ESAIM: COCV 17 (2011) 410–445
DOI: 10.1051/cocv/2010014

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

REALIZATION THEORY FOR LINEAR AND BILINEAR SWITCHED SYSTEMS: A FORMAL POWER SERIES APPROACH PART I: REALIZATION THEORY OF LINEAR SWITCHED SYSTEMS *

Mihály Petreczky[1]

Abstract. The paper represents the first part of a series of papers on realization theory of switched systems. Part I presents realization theory of linear switched systems, Part II presents realization theory of bilinear switched systems. More precisely, in Part I necessary and sufficient conditions are formulated for a family of input-output maps to be realizable by a linear switched system and a characterization of minimal realizations is presented. The paper treats two types of switched systems. The first one is when all switching sequences are allowed. The second one is when only a subset of switching sequences is admissible, but within this restricted set the switching times are arbitrary. The paper uses the theory of formal power series to derive the results on realization theory.

Mathematics Subject Classification. 93B15, 93B20, 93B25, 93C99.

Received July 29, 2005. Revised January 3rd, 2008 and September 12, 2009.
Published online March 31, 2010.

1. Introduction

In Part I and Part II of the current series of papers we develop realization theory for the linear switched systems and bilinear switched systems. Realization theory is one of central topics of systems theory. In addition to its theoretical relevance, realization theory is potentially useful for control, model reduction, and systems identification. Switched systems are one of the best studied subclasses of hybrid systems, see [15] for a survey.

Problem statement. We address the following problems.

(1) **Existence and minimality: arbitrary switching.** Find conditions for the existence and minimality of a linear (bilinear) switched system realizing a given set of input-output maps Φ.
(2) **Existence and minimality: constrained switching.** Assume that a set of admissible switching sequences is defined. Assume that the switching times of the admissible switching sequences are arbitrary. Consider a set of input-output maps Φ defined only for the admissible sequences. Find conditions for the existence and minimality of a linear (bilinear) switched system realizing Φ.

The motivation of the Problem 2 is the following. Assume that the switching is controlled by a finite automaton and the discrete modes are the states of this automaton. Assume that the automaton is driven by external

Keywords and phrases. Hybrid systems switched linear systems, switched bilinear systems, realization theory, formal power series, minimal realization.

* *The research presented in this paper is part of the author's Ph.D. thesis [20] and it was carried out during the author's stay at Centrum voor Wiskunde en Informatica (CWI), Amsterdam, The Netherlands.*

[1] Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. M.Petreczky@maastrichtuniversity.nl

events, which can trigger a discrete-state transition at any time. Then the traces of this automaton combined with arbitrary switching times give us the admissible switching sequences. If we can solve Problem 2 for which the corresponding set of admissible sequences of discrete modes is a regular language, then we can solve the realization problem for the hybrid systems sketched above, if the automaton is known in advance.

Contribution. First, the paper presents a complete realization theory for linear and bilinear switched systems. Second, the paper demonstrates the usefulness of the theory of rational formal power series in studying hybrid systems. More precisely, in this series of papers we prove the following.

- A linear (bilinear) switched system is a minimal realization of a set of input-output maps if and only if it is observable and semi-reachable from the set of states which induce the input-output maps of the given set. Minimal linear (bilinear) switched systems are unique up to similarity. Each linear (bilinear) switched system can be transformed to a minimal one realizing the same set of input-output maps.
- A set of input/output maps is realizable by a linear (bilinear) switched system if and only if it has a *generalized kernel representation* (*generalized Fliess-series expansion*) and the rank of its Hankel-matrix is finite. A minimal realization can be constructed from the columns of the Hankel-matrix.
- Consider a set of input-output maps Φ defined on some subset of switching sequences for which the switching times are arbitrary, and the sequence of discrete modes belong a regular language L. Then Φ has a realization by a linear (bilinear) switched system if and only if Φ has a generalized kernel representation (has a generalized Fliess-series expansion) and its Hankel-matrix is of finite rank. Again, there exists a procedure to construct a realization from the columns of the Hankel-matrix. The procedure yields an observable and semi-reachable realization of Φ. But this realization is not a realization with the smallest state-space dimension possible.

It turns out that realization theory of both linear and bilinear switched systems can be reformulated in terms of the theory of rational formal power series. Exactly this similarity prompted us to treat linear and bilinear switched systems within a single series of papers. Rational formal power series were introduced several decades ago in computer science and control theory, see [1,5,14,23–25]. For the purposes of this paper, we had to extend the existing results, which deal with a *single formal power series*, to *families of formal power* series.

Prior work. For realization theory for hybrid systems other than switched systems, see [17,19]. The paper [16] developed realization theory for linear switched systems using elementary techniques, but results of this paper are more general. The papers [18,21] can be viewed as short versions of parts of the current paper, but they do not contain detailed proofs. The current paper contains all the results of [18,21] and also provides all the proofs. The thesis [20] contains all the results and the proofs of the paper.

Relationship with nonlinear realization theory. The approach to the realization theory taken in this paper was inspired by the realization theory of nonlinear systems [3,5–7,12,13,22,24,25,27]. In particular, realization theory of bilinear systems was presented in [3,9,10,24,25] and was again based on formal power series in non-commuting variables. Intuitively, the reason why formal power series are applicable for both nonlinear and switched systems is that switched systems can be viewed as nonlinear systems whose inputs are the switching sequences and continuous-valued input functions. Unfortunately, the existing results on realization theory of nonlinear systems did not seem to be directly applicable to switched systems. First, the classes of systems for which nonlinear realization theory exists are different from bilinear and linear switched systems. Second, the existing results do not seem to include the case of constrained switching. Third, to the best of our knowledge, the existing results do not deal with families of input-output maps, except [28], where sufficient conditions for realizability of families of input-output maps by rational control systems were presented. However, in [28] minimality was not addressed and the class of control systems considered is very different from switched systems.

