a/2,\lambda} = \frac{\xi\Big(\frac{2\lambda}{a} - 1\Big)}{\sqrt{2}}\, \rho^\star_{0,0}$$

is an increasing function of λ from $[\mu, +\infty)$ to $[\rho^\star_{0,0}/\sqrt{2}, +\infty)$.

Remark 11: The value λ such that $\rho^\star_{a/2,\lambda} = \rho^\star_{0,0}$ is computed by solving $\sqrt{2}\,\xi(2\lambda/a - 1) = 1$, which yields $\lambda/a \approx 0.58$.

VI. Conclusion

Bounds on the convergence radius and truncation errors of Volterra series expansions have been proposed for SI nonlinear systems that are analytic in state and affine in input. Results have been illustrated on several examples. The main advantages of the method are that: 1) computable bounds are given (rather than only existence or theoretical results); 2) the corresponding algorithm is adapted to both exact and numerical computations; 3) results are available for several norms that are adapted to address stable and unstable systems, bounded or exponentially damped input-to-state results, and finite- or infinite-time horizons; 4) the hypothesis required by the study on an infinite horizon is weak, that is, the system must have a stable linear part. When simulating systems on a finite-time horizon T using Volterra series, our method shows that relaxing the stability condition on the linear part still makes sense for sufficiently small inputs. However, in this case, the convergence radius quickly decreases to 0 with T. The main limitations of the method are that: 1) results are available only for the "analytic linear systems," described in Section II-B; 2) the bound given by the algorithm can yield an underestimated convergence radius [typically, when the accuracy of inequality (21) becomes poor]. Our results bring a useful contribution in all the applications where Volterra series expansion with a guaranteed precision are needed (e.g., simulation and model order reduction). It can also be used for the characterization of stability domains of nonlinear systems, as well as, e.g., the optimization of parameterized stabilizing controllers through the maximization of the convergence radius.

The extension of these results to the multiple-input case is under study. In the near future, we also plan to generalize the above results to systems that are, in addition to the above assumptions, analytic in input and have nonzero initial conditions. Another extension will consist of generalizing these results to some classes of infinite-dimensional systems, such as boundary and distributed controlled PDE systems solved using Volterra series (see, e.g., [15] and [18]).

APPENDIX A
LEMMA A

Let $A(X) = \sum_{k=1}^{+\infty} a_k X^k$ and $B(X) = \sum_{k=1}^{+\infty} b_k X^k$ be analytic functions at $X = 0$ with nonnegative coefficients. Let $\beta \in \mathbb{R}_+^*$. Define $F(X) = (\beta + B(X))/(1 - A(X))$, and let $r \in \mathbb{R}_+^* \cup \{+\infty\}$ be the radius of convergence of F at $x = 0$. Then, the following results hold.

(i) At $x = 0$, F is nonzero and analytic with nonnegative Taylor coefficients.

(ii) Equation $x F'(x) - F(x) = 0$ has either one solution denoted σ (case 1) or zero solution (case 2) in $]0, r[$.

(iii) There exists a unique function $z \mapsto \Psi(z)$, analytic at $z = 0$ such that $\Psi(z) = z F\big(\Psi(z)\big)$. Its convergence radius ρ_Ψ at $z = 0$ is such that

$$\text{(case 1)} \qquad \rho_\Psi = \rho^\star = \frac{\sigma}{F(\sigma)} \tag{36}$$

$$\text{(case 2)} \qquad \rho_\Psi \geq \rho^\star = \lim_{x \to r^-} \frac{x}{F(x)}. \tag{37}$$

Proof:

Assertion (i): If $A = 0$, (i) is straightforward. Otherwise, A has at least one positive Taylor coefficients so that, for all $z \in \mathbb{C}$ such that $|z| < r$, $|A(z)| < A\big(|z|\big) < \lim_{x \to r^-}(x) \leq 1$ and $F(z) = \big(\beta + B(z)\big) \sum_{n=0}^{+\infty} \big(A(z)\big)^n$, which proves (i).

Assertion (ii): Define $H(x) = x F'(x) - F(x)$ for $x \in [0, r[$. If F is affine, then $H(x) = -\beta$ so that $x F'(x) - F(x) = 0$ has no solution. Otherwise, H is a strictly increasing function on $]0, r[$ from $H(0) < 0$ to $\ell = \lim_{x \to r^-} H(x) \in \mathbb{R} \cup \{+\infty\}$ since, for all $x \in]0, r[$, $H'(x) = x F''(x) > 0$. Therefore, if $\ell > 0$, then $x F'(x) - F(x) = 0$ has a unique solution on $[0, r[$ (case 1); otherwise ($\ell \leq 0$), it has no solution (case 2).

