2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

The Dimension Estimation Problem for Nonlinear Systems

Alberto Padoan and Alessandro Astolfi

***Abstract*— We consider the dimension estimation problem for nonlinear systems, that is, the problem of finding the least possible dimension of the state space of a (minimal) realization of a nonlinear system from given input-output measurements. Exploiting tools from nonlinear realization theory and free Lie algebras, we develop algorithms to solve the dimension estimation problem both for bilinear and for general nonlinear systems, provided sufficient measurements are available. Simple worked-out examples illustrate the theory.**

I. INTRODUCTION

Consider a continuous-time, multi-input, multi-output, nonlinear system described by equations of the form

$$\dot{x} = f(x) + g(x)u, \quad y = h(x), \tag{1}$$

in which $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ denote the state, the input and the output of the system at time $t \in \mathbb{R}_+$ with initial condition $x(0) = x_0 \in \mathbb{R}^n$ and the mappings $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, and $h : \mathbb{R}^n \to \mathbb{R}^p$ are analytic, respectively.

The dimension estimation problem consists in finding the dimension of the state space of a (minimal) representation of the form (1) able to generate a prescribed input-output behaviour from given measured input-output data. Conceptually, the dimension of the state space can be thought of as a possible measure of the complexity of a system. Therefore determining the (least possible) dimension of the state space of a system from experimental observations is of paramount importance for modelling.

The properties of a class of systems play an important role in the dimension estimation problem. For example, in the case of linear systems the dimension of the state space can be recovered counting the non-zero singular values of a matrix using subspace identification methods [1 – 5], provided sufficient data are available. Understanding the properties of nonlinear systems, however, is more difficult in general, even under simplifying assumptions [6 , 7].

The main goal of this paper is to show that the dimension estimation problem and, more generally, the system identification problem for nonlinear systems may be solved with a geometric approach resorting to nonlinear realization theory [8 – 10]. This, in turn, is envisaged to be the starting point for the development of a truly nonlinear analogue of subspace identification methods. Historicallly, realization theory for linear systems has served as the foundation for the development of subspace identification methods. Motivated by the evolution of the theory in the linear case, we take a step in this direction for nonlinear systems.

This work belongs to a series of efforts devoted to the development of a system identification theory based on the differential-geometric approach to nonlinear systems [11 – 15]. In this paper, the algebraic formalism of nonlinear realization theory [8 – 10] and tools from free Lie algebras [16 , 17] are used to solve the dimension estimation problem both for bilinear and general nonlinear systems assuming that the prescribed input-output behaviour is specified by the coefficients of a generating series or, equivalently, by a Fliess operator [18 , 19]. The proposed framework allows to determine the dimension of a system from the rank of a Hankel matrix, thus generalizing a well-known result from linear systems theory [20]. Worked-out examples illustrate the applicability of our framework. Note that the class of systems considered does not include systems defined on manifolds, for which the notion of dimension is a local concept.

The rest of this work is organized as follows. Section II provides basic definitions and preliminary results. Section III contains the main results of the paper, where the dimension estimation problem is solved both for bilinear and general nonlinear systems. Section IV summarises our findings.

Notation: $\mathbb{N}$ and $\mathbb{Z}_+$ denote the set of non-negative integer numbers and the set of positive integer numbers, respectively. $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{R}^{p\times m}$ denote the set of real numbers, the set of n-dimensional vectors with real entries and the set of $p \times m$-dimensional matrices with real entries, respectively. $\mathbb{R}_+$ denotes the set of non-negative real numbers. M' denotes the transpose of the matrix $M \in \mathbb{R}^{p\times m}$. $\|v\|_\infty$ denotes the infinity norm of the vector $v \in \mathbb{R}^n$, defined as $\|v\|_\infty = \max_{1\le i\le n} |v_i|$ for all $v \in \mathbb{R}^n$. $\varphi^{(k)}$ denotes the time derivative of order $k \in \mathbb{Z}_+$ of the function φ, provided it exists. The Lie derivative of the function h along the vector field f is defined as $L_f h = \frac{\partial h}{\partial x} f$. The Lie derivative of order $k \in \mathbb{Z}_+$ of the function h along the vector field f is defined recursively as $L_f^k h = L_f(L_f^{k-1} h)$, with $L_f^0 h = h$. If the function h is vector-valued then its Lie derivatives are defined in a component-wise fashion.

II. PRELIMINARIES

This section recalls some notions from nonlinear control theory [8 , 9] and (non-commutative) formal power series [16 , 17].