Outline. The current paper represents the first part of a series of papers. In Part I we present realization theory for linear switched systems. In Part II we present realization theory for bilinear switched systems. The outline of the paper is the following. Section 2 describes some properties and concepts related to switched systems which are used in the rest of the paper. In Section 3 we present the main results on linear switched systems. Section 4 contains the necessary extension of the classical results on formal power series. In Section 5 the proof of the results on realization theory of linear switched systems is presented. In Appendix A we present the proof of certain technical results on linear switched systems. In Appendix B we present the proofs of the results on formal power series presented in Section 4.

2. Switched systems

We will start with fixing some notation and terminology which will be used throughout the paper.

2.1. Notation and terminology

Denote by T the time-axis, *i.e.* $T = [0, +\infty) \subseteq \mathbb{R}$ is the set of non-negative reals. Denote by $PC(T, \mathbb{R}^m)$, $m > 0$ the class of piecewise-continuous maps from T to $\mathbb{R}^m$, *i.e.* for $f \in PC(T, \mathbb{R}^m)$, f has finitely many points of discontinuity on each finite interval $[0, t]$, $t \in T$, and at each point of discontinuity the right- and left-hand side limits exist and they are finite. Denote by $\mathbb{N}$ the set of natural number including 0. We identify any constant function with its value. For any function g the range of g will be denoted by $\mathrm{Im} f$. For two functions f and g, $g \circ f$ denotes the composition of g and f, *i.e.* $g \circ f(a) = g(f(a))$ for any a in the domain of f. If $\mathcal{X}$ is a vector space and $Z \subseteq \mathcal{X}$, then $\mathrm{Span} Z$ denotes the linear span of elements of Z. If F_1 and F_2 are two linear maps, then $F_1 F_2$ denotes the composition $F_1 \circ F_2$. If $x \in \mathcal{X}$, then $F_1 x$ denotes the value $F_1(x)$. For any $m > 0$, e_j denotes the jth unit vector of $\mathbb{R}^m$, *i.e.* $e_j = (\delta_{1j}, \delta_{2j}, \ldots, \delta_{mj})$ where δ_{ij} is the Kronecker symbol, *i.e.* $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$, for all $i, j = 1, \ldots, m$. The cardinality of a set A is denoted by $|A|$.

We use the notation of [11] for infinite matrices. Let I and J be two arbitrary sets. A (real) matrix M with column index set J and row index set I is simply 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 defined as $M_{i,j} = M(i, j)$. For a matrix $M \in \mathbb{R}^{I \times J}$, the columns of M are maps of the form $I \to \mathbb{R}$, *i.e.* the column of M indexed by $j \in J$, denote by $M_{.,j}$, is the map $I \ni i \mapsto M_{i,j} \in \mathbb{R}$. The set of maps of the form $I \to \mathbb{R}$ is denoted by $\mathbb{R}^I$. Notice that $\mathbb{R}^I$ forms a vector space with respect to point-wise addition and multiplication by scalar, *i.e.* if $f, g \in \mathbb{R}^I$ and $\alpha, \beta \in \mathbb{R}$, then $\alpha f + \beta g \in \mathbb{R}^I$ is defined by $(\alpha f + \beta g)(i) = \alpha f(i) + \beta g(i)$ for all $i \in I$. The *rank of* M, denoted by $\mathrm{rank}\, M \in \mathbb{N} \cup \{\infty\}$, is the dimension of the linear subspace of $\mathbb{R}^I$ spanned by the columns of M.

Notation 2.1 (high-order partial derivatives)**.** Let $\phi : \mathbb{R}^k \to \mathbb{R}^{p \times m}$ be a smooth map. Consider a k tuple $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \in \mathbb{N}^k$. We denote by $D^\alpha \phi$ the following partial derivative

$$D^\alpha \phi = \frac{\mathrm{d}^{\alpha_1}}{\mathrm{d}t_1^{\alpha_1}} \frac{\mathrm{d}^{\alpha_2}}{\mathrm{d}t_2^{\alpha_2}} \cdots \frac{\mathrm{d}^{\alpha_k}}{\mathrm{d}t_k^{\alpha_k}} \phi(t_1, t_2, \ldots, t_k)|_{t_1 = t_2 = \ldots = t_k = 0}.$$

If $m = 1$, then ϕ can be viewed as a map of the form $\phi : \mathbb{R}^k \to \mathbb{R}^p$ and the notation above still applies.

Notation 2.2 (time shift)**.** For $f \in PC(T, \mathbb{R}^m)$ and for any $t \in T$ denote by $\mathrm{Shift}_t f$ the map defined by $\mathrm{Shift}_t(f) : T \ni \tau \mapsto f(t + \tau)$. Notice that $\mathrm{Shift}_t(f) \in PC(T, \mathbb{R}^m)$.

