Spaces of nonlinear and hybrid systems representable by recognizable formal power series

Mihály Petreczky and Ralf Peeters
Maastricht University P.O. Box 616, 6200 MD Maastricht, The Netherlands
{M.Petreczky, Ralf.Peeters}@maastrichtuniversity.nl

***Abstract*— The paper presents the manifold structure of the spaces of those nonlinear and hybrid system which can be encoded by rational formal power series. The latter class contains bilinear systems, linear multidimensional systems, linear switched and hybrid systems and jump-markov linear systems.**

I. Introduction

In this paper we present results on topology and geometry of the spaces of those control systems, input-output behavior of which can be described by rational formal power series.

The motivation for studying topology and geometry of space of systems is that it helps to design and analyse model reduction and systems identification algorithms. More specifically, for model reduction we need to define when two systems are close, i.e. we need a topology and metric on the spaces of systems. In parametric system identification one tries to find an optimum of a functional, defined on the space of systems. In order to derive and analyse algorithms for finding such an optimum, it is useful to know the geometry of the space of parametrizations (systems). For instance, if the space of systems has a Riemannian manifold structure, then optimization techniques for Riemannian manifolds can be used. The importance of topology and geometry of system spaces in model reduction and systems identification has been demonstrated for linear systems, see [13], [6].

General requirement on structure of spaces of systems: There is a number of conditions which the topology and geometry of the space of systems should satisfy, in order to be useful for model reduction and systems identification.

System=input-output behavior We identify a system with its input-output behavior, i.e. two state-space representations are considered equivalent if they have the same input-output behavior. The space of systems is then the set of equivalence classes of state-space representations.

The system depends continuously (smoothly) on the parameters of the state-space representation If the parameters of the state-space representations, viewed as real vectors, are close, then the corresponding input-output maps (or equivalence classes of state-space representations) are close as well.

Computability The topology and the differentiable structure of the space of systems should be computable.

Contribution of the paper: In this paper we investigate the space of *rational families of formal power series (**RFFPS** for short)* , or, which is the same, the space of equivalence classes of minimal rational representations of **RFFPSs** modulo isomorphism. For the definition of **RFFPSs** and their representations see [15], [14] and the references therein.

We prove that the set M_n of those **RFFPSs** which admit a minimal representation of a fixed order n forms a Nash-submanifold, in particular, it is an analytic manifold. Moreover, the elements of M_n depend in a smooth and semi-algebraic manner on the parameters of the corresponding state-space representation. In addition, we show that the subset SM_n of M_n, formed by *square summable* **RFFPSs** (see [17] for the definition), is an open Nash-submanifold of M_n. Moreover, the topology of SM_n as a manifold coincides with the topology of SM_n as a subset of the Hilbert-space of square summable **RFFPSs**. In turn, square summable **RFFPSs** are important, because **(a)** they often correspond to stable systems, and **(b)** the distance between square summable **RFFPSs** determined by the Hilbert-space structure is computable. We also provide an explicit computable construction of the coordinate chart of M_n.

Motivation : Several classes of systems have the property that their input-output behavior can be encoded as a **RFFPS**. These classes include linear and bilinear hybrid and switched systems (see [15] for an overview), jump-markov linear systems [17], [18], bilinear systems [19], [8] and hidden-markov models. Moreover, for these systems, there is a correspondence between state-space realizations and representations of the corresponding **RFFPSs**. Hence, the results on the structure of M_n and SM_n can be translated to similar results on the structure of the spaces of such systems.

Previous work: The structure of the spaces of linear systems has been extensively investigated, for an overview see [13], [6]. The results of this paper are extensions of the known results for linear systems. To the best of our knowledge, spaces of **RFFPSs** were not investigated so far. However, there is a one-to-one relationship between bilinear systems and rational formal power series. In [21] it is shown that the space of bilinear systems forms a quasi-affine variety, and in [22] it is shown that the same space is an analytic manifold. However, [22] does not provide an explicit construction of the charts of the manifold of bilinear system and it does not relate the topology of the manifold to the topology of square-summable **RFFPSs**.

Outline: In §II we present the notation and terminology of the paper. In §III we review the basics on **RFFPSs**.

ISBN 978-963-311-370-7

The main results are presented in §IV. We illustrate the relevance of these results by sketching in §V their application to bilinear and switched systems. In §VII we present the proof of the main result. The proof is based on the notion of nice selection, which is presented in §VI.

II. Preliminaries

Automata theory We use standard notation of automata theory, see [5], [4]. For a finite set X, called the *alphabet*, denote by X^* the set of finite sequences (also called strings or words) of elements of X. The length of a word w is denoted by $|w|$, i.e. $|w| = k$. We denote by ϵ the *empty sequence (word)*. In addition, we define $X^+ = X^* \setminus \{\epsilon\}$. Denote by $\mathbb{N}$ the set of natural numbers including 0.

Infinite matrices We use the notation of [10] for matrices indexed by sets other than natural numbers. Let I and J be two arbitrary sets. A (real) matrix M with columns indexed by J and rows are indexed by I is a map $M : I \times J \to \mathbb{R}$. The set of all such matrices is denoted by $\mathbb{R}^{I\times J}$. The entry of M indexed by the row index $i \in I$ and column index $j \in J$ is denoted by $M_{i,j}$ and it is defined as $M(i, j)$. In the sequel, $\mathbb{R}^{I\times n}$ ($\mathbb{R}^{n\times J}$) denotes the set $\mathbb{R}^{I\times\{1,2,\dots,n\}}$ (resp. $\mathbb{R}^{\{1,\dots,n\}\times J}$), if n is an integer. If K is a set, then the set of all maps of the form $K \to \mathbb{R}$ is denoted by $\mathbb{R}^K$. The column of M indexed by $j \in J$ is denoted by $M_{.,j}$ and is defined as a map $M_{.,j} \in \mathbb{R}^I$, $M_{.,j}(i) = M_{i,j}$, $i \in I$. Similarly, the row if M indexed by $i \in I$ is denoted by $M_{i,.}$ and is defined as $M_{i,.} \in \mathbb{R}^I$. $M_{i,.}(j) = M_{i,j}$ for all $j \in J$. Notice that $\mathbb{R}^K$ is a vector space with respect to point-wise addition and multiplication by scalar. Consider a matrix $M \in \mathbb{R}^{I\times J}$. We denote by $\mathrm{Im} M$ the linear subspace of $\mathbb{R}^I$ spanned by the columns of M. The *rank of* M, denoted by $\mathrm{rank}\ M \in \mathbb{N} \cup \{\infty\}$, is the dimension of $\mathrm{Im} M$.

Real algebraic geometry A subset $S \subseteq \mathbb{R}^n$ is *semi-algebraic* [2] if it is of the form

$$S = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \bigvee_{i=1}^{d} \bigwedge_{j=1}^{m_i} (P_{i,j}(x_1, \dots, x_n) \ \ \epsilon_{i,j} \ \ 0)\},$$

where for each $i = 1, \dots, d$ and $j = 1, \dots, m_i$ the symbol $\epsilon_{i,j} \in \{<, >, \leq, \geq, =\}$ and $P_{i,j} \in \mathbb{R}[X_1, \dots, X_n]$. Here $\bigvee$ stands for the *logical or* operator and $\bigwedge$ stands for the *logical and* operator. Let $A \subseteq \mathbb{R}^n, B \subseteq \mathbb{R}^m$ be two semi-algebraic sets. A map $f : A \to B$ is called *semi-algebraic* if its graph is a semi-algebraic subset of $\mathbb{R}^{n+m}$. By a *Nash function* we mean a smooth semi-algebraic function. A *Nash submanifold* M of $\mathbb{R}^n$ is a semi-algebraic subset of $\mathbb{R}^n$, such that M is a regular submanifold of $\mathbb{R}^n$, the coordinate neighbourhoods of M are semi-algebraic sets, and the coordinate functions are Nash functions.