Let $m \in \mathbb{Z}_+$ and consider a finite nonempty set of symbols $Z = \{z_0, z_1, \ldots, z_m\}$. Every element $z_i \in Z$ is referred to as

A. Padoan is with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK. (Email: a.padoan@eng.cam.ac.uk). A. Astolfi is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK and with the Dipartimento di Ingegneria Civile e Ingegneria Informatica, Università di Roma “Tor Vergata”, Via del Politecnico 1, 00133, Rome, Italy. (Email: a.astolfi@imperial.ac.uk). This work has been partially supported by the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 739551 (KIOS CoE).

a *letter* and every finite sequence $z_{i_k}\cdots z_{i_1}$ of zero or more elements of Z is referred to as a *word*. The *length* of a word w, denoted by $|w|$, is the number of letters that compose w. The *empty word*, denoted by 1, is the unique word of length zero. The set of all words of all lengths is denoted by $Z^\star$. The set $Z^\star$ is referred to as the *free monoid generated by* Z, since it is a monoid under the concatenation product $(v,w)\mapsto vw$ with identity element 1 (usually omitted in a product). Note that the order of products is important since, in general, the concatenation product is non-commutative.

A *formal power series* (also abbreviated as *series*) over Z with coefficients in $\mathbb{R}^p$ is a mapping of the form $s: Z^\star \to \mathbb{R}^p$, conventionally denoted by an infinite sum of the form

$$s = \sum_{w\in Z^\star} s_w w,$$

in which $s_w\in\mathbb{R}^p$ is the image of the word $w\in Z^\star$ under s and is referred to as the *coefficient* of w in s. The set of formal power series over Z with coefficients in $\mathbb{R}^p$ is denoted by $\mathbb{R}^p\langle\langle Z\rangle\rangle$. A *polynomial* is a series with finitely many non-zero coefficients. The set of polynomials over Z with coefficients in $\mathbb{R}^p$ is denoted by $\mathbb{R}^p\langle Z\rangle$.

Formal power series can be combined through various operations [9, 16, 17]. With scalar product and sum, the sets $\mathbb{R}\langle Z\rangle$ and $\mathbb{R}^p\langle\langle Z\rangle\rangle$ assume the structure of vector spaces over $\mathbb{R}$. The set $\mathbb{R}\langle Z\rangle$ can be also regarded as a ring when equipped with sum and Cauchy product and as a Lie algebra over $\mathbb{R}$ equipped with Lie bracket. The smallest Lie subalgebra of $\mathbb{R}\langle Z\rangle$ containing Z is denoted by $L(Z)$, with the set Z identified with the set of degree one monomials. Every element of $L(Z)$ is referred to as a *Lie polynomial*.

A formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ is said to be *convergent* if there exist $K\in\mathbb{R}_+$ and $M\in\mathbb{R}_+$ such that $\|s_w\|_\infty \le KM^{|w|}|w|!$ for every $w\in Z^\star$. In this case, $T\in\mathbb{R}_+$ is said to be *admissible* if $0<T<M^{-1}(m+1)^{-1}$. A convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ is defined as

$$s_w = L_w h(x_0), \quad w\in Z^\star, \tag{2}$$

in which, by a convenient abuse of notation, $L_1 h(x_0) = h(x_0)$, and[1] $L_w h(x_0) = L_{g_{i_0}}\cdots L_{g_{i_k}} h(x_0)$, for every $w = z_{i_k}\cdots z_{i_0}\in Z^\star$, is said to be a *generating series* for system (1), because it completely specifies the output of the system provided that some assumptions hold [9]. Conversely, system (1) is said to be a *realization* at $x_0\in\mathbb{R}^n$ of the convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ if the mappings which describe the system satisfy (2). A realization (1) at $x_0\in\mathbb{R}^n$ of the convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ is said to be *bilinear* if there exist constant matrices $A\in\mathbb{R}^{n\times n}$, $B_1,\ldots,B_m\in\mathbb{R}^{n\times n}$ and $C\in\mathbb{R}^{p\times n}$ such that $f(x)=Ax$, $g_1(x)=B_1x$, ..., $g_m(x)=B_mx$ and $h(x)=Cx$.

III. Main results

The *Hankel operator* associated with a convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ is the morphism of real vector spaces $\mathcal{H}_s:\mathbb{R}\langle Z\rangle\to\mathbb{R}^p\langle\langle Z\rangle\rangle$ uniquely specified by the property $[\mathcal{H}_s(w)]_v = s_{vw}$, for all $v,w\in Z^\star$. The *Hankel rank* $\rho_H(s)$ of a convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ is defined as $\rho_H(s)=\dim\mathcal{H}_s(\mathbb{R}\langle Z\rangle)$, *i.e.* as the dimension of the image under $\mathcal{H}_s$ of the whole space $\mathbb{R}\langle Z\rangle$. The *Lie rank* $\rho_L(s)$ of a convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ is defined as $\rho_L(s)=\dim\mathcal{H}_s(L(Z))$, *i.e.* as the dimension of the image of the restriction of $\mathcal{H}_s$ to the subspace $L(Z)$.