The notation described below is standard in automata theory, see [4,8]. Consider a (possibly infinite) set X. Denote by X^* the set of finite sequences (referred to as *words or strings*) of elements of X. The length of a word of $w \in X^*$ is denoted by $|w|$. The *empty sequence (word)* is denoted by ϵ. A word $w \in X^*$ can always be written as $w = a_1 a_2 \ldots a_k$ for some $a_1, a_2, \ldots, a_k \in X$ and $k \geq 0$; if $k = 0$ then by convention $w = \epsilon$. Note that $|\epsilon| = 0$. We denote by X^+ the set of of non-empty words over X, *i.e.* $X^+ = X^* \setminus \{\epsilon\}$. For two words $v = v_1 v_2 \ldots v_k \in X^*$, and $w = w_1 w_2 \ldots w_m \in X^*$, $v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_m \in X$, define the concatenation $vw \in X^*$ of v and w as the the word $vw = v_1 v_2 \ldots v_k w_1 w_2 \ldots w_m$. In particular, if $v = \epsilon$, then $vw = w$ and if $w = \epsilon$, then $vw = v$. If $w \in X^+$, then w^k denotes the word $\underbrace{ww \ldots w}_{k\text{-times}}$. Here $w^0 = \epsilon$. If X is *finite*, we call any subset $L \subseteq X^*$ a *language*. A language is regular if it can be recognized by a finite-state automaton, see [4,8].

2.2. Definition of switched systems

Below we will present the definition of switched systems and some basic system theoretic notions.

Definition 2.1 (switched systems)**.** A switched system Σ is a control system of the form

$$\dot{x}(t) = f_{q(t)}(x(t), u(t)) \text{ and } y(t) = h_{q(t)}(x(t)). \tag{2.1}$$

Here $x(t) \in \mathbb{R}^n$, $n > 0$ is the continuous state at time $t \in T$, $y(t) \in \mathbb{R}^p$, $p > 0$ is the continuous output at time t, $q(t) \in Q$ is the discrete mode at time t and $u(t) \in \mathbb{R}^m$, $m > 0$ is the continuous input at time t. Consequently, $\mathcal{X} = \mathbb{R}^n$ is the continuous state-space, $\mathcal{Y} = \mathbb{R}^p$ is the continuous output-space, $\mathcal{U} = \mathbb{R}^m$ is the continuous input-space, and Q is the finite set of discrete modes (discrete states). For each discrete mode $q \in Q$, the vector field $f_q : \mathcal{X} \times \mathcal{U} \to \mathcal{X}$ is smooth in both variables x and u, and globally Lipschitz in x, and the *readout map* $h_q : \mathcal{X} \to \mathcal{Y}$ is smooth. The dimension of Σ, denoted by $\dim \Sigma$, is the dimension $\dim \mathcal{X}$ of $\mathcal{X}$.

Notation 2.3. In the rest of the paper we use the symbols $\mathcal{U} = \mathbb{R}^m$, $\mathcal{Y} = \mathbb{R}^p$ and Q to denote the continuous-valued inputs, outputs and the set of discrete modes respectively.

In the rest of the section, Σ *denotes a switched system of the form* (2.1). Informally, the *state trajectory* $x : T \to \mathcal{X}$ is a continuous and piecewise-differentiable function which satisfies the differential equation (2.1) for a given initial state $x(0) = x_0$, input $u \in PC(T, \mathcal{U})$ and piecewise-constant *switching signal* $q(.) : T \to Q$. The output signal $y(t)$ is obtained from $x(t)$ by applying the readout map $h_{q(t)}$. That is, *both the switching signals and the piecewise-continuous inputs are viewed as the inputs to the switched system* Σ. Below we define state- and input-output behavior of switched systems more rigorously. To this end, we need the following notation. *In the rest of the section,* Σ *denotes a switched system of the form* (2.1).

Definition 2.2 (switching sequences)**.** A *switching sequence* is a sequence of the form $w = (q_1, t_1)(q_2, t_2)\dots(q_k, t_k)$, where $q_1, \dots, q_k \in Q$ are discrete modes and $t_1, \dots, t_k$ denote the *switching times* and $k \geq 0$. The set of all switching sequences are denoted by $(Q \times T)^*$. If $k = 0$ above, then we say that w is an empty sequence and we denote it by ϵ. We denote the set of all *non-empty switching sequences* by $(Q \times T)^+$.

The interpretation of the sequence $w = (q_1, t_1)(q_2, t_2)\dots(q_k, t_k)$ is the following. From time instance 0 to time instance t_1 the active discrete mode is q_1, *i.e.* the value of the switching signal is q_1, from t_1 to $t_1 + t_2$ the value of the switching signal is q_2, from $t_1 + t_2$ to $t_1 + t_2 + t_3$ the value of the switching signal is q_3, and so on. Next we define the state and output trajectories of switched systems.

Definition 2.3 (state and output trajectories)**.** Let $u \in PC(T, \mathcal{U})$ be an input and $w = (q_1, t_2)(q_2, t_2)\dots(q_k, t_k) \in (Q \times T)^*$ be a switching sequence. The *state of* Σ *reached from the state* $x_0 \in \mathcal{X}$ *with the inputs* u *and* w is denoted by $x_\Sigma(x_0, u, w)$ and it is defined as follows. If $k = 0$, *i.e.* $w = \epsilon$, then $x_\Sigma(x_0, u, w) = x_0$. If $k > 0$, then

$$x_\Sigma(x_0, u, w) = F(q_k, \mathrm{Shift}_{\sum_{i=1}^{k-1} t_i}(u), t_k) \circ F(q_{k-1}, \mathrm{Shift}_{\sum_{i=1}^{k-2} t_i}(u), t_{k-1}) \circ \dots \circ F(q_1, u, t_1)(x_0). \tag{2.2}$$

Recall from Notation 2.2 that $\mathrm{Shift}_t(u)$ denotes the shift of u by time t. The function $F(q, u, t) : \mathcal{X} \to \mathcal{X}$ maps x_0 to the solution $x(t)$ of the differential equation $\dot{x}(t) = f_q(x(t), u(t))$ at time t with the initial condition $x(0) = x_0$.