III. Formal Power Series

The section presents basic results on formal power series. The material of this section is an extension of the classical theory of [1], [12]. For the current formalism, see [15], [14].

Let X be a finite set, which we will refer to as alphabet. A *formal power series* S with coefficients in $\mathbb{R}^p$ is a map

$$S : X^* \to \mathbb{R}^p$$

We denote by $\mathbb{R}^p \ll X^* \gg$ the set of all formal power series with coefficients in $\mathbb{R}^p$. The set $\mathbb{R}^p \ll X^* \gg$ is a vector space with point-wise addition and multiplication, i.e. if $\alpha, \beta \in \mathbb{R}$, $S, T \in \mathbb{R}^p \ll X^* \gg$, then the linear combination $\alpha S + \beta T$ is defined by $\forall w \in X^*$, $\alpha S(w) + \beta T(w)$.

Definition 1: Let J be an arbitrary (possibly infinite) set. A *family of formal power series in* $\mathbb{R}^p \ll X^* \gg$ *indexed by* J, abbreviated as **RFFPS** is a collection

$$\Psi = \{S_j \in \mathbb{R}^p \ll X^* \gg \mid j \in J\} \tag{1}$$

Let J be an arbitrary set and let $p > 0$. A *rational representation of type* p-J *over the alphabet* X is a tuple

$$R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in X}, B, C) \tag{2}$$

where $\mathcal{X}$ is a finite dimensional vector space over $\mathbb{R}$, for each letter $\sigma \in X$, $A_\sigma : \mathcal{X} \to \mathcal{X}$ is a linear map, $C : \mathcal{X} \to \mathbb{R}^p$ is a linear map, and $B = \{B_j \in \mathcal{X} \mid j \in J\}$ is a family of elements $\mathcal{X}$ indexed by J. If p and J are clear from the context we will refer to R simply as a *rational representation*. We call $\mathcal{X}$ the *state-space*, the maps A_σ, $\sigma \in X$ the *state-transition maps*, and the map C is called the *readout map* of R. The family B will be called the *indexed set of initial states of* R. The dimension $\dim \mathcal{X}$ of the state-space is called the *dimension* of R and it is denoted by $\dim R$. If $\mathcal{X} = \mathbb{R}^n$, then we identify the linear maps A_σ, $\sigma \in X$ and C with their matrix representations in the standard Eucledian bases, and we call them the *state-transition matrices* and the *readout matrix* respectively.

Notation 1: Let $A_\sigma : \mathcal{X} \to \mathcal{X}, \sigma \in X$ be linear maps and let $w \in X^*$. If $w = \epsilon$, then let A_ϵ be the identity map. If $w = v\sigma$ for some word $v \in X^*$ and letter $\sigma \in X$, then $A_{v\sigma} = A_\sigma A_v$.

Let Ψ be a **RFFPS** of the form (1). The representation R from (2) is said to be a *representation of* Ψ, if

$$\forall j \in J, \forall w \in X^* : S_j(w) = C A_w B_j \tag{3}$$

We say that the family Ψ is *rational*, if there exists a representation R such that R is a representation of Ψ. A representation R_{min} of Ψ is called *minimal* if for each representation R of Ψ, $\dim R_{min} \leq \dim R$. Define the subspaces

$$W_R = \mathrm{Span}\{A_w B_j \in \mathcal{X} \mid w \in X^*, |w| \leq n, j \in J\}$$
$$O_R = \bigcap_{w\in X^*, |w|\leq n} \ker C A_w$$

We will say that the representation R is *reachable* if $\dim W_R = \dim R$, and we will say that R is *observable* if $O_R = \{0\}$. Let $R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in X}, B, C)$, $\widetilde{R} = (\widetilde{\mathcal{X}}, \{\widetilde{A}_\sigma\}_{\sigma\in X}, \widetilde{B}, \widetilde{C})$ be two p-J rational representations. A linear isomorphism $T : \mathcal{X} \to \widetilde{\mathcal{X}}$ is called a *representation isomorphism*, and is denoted by $T : R \to \widetilde{R}$, if

$$T A_\sigma = \widetilde{A}_\sigma T, \forall \sigma \in X, \quad T B_j = \widetilde{B}_j, \forall j \in J, \quad C = \widetilde{C} T$$

If $T : R \to \widetilde{R}$ is an isomorphism, then $\widetilde{R}$ and R are representations of the same **RFFPS**, and R is observable (reachable) if and only if $\widetilde{R}$ is observable (reachable).

Remark 1: Let R be representation of Ψ of the form (2), and consider a vector space isomorphism $T : \mathcal{X} \to \mathbb{R}^n$, $n = \dim R$. Then $TR = (\mathbb{R}^n, \{TA_\sigma T^{-1}\}_{\sigma\in X}, TB, CT^{-1})$, where $TB = \{TB_j \in \mathbb{R}^n \mid j \in J\}$ is also a representation of Ψ. Moreover, $TA_\sigma T^{-1}$, $\sigma \in X$, CT^{-1} and TB_j, $j \in J$ can naturally be viewed as $n\times n$, $p\times n$ and $n\times 1$ matrices. Moreover, $T : R \to TR$ is a representation isomorphism. That is, we can always replace a representation of Ψ with an isomorphic representation, state-space of which is $\mathbb{R}^n$ for some n, and the parameters of which are matrices and real vectors.

Below we state the main results on existence and minimality of representations of **RFFPS** . We start with the definition of the concept of *Hankel matrix* of a **RFFPS** . Let Ψ be a **RFFPS** of the form (1).

Definition 2 (Hankel-matrix): Define the *Hankel-matrix* $H_\Psi \in \mathbb{R}^{(X^*\times I)\times(X^*\times J)}$, $I = \{1,\dots,p\}$ of Φ as the infinite matrix, the rows of which are indexed by pairs (v,i) where $v \in X^*$ and $i = 1,\dots,p$, and the columns of which are indexed by pairs (w,j) where $w \in X^*$ and $j \in J$. The entry $(H_\Psi)_{(v,i),(w,j)}$ of H_Ψ indexed with the row index (v,i) and the column index (w,j) is

$$(H_\Psi)_{(v,i)(w,j)} = (S_j(wv))_i \tag{4}$$

where $(S_j(wv))_i$ denotes the ith entry of the vector $S_j(wv) \in \mathbb{R}^p$. The rank of H_Ψ is understood as the dimension of the linear space spanned by the columns of H_Ψ, and it is denoted by $\operatorname{rank} H_\Psi$.

Theorem 1 (Existence and minimality, [14]): The family Ψ is rational, if and only if the rank of the Hankel-matrix H_Ψ is finite.

Assume that R_{min} is a representation of Ψ. Then R_{min} is a minimal representation of Ψ, if and only if R_{min} is reachable and observable. If R_{min} is minimal, then $\operatorname{rank} H_\Psi = \dim R_{min}$. In addition, all minimal representations of Ψ are isomorphic.

Next, we recall from [17] the notion of square summable formal power series. In the sequel, *we assume that J is a finite set* and Ψ is of the form (1). Consider a formal power series $S \in \mathbb{R}^p \ll X^* \gg$, and denote by $||\cdot||_2$ the Euclidean norm in $\mathbb{R}^p$. Consider the sequence,

$$L_n = \sum_{k=0}^{n} \sum_{\sigma_1\in X} \cdots \sum_{\sigma_k\in X} ||S(\sigma_1\sigma_2\cdots\sigma_k)||_2^2. \tag{5}$$

The series S is called *square summable*, if the $\lim_{n\to+\infty} L_n$ exists and is finite. We call Ψ *square summable*, if for each $j \in J$, the formal power series S_j is square summable.