It is well-known [9] that a convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ admits a bilinear realization if and only if its Hankel rank is finite, in which case $\rho_H(s)$ coincides with the dimension of all minimal bilinear realizations of s. Similarly, a convergent formal power series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ admits a (possibly nonlinear) realization if and only if its Lie rank is finite, in which case $\rho_L(s)$ coincides with the dimension of all minimal realizations of s. Thus *computing the dimension of a system admitting a bilinear realization (nonlinear realization) amounts to computing the Hankel rank (Lie rank) of its generating series*. For this reason, in the sequel we focus on computing these quantities assuming that sufficiently many coefficients of a given convergent formal power series are known. Note that this assumption is *not* restrictive, since the coefficients of the generating series of a given system can be determined using input-output measurements, as discussed below in more detail.

A. Computation of the coefficients of a generating series

The coefficients of a generating series can be regarded as a nonlinear analogue of the Markov parameters of a linear system, which can be determined by evaluating the derivatives of the impulse response of the system at zero [21]. For completeness, we describe a methodology for computing the coefficients of the generating series $s\in\mathbb{R}^p\langle\langle Z\rangle\rangle$ of system (1) from the knowledge of its input and its output (together with their derivatives up to a given order) over some interval $[0,T]$, with $T\in\mathbb{R}_+$ admissible for s. The methodology relies on the fact that the observation space of system (1), defined as

$$\mathcal{O} = \operatorname{span}_{\mathbb{R}}\{\, L_w h,\ w\in Z^\star\,\},$$

coincides with the linear space [22, 23]

$$\widetilde{\mathcal{O}} = \operatorname{span}_{\mathbb{R}}\{\, y^{\mu_1\cdots\mu_{k-1}},\ \mu_i\in\mathbb{R}^m,\ k\in\mathbb{N}\,\},$$

in which $y^{\mu_1\cdots\mu_{k-1}}(x) = y_x^{(k)}(0)$ for every $\mu_1,\ldots,\mu_{k-1}$ and $x\in\mathbb{R}^n$, where y_x is the family of outputs corresponding to the initial state x and the family of (infinitely many times differentiable) inputs u such that $u^{(j)}(0)=\mu_j$ for every integer $0\le j\le k-1$. In other words, the coefficients of the generating series s can be computed as linear combinations of derivatives of the output corresponding to derivatives of the input with prescribed values, as illustrated by the following example.

Example 1. Consider system (1), with $m=1$, and suppose we wish to compute the coefficients of the generating series of system (1) corresponding to every word $w\in Z^\star$ such that $0\le|w|\le 2$. These coefficients are retrieved using the identities (3), which can be determined by direct

[1]Throughout the paper h_i and g_j denote the i-th component of h and the j-th column of g, respectively. For convenience, g_0 is used to denote the vector field f.

$$
\begin{aligned}
y &= h(x) && \text{(3a)}\\
y^{\mu_0} &= L_f h(x) + L_g h(x)\mu_0 && \text{(3b)}\\
y^{\mu_0\mu_1} &= L_f^2 h(x) + (L_g L_f h(x) + L_f L_g h(x))\mu_0 + L_g^2 h(x)\mu_0^2 + L_g h(x)\mu_1 && \text{(3c)}\\
y^{\mu_0\mu_1\mu_2} &= L_f^3 h(x) + (L_f^2 L_g h(x) + L_f L_g L_f h(x) + L_g L_f^2 h(x))\mu_0 + (L_f L_g^2 h(x) + L_g L_f L_g h(x) + L_g^2 L_f h(x))\mu_0^2 \\
&\quad + L_g^3 h(x)\mu_0^3 + 2L_g^2 h(x)\mu_0\mu_1 + (L_g L_f h(x) + 2L_f L_g h(x))\mu_1 + L_g h(x)\mu_2 && \text{(3d)}
\end{aligned}
$$

computation, as follows. For $|w| = 0$, it follows directly from (3a) that $s_1 = y(0)$. For $|w| = 1$, setting $\mu_0 = 0$ in (3b) gives $s_{z_0} = y^0$, while selecting $\mu_0 = 1$ a simple substitution gives $s_{z_1} = -y^0 + y^1$. For $|w| = 2$, it follows from (3c) that $s_{z_0^2} = y^{00}$. Exploiting (3d), a lengthier, but elementary computation yields

$$
s_{z_0 z_1} = 2y^{01} - 2y^{00} - y^{010} + y^{000}, \quad \text{(4a)}
$$

$$
s_{z_1 z_0} = -y^{01} + y^{00} + y^{010} + y^{000}. \quad \text{(4b)}
$$

Finally, combining (3c) and (3d) with (4) yields

$$
s_{z_1^2} = y^{11} - 2y^{01} + y^{00} - 2y^{000}.
$$

Note that all the coefficients have been expressed in terms of *measurable* quantities. ▲