Assume w is non-empty, *i.e.* $k > 0$. The *output generated by* Σ *if started from initial state* x_0 *and fed with the inputs* u *and* w is denoted by $y_\Sigma(x_0, u, w) \in \mathcal{Y}$, and it is defined by $y_\Sigma(x, u, w) = h_{q_k}(x_\Sigma(x, u, w))$.

Definition 2.4 (input-output maps)**.** Consider a state $x_0 \in \mathcal{X}$ of Σ. Define the *input-output map of* Σ *induced by the state* x_0 as the map $y_\Sigma(x_0, ., .) : PC(T, \mathcal{U}) \times (Q \times T)^+ \to \mathcal{Y}$ such that for all input $u \in PC(T, \mathcal{U})$ and switching sequence $w \in (Q \times T)^+$, $y_\Sigma(x_0, ., .)(u, w) = y_\Sigma(x_0, u, w)$.

The *reachable set* of the system Σ from a set of initial states $\mathcal{X}_0 \subseteq \mathcal{X}$ is defined by

$$\mathrm{Reach}(\Sigma, \mathcal{X}_0) = \{x_\Sigma(x_0, u, w) \in \mathcal{X} \mid u \in PC(T, \mathcal{U}),\ w \in (Q \times T)^*,\ x_0 \in \mathcal{X}_0\}. \tag{2.3}$$

That is, $\mathrm{Reach}(\Sigma, \mathcal{X}_0)$ is the set of all those states which can be reached from an initial state in $\mathcal{X}_0$ by applying some continuous-valued input and some finite switching sequence.

Definition 2.5 (reachability and semi-reachability)**.** Σ is said to be *reachable* from $\mathcal{X}_0$ if $\mathrm{Reach}(\Sigma, \mathcal{X}_0) = \mathcal{X}$. Σ is *semi-reachable* from $\mathcal{X}_0$ if $\mathcal{X}$ is the smallest vector space containing $\mathrm{Reach}(\Sigma, \mathcal{X}_0)$.

I.e., Σ is semi-reachable from $\mathcal{X}_0$ if the linear span of the states reachable from $\mathcal{X}_0$ yields the whole state-space.

Definition 2.6. Two states $x_1 \neq x_2 \in \mathcal{X}$ of Σ are *indistinguishable* if the input-output maps induced by x_1 and x_2 coincide, *i.e.* $y_\Sigma(x_1, ., .) = y_\Sigma(x_2, ., .)$. Σ is *observable* if it has no pair of indistinguishable states.

In other words, $x_1 \neq x_2$ are indistinguishable, if and only if for all continuous-valued inputs $u \in PC(T, \mathcal{U})$ and switching sequences $w \in (Q \times T)^+$, $y_\Sigma(x_1, u, w) = y_\Sigma(x_2, u, w)$.

From the discussion above it follows that the potential input-output maps of switched systems are maps of the form $f : PC(T, \mathcal{U}) \times (Q \times T)^+ \to \mathcal{Y}$. Below we define the class of input-output maps of interest formally.

Definition 2.7 (input-output maps: arbitrary switching)**.** An *input-output map defined for arbitrary switching* is a map of the form $f : PC(T, \mathcal{U}) \times (Q \times T)^+ \to \mathcal{Y}$. The set of all such maps will be denoted by $F(PC(T, \mathcal{U}) \times (Q \times T)^+, \mathcal{Y})$. A *family of input-output maps defined for arbitrary switching* (*family of input-output maps* in short) is just a (possibly infinite) subset of $F(PC(T, \mathcal{U}) \times (Q \times T)^+, \mathcal{Y})$.

In this paper we will be concerned with realizations of families of input-output maps. We formalize this notion as follows. Consider a family $\Phi \subseteq F(PC(T, \mathcal{U}) \times (Q \times T)^+, \mathcal{Y})$ of input-output maps.

Definition 2.8 (realization of input-output maps: arbitrary switching)**.** The family Φ is said to be *realized* by a switched system Σ if there exists a map $\mu : \Phi \to \mathcal{X}$, which maps each input-output map f from Φ to a state $\mu(f)$ of Σ, such that $f = y_\Sigma(\mu(f), ., .)$, *i.e.* for each $f \in \Phi$, $u \in PC(T, \mathcal{U})$, $w \in (Q \times T)^+$,

$$y_\Sigma(\mu(f), u, w) = f(u, w). \tag{2.4}$$

One can think of the map μ as a way to determine the corresponding initial state for each element of Φ. In the sequel we will mainly deal with pairs (Σ, μ) where Σ is a switched system of the form (2.1) and $\mu : \Phi \to \mathcal{X}$ is a map assigning to each input-output map f a state of Σ. This prompts us to introduce the notion of a switched system realization.

Definition 2.9 (switched system realizations)**.** We refer to the pair (Σ, μ), where $\mu : \Phi \to \mathcal{X}$ is a map mapping elements of Φ to the states of Σ, as *realizations*. A realization (Σ, μ) is a *realization of the family of input-output maps* Φ, if (2.4) holds for all $f \in \Phi$, $u \in PC(T, \mathcal{U})$ and $w \in (Q \times T)^+$.

Note that not any realization (Σ, μ) with $\mu : \Phi \to \mathcal{X}$ is a realization of Φ.

Definition 2.10 (observability and semi-reachability of realizations: arbitrary switching)**.** The realization (Σ, μ) is *semi-reachable* , if Σ is semi-reachable from the range $\mathrm{Im}\mu$ of μ; (Σ, μ) is *observable*, if Σ is observable.

In this paper we also investigate realization theory for input-output maps which are defined only for a subset of switching sequences. In order to state the problem formally, we need additional notation and terminology. Let $L \subseteq Q^+$ be the set of *admissible sequences of discrete modes*. The set L contains all those sequences of discrete modes along which the switched system is allowed to switch. Note that *L can be viewed as a language over the finite alphabet Q formed by the discrete modes*. In order to make the discussion of results easier, we will introduce a separate term for denoting the set of switching sequences which are admissible according to L.