Next we characterize square summability of a **RFFPS** in terms of its representation. Let R be a representation of Ψ of the form (2). We call R *stable*, if the matrix $\widetilde{A}_R = \sum_{\sigma\in X} A_\sigma \otimes A_\sigma$, is stable, i.e the eigenvalues λ of $\widetilde{A}_R$ lie inside the unit disk ($|\lambda| < 1$).

Theorem 2 ([17]): A **RFFPS** Ψ is square summable if and only if all minimal representations of Ψ are stable.

Notice the analogy with the case of linear systems, where the minimal realization of a stable transfer matrix is also stable. Consider the set $\mathcal{P}_s$ of *square summable* ***RFFPSs*** which are indexed by the elements of a fixed set J. It is clear that $\mathcal{P}_s$ is a vector space, if we define addition and multiplication by a scalar as follows. Let Ψ_1 be as in (1), let $\Psi_2 = \{T_j \in \mathbb{R}^p \ll X^* \gg \mid j \in J\}$. For each $\alpha,\beta \in \mathbb{R}$, let $\alpha\Psi_1 + \beta\Psi_2 = \{\alpha T_j + \beta S_j \in \mathbb{R}^p \ll X^* \gg \mid j \in J\}$. Now, for $\Psi_1, \Psi_2 \in \mathcal{P}_s$,, define the bilinear map

$$< \Psi_1, \Psi_2 >= \sum_{j\in J} \sum_{w\in X^*} S_j(w)^T T_j(w) \tag{6}$$

Lemma 1 ([17]): The map $< \cdot,\cdot >_J$ is a scalar product and $(\mathcal{P}_s, < \cdot,\cdot >)$ is a Hilbert space.

The following theorem gives a formula the scalar product of two families of formal power series in terms of the corresponding representations.

Theorem 3 ([17]): Consider two stable representations $R_i = (\mathbb{R}^{n_i}, \{A_{i,\sigma}\}_{\sigma\in X}, C_i, B_i)$, $i = 1,2$, and assume that for $i = 1,2$, R_i is a representation of **RFFPS** Ψ_i and that $B_i = \{B_{i,j} \in \mathbb{R}^{n_i} \mid j \in J\}$. Then there exists a unique solution $P \in \mathbb{R}^{n_1\times n_2}$ to the Sylvester equation

$$P = \sum_{\sigma\in X} A_{1,\sigma}^T P A_{2,\sigma} + C_1^T C_2 \tag{7}$$

and the scalar product $< \Psi_1, \Psi_2 >$ can be written as

$$< \Psi_1, \Psi_2 >= \sum_{j\in J} B_{1,j}^T P B_{2,j}. \tag{8}$$

IV. Main result

Below we state the main result of the paper. In the sequel, the integer p and the set J are fixed. Note that the set J is assumed to *finite*.

Definition 3 (Space of formal power series): Denote by M_n the set of all **RFFPSs** Ψ such that Ψ admits a minimal p-J representation of dimension n. Denote by SM_n the subset of of all square summable elements of M_n.

In order to present the main result, we will need the following notation.

Notation 2: Denote the set of all words over X of length at most N by $X^{\leq N}$, and let $\mathbb{R}^p \ll X^{\leq N} \gg$ be the set of functions $T : X^{\leq N} \to \mathbb{R}^p$.

It is easy to see that $T \in \mathbb{R}^p \ll X^{\leq N} \gg$ can be identified with a vector in $\mathbb{R}^{pM(N)}$ where $M(N) = \frac{|X|^{N+1}-1}{|X|-1}$ is the number of all words over X of length at most N. Similarly, a family $\mathrm{K} = \{T_j \in \mathbb{R}^p \ll X^{\leq N} \gg \mid j \in J\}$ can be identified with a vector in $\mathbb{R}^{p|J|M(N)}$. That is, the set of all such families can be identified with the space $\mathbb{R}^{p|J|M(N)}$. Define the map

$$\eta_N : \mathbb{R}^p \ll X^* \gg \to \mathbb{R}^p \ll X^{\leq N} \gg$$

which maps any power series to its restriction to $X^{\leq N}$, i.e for each $T \in \mathbb{R}^p \ll X^* \gg$, $\eta_N(T)(v) = T(v)$ for all $v \in X^{\leq N}$. We can extend η_N in a natural way to act on families of formal power series as follows. Define the map

$$\widetilde{\eta}_N : \mathrm{M}_n \to \mathbb{R}^{p|J|M(N)}$$

such that if Ψ is of the form (1), then

$$\widetilde{\eta}_N(\Psi) = \{\eta_N(S_j) \mid j \in J\}$$

By abuse of notation, *we denote* $\widetilde{\eta}_N$ *by* η_N.

Lemma 2 ([15]): If $2n+1 \leq N$, then the map $\eta_N : M_n \to \mathbb{R}^{p|J|M(N)}$ is injective.

Prompted by the lemma above, in the sequel we will use the following convention.

Convention 1: In the sequel, we will identify the set M_n with $\eta_{2n+1}(\mathrm{M}_n)$ and the set SM_n with the set $\eta_{2n+1}(\mathrm{SM}_n)$. In addition, we identify η_N with the map $\eta_N \circ \eta_{2n+1}^{-1}$.

Finally, we will define a map relating representations and elements of M_n. Recall from Remark 1 that if R is a representation of Ψ of dimension n, then R can be always replaced by an isomorphic representation, for which state-space is $\mathbb{R}^n$, the state-transition maps, the readout map are $n \times n$ and $p \times n$ matrices respectively and the initial states can be identified with a $|J|$-tuple of $n \times 1$ vectors. That is, there is one-to-one correspondence between such representations and elements of $\mathbb{R}^{|X|n^2+n|J|+np}$.

Definition 4: Denote by $L(n)$ the subset of $\mathbb{R}^{(|X|n+|J|+p)n}$ which corresponds to minimal representations R of the form (2) such that $\mathcal{X} = \mathbb{R}^n$, $A_\sigma \in \mathbb{R}^{n\times n}$, $\sigma \in X$, $B_j \in \mathbb{R}^n$, $j \in J$, $C \in \mathbb{R}^{p\times n}$.

Lemma 3: $L(n)$ is an open Nash-submanifold of $\mathbb{R}^{(|X|n+|J|+p)n}$ of dimension $(|X|n+|J|+p)n$.

Consider any $p-J$ representation R of the form (2), such that $R \in \mathbb{R}^{|X|n^2+|J|n+np}$. Then R determines a **RFFPS** Ψ_R of the form (1), such that $S_j(w) = CA_wB_j$ for all $w \in X^*$, $j \in J$. It then follows that R is a representation of Ψ_R.

Moreover, for any $N \in \mathbb{N}$, we can define

$$IO_N : \mathbb{R}^{|X|n^2+n|J|+np} \ni R \mapsto \eta_{2N+1}(\Psi_R) \in \mathbb{R}^{p|J|M(2N+1)}$$

i.e. IO_N maps R to the projection of the corresponding **RFFPS** . It is clear that each entry of $IO_N(R)$ is an entry of the vector $S_j(w) = CA_wB_j$ for some $w \in X^{\leq 2N+1}$, $j \in J$, and hence it is a polynomial in the parameters of R. That is, IO_N *is a polynomial map*. In the sequel we will be mostly interested in the case $N = n$. Notice that then $IO_n(L(n)) = \mathrm{M}_n$. The main result can be stated as follows.