Assertion (iii): In case 1, the hypotheses of the singular inversion theorem (see e.g., [12, Proposition IV.5 and Theorem VI.6]) are met, and its application proves (iii). In case 2, (iii) is a direct consequence of the analytic inversion lemma (see, e.g., in [12, Lemma 4.2]). ■

APPENDIX B
PROOF OF THEOREM 3

Proof: Let $u \in \mathcal{D}^\star$. From Theorem 2, the Volterra series (6), denoted x, belongs to $\mathcal{X}$. Let us prove that x is absolutely continuous and that its time derivative belongs to $\mathcal{X}$.

From the kernel recursive construction formula, it follows by induction that for all $m \geq 2$, h_m has a partial derivative w.r.t. t that belongs to $\mathcal{V}^m$, expressed as

$$\frac{\partial}{\partial t} h_m(t, \tau) = A h_m(t, \tau) + v_m(t, t, \tau).$$

Moreover, denoting $\widehat{\tau}_k(t) = (\tau_1, \ldots, \tau_{k-1}, t, \tau_{k+1}, \ldots, \tau_m)$ and $\mathrm{d}\widehat{\tau}_k = \mathrm{d}\tau_1 \cdots \mathrm{d}\tau_{k-1}\, \mathrm{d}\tau_{k+1} \cdots \mathrm{d}\tau_m$, we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} &\left(\int_{[0,t]^m} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \right) \\ &= \int_{[0,t]^m} \frac{\partial}{\partial t} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \\ &\quad + \sum_{k=1}^{m} \int_{[0,t]^{m-1}} h_m\big(t, \widehat{\tau}_k(t)\big) \left[\Pi_m u\right]\big(\widehat{\tau}_k(t)\big)\, \mathrm{d}\widehat{\tau}_k \\ &= \int_{[0,t]^m} \frac{\partial}{\partial t} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \\ &\quad + \int_{[0,t]^{m-1}} w_m\big(t, \widehat{\tau}_m(t)\big) \left[\Pi_m u\right]\big(\widehat{\tau}_m(t)\big)\, \mathrm{d}\widehat{\tau}_m. \end{aligned}$$

From (14) and (15), we have

$$\begin{aligned} \left\| \frac{\partial}{\partial t} h_m \right\|_{\mathcal{V}^m} &\leq \|A\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \psi_m + \sum_{k=2}^{m} \|P_k\|_{\mathcal{ML}_k(\mathbb{X},\mathbb{X})} \sum_{p \in \mathbb{M}_m^k} \prod_{\ell=1}^{k} \psi_{p_\ell} \\ &\leq \|A\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} \psi_m + \sum_{k=2}^{m} \frac{\mathcal{P}_k}{\kappa} \sum_{p \in \mathbb{M}_m^k} \prod_{\ell=1}^{k} \psi_{p_\ell} \\ &\leq \left(\|A\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} + \frac{1}{\kappa} \right) \psi_m \end{aligned}$$

so that

$$\begin{aligned} &\left\| \int_{[0,t]^m} \frac{\partial}{\partial t} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \right\|_{\mathcal{X}} \\ &\qquad \leq \left(\|A\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} + \frac{1}{\kappa} \right) \psi_m \|u\|_{\mathcal{X}}^m. \end{aligned}$$

In the same way

$$\left\| \int_{[0,t]^{m-1}} w_m\big(t, \widehat{\tau}_m(t)\big) \left[\Pi_m u\right]\big(\widehat{\tau}_m(t)\big)\, \mathrm{d}\widehat{\tau} \right\|_{\mathcal{X}} \leq \frac{1}{\kappa} \psi_m \|u\|_{\mathcal{U}}^m$$

and finally

$$\begin{aligned} &\left\| \frac{\mathrm{d}}{\mathrm{d}t} \left(\int_{[0,t]^m} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \right) \right\|_{\mathcal{X}} \\ &\qquad \leq \left(\|A\|_{\mathcal{L}(\mathbb{X},\mathbb{X})} + \frac{2}{\kappa} \right) \psi_m \|u\|_{\mathcal{X}}^m. \end{aligned}$$

The series $\sum_{m \in \mathbb{N}^*} (\mathrm{d}/\mathrm{d}t) \left(\int_{[0,t]^m} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \right)$ is therefore convergent in norm in the Banach space $\mathcal{X}$, and we have

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = \sum_{m \in \mathbb{N}^*} \frac{\mathrm{d}}{\mathrm{d}t} \left(\int_{[0,t]^m} h_m(t, \tau) \left[\Pi_m u\right](\tau)\, \mathrm{d}\tau \right). \tag{38}$$

This proves that property (i) of Definition 1 is satisfied.