Definition 2.11. Define the subset of *admissible switching sequences* $TL \subseteq (Q \times T)^+$ *associated with* L by

$$TL = \{(q_1, t_1)(q_2, t_2) \dots (q_k, t_k) \in (Q \times T)^+ \mid q_1 q_2 \dots q_k \in L,\, k > 0,\, t_1, \dots, t_k \in T,\, q_1, \dots, q_k \in Q\}. \tag{2.5}$$

That is, TL is the set of those switching sequences, for which the sequence of discrete modes belongs to L and the switching times are arbitrary. If $L = Q^+$ then $TL = (Q \times T)^+$, *i.e.* any switching sequences is admissible.

Next, we formulate the counterparts of Definitions 2.7–2.10 *i.e.* we define the concept of input-output map, realization by a switched system, switched system realization, etc. for the case of constrained switching. For $L = Q^+$ the new definitions are equivalent to the ones for arbitrary switching.

Definition 2.12 (input-output maps: constrained switching)**.** The *input-output maps with the switching constraint* L are maps the form $f : PC(T,\mathcal{U}) \times TL \to \mathcal{Y}$, where TL is the set of admissible switching sequences from (2.5). We denote the set of all such input-output maps by $F(PC(T,\mathcal{U}) \times TL, \mathcal{Y})$. A *family of input-output maps with the switching constraint* L is an arbitrary subset of $F(PC(T,\mathcal{U}) \times TL, \mathcal{Y})$.

Let $\Phi \subseteq F(PC(T,\mathcal{U}) \times TL, \mathcal{Y})$ be a family of input-output maps with the switching constraint L.

Definition 2.13 (realization by switched systems: constrained switching)**.** The switched system Σ *realizes* Φ *with constraint* L if there exists a map $\mu : \Phi \to \mathcal{X}$ such that for each $f \in \Phi$, the restriction of the input-output map $y_\Sigma(\mu(f), ., .)$ to the set TL coincides with f, *i.e* for each $f \in \Phi$, $u \in PC(T,\mathcal{U})$ and for all $w \in TL$,

$$y_\Sigma(\mu(f), u, w) = f(u, w). \tag{2.6}$$

Definition 2.14 (switched system realizations: constrained switching)**.** We refer to pairs (Σ, μ), where Σ is a switched system and $\mu : \Phi \to \mathcal{X}$ is a map as *realizations*. We will say that (Σ, μ) is a *realization of* Φ *with constraint* L, if (2.6) holds for all $f \in \Phi$, $u \in PC(T,\mathcal{U})$ and $w \in TL$.

Definition 2.15. The realization (Σ, μ) with $\mu : \Phi \to \mathcal{X}$ is *semi-reachable*, if it is semi-reachable from the range Imμ of μ according to Definition 2.5; (Σ, μ) is observable, if Σ is observable according to Definition 2.6.

Notice that in Definition 2.15 semi-reachability and observability of Σ is understood as a property involving *all (including non-admissible) switching sequences*. Just as before, the map μ can be thought of as a way to specify initial states of the system Σ.

Remark 2.1 (abuse of terminology)**.** Note that if $L = Q^+$, then Definitions 2.12, 2.13, 2.14 and 2.15 are equivalent to Definitions 2.7, 2.8, 2.9 and 2.10 respectively. This leads us to adopt the following abuse of terminology. In the sequel we will not specify explicitly whether we mean realization with constrained or arbitrary switching as long as it is clear from the context.

3. Main results on realization theory for linear switched systems

The purpose of this section is to present formally the main results of the paper on realization theory of linear switched systems. In Section 3.1 we will present the definition and some basic properties of linear switched systems. In Section 3.2 we will describe the main results on minimality of linear switched systems. In Section 3.3 we will state the necessary and sufficient conditions for existence of a linear switched systems realization of a family of input-output maps.

3.1. Definition and basic properties of linear switched systems

Informally, a linear switched system is a switched system, such that for each discrete mode, the underlying continuous system is a finite dimensional linear time-invariant system.

Definition 3.1 (linear switched systems)**.** A *linear switched system* Σ is a switched system of the form (2.1), such that for each discrete mode $q \in Q$, there exist matrices $A_q \in \mathbb{R}^{n\times n}$, $B_q \in \mathbb{R}^{n\times m}$ and $C_q \in \mathbb{R}^{p\times n}$, such that

$$f_q(x, u) = A_q x + B_q u \text{ and } h_q(x) = C_q x. \tag{3.1}$$

We use the following notation for linear switched systems above; $\Sigma = (\mathcal{X}, \mathcal{U}, \mathcal{Y}, Q, \{(A_q, B_q, C_q) \mid q \in Q\})$.

Since linear switched systems are switched systems, we will use the same notation and definitions, *i.e.* the same notion of state and output trajectory, realization, observability, semi-reachability, etc., as described in Section 2. Next we define the notion of minimality for linear switched systems. To this end, recall that the dimension of a linear switched system equals the dimension of its state-space. Let Φ be a family of input-output maps defined either for arbitrary or for constrained switching. In the sequel, a *linear switched system realization* means a switched system realization (Σ, μ) such that Σ is a linear switched system.

Definition 3.2 (minimality)**.** A linear switched system realization (Σ, μ) is *a minimal realization of* Φ if (Σ, μ) is a realization of Φ and for any linear switched system realization $(\hat{\Sigma}, \hat{\mu})$, of Φ, it holds that $\dim \Sigma \leq \dim \hat{\Sigma}$. The linear switched system Σ is a minimal realization of Φ, if (Σ, μ) is a minimal realization of Φ for some μ.