Theorem 4 (Main result): **Manifold structure of** M_n**.** M_n is a regular Nash-submanifold of $\mathbb{R}^{p|J|M(2n+1)}$ of dimension $D(n) = n(|J|+p) + n^2(|X|-1)$.

Manifold structure of SM_n**.** The space SM_n is a semi-algebraic open subset of M_n, where M_n is considered with the topology of the corresponding manifold. Moreover, the subset topology of SM_n as a submanifold of M_n coincides with the topology of SM_n as a Hilbert-space with the scalar product from (6).

Embedding of M_n**.** For each $N \geq 2n+1$, the map $\eta_N : \mathrm{M}_n \to \mathbb{R}^{p|J|M(N)}$ is an injective Nash map.

Relationship between representations and M_n**.** The restriction of the map IO_n to $L(n)$ is a smooth semi-algebraic map from $L(n)$ to M_n.

Before proceeding to the proof of the theorem we would like to explain the significance of the above results.

Consequence of Theorem 4

The topology of M_n **is the natural topology** The theorem implies that two families $\Psi_1 = \{S_j \mid j \in J\}$ and $\Psi_2 = \{T_j \mid j \in J\}$ are close, if the values of the corresponding formal power series are close, i.e. for each $j \in J$ and $w \in X^*$, $S_j(w)$ and $T_j(w)$ are close. That is, in a sense the topology of M_n as a manifold is the natural topology.

The Hilbert-space topology of SM_n **is compatible with the topology of** M_n**.** The theorem implies that the natural topology of SM_n and its Hilbert-space topology coincide. This is important, because in many applications the Hilbert-space topology of SM_n corresponds to the operator norm topology for input-output maps of the corresponding system.

The Hilbert-space distance and the Eucledian distance in SM_n **are equivalent**. In addition to the Hilbert-space distance, we can also define the following distance on SM_n; $d(\Psi_1, \Psi_2) = ||\eta_{2n+1}(\Psi_1) - \eta_{2n+1}(\Psi_2)||_2$, i.e. we just take the usual Eucledian distance in $\mathbb{R}^{p|J|M(2n+1)}$. The results of the theorem imply that the latter distance induces the same topology as the Hilbert-space distance. In particular, it means that if we can approximate a **RFFPS** in one distance, we can do so in the other one.

Computability of the coordinate charts That M_n is a Nash submanifold implies that the differentiable structure of M_n is computable. This is important for developing parametric system identification and model reduction algorithms.

Relationship with parameters of representations The last statement of the theorem regarding the map IO_n implies that if $R \in L(n)$, then the corresponding **RFFPS** Ψ_R depends on R is a smooth way. In other words, the manifold structure of IO_n is consistent with the natural manifold structure of all minimal representations of dimension n.

M_n **as set of equivalence classes** Recall that if Ψ is a **RFFPS** from M_n, then Ψ admits a minimal representation of dimension n, moreover, all minimal representations of Ψ are isomorphic. That is, there is a one-to-one correspondence between elements of M_n and equivalence classes of minimal $p-J$ representations of dimension n, where two representations are considered equivalent if they are isomorphic. Since isomorphisms correspond to non-singular matrices, we can define an action of the Lie-group $GL(n)$ on the manifold $L(n)$, such that the set $L(n)/GL(n)$ corresponds to M_n. The theorem above implies that $L(n)/GL(n)$ is an analytic manifold. The latter was proven in [22] by showing that the action of $GL(n)$ on $L(n)$ is a proper action of a analytic Lie-group on an analytic manifold and use general theory of Lie-group actions [3].

V. Application to hybrid and non-linear systems

Below we present several applications of the above results to hybrid and non-linear systems.

Bilinear systems: Below, we consider SISO discrete-time bilinear systems, [20], [9]. The discussion below remains valid for MIMO systems and partially for continuous-time systems. Consider a bilinear system [20], [9].

$$\Sigma \begin{cases} x(t+1) = A_0x(t) + (A_1x(t))u(t),\ x(0) = x_0 \\ y(t) = Cx(t) \end{cases} \tag{9}$$

where $A_0, A_1 \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{1\times n}$, $x_0 \in \mathbb{R}^n$. It is well-known [20] that the input-output map $\mathbf{Y}$ of Σ can be encoded as the following **RFFPS** $\Psi_{\mathbf{Y}}$ with $J = \{0\}$, $X = \{0, 1\}$, $p = 1$, and $S_0(w) = S(w) = CA_w x_0$, $w \in X^*$, formed by the Volterra-kernels of Σ. There is a one-to-one correspondence between representations $R = (\mathbb{R}^n, \{A_0, A_1\}, \{x_0\}, C)$, where $p = 1$, $J = \{0\}$ and $X = \{0, 1\}$, and systems (9), see [20], [9].

With the identification above, M_n corresponds to the set of all input-output maps $\mathbf{Y}$ which can be realized by a minimal bilinear system of dimension n. Alternatively, M_n is the set of equivalence classes of minimal bilinear systems, where two systems are considered equivalent, if there exists a linear isomorphism between them (see [20], [9] for definition). Hence, the results of the paper imply that the space of bilinear systems has the structure of a Nash submanifold.

Hybrid systems: switched systems: For simplicity, we will discuss the application of our results to SISO discrete-time linear switched systems. However, the discussion below also holds for MIMO systems, and to a large extent for continuous-time systems as well. Recall from [16] that a SISO linear switched system (abbreviated by LSS),

$$\Sigma \begin{cases} x(t+1) = A_{q_t} x(t) + B_{q_t} u(t) \text{ and } x(0) = 0 \\ y(t) = C_{q_t} x(t) \end{cases} \tag{10}$$

Here $Q = \{1, \ldots, \mathcal{D}\}$ is the finite set of discrete modes, $q_t \in Q$ is the switching signal, $u(t) \in \mathbb{R}$ is the continuous input, $y(t) \in \mathbb{R}$ is the output and $A_q \in \mathbb{R}^{n \times n}$, $B_q \in \mathbb{R}^{n \times 1}$, $C_q \in \mathbb{R}^{1 \times n}$ are the matrices of the linear system in mode $q \in Q$. Finally, $x(0) = 0$ is the initial continuous state. Note that the switching signal $q(t)$ and the continous inputs $u(t)$ are both inputs, and $y(t)$ is the output of the system. If we denote by $\mathbf{Y}(w)$ the output of the system which corresponds to the sequence $w = (q_0, u(0)) \cdots (q_t, u(t))$, $t \geq 0$, then

$$\mathbf{Y}(w) = \sum_{k=0}^{t-1} S_{q_k, q_t}(q_{k+1} \cdots q_{t-1}) u(k)$$
$$S_{q_k, q_t}(q_{k+1} \cdots q_{t-1}) = C_{q_t} A_{q_{t-1}} \cdots A_{q_{k+1}} B_{q_k}$$

We can view $\mathbf{Y}$ as a map which maps the sequence $(u(t))_{t=1}^{\infty}$ to the map $y : Q^+ \to \mathbb{R}$, where $y(q_0 \cdots q_t) = \mathbf{Y}((q_0, u(0)) \cdots (q_t, u(t))$. Let $l^{\infty}(Q^+)$ be the space of all those maps $f : Q^+ \to \mathbb{R}$ such that $||f|| = \sup_{w \in Q^+} |f(w)| < +\infty$. It is clear that $l^{\infty}(Q^+)$ is a Banach-space with the norm $||f|| = \sup_{w \in Q^+} |f(w)|$. If the system Σ is *stable* in the sense that the matrix $\widetilde{A}_{\Sigma} = \sum_{q \in Q} A_q \otimes A_q$ is stable, then $\mathbf{Y}$ is a linear map from l^2 to $l^{\infty}(Q^+)$ and we can define its induced norm,

$$||\mathbf{Y}|| = \sup_{u \in l^2, ||u|| = 1} ||Y(u)|| \tag{11}$$

Recall the one-to-one correspondence between $\mathbf{Y}$ and **RFFPS** $\Psi_{\mathbf{Y}} = \{S_j \in \mathbb{R}^{|Q|} \ll Q^* \gg | \ j \in J = Q\}$, where

$$S_q(v) = \begin{bmatrix} S_{q,1}(v) & \ldots S_{q,\mathcal{D}}(v) \end{bmatrix}^T$$

In addition, recall from [16], that Σ can be associated with a representation R_{Σ} of the form (2), where $\mathcal{X} = \mathbb{R}^n$, $X = Q$, the state-transition matrices A_q of R are the same as the matrices of Σ, $C = \begin{bmatrix} C_1 & \ldots & C_{\mathcal{D}} \end{bmatrix}^T$, $B = \{B_q \mid q \in Q\}$. Conversely, for any $\mathcal{D} - Q$ representation R of the form (2), such that the state-space of R is $\mathbb{R}^n$, we can define a LSS Σ_R, such that the representation R_{Σ_R} equals R.