Remark 1. The representation of the elements of the observation space $\mathcal{O}$ in terms of elements of the linear space $\widetilde{\mathcal{O}}$ applies to *every* system described by analytic mappings [23]. This universal property implies that the coefficients of the generating series of a system can be computed by performing a number of input-output experiments and by subsequently substituting the values of input and output derivatives in expressions analogous to those obtained in Example 1. △

Remark 2. The methodology illustrated in Example 1 suggests how to design inputs, since one only needs to know their values and the values of their derivatives up to a certain order. However, it should be noted that for this methodology to work one needs to perform multiple experiments with the system starting from the *same* initial condition, which could be restrictive for some applications. △

B. Hankel rank computation

The computation of the Hankel rank of a convergent formal power series admitting a bilinear realization is conceptually simple and can be performed as follows. Representing the elements of the vector spaces $\mathbb{R}\langle Z\rangle$ and $\mathbb{R}^p\langle\langle Z\rangle\rangle$ as vectors, the (infinitely many) entries of which are real numbers and real vectors with p components indexed by the elements of $Z^\star$, it is possible to derive a matrix representation of the operator $\mathcal{H}_s$. For example, fixing two (possibly distinct) monomial orderings $v_1 < v_2 < v_3 < \dots$ and $w_1 < w_2 < w_3 < \dots$ over the elements of $Z^\star$, a matrix representation of the operator $\mathcal{H}_s$ has the form

$$
H_s = \begin{bmatrix} s_{v_1 w_1} & s_{v_1 w_2} & s_{v_1 w_3} & \cdots \\ s_{v_2 w_1} & s_{v_2 w_2} & s_{v_2 w_3} & \cdots \\ s_{v_3 w_1} & s_{v_3 w_2} & s_{v_3 w_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.
$$

The matrix H_s is referred to as the *Hankel matrix* of the series s. The rank[2] of the Hankel matrix H_s coincides, by definition, with the Hankel rank of s, *i.e.* $\rho_H(s) = \operatorname{rank} H_s$.

Remark 3. A direct consequence of Remark 1 is that the Hankel matrix associated with *any* generating series can be constructed using exclusively input-output data. For example, consider system (1), with $m = 1$, and the corresponding generating series (2). Then, ordering the elements of $Z^\star$ according to the (graded) lexicographic ordering

$$
1 < z_0 < z_1 < z_0^2 < z_0 z_1 < z_1 z_0 < z_1^2 < z_0^3 < z_0^2 z_1 < \dots,
$$

and using the results of Example 1, the Hankel matrix H_s associated with the generating series (2) can be represented as in Fig. 1. The Hankel matrix associated with a given generating series can be then found by substituting the values of input and output derivatives in the matrix representation in Fig. 1. Similar considerations can be performed for the Lie-Hankel matrix, which is defined later in the paper. △

Since a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z\rangle\rangle$ admits a bilinear realization if and only if $\rho_H(s) = n$ for some $n \in \mathbb{Z}_+$, we infer that for a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z\rangle\rangle$ admitting a bilinear realization there exist $M \in \mathbb{Z}_+$ and $N \in \mathbb{Z}_+$ such that

$$
\operatorname{rank} H_s^{M+i,N+j} = \operatorname{rank} H_s^{M,N} = n,
$$

for every $i, j \in \mathbb{N}$, in which $H_s^{M,N} \in \mathbb{R}^{pM\times N}$ is the matrix formed by the first M block-rows and N columns of the Hankel matrix H_s. In other words, if the convergent formal power series s admits a bilinear realization then the dimension of the system can be computed as the rank of a sufficiently large (top-left) submatrix of H_s. We emphasise, however, that it may *not* be possible to determine the dimension of the system in case the amount of data available is not sufficient, *i.e.* if the condition $n \geq \min\{n_r, n_c\}$ holds, where $n_r \in \mathbb{Z}_+$ and $n_c \in \mathbb{Z}_+$ are the number of rows and the number of columns of the largest top-left submatrix of H_s available. In this case, the dimension of the system cannot be correctly determined unless more coefficients of the formal power series s are computed. The main ideas of our discussion are summarized in pseudo-code in Algorithm 1, which takes as input finitely many coefficients of a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z\rangle\rangle$ admitting a bilinear realization and returns

[2] The rank of a matrix (even an infinite one) can be defined as the dimension of its largest non-vanishing sub-determinant. With such definition, the rank of a matrix coincides with the dimension of the space generated by its rows (and by its columns).

$$
H_s : \begin{array}{c|cccc}
 & 1 & z_0 & z_1 & \cdots \\
\hline
1 & y(0) & y^0 & -y^0 + y^1 & \cdots \\
z_0 & y^0 & y^{00} & 2y^{01} - 2y^{00} - y^{010} + y^{000} & \cdots \\
z_1 & -y^0 + y^1 & -y^{01} + y^{00} + y^{010} + y^{000} & y^{11} - 2y^{01} + y^{00} - 2y^{000} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

Fig. 1. The Hankel matrix associated with the generating series (2).

the dimension of all minimal bilinear realizations of s or the flag "Insufficient data", in case more coefficients are needed.