That is, a linear switched system is a minimal realization of Φ if it has the smallest state-space dimension among all the linear switched systems which are realizations of Φ. Notice that a linear switched system can be a minimal realization for Φ and can fail to be a minimal realization for another family of input-output maps.

Definition 3.3 (linear switched system morphism)**.** Consider the linear switched systems

$$\Sigma_1 = (\mathcal{X}, \mathcal{U}, \mathcal{Y}, Q, \{(A_q, B_q, C_q) \mid q \in Q\}) \quad \text{and} \quad \Sigma_2 = (\mathcal{X}_a, \mathcal{U}, \mathcal{Y}, Q, \{(A^a_q, B^a_q, C^a_q) \mid q \in Q\}).$$

Assume that Φ is a family of input-output maps and $\mu_1 : \Phi \to \mathcal{X}$, $\mu_2 : \Phi \to \mathcal{X}_a$. A *linear switched system morphism* S from (Σ_1, μ_1) to (Σ_2, μ_2), denoted by $S : (\Sigma, \mu_1) \to (\Sigma_2, \mu_2)$, is a linear map $S : \mathcal{X} \to \mathcal{X}_a$ such that

$$\forall q \in Q : A^a_q S = S A_q, \quad B^a_q = S B_q, \quad C^a_q S = C_q \quad \text{and} \quad \forall f \in \Phi : S\mu_1(f) = \mu_2(f). \tag{3.2}$$

The linear switched morphism S is called surjective, injective or isomorphism, if it is surjective, injective, respectively isomorphism as a linear map. The linear switched systems realizations (Σ_1, μ_1) and (Σ_2, μ_2) are said to be *algebraically similar* or *isomorphic* if there exists an isomorphism $S : (\Sigma_1, \mu_1) \to (\Sigma_2, \mu_2)$.

Finally, we recall from [26] some basic fact on linear switched systems.

Theorem 3.1 (state- and output-trajectory [26])**.** *For any linear switched system Σ of the form (3.1), the state and output trajectories are of the following form. For each input $u \in PC(T, \mathcal{U})$, initial state $x_0 \in \mathcal{X}$ and switching sequence $w = (q_1, t_1)(q_2, t_2) \dots (q_k, t_k) \in (Q \times T)^+$, $q_1, q_2, \dots, q_k \in Q$, $t_1, t_2, \dots, t_k \in T$, $k > 0$,*

$$\begin{aligned} x_\Sigma(x_0, u, w) &= \mathrm{e}^{A_{q_k} t_k} \mathrm{e}^{A_{q_{k-1}} t_{k-1}} \dots \mathrm{e}^{A_{q_1} t_1} x_0 + \int_0^{t_k} \mathrm{e}^{A_{q_k}(t_k - s)} B_{q_k} u\left(\sum_{i=1}^{k-1} t_i + s\right) \mathrm{d}s \\ &\quad + \mathrm{e}^{A_{q_k} t_k} \int_0^{t_{k-1}} \mathrm{e}^{A_{q_{k-1}}(t_{k-1} - s)} B_{q_{k-1}} u\left(\sum_{i=1}^{k-2} t_i + s\right) \mathrm{d}s + \dots \\ &\quad + \mathrm{e}^{A_{q_k} t_k} \dots \mathrm{e}^{A_{q_2} t_2} \int_0^{t_1} \mathrm{e}^{A_{q_1}(t_1 - s)} B_{q_1} u(s) \mathrm{d}s \end{aligned} \tag{3.3}$$

$$\begin{aligned} y_\Sigma(x_0, u, w) &= C_{q_k} x_\Sigma(x, u, w) = C_{q_k} \mathrm{e}^{A_{q_k} t_k} \dots \mathrm{e}^{A_{q_1} t_1} x_0 + \int_0^{t_k} C_{q_k} \mathrm{e}^{A_{q_k}(t_k - s)} B_{q_k} u\left(\sum_{i=1}^{k-1} t_i + s\right) \mathrm{d}s \\ &\quad + C_{q_k} \mathrm{e}^{A_{q_k} t_k} \int_0^{t_{k-1}} \mathrm{e}^{A_{q_{k-1}}(t_{k-1} - s)} B_{q_{k-1}} u\left(\sum_{i=1}^{k-2} t_i + s\right) \mathrm{d}s + \dots \\ &\quad + C_{q_k} \mathrm{e}^{A_{q_k} t_k} \dots \mathrm{e}^{A_{q_2} t_2} \int_0^{t_1} \mathrm{e}^{A_{q_1}(t_1 - s)} B_{q_1} u(s) \mathrm{d}s. \end{aligned} \tag{3.4}$$

Remark 3.1. Notice that (3.3) and (3.5) imply that the state- and output-trajectory of a linear switched system are a sum of products of matrix exponentials. This implies that the derivatives of the state- and output-trajectories with respect to the switching times are products of the system matrices. In turn, the latter observation will be crucial for developing realization theory for linear switched systems.

Theorem 3.2 ([26])**.**

Reachability: *The set of states reachable from the zero initial state is the linear span of the columns of the matrices of the form $A_{q_k}A_{q_{k-1}}\dots A_{q_1}B_{q_0}$, B_{q_0}, that is,*

$$\mathrm{Reach}(\Sigma,\{0\}) = \mathrm{Span}\{A_{q_k}A_{q_{k-1}}\dots A_{q_1}B_{q_0}u, B_{q_0}u \mid u\in\mathcal{U},\, q_0,q_1,\dots,q_k\in Q,\, k>0\}. \tag{3.5}$$

Observability: *Let O_Σ be the following intersection of kernels of $C_qA_{q_k}A_{q_{k-2}}\dots A_{q_1}$, C_q,* i.e.

$$O_\Sigma = \bigcap_{q\in Q}(\ker C_q \cap \bigcap_{q_1,q_2,\dots,q_k\in Q, k>0} \ker C_qA_{q_k}A_{q_{k-1}}\dots A_{q_1}).$$

O_Σ is called the observability kernel *of Σ. Then Σ is observable if and only if $O_\Sigma=\{0\}$.*

Next, we present an algebraic characterization for semi-reachability of linear switched systems.