By [16], $\mathbf{Y}$ is the input-output map of a LSS Σ if and only if R_{Σ} is a representation of $\Psi_{\mathbf{Y}}$. Hence, there is a one-to-one correspondence between minimal LSS realizations of $\mathbf{Y}$ and minimal representations of $\Psi_{\mathbf{Y}}$. Note that Σ is stable if and only if R is stable, and hence $\Psi_{\mathbf{Y}}$ is square summable. Moreover, if $||\Psi_{\mathbf{Y}}||$ denotes the Hilbert-space norm of $\Psi_{\mathbf{Y}}$,

$$\frac{1}{\mathcal{D}\sqrt{M(N)}} ||\eta_N(\Psi_{\mathbf{Y}})||_2 \leq ||\mathbf{Y}|| \leq ||\Psi_{\mathbf{Y}}|| \tag{12}$$

where $M(N)$ is the number of all words over Q of length at most N. Hence, *we can identify input-output maps which are realizable by minimal LSSs of dimension n with the set M_n, where $X = Q$, $p = \mathcal{D}$ and $J = Q$*. In addition, the set of input-output maps which are realizable by a stable minimal LSSs of dimension n can be identified with the set SM_n. Moreover, Theorem 4 in combination with (12) implies that the topology of the set of input-output maps induced by the operator norm and the topology of SM_n are equivalent. Hence, we immediately get that the set of all input-output maps of minimal switched systems of order n form a manifold of dimension $(\mathcal{D} - 1)n^2 + 2\mathcal{D}n)$. Moreover, the input-output maps which admit a stable realization form an open subset of this manifold. In addition, the input-output maps in the topology of the above manifolds depend continuously on the parameters of minimal state-space realizations, where the latter parameters are viewed in the usual Eucledian topology.

Hidden Markov Models and Jump-Linear Systems: Recall from [11] that there is a one-to-one correspondence between stable rational representations and stationary identically distributed jump-linear systems (abbreviated as iid-JLS). There is a one-to-one correspondence between the second-order moments of the output process of iid-JLS and square summable **RFFPS** over an alphabet X, with $|J| = 1$ and $p = 1$. In [11] a distance measure for the output process of iid-JLSs was proposed. That distance coincides with the Hilbert-space distance of square-summable **RFFPSs**. Hence, the results of the paper give a characterization of the manifold structure of equivalence classes of minimal iid-JLS, where two systems are equivalent, if the second-order moments of their output processes are the same. Simillar correspondence can be formulated for mode general jump-markov linear systems, using the results of [18], [17]. Note that in [11], a transformation from Hidden-Markov models (HMM) to iid-JLS was defined, such that the probability distribution of the output of a HMM corresponds to the second-order moments of the output of the iid-JLS. Hence, the results of the paper have implications for the structure of the set of probability distributions of HMMs.

VI. Nice selection and local coordinates of M_n

In this section we prepare the ground for the proof of Theorem 4. The proof relies on partial-realization theory of formal power series and the notion of nice selection.

The former has appeared in a slightly different form [19], [20], [7], see [15] for the current setting. The notion of nice selection is an extension of related notion from linear theory.

Partial-realization theory for RFFPSs : Below we recall from [15] partial-realization theory for **RFFPSs** . In the sequel, Ψ is of the form (1), and H_Ψ is its Hankel-matrix.

Definition 5: Let $\mathbb{N} \ni M, K > 0$ and define the matrix $H_{\Psi,M,K} \in \mathbb{R}^{I_M \times J_K}$, such that

$$I_M = \{(v,i) \mid v \in X^{\leq M}, i = 1,\dots,p\}$$
$$J_K = \{(w,j) \mid j \in J, w \in X^{\leq K}\} \quad (13)$$
$$\forall l \in J_K, k \in I_M : (H_{\Psi,M,K})_{k,l} = (H_\Psi)_{k,l}$$

That is, $H_{\Psi,M,K}$ is the left upper $I_M \times J_K$ block matrix of H_Ψ. If J is finite, then $H_{\Psi,M,K}$ is a *finite matrix*.

Theorem 5 ([15]): If $\mathrm{rank}\ H_{\Psi,N,N} = \mathrm{rank}\ H_\Psi$, then there exists a minimal representation R_N of Ψ of the form (2), such that state-space $\mathcal{X} = \mathrm{Im} H_{\Psi,N,N+1}$, and the following holds. Let $\mathbf{C}_{w,j}$ be the column of $\mathrm{Im}H_{\Psi,N,N}$ indexed by $(w,j) \in J_N$. Then for each $(w,j) \in J_N$,

$$A_\sigma(\mathbf{C}_{w,j}) = \mathbf{C}_{w\sigma,j}$$
$$C(\mathbf{C}_{w,j}) = \begin{bmatrix} \mathbf{C}_{w,j}((\epsilon,1)), & \cdots, & \mathbf{C}_{w,j}((\epsilon,p)) \end{bmatrix}^T$$

and $\forall j \in J : B_j = \mathbf{C}_{\epsilon,j}$ [1]. If $\mathrm{rank}\ H_\Psi \leq N$, then $\mathrm{rank}\ H_{\Psi,N,N} = \mathrm{rank}\ H_\Psi$.

Nice selection: In the rest of the paper we assume the following.

Assumption 1: Assume that $J = \{1,2,\dots,m\}$. and fix a complete ordering on the finite sets $X^{\leq n} \times J$ and $X^{\leq n} \times I$.

Definition 6 (Column nice selection): A finite subset α of $X^{\leq n} \times J$ is called a nice column selection, if α has precisely n elements and α has the following property; if $(w\sigma, j) \in \alpha$ for $j \in J$, some word $w \in X^*$ and letter $\sigma \in X$, then $(w,j) \in \alpha$. Using the ordering from Assumption 1, we order the elements of α as $(w_1,j_1) < (w_2,j_2) < \cdots < (w_n,j_n)$.

Definition 7 (Row nice selection): A finite subset β of $X^{\leq n} \times I$ is called a nice row selection, if β has precisely n elements and β has the following property; if $(\sigma w, i) \in \beta$ for some word $w \in X^*$, letter $\sigma \in X$ and $i \in I$, then $(w,i) \in \beta$. Using the ordering from Assumption 1, we order the elements of β as $(v_1,i_1) < (v_2,i_2) < \cdots < (v_n,i_n)$.

The purpose of nice selections is to describe indices of $H_{\Psi,n,n}$, such that the corresponding sub-matrix of $H_{\Psi,n,n}$ has rank n. Below we present a number of results which employ the concept of nice selection.