Equations (6) and (38) prove that the infinite sum $\sum_{m \in \mathbb{N}^*}$ in the series defining x and the time derivative $\mathrm{d}/\mathrm{d}t$ can be commuted. Hence, from Proposition 1, x is a solution of (1)–(5). This proves property 1 of Definition 1 and concludes the proof of the theorem. ■

APPENDIX C
PROOF OF THEOREM 5

Let $\mu \in [0,a)$, $\lambda \geq \mu$, $\{h_m\}_{m\in\mathbb{N}^*} \in \mathcal{VS}_{\lambda,\mu}$, $u \in \mathcal{U}_\lambda$.

Extended Theorem 1: the adaptation of the proof relies on the following inequalities:

$$\sup_{t\in\mathbb{R}_+}\left(\mathrm{e}^{\mu t}\left\|\int_{[0,t]^m} h_m(t,\tau)\left[\Pi_m u\right](\tau)\,\mathrm{d}\tau\right\|_{\mathbb{X}}\right)$$
$$\leq \sup_{t\in\mathbb{R}_+}\left(\int_{[0,t]^m}\left\|h_m(t,\tau)\right\|_{\mathbb{X}}\mathrm{e}^{-m\overline{\tau}}\left(\|u\|_{\mathcal{U}_\lambda}\right)^m\mathrm{d}\tau\right)$$
$$\leq \|h_m\|_{\mathcal{V}^m_{\lambda,\mu}}\left(\|u\|_{\mathcal{U}}\right)^m$$

and on the use of the gain bound function $\varphi_{\lambda,\mu}$ derived from the $\mathcal{V}^m_{\lambda,\mu}$-norms.

Extended Theorem 2: Only the first step of the proof needs to be modified as follows. Equations (14) and (15) are adapted using $\|h_1\|_{\mathcal{V}^1_{\lambda,\mu}}$ and $\mathcal{P}_{k,\mu}$, $\mathcal{Q}_{k,\mu}$ (see Definition 7). Moreover, the left-hand side of (16) becomes $\mathrm{e}^{\mu t}\int_{[0,t]^m}\|h_m(t,\tau)\|_{\mathbb{X}}\mathrm{e}^{-\lambda\overline{\tau}}\,\mathrm{d}\tau$, and (17) becomes

$$\mathcal{A}_p(t) = \mathrm{e}^{\mu t}\int_{[0,t]^m}\int_{\max(\tau)}^{t}\widetilde{\mathcal{A}}_p(t,\theta,\tau)\,\mathrm{d}\theta\,\mathrm{e}^{-\lambda\overline{\tau}}\,\mathrm{d}\tau.$$

Equation (19) becomes $\mathcal{B}_q(t) = \mathrm{e}^{\mu t}\int_{[0,t]^m}\widetilde{\mathcal{B}}_q(t,\tau)\,\mathrm{e}^{-\lambda\overline{\tau}}\,\mathrm{d}\tau$, and (22) becomes

$$\int_{[0,t]^{p_\ell}}\|h_{p_\ell}(\theta,\eta)\|_{\mathbb{X}}\mathrm{e}^{-\lambda\overline{\eta}}\,\mathrm{d}\eta$$
$$\leq \int_{\mathbb{R}_+^{p_\ell}}\|h_{p_\ell}(\theta,\eta)\|_{\mathbb{X}}\mathrm{e}^{-\lambda\overline{\eta}}\,\mathrm{d}\eta$$
$$\leq \mathrm{e}^{-\mu\theta}\sup_{\tau\in\mathbb{R}_+}\left(\mathrm{e}^{\mu\tau}\int_{\mathbb{R}_+^{p_\ell}}\|h_{p_\ell}(\tau,\eta)\|_{\mathbb{X}}\mathrm{e}^{-\lambda\overline{\eta}}\,\mathrm{d}\eta\right)$$
$$\leq \mathrm{e}^{-\mu\theta}\,\|h_{p_\ell}\|_{\mathcal{V}^{p_\ell}_{\lambda,\mu}}$$

so that (23) is valid with $\mathcal{P}_{k,\mu}$. Similar modifications are performed to obtained the new version of (24).

Extended Theorem 3: The proof is straightforward by including the exponential weights in the formula of the original proof.

Extended Theorem 4: The proof is unchanged (using the new spaces $\mathcal{U}_\lambda$, $\mathcal{X}_\lambda$, $\mathcal{V}^m_{\lambda,\mu}$ and the new quantities σ_λ, $\rho^\star_\lambda$).