Algorithm 1

Construct the matrix $H_s^{n_r,n_c}$.
Compute $r = \operatorname{rank} H_s^{n_r,n_c}$.
if $r < \min\{n_r, n_c\}$ **then**
    return $n = r$.
else
    return "Insufficient data".
end if
stop

The following example illustrates how to compute the Hankel rank of a convergent formal power series admitting a bilinear realization considering a system first studied in [24].

Example 2. Consider the system

$$\dot{x} = Ax + Bxu, \quad y = Cx \tag{5}$$

with $x(t) \in \mathbb{R}^2$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $x(0) = [\,0 \;\; 1\,]'$ and

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad C' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{6}$$

Let $Z = \{z_0, z_1\}$ and note that the generating series of system (5) is

$$s = \sum_{w \in Z^\star} (|w|_{z_0} - |w|_{z_1}) w, \tag{7}$$

in which $|w|_{z_0}$ and $|w|_{z_1}$ denote the number of occurrences of the letters z_0 and z_1, respectively. Ordering the elements of $Z^\star$ according to the (graded) lexicographic ordering

$$1 < z_0 < z_1 < z_0^2 < z_0 z_1 < z_1 z_0 < z_1^2 < z_0^3 < z_0^2 z_1 < \ldots,$$

the Hankel matrix H_s of the generating series (7) is

$$
H_s : \begin{array}{c|cccccccccc}
 & 1 & z_0 & z_1 & z_0^2 & z_0 z_1 & z_1 z_0 & z_1^2 & z_0^3 & z_0^2 z_1 & \cdots \\
\hline
1 & 0 & 1 & -1 & 2 & 0 & 0 & -2 & 3 & 2 & \cdots \\
z_0 & 1 & 2 & 0 & 3 & 1 & 1 & -1 & 4 & 2 & \cdots \\
z_1 & -1 & 0 & -2 & 1 & -1 & -1 & -3 & 2 & 0 & \cdots \\
z_0^2 & 2 & 3 & 1 & 4 & 2 & 2 & 0 & 5 & 3 & \cdots \\
z_0 z_1 & 0 & 1 & -1 & 2 & 0 & 0 & -2 & 3 & 1 & \cdots \\
z_1 z_0 & 0 & 1 & -1 & 2 & 0 & 0 & -2 & 3 & 1 & \cdots \\
z_1^2 & -2 & -1 & -3 & 0 & -2 & -2 & -4 & 1 & -1 & \cdots \\
z_0^3 & 3 & 4 & 2 & 5 & 3 & 3 & 1 & 6 & 4 & \cdots \\
z_0^2 z_1 & 1 & 2 & 0 & 3 & 1 & 1 & -1 & 4 & 2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

The use of a numerical computing software gives $\operatorname{rank} H_s^{i+2,j+2} = 2$, for every $i, j \in \mathbb{N}$. Hence $\rho_H(s) = 2$. Consistently, the series (7) is realized by the (minimal) bilinear realization (5). ▲

C. Lie rank computation

The computation of the Lie rank of a convergent formal power series admitting a (nonlinear) realization is a somewhat delicate matter. It is clear that a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z \rangle\rangle$ admitting a realization is such that $\rho_L(s) = n$ for some $n \in \mathbb{Z}_+$ and, by definition, $\rho_L(s) = \dim \mathcal{H}_s(L(Z))$. Since the Lie algebra $L(Z)$ is a subspace of the vector space $\mathbb{R}\langle Z \rangle$, the Lie rank can be computed as the rank of a matrix representation of the restriction of the Hankel operator to the Lie algebra $L(Z)$. Instrumental to the definition of such representation are *Hall polynomials*[3] due to the following property [17].

Theorem 1. Let $H(Z)$ be a Hall set. Then the Hall polynomials form a basis of the Lie algebra $L(Z)$ (viewed as a vector space over $\mathbb{R}$). Moreover, decreasing products of Hall polynomials of the form $p_{\mathtt{h}_1} \cdots p_{\mathtt{h}_n}$, with $\mathtt{h}_i \in H(Z)$ and $\mathtt{h}_1 \geq \cdots \geq \mathtt{h}_n$, form a basis of the algebra $\mathbb{R}\langle Z \rangle$ (viewed as a vector space over $\mathbb{R}$).