Definition 8 (Hankel sub-matrix): Assume that α is a subset of $X^{\leq n} \times J$ and β is a subset of $X^{\leq n} \times I$, and α has r elements and β has q elements. Denote by $H_{\Psi,\beta,\alpha}$ the following $q \times r$ sub-matrix of $H_{\Psi,n,n}$; the element of $H_{\Psi,\beta,\alpha}$ indexed by (k,l) is $(H_{\Psi,n,n})_{(v_k,i_k),(w_l,j_l)}$. Here, we used the ordering $\alpha = \{(w_1,j_1) < \cdots < (w_r,j_r)\}$ and $\beta = \{(v_1,i_1) < \cdots < (v_q,i_q)\}$, using the ordering of $X^{\leq n} \times J$ and $X^{\leq n} \times I$ fixed in Assumption 1.

If α and β are nice selections, the the definition above applies and $H_{\Psi,\beta,\alpha}$ is a sub-matrix of H_Ψ. Using this submatrix, we define the following minimal representation $R_{\alpha,\beta}$ of Ψ.

[1]Recall that $C_{w,j}((\epsilon,i))$ is the entry of the column $\mathbf{C}_{w,j}$ indexed by (ϵ,i).

Construction 1: Consider the nice column selection α and nice row selection β. Assume that $\mathrm{rank}\ H_{\Psi,\beta,\alpha} = n$. Define $T_{\alpha,\beta} : \mathrm{Im}H_{\Psi,n,n} \to \mathbb{R}^n$ by

$$T_{\alpha,\beta}(S) = H^{-1}_{\Psi,\beta,\alpha} \begin{bmatrix} S((v_1,i_1)) & \cdots & S((v_n,i_n)) \end{bmatrix}^T$$

for all $S \in \mathrm{Im}H_{\Psi,n,n}$. For simplicity, we denote $T_{\alpha,\beta}$ by T. Consider the isomorphic copy

$$R_{\alpha,\beta} = TR_n = (\mathbb{R}^n, \{TA_\sigma T^{-1}\}_{\sigma \in X}, TB, CT^{-1})$$

of the representation R_n from Theorem 5, where $TB = \{TB_j \in \mathbb{R}^n \mid j \in J\}$. Then the elements of $R_{\alpha,\beta}$ are as follows. Let $\alpha(\sigma) = \{(w\sigma,j) \mid (w,j) \in \alpha\}$, $\gamma = \{(\epsilon,1),(\epsilon,2),\dots(\epsilon,p)\}$. Then

$$TA_\sigma T^{-1} = H^{-1}_{\Psi,\beta,\alpha} H_{\Psi,\beta,\alpha(\sigma)} \text{ for all } \sigma \in X$$
$$CT^{-1} = H_{\Psi,\gamma,\alpha} \text{ and } TB_j = H_{\Psi,\beta,\{(\epsilon,j)\}} \text{ for all } j \in J.$$

Lemma 4: The representation $R_{\alpha,\beta}$ from Construction 1 is isomorphic to the representation R_n from Theorem 5.

Definition 9: Consider a nice column selection α and a nice row selection β. Define $V_{\alpha,\beta}$ as the subset of all those families $\Psi \in \mathrm{M}_n$, for which $\mathrm{rank}\ H_{\Psi,\beta,\alpha} = n$.

The motivation for introducing the sets $V_{\alpha,\beta}$ is that they will be the coordinate neighbourhoods of M_n. We conclude with the definition of α-reachability and β-observability for rational representations.

Definition 10 (β-observability, α-reachability): Let α be a nice column selection and β be a nice row selection. Consider a representation $R \in \mathbb{R}^{|X|n^2+n|J|+np}$. Denote by $O_{R,\beta}$ the following $n \times n$ matrix

$$O_{R,\beta} = \begin{bmatrix} (C_{i_1,.}A_{v_1})^T & \dots & (C_{i_n,.}A_{v_n})^T \end{bmatrix}^T$$

where $C_{k,.}$ denotes the kth row of C for some $k = 1,\dots,p$. Denote by $W_{R,\alpha}$ the following $n \times n$ matrix

$$W_{R,\alpha} = \begin{bmatrix} A_{w_1}B_{j_1} & \dots & A_{w_n}B_{j_n} \end{bmatrix}$$

We say that R is α-reachable, if $\mathrm{rank}\ W_{R,\alpha} = n$, and we say that R is β-observable, if $\mathrm{rank}\ O_{R,\beta} = n$.

Remark 2: If R is β-observable, then R is observable. Similarly, if R is α-reachable, then R is reachable.

The relationship between the set $V_{\alpha,\beta}$ and the notions of α-reachability and β-observability is as follows.

Lemma 5: If R is a representation of $\Psi \in \mathrm{M}_n$, then

$$O_{R,\beta} W_{R,\alpha} = H_{\Psi,\beta,\alpha}$$

In particular, $\eta_{2n+1}(\Psi) \in V_{\alpha,\beta}$ if and only if Ψ has a minimal representation R of dimension n, such that R is α-reachable and β-observable.

In other word, the choice of a full-rank minor of $H_{\Psi,n,n}$ is directly related to the choice of a basis in the observability and reachability subspaces of a minimal representation of Ψ.

Coordinate charts of M_n*:* Below we define coordinate functions $\phi_{\alpha,\beta} : V_{\alpha,\beta} \to \mathbb{R}^{D(n)}$, where $D(n)$ is the dimension of M_n defined in Theorem 4, such that the collection $(V_{\alpha,\beta}, \phi_{\alpha,\beta})$ with α ranging through all nice column selections and β ranging through all nice row selections form a semi-algebraic and smooth system of coordinate charts.

Lemma 6: For any $\Psi \in \mathrm{M}_n$, there exists a nice column selection α and a nice row selection β such that $\Psi \in V_{\alpha,\beta}$.

Lemma 7: The set $V_{\alpha,\beta}$ is semi-algebraic open subset of M_n, if M_n is considered with the subset topology.

Next, we define a coordinate function on $V_{\alpha,\beta}$.

Definition 11: Define the map $\psi_{\alpha,\beta} : V_{\alpha,\beta} \to L(n)$ by $\psi_{\alpha,\beta}(\Psi) = R_{\alpha,\beta}$, where $R_{\alpha,\beta}$ is the representation from Construction 1.

It turns out the the representation $\phi_{\alpha,\beta}(\Psi)$ has a very specific structure, analogous to the controllability canonical form for linear systems. More specifically, there is a small number of parameters of $\phi_{\alpha,\beta}(\Psi)$ which depend on Ψ, all the other parameters depend only on the nice selections α and β. In order to present this structure, we need additional notation.

Notation 3: For each $\sigma \in X$ and denote by E_σ the set of all indices $i = 1, 2 \ldots, n$ such that $(w_i\sigma, j_i) \notin \alpha$. For each $i \notin E_\sigma$, define $\delta(i,\sigma) = k$, where $(w_i\sigma, j_i) = (w_k, j_k) \in \alpha$ for some $k = 1, \ldots, n$. Let J_c be the set of all $j \in J$ such that $(\epsilon, j) \notin \alpha$.

Using the notation above, we can state the following result about the parameters of $\psi_{\alpha,\beta}(\Psi)$.

Theorem 6 (Canonical form): Assume that $\Psi \in V_{\alpha,\beta}$ and the representation $R = \psi_{\alpha,\beta}(\Psi)$ is of the form (2). Then the following holds.