Proposition 3.1 (semi-reachability)**.** *Consider the set $\mathcal{X}_0\subseteq\mathcal{X}$ and the linear space*

$$WR(\mathcal{X}_0) = \mathrm{Span}\{x_0, A_{q_k}A_{q_{k-1}}\dots A_{q_1}x_0 \mid q_1,\dots,q_k\in Q,\, k>0,\, x_0\in\mathcal{X}_0 \text{ or } x_0=B_qu,\, q\in Q,\, u\in\mathcal{U}\}.$$

With the notation above, Σ is semi-reachable from the set of initial states $\mathcal{X}_0$ if and only if

$$\dim\Sigma = \dim WR(\mathcal{X}_0). \tag{3.6}$$

In particular, the realization (Σ,μ) with $\mu:\Phi\to\mathcal{X}$ is semi-reachable if and only if (3.6) holds for $\mathcal{X}_0=\mathrm{Im}\mu$.

The result of Proposition 3.1 is new and its proof can be found in Appendix A.2.

Corollary 3.1. *Σ is semi-reachable from $\{0\}$, if and only if it is reachable from the zero initial state.*

Remark 3.2 (algorithm)**.** If $\mathrm{Im}\mu$ is finite, then semi-reachability of (Σ,μ) can be checked numerically. Similarly, observability of Σ can be checked numerically.

Notice that semi-reachability (observability) of (Σ,μ) does not imply reachability (observability) of any of the linear subsystems. For a counter-example, see Example 3.1 below.

Example 3.1. Consider Σ of the form (3.1) with two discrete modes $Q=\{q_1,q_2\}$ with scalar inputs and outputs, *i.e.* $\mathcal{Y}=\mathcal{U}=\mathbb{R}$, with state-space $\mathcal{X}=\mathbb{R}^3$ and with the matrices A_{q_i},B_{q_i},C_{q_i}, $i=1,2$ defined as follows:

$$A_{q_1}=\begin{bmatrix}0&1&0\\0&0&0\\0&0&1\end{bmatrix},\quad B_{q_1}=\begin{bmatrix}0\\1\\0\end{bmatrix},\quad C_{q_1}=\begin{bmatrix}1\\1\\0\end{bmatrix}^T,\quad A_{q_2}=\begin{bmatrix}0&0&0\\0&0&0\\0&1&0\end{bmatrix},\quad B_{q_2}=\begin{bmatrix}0\\0\\0\end{bmatrix},\quad C_{q_2}=\begin{bmatrix}0\\0\\1\end{bmatrix}^T.$$

Let $\Phi=\{y_\Sigma(0,.)\}$ be the singleton set consisting of the input-output map of the system Σ induced by the zero initial state $0\in\mathbb{R}^3$. Let $\mu: y_\Sigma(0,.,.)\mapsto 0$ be the initial state assigning map. Then it is easy to see that (Σ,μ) is a realization of Φ. Using the linear algebraic conditions, it is easy to see that (Σ,μ) is semi-reachable from $\{0\}$ and it is observable, yet none of the linear subsystems are reachable or observable.

3.2. Minimality of linear switched systems

Minimality: the case arbitrary switching. Assume that $\Phi \subseteq F(PC(T,\mathcal{U}) \times (Q \times T)^+, \mathcal{Y})$ is a family on input-output maps defined for arbitrary switching.

Theorem 3.3 (minimality)**.** *If (Σ, μ) is a linear switched system realization of Φ, then the following are equivalent.*

(i) *(Σ, μ) is a minimal linear switched system realization of Φ.*
(ii) *The realization (Σ, μ) is semi-reachable and it is observable.*
(iii) *The state-space dimension of Σ equals the rank of the Hankel-matrix of Φ,* i.e. $\dim \Sigma = \operatorname{rank} H_\Phi$. *The Hankel-matrix H_Φ of Φ and its rank will formally be defined later on, in Definition 3.6.*
(iv) *If $(\hat{\Sigma}, \hat{\mu})$ is a semi-reachable linear switched system realization of Φ, then there exists a surjective linear switched system morphism $T : (\hat{\Sigma}, \hat{\mu}) \to (\Sigma, \mu)$.*

In addition, all minimal linear switched system realizations of Φ are algebraically similar.

The proof of Theorem 3.3 will be presented in Section 5.2.

Corollary 3.2. *A linear switched system realization of Φ is minimal if and only if it is semi-reachable and observable. All minimal linear switched system realizations of Φ are isomorphic.*

Remark 3.3. Any linear switched system realization of Φ can effectively be transformed to a minimal one, see [20]. If (Σ, μ) is a minimal realization, then, in general, it does not follow that any of its linear subsystems is minimal. For a counter example see Example 3.1.