Another important ingredient is the $\mathbb{R}\langle Z \rangle$-module structure of $\mathbb{R}^p\langle\langle Z \rangle\rangle$ [9]. With this algebraic structure, the Hankel operator associated with a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z \rangle\rangle$ can be equivalently defined as the morphism of $\mathbb{R}\langle Z \rangle$-modules $\mathcal{H}_s : \mathbb{R}\langle Z \rangle \to \mathbb{R}^p\langle\langle Z \rangle\rangle$ uniquely specified by the property

$$\mathcal{H}_s(p) = p \cdot s, \tag{8}$$

for every $p \in \mathbb{R}\langle Z \rangle$.

We now construct a matrix representation of the restriction of the Hankel operator to the Lie algebra $L(Z)$. Let $H(Z)$ be a Hall set and let the elements of $L(Z)$ and of $\mathbb{R}^p\langle\langle Z \rangle\rangle$ (viewed as $\mathbb{R}\langle Z \rangle$-modules) be represented by vectors, the (infinitely many) entries of which are real numbers and real vectors with p components indexed as follows. The entries of a Lie polynomial $p \in L(Z)$ are indexed by Hall words[4] of increasing length, with entries corresponding to Hall words of the same length ordered by the total order $\leq$ of $H(Z)$. By Theorem 1, this representation completely specifies every Lie polynomial in $L(Z)$. The entries of a formal power series $s \in \mathbb{R}^p\langle\langle Z \rangle\rangle$ are indexed according to a fixed monomial

[3]For reasons of space, we do not discuss the construction of Hall sets and Hall polynomials. The reader is referred to [17, 25] for more detail.

[4]A Hall word is the foliage of a Hall tree [17, p.89].

ordering $v_1 < v_2 < v_3 < \dots$ over the elements of $Z^\star$. In view of the property (8), the restriction of the operator $\mathcal{H}_s$ to the subspace $L(Z)$ has a matrix representation of the form

$$L_s = \begin{bmatrix} (p_{\mathtt{h}_1} \cdot s)_{v_1} & (p_{\mathtt{h}_2} \cdot s)_{v_1} & (p_{\mathtt{h}_3} \cdot s)_{v_1} & \cdots \\ (p_{\mathtt{h}_1} \cdot s)_{v_2} & (p_{\mathtt{h}_2} \cdot s)_{v_2} & (p_{\mathtt{h}_3} \cdot s)_{v_2} & \cdots \\ (p_{\mathtt{h}_1} \cdot s)_{v_3} & (p_{\mathtt{h}_2} \cdot s)_{v_3} & (p_{\mathtt{h}_3} \cdot s)_{v_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The matrix L_s is referred to as the *Lie-Hankel matrix* of the series s. The rank of the Lie-Hankel matrix L_s coincides, by definition, with the Lie rank of s, *i.e.* $\rho_L(s) = \operatorname{rank} L_s$.

Since a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z \rangle\rangle$ admits a (possibly nonlinear) realization if and only if $\rho_L(s) = n$ for some $n \in \mathbb{Z}_+$, we infer that for a convergent formal power series $s \in \mathbb{R}^p\langle\langle Z \rangle\rangle$ admitting a realization there exist $M \in \mathbb{Z}_+$ and $N \in \mathbb{Z}_+$ such that

$$\operatorname{rank} L_s^{M+i,N+j} = \operatorname{rank} L_s^{M,N} = n,$$

for every $i, j \in \mathbb{N}$, in which $L_s^{M,N} \in \mathbb{R}^{pM \times N}$ is the matrix formed by the first M block-rows and N columns of the Lie-Hankel matrix L_s. Algorithm 1 can be thus modified to estimate the dimension of general nonlinear systems, as summarized in pseudo-code in Algorithm 2.

Algorithm 2

Construct the matrix $L_s^{n_r,n_c}$.
Compute $r = \operatorname{rank} L_s^{n_r,n_c}$.
if $r < \min\{n_r, n_c\}$ **then**
  return $n = r$.
else
  return "Insufficient data".
end if
stop