- For each $\sigma \in X$, and for each $k = 1, \ldots, n$, the kth column of the matrix A_σ is of the form

$$(A_\sigma)_{.,k} = \begin{cases} e_{\delta(k,\sigma)} & \text{if } k \notin E_\sigma \\ x_{\sigma,k} & \text{if } k \in E_\sigma \end{cases} \qquad (14)$$

  where e_r denotes the rth unit vector of $\mathbb{R}^n$ and

$$x_{\sigma,k} = H_{\beta,\alpha}^{-1}\left[H_{(v_1,i_1),(w_k\sigma,j_k)}, \ldots H_{(v_n,i_n),(w_k\sigma,j_k)}\right]^T$$

- The C matrix is as described in Construction 1
- For each $j \notin J_c$, $B_j = e_k$ where $(w_k, j_k) = (\epsilon, j)$ for some $k = 1, 2, \ldots, n$. For each $j \in J_c$,

$$B_j = H_{\beta,\alpha}^{-1}\left[H_{(v_1,i_1),(\epsilon,j)}, \ldots H_{(v_n,i_n),(\epsilon,j)}\right]^T$$

Here $H = H_{\Psi,n,n}$ and $H_{\beta,\alpha} = H_{\Psi,\beta,\alpha}$.

The theorem above implies that in order to store $\psi_{\alpha,\beta}(\Psi)$, it is enough to store the matrix C, those columns of A_σ index of which belongs to E_σ and those vectors B_j, for which $j \in J_c$. In total, on needs $pn + |J_c|n + n\sum_{\sigma\in X}|E_\sigma|$ parameters to encode the representation $\psi_{\alpha,\beta}(\Psi)$. In fact, the latter number equals $D(n)$ for all α and β.

Lemma 8: $D(n) = n(\sum_{\sigma\in X}|E_\sigma|) + pn + |J_c|n$.

That is, a representation of Ψ corresponding to the nice selections α and β can be encoded by $D(n)$ parameters. With the notation above, we can define the map

$$\phi_{\alpha,\beta} : V_{\alpha,\beta} \to \mathbb{R}^{D(n)}$$

such that for all $\Psi \in V_{\alpha,\beta}$,

$$\phi_{\alpha,\beta}(\Psi) = (\{x_{\sigma,i}\}_{\sigma\in X, i\in E_\sigma}, \{B_j \mid j \in J_c\}, C) \in \mathbb{R}^{D(n)}$$

where $x_{\sigma,i}, B_j$ and C are as in Theorem 6. Here we view the collection of parameters which describe the representation R of Theorem 6 as a vector with $D(n)$ entries.

Conversely, consider a vector $w \in \mathbb{R}^{D(n)}$. We can always view the vector w as the collection of real numbers and vectors $w = (\{x_{\sigma,i}\}_{\sigma\in X, i\in E_\sigma}, \{B_j \mid j \in J_c\}, C)$. Here $x_{\sigma,i} \in \mathbb{R}$, $B_j \in \mathbb{R}^n$, $j \in J$, $C \in \mathbb{R}^{p\times n}$. Using this interpretation, we can associate with w the rational representation R_w of the form (2), such that for each $\sigma \in X$, A_σ is defined as in (14) and for each $j \notin J_c$, $B_j = e_i$, where $(j_i, w_i) = (j, \epsilon)$ and e_i is the ith unit vector of $\mathbb{R}^n$. Finally, recall that every $p - J$ representation can be identified with a vector in $\mathbb{R}^{|X|n^2+n|J|+pn}$. Hence, we obtain the map

$$\nu_{\alpha,\beta} : \mathbb{R}^{D(n)} \ni w \mapsto \mathbb{R}^{|X|n^2+n|J|+pn}$$

Note that the entries of w enter the parameters of R_w in a linear manner, hence, the map $\nu_{\alpha,\beta}$ above is linear. Also notice that $\nu_{\alpha,\beta} \circ \phi_{\alpha,\beta}(\Psi) = \psi_{\alpha,\beta}(\Psi) \in L(n)$.

Next, we define the following subspace of $\mathbb{R}^{D(n)}$ which characterizes the image of $\phi_{\alpha,\beta}$.

Lemma 9: For every $w \in \mathbb{R}^{D(n)}$, $R = \nu_{\alpha,\beta}(w)$ is α-reachable, in fact $W_{R,\alpha} = I_n$. In addition, $R = \nu_{\alpha,\beta}(w)$ is β-observable if and only if $IO_n(R) \in V_{\alpha,\beta}$.

Corollary 1: If $\Psi \in V_{\alpha,\beta}$, then the representation $\psi_{\alpha,\beta}(\Psi)$ is α-reachable and β-observable.

Lemma 10: The map $\phi_{\alpha,\beta}$ is a one-to-one Nash map from $V_{\alpha,\beta}$ onto $\nu_{\alpha,\beta}^{-1}(W_{\alpha,\beta})$, where $W_{\alpha,\beta}$ is the subset of α-reachable and β-observable representations from $\mathbb{R}^{|X|n^2+|J|n+pn}$. The set $\nu_{\alpha,\beta}^{-1}(W_{\alpha,\beta})$ is an open semi-algebraic subset of $\mathbb{R}^{D(n)}$, and $\phi_{\alpha,\beta}$ is a Nash diffeomorphism from $V_{\alpha,\beta}$ onto $\nu_{\alpha,\beta}^{-1}(W_{\alpha,\beta})$, if $V_{\alpha,\beta}$ is viewed with the subset topology of $\mathbb{R}^{p|J|M(2n+1)}$.

The above result indicates the the pairs $(V_{\alpha,\beta}, \phi_{\alpha,\beta})$ can be viewed as coordinate functions of M_n. In order to finish the construction, we need to show that the coordinate transformation between various charts is consistent.

Lemma 11: Consider nice column selections α, γ, and nice row selections β, η. The map $\phi_{\alpha,\beta} \circ \phi_{\gamma,\eta}^{-1} : \phi_{\gamma,\eta}(V_{\alpha,\beta} \cap V_{\gamma,\eta}) \to \phi_{\alpha,\beta}(V_{\alpha,\beta} \cap V_{\gamma,\eta})$ is a Nash diffeomorphism.

VII. PROOF OF THEOREM 4

In this section we present the proof of Theorem 4. We proceed with the claims of Theorem 4 one by one.

Manifold structure of M_n*:* Consider M_n with the subset topology. Then M_n is second countable and Hausdorff, i.e. it is a topological manifold. Moreover, by Lemma 6, M_n is the union of the sets $V_{\alpha,\beta}$, where α goes through the set of all nice column selections and β goes through the set of all nice row selections. Notice that the set of all nice column or row selections is finite. By Lemma 7, each $V_{\alpha,\beta}$ is a semi-algebraic set which is open in the relative topology of M_n. Hence, M_n is a semi-algebraic set itself. Finally, consider the family of coordinate charts $(V_{\alpha,\beta}, \phi_{\alpha,\beta})$. By Lemma 10, $\phi_{\alpha,\beta}$ is a Nash diffeomorphism from $V_{\alpha,\beta}$ into $\mathbb{R}^{D(n)}$. Finally, Lemma 11 implies that the coordinate charts are consistent. Hence, M_n is an analytic manifold, in fact it is a Nash-submanifold of $\mathbb{R}^{p|J|M(2n+1)}$.

Manifold structure of SM_n*:* It is enough to show that for any nice column selection α and row selection β, the intersection $\mathrm{SM}_n \cap V_{\alpha,\beta}$ is open in M_n. To this end, consider the map $\psi_{\alpha,\beta} : V_{\alpha,\beta} \rightarrow L(n)$. This map is smooth semi-algebraic, in particular, it is continuous. Recall the notion of stable representation and denote by $S(n)$ the set of all stable and minimal representations in $\mathbb{R}^{|X|n^2+np+n|J|}$. The claim that $\mathrm{SM}_n \cap V_{\alpha,\beta}$ is open follows from the following results.