ACKNOWLEDGMENT

The authors would like to thank L. Smith for proofreading.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1970.
[2] S. Boyd and L. Chua, "Fading memory and the problem of approximating nonlinear operators with Voltera series," *IEEE Trans. Circuits Systems*, vol. CAS-32, no. 11, pp. 1150–1161, Nov. 1985.
[3] S. Boyd, L. O. Chua, and C. A. Desoer, "Analytical foundations of Volterra series," *J. Math. Control Inf.*, vol. 1, pp. 243–282, 1984.
[4] S. P. Boyd, "Volterra series: Engineering fundamentals," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, CA, 1985.
[5] R. W. Brockett, "Volterra series and geometric control theory," *Automatica*, vol. 12, pp. 167–176, 1976.
[6] R. W. Brockett, "Convergence of Volterra series on infinite intervals and bilinear approximations," in *Nonlinear Systems and Applications*, V. Lakshmikanthan, Ed. New York: Academic, 1977, pp. 39–46.
[7] F. Bullo, "Series expansions for analytic systems linear in control," *Automatica*, vol. 38, pp. 1425–1432, 2002.
[8] J. L. Casti, *Nonlinear System Theory*. Orlando, FL: Academic, 1985, vol. 175, Mathematics in Science and Engineering.
[9] P. E. Crouch and P. C. Collingwood, "The observation space and realizations of finite volterra series," *SIAM J. Control Optim.*, vol. 25, no. 2, pp. 316–333, 1987.
[10] M. Fliess, "Fonctionnelles causales non linéaires et indéterminées non commutatives," *Bull. S.M.F.*, vol. 109, pp. 3–40, 1981.
[11] M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarrigue, "An algebraic approach to nonlinear functional expansions," *IEEE Trans. Circuits Syst.*, vol. CAS-30, no. 8, pp. 554–570, Aug. 1983.
[12] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[13] E. G. Gilbert, "Functional expansions for the response of nonlinear differential systems," *IEEE Trans. Autom. Control*, vol. AC-22, no. 6, pp. 909–921, Dec. 1977.
[14] W. S. Gray and Y. Wang, "Fliess operators on LP spaces: Convergence and continuity," *Syst. Control Lett.*, vol. 46, no. 2, pp. 67–74, 2002.
[15] T. Hélie and M. Hasler, "Volterra series for solving weakly non-linear partial differential equations: Application to a dissipative Burgers' equation," *Int. J. Control*, vol. 77, pp. 1071–1082, 2004.
[16] T. Hélie and B. Laroche, "On the convergence of Volterra series of finite dimensional quadratic MIMO systems," *Int. J. Control*, vol. 81–3, pp. 358–370, 2008.
[17] T. Hélie and B. Laroche, "Computation of convergence radius and error bounds of Volterra series for single input systems with a polynomial nonlinearity," in *Proc. IEEE Conf. Decision Control*, Shangai, China, 2009, vol. 48, pp. 1–6.
[18] T. Hélie and D. Roze, "Sound synthesis of a nonlinear string using Volterra series," *J. Sound Vibration*, vol. 314, pp. 275–306, 2008.
[19] A. Isidori, *Nonlinear Control Systems*, 3rd ed. New York: Springer, 1995.
[20] X. J. Jing, Z. Q. Lang, and S. A. Billings, "Magnitude bounds of generalized frequency response functions for nonlinear Volterra systems described by Narx model," *Automatica*, vol. 44, pp. 838–845, 2008.
[21] F. Lamnabhi-Lagarrigue, *Analyse des Systèmes Non Linéaires*. London, U.K.: Hermès, 1994.
[22] Z. K. Peng and Z. Q. Lang, "On the convergence of the Volterra-series representation of the Duffing's oscillators subjected to harmonic excitations," *J. Sound Vibration*, vol. 305, pp. 322–332, 2007.
[23] W. J. Rugh, *Nonlinear System Theory, The Volterra/Wiener Approach*. Baltimore, MD: Johns Hopkins Univ. Press, 1981.
[24] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York: Wiley-Interscience, 1989.
[25] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*. New York: Dover, 1959.

Thomas Hélie received the Dipl. Ing. degree from the Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest, France, in 1997, the M.S. degree in acoustics, signal processing, and informatics applied to music from Université Paris 6, Paris, France, in 1999, and the M.S. and Ph.D. degrees in and automatic and signal processing from the Université Paris-Sud, Paris, France, in 1999 and 2002, respectively.

After post-doctoral research with the Laboratory of NonLinear Systems, Swiss Federal Institute of Lausanne, Lausanne, Switzerland, in 2003, and a Lecturer position with the Université Paris-Sud in 2004, he has been a researcher with the National Research Council in the Analysis/Synthesis Team, IRCAM-CNRS UMR 9912, Paris, France, since 2004. His research topics include physics of musical instruments, physical modeling, nonlinear dynamical systems, and inversion processes.

Béatrice Laroche received the M.Sc. degree in numerical analysis from University Paris VI, Paris, France, in 1985, and the Ph.D. degree in mathematics and automatic control from the École des Mines de Paris, Paris, France, in 2000.

She spent some years in research and development of flight control systems in industry. She is now an Associate Professor with the Physical Science Department, Control Theory, University Paris-Sud, Paris, France. Her scientific interests are in PDE control, modeling of nonlinear systems, model order reduction, and biological systems.