Remark 4 (Construction of the Lie-Hankel matrix). Every Hall polynomial $p_{\mathtt{h}}$ is a homogeneous polynomial with integer coefficients [17]. This implies that the image of a Hall polynomial under the Hankel operator $\mathcal{H}_s$ associated with a convergent formal power series s can be computed adding the image of its homogeneous components, weighted by the corresponding coefficients. This, in turn, allows to compute the Lie-Hankel matrix L_s as the product $L_s = H_s K$, where H_s is the Hankel matrix associated with the formal power series s and K is a matrix defined as follows. The rows of the matrix K are indexed by the monomial ordering $w_1 < w_2 < w_3 < \dots$ over the elements of $Z^\star$ used to index the columns of the Hankel matrix H_s. The columns of the matrix K are indexed by Hall words of increasing length, with entries corresponding to Hall words of the same length ordered by the total order $\leq$ of $H(Z)$. The entries of the matrix K are defined as $K_{w\mathtt{h}} = [p_{\mathtt{h}}]_w$, for every $w \in Z^\star$ and every $\mathtt{h} \in H(Z)$, *i.e.* the entry of K indexed by $(w, \mathtt{h})$ is the coefficient of the word w in the Hall polynomial $p_{\mathtt{h}}$. △

The following academic example illustrates how to compute the Lie rank of a convergent formal power series admitting a realization.

Example 3. Consider the system (1), with $x(t) \in \mathbb{R}^2$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, $x(0) = [\,0\ 0\,]'$ and

$$f(x) = [\,x_2 \;\; x_1^2\,]', \quad g(x) = [\,0 \;\; 1\,]', \quad h(x) = x_1. \tag{9}$$

Let $Z = \{z_0, z_1\}$ and note that the generating series of system (9) is

$$s = z_0 z_1 + 2 z_0^3 z_1 z_0 z_1 + 4 z_0^4 z_1^2 + \dots. \tag{10}$$

Fig. 2 displays the Hankel matrix associated with the generating series (10), with the elements of $Z^\star$ ordered according to the (graded) lexicographic ordering

$$1 < z_0 < z_1 < z_0^2 < z_0 z_1 < z_1 z_0 < z_1^2 < z_0^3 < z_0^2 z_1 < \dots.$$

We conjecture that the Hankel rank of the generating series (10) is infinite, since experimental evidence shows that $\operatorname{rank} H_s^{i,i}$ grows unbounded as $i \to \infty$. To compute the Lie-Hankel matrix we first compute a sufficiently large number of Hall polynomials. The Hall polynomials of degree less than five are $p_{z_0} = z_0$, $p_{z_1} = z_1$, $p_{z_0 z_1} = z_1 z_0 - z_0 z_1$, $p_{z_0^2 z_1} = z_1^2 z_0 - 2 z_1 z_0 z_1 + z_1 z_0^2$ and $p_{z_0 z_1^2} = z_0^2 z_1 - 2 z_0 z_1 z_0 + z_0 z_1^2$. In view of Remark 4, we then define the matrix K as

$$K : \quad \begin{array}{c|cccccc} & p_{z_0} & p_{z_1} & p_{z_0 z_1} & p_{z_0^2 z_1} & p_{z_0 z_1^2} & \cdots \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ z_1 & 0 & 1 & 0 & 0 & 0 & \cdots \\ z_0^2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0 z_1 & 0 & 0 & 1 & 0 & 0 & \cdots \\ z_1 z_0 & 0 & 0 & -1 & 0 & 0 & \cdots \\ z_1^2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0^3 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0^2 z_1 & 0 & 0 & 0 & 1 & 0 & \cdots \\ z_0 z_1 z_0 & 0 & 0 & 0 & -2 & 0 & \cdots \\ z_0 z_1^2 & 0 & 0 & 0 & 0 & 1 & \cdots \\ z_1 z_0^2 & 0 & 0 & 0 & 1 & 0 & \cdots \\ z_1 z_0 z_1 & 0 & 0 & 0 & 0 & -2 & \cdots \\ z_1^2 z_0 & 0 & 0 & 0 & 0 & 1 & \cdots \\ z_1^3 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Finally, the matrix L_s is computed considering the product $L_s = H_s K$, which gives

$$L_s : \quad \begin{array}{c|cccccc} & p_{z_0} & p_{z_1} & p_{z_0 z_1} & p_{z_0^2 z_1} & p_{z_0 z_1^2} & \cdots \\ \hline 1 & 0 & 0 & 1 & 0 & 0 & \cdots \\ z_0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ z_1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0^2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0 z_1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_1 z_0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_1^2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0^3 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0^2 z_1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0 z_1 z_0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_0 z_1^2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_1 z_0^2 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_1 z_0 z_1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_1^2 z_0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ z_1^3 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

$$
H_s: \quad \begin{array}{c|cccccccccccccccc}
 & 1 & z_0 & z_1 & z_0^2 & z_0z_1 & z_1z_0 & z_1^2 & z_0^3 & z_0^2z_1 & z_0z_1z_0 & z_0z_1^2 & z_1z_0^2 & z_1z_0z_1 & z_1^2z_0 & z_1^3 & \cdots \\
\hline
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_0^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_0z_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1z_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_0^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 2 & 0 & 0 & \cdots \\
z_0^2z_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_0z_1z_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_0z_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1z_0^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1z_0z_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1^2z_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
z_1^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

Fig. 2. The Hankel matrix associated with the generating series (10).