Lemma 12: The set $S(n)$ is an open subset of $L(n)$ in the subset topology of $L(n)$.

Lemma 13: $\mathrm{SM}_n \cap V_{\alpha,\beta} = \psi^{-1}_{\alpha,\beta}(S(n))$

Indeed, Lemma 12, 13 and the continuity of $\psi_{\alpha,\beta}$ implies that $\mathrm{SM}_n \cap V_{\alpha,\beta}$ is open in the subset topology of M_n.

Next, we show that the topology of SM_n as a subset of M_n and the topology of SM_n as a Hilbert-space coincide. More precisely, we show that a subset of SM_n is open in its Hilbert-space topology if and only this subset is open in the subset topology. To this end, it is enough to show that any ball in the Hilbert-space topology of SM_n is contained in a ball defined for the Eucledian metric inherited from $\mathbb{R}^{p|J|M(2n+1)}$ and vice versa. The latter follows from the following statements.

Lemma 14: For any $\Psi \in \mathrm{SM}_n$, let $B_r(\Psi)$ be the open ball of radius r in SM_n, if SM_n is considered with the norm of $\mathbb{R}^{p|J|M(2n+1)}$. Let $B^H_r(\Psi)$ be the open ball of radius r in SM_n, if SM_n is viewed as a subset of the Hilbert-space, i.e.

$$B_r(\Psi) = \{\Psi_1 \in \mathrm{SM}_n \mid ||\eta_{2n+1}(\Psi_1) - \eta_{2n+1}(\Psi)||_2 < r\}$$
$$B^H_r(\Psi) = \{\Psi_1 \in \mathrm{SM}_n \mid ||\Psi_1 - \Psi|| < r\}$$

where $||\Phi|| = \sqrt{<\Phi,\Phi>}$ for any square summable **RFFPS** Φ. Then $B^H_r(\Psi) \subseteq B_r(\Psi)$.

Lemma 15: The scalar product $< .,. >$: $\mathrm{SM}_n \times \mathrm{SM}_n \rightarrow \mathbb{R}$ is a Nash function. In particular, $< .,. >$ is continious if SM_n is considered with the subset topology.

The proof of Lemma 15 relies on noticing that

1) the formula of Theorem 2 implies that $< \Psi_1, \Psi_1 >$ is smooth and semi-algebraic in the parameters of the minimal stable representations of Ψ_1, Ψ_2, and
2) For any nice column selection α, γ and nice row selection β, η, the map $V_{\alpha,\beta} \times V_{\gamma,\eta} \ni (\Psi_1, \Psi_2) \mapsto (\psi_{\alpha,\beta}(\Psi_1), \psi_{\gamma,\eta}(\Psi_2)) \in S(n) \times S(n)$, mapping pairs **RFFPSs** to the corresponding representations, is a Nash function.

Corollary 2: For any $\Psi \in \mathrm{SM}_n$, for any $r > 0$ there exists $\delta > 0$ such that $B_\delta(\Psi) \subseteq B^H_r(\Psi)$.

Embedding of M_n*:* The restriction of the map $\eta_N : \mathrm{M}_n \rightarrow \mathbb{R}^{p|J|M(N)}$ to $V_{\alpha,\beta}$ for some nice selections α and β can be written as follows; $\eta_N(\Psi) = IO_N \circ \psi_{\alpha,\beta} \circ \eta_{2n+1}(\Psi)$. Since $\psi_{\alpha,\beta} = \nu_{\alpha,\beta} \circ \phi_{\alpha,\beta}$ is a Nash map, we get that the restriction of η_N to $V_{\alpha,\beta}$ is a Nash map, i.e. it is smooth and semi-algebraic. Since the open semi-algebraic sets $V_{\alpha,\beta}$ cover M_n, we get that η_N is a Nash-map.

Relationship between representations and M_n*:* Note that the map $IO_n : L(n) \rightarrow \mathrm{M}_n$ is smooth, and it is semi-algebraic, if viewed as a map $IO_n : L(n) \rightarrow \mathbb{R}^{p|J|M(2n+1)}$. Since M_n is a regular submanifold of $\mathbb{R}^{p|J|M(2n+1)}$ it then follows that IO_n is a smooth map with the differentiable structure of M_n.

VIII. Conclusions

We have presented a characterization of the manifold structure of spaces of formal power series and their representations. We have also shown that the obtained results are relevant for characterizing the space of non-linear and hybrid systems. Future research is directed to application of the presented results to model reduction and system identification of control systems which can be encoded by rational formal power series.

References

[1] J. Berstel and C. Reutenauer. *Rational series and Their Languages*. Springer-Verlag, 1984.
[2] M. Coste, J. Bochnak, and M-F. Roy. *Real Algebraic Geometry*. Ergebnisse der Math. Springer Verlag, 1998.
[3] J. J. Duistermaat and Johan A. C. Kolk. *Lie Groups*. Springer Verlag, 2000.
[4] Samuel Eilenberg. *Automata, Languages and Machines*. Academic Press, New York, London, 1974.
[5] F. Gécseg and I Peák. *Algebraic theory of automata*. Akadémiai Kiadó, Budapest, 1972.
[6] B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems*, volume 63,64 of *CWI Tract*. CWI, Amsterdam, 1989.
[7] A. Isidori. Direct construction of minimal bilinear realizations from nonlinear input-output maps. *IEEE Transactions on Automatic Control*, pages 626–631, 1973.
[8] Alberto Isidori. *Nonlinear Control Systems*. Springer Verlag, 1989.
[9] Alberto Isidori, Paolo D'Alessandro, and Antonio Ruberti. Realization and structure theory of bilinear dynamical systems. *SIAM J. Control*, 12(3), 1974.
[10] Nathan Jacobson. *Lectures in Abstract Algebra*, volume II: linear algebra. D. van Nostrand Company, Inc. New York, 1953.
[11] Georgios Kotsalis. *Model Reduction for Hidden Markov Models*. PhD thesis, Massachusets Institute of Technology, 2006.
[12] W. Kuich and A. Salomaa. *Semirings, Automata, Languages*. Springer-Verlag, 1986.
[13] Ralf Peeters. *System Identification Based on Riemannian Geometry: Theory and Algorithms*. PhD thesis, Free University, Amsterdam, 1994.
[14] M. Petreczky. Realization theory of linear and bilinear switched systems: A formal power series approach. To appear in ESAIM Control, Optimization and Calculus of Variations.
[15] M. Petreczky. *Realization Theory of Hybrid Systems*. PhD thesis, Vrije Universiteit, Amsterdam, 2006.
[16] M. Petreczky, L. Bako, and J.H. van Schuppen. Identifiability of discrete-time linear switched systems. In *Hybrid Systems: Computation and Control*. ACM, 2010.
[17] M. Petreczky and R. Vidal. Metrics and topology for nonlinear and hybrid systems. In *Hybrid Systems: Computation and Control*, 2007.
[18] M. Petreczky and R. Vidal. Realization theory of stochastic jump-markov linear systems. In *Proc. 46th IEEE Conference on Decision and Control*, 2007.
[19] Eduardo D. Sontag. *Polynomial Response Maps*, volume 13 of *Lecture Notes in Control and Information Sciences*. Springer Verlag, 1979.
[20] Eduardo D. Sontag. Realization theory of discrete-time nonlinear systems: Part I – the bounded case. *IEEE Trans. on Circuits and Systems*, CAS-26(4), 1979.
[21] Eduardo D. Sontag. A remark on bilinear systems and moduli spaces of instantons. *Systems and Control Letters*, 9(5):361–367, 1987.
[22] Hector J. Sussmann. Minimal realizations and canonical forms for bilinear systems. *Journal of the Franklin Institute*, 301(1):593–604, 1976.