Direct inspection yields $\operatorname{rank} L_s^{i+3,j+3} = 2$, for every $i, j \in \mathbb{N}$, which implies $\rho_L(s) = 2$. Consistently, the series (10) is realized by the (minimal) nonlinear realization (9). ▲

IV. CONCLUSION

The dimension estimation problem has been studied and solved both for bilinear and general nonlinear systems using tools from nonlinear realization theory and free Lie algebras. The theory has been illustrated with two academic examples. Future research should investigate the accuracy of the proposed algorithms and implementative aspects, such as sampling the needed quantities or the effect of noise.

REFERENCES

[1] L. Ljung, *System identification - theory for the user (2nd edition)*. Upper Saddle River, NJ: Prentice Hall, 1999.

[2] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: theory - implementation - applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

[3] T. Katayama, *Subspace methods for system identification*. London, U.K.: Springer, 2006.

[4] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge, U.K.: Cambridge University Press, 2007.

[5] D. Bauer, "Order estimation for subspace methods," *Automatica*, vol. 37, no. 10, pp. 1561–1573, 2001.

[6] Y. Wang and E. D. Sontag, "Orders of input/output differential equations and state-space dimensions," *SIAM J. Control Optim.*, vol. 33, no. 4, pp. 1102–1126, 1995.

[7] F. Xiong, O. I. Camps, and M. Sznaier, "Low order dynamics embedding for high dimensional time series," in *Proc. IEEE Conf. Computer Vision*, New York, NY, USA, 2011, pp. 2368–2374.

[8] H. Nijmeijer and A. Van der Schaft, *Nonlinear dynamical control systems*. New York: Springer-Verlag, 1990.

[9] A. Isidori, *Nonlinear control systems (3rd edition)*. New York, NY, USA: Springer-Verlag, 1995.

[10] M. Fliess and M. Hazewinkel, Eds., *Algebraic and geometric methods in nonlinear control theory*, ser. Mathematics and its applications. Dordrecht, Holland: D. Reidel Publishing Company, 1986.

[11] A. Padoan and A. Astolfi, "Dimension estimation for autonomous nonlinear systems," in *Proc. 54th Conf. Decision Control*, Osaka, Japan, 2015, pp. 103–108.

[12] ——, "Towards deterministic subspace identification for autonomous nonlinear systems," in *Proc. 54th Conf. Decision Control*, Osaka, Japan, 2015, pp. 127–132.

[13] ——, "Nonlinear system identification for autonomous systems via functional equations methods," in *Proc. Amer. Control Conf.*, Boston, MA, USA, 2016, pp. 2378–5861.

[14] A. Padoan, G. Scarciotti, and A. Astolfi, "A geometric characterisation of the persistence of excitation condition for the solutions of autonomous systems," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 5666 – 5677, 2017.

[15] A. Padoan, "Input-output Schröder equations: a new look at some old problems," *IEEE Trans. Autom. Control*, 2017, under review.

[16] J. Berstel and C. Reutenauer, *Noncommutative rational series with applications*. Cambridge, U.K.: Cambridge University Press, 2011.

[17] C. Reutenauer, *Free Lie algebras*. Oxford, U. K.: Oxford University Press, 1993.

[18] Y. Wang and E. D. Sontag, "Generating series and nonlinear systems: analytic aspects, local realizability, and i/o representations," vol. 4, no. 3, pp. 299–322, 1992.

[19] W. S. Gray and Y. Wang, "Fliess operators on L_p spaces: convergence and continuity," *Syst. Control Lett.*, vol. 46, no. 2, pp. 67–74, 2002.

[20] A. C. Antoulas, *Approximation of large-scale dynamical systems*. Philadelphia, PA: SIAM, 2005.

[21] T. Kailath, *Linear systems*. Englewood Cliffs, NJ, USA: Prentice Hall, 1980.

[22] Y. Wang and E. D. Sontag, "On two definitions of observation spaces," *Syst. Control Lett.*, vol. 13, no. 4, pp. 279–289, 1989.

[23] Y. Wang, "Algebraic differential equations and nonlinear control systems," Ph.D. dissertation, Rutgers University, New Brunswick, 1990.

[24] C. Reutenauer, "Noncommuting variables: I. Finite automata, II. Rational generating series, III. Realization of bilinear systems," *Syst. Control Encyclopedia*, pp. 3268–3272, 1987.

[25] A. Padoan, "Geometric methods for modelling and approximation of nonlinear systems," PhD Thesis, Imperial College London, 2018.