Minimality: constrained switching. Let $L \subseteq Q^+$ be the set of admissible sequences of discrete modes. Let $\Phi \subseteq F(PC(T,\mathcal{U}) \times TL, \mathcal{Y})$ be a family of input-output maps with the switching constraint L. Let $\operatorname{comp}(L) \subseteq Q^+$ be the set of the sequences which end in a letter such that no word in L ends with that letter, *i.e.*

$$\operatorname{comp}(L) = \{q_1 \ldots q_k \in Q^+ \mid q_1, \ldots, q_k \in Q,\ k \geq 1,\ \forall v \in Q^* : vq_k \notin L\}. \tag{3.7}$$

Intuitively, the language $\operatorname{comp}(L)$ contains those sequences which can never be observed if the switching system is run with constraint L. If we apply Definition 2.11 to $\operatorname{comp}(L)$ instead of L, we obtain the set $T(\operatorname{comp}(L))$ of all the switching sequences for which the sequence of discrete modes belong to $\operatorname{comp}(L)$, *i.e.* $T(\operatorname{comp}(L)) = \{(q_1, t_1) \ldots (q_k, t_k) \in (Q \times T)^+ \mid q_1, \ldots, q_k \in Q,\ t_1, \ldots, t_k \in T,\ q_1 \ldots q_k \in \operatorname{comp}(L),\ k \geq 1\}$.

Theorem 3.4 (quasi-minimality)**.** *Assume that L is a regular language and that Φ has a realization by a linear switched system. Then there exists a linear switched system realization (Σ, μ) of Φ, such (Σ, μ) is semi-reachable and it is observable, and for any $f \in \Phi$, input $u \in PC(T,\mathcal{U})$ and switching sequence $w \in T(\operatorname{comp}(L))$,*

$$y_\Sigma(\mu(f), u, w) = 0. \tag{3.8}$$

Moreover, there is a constant $M > 0$, determined by L, such that for any linear switched system realization $(\widetilde{\Sigma}, \widetilde{\mu})$ of Φ,

$$\dim \Sigma \leq M \cdot \dim \widetilde{\Sigma}. \tag{3.9}$$

The proof of Theorem 3.4 is be presented in Section 5.3. A realization (Σ, μ) of Φ which has the properties described in Theorem 3.4 will be called a *quasi-minimal realization of* Φ. Notice that the dimension of Σ from Theorem 3.4 is at most a factor M bigger than the smallest dimension of a linear switched realization of Φ.

Remark 3.4 (algorithms)**.** Based on the size of the finite state automata which accepts L, it is possible to give an upper bound for M, and any realization of Φ can effectively be transformed to a quasi-minimal one, see [20].

Example 3.2. In fact, the result of the Theorem 3.4 is sharp in the following sense. One can construct the following input-output y map and language L and realizations Σ_1 and Σ_2 such that the following holds. Both Σ_1 and Σ_2 realize y from the initial state zero and they are both reachable from zero and observable, but $\dim\Sigma_1 = 1$ and $\dim\Sigma_2 = 2$. Let $Q = \{1, 2\}$, $L = \{q_1^k q_2 \mid k > 0\}$, $\mathcal{Y} = \mathcal{U} = \mathbb{R}$. Define $f : PC(T, \mathcal{U}) \times TL \to \mathcal{Y}$ by

$$f(u,(q_1,t_1)(q_1,t_2)\dots(q_1,t_m)(q_2,t_{m+1})) = \int_0^{t_{m+1}} \mathrm{e}^{2(t_{m+1}-s)} u\left(s + \sum_1^m t_i\right)\mathrm{d}s + \int_0^{\sum_1^m t_i} \mathrm{e}^{2t_{m+1}} \mathrm{e}^{\sum_1^m t_i - s} u(s)\mathrm{d}s.$$

Define the linear switched system $\Sigma_1 = (\mathbb{R}, \mathbb{R}, \mathbb{R}, Q, \{(A_{1,q}, B_{1,q} C_{1,q}) \mid q \in \{q_1, q_2\}\})$ by

$$A_{1,q_1} = 1,\ B_{1,q_1} = 1,\ C_{1,q_1} = 1,\ A_{1,q_2} = 2,\ B_{1,q_2} = 1,\ C_{1,q_2} = 1.$$

Define the linear switched system $\Sigma_2 = (\mathbb{R}^2, \mathbb{R}, \mathbb{R}, Q\{(A_{2,q}, B_{2,q}, C_{2,q}) \mid q \in Q\})$ by

$$A_{2,q_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ B_{2,q_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ C_{2,q_1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}^T \qquad A_{2,q_2} = \begin{bmatrix} 0 & 0 \\ 2 & 2 \end{bmatrix},\ B_{2,q_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ C_{2,q_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T.$$

Both Σ_1 and Σ_2 are reachable and observable as linear switched systems, therefore they are the minimal realizations of the input-output maps $y_{\Sigma_1}(0, ., .)$ and $y_{\Sigma_2}(0, ., .)$ respectively, defined for all switching sequences. It is also easy to see that (Σ_i, μ_i), $i = 1, 2$ is a realization of f, where $\mu_i : f \mapsto 0 \in \mathcal{X}_i$, $i = 1, 2$.

3.3. **Existence of a realization**

First, we introduce the notion of *generalized kernel representation*. We then present necessary and sufficient conditions for existence of a realization, first for arbitrary switching, then for the case of constrained switching.

3.3.1. *Generalized kernel representation*

Let $L \subseteq Q^+$ be the set of admissible sequences of discrete modes and let $\Phi \subseteq F(PC(T, \mathcal{U}) \times TL, \mathcal{Y})$ be a family of input-output maps with the switching constraint L. Informally, Φ has a generalized kernel representation, if

(1) there exists an input-output map y^Φ such that for all $f \in \Phi$, $f(u, w) = f(0, w) + y^\Phi(u, w)$; and
(2) each element f of Φ is affine in continuous inputs and analytic in switching times for all constant inputs.

In order to define the notion of generalized kernel representation formally, we need the following notation:

$$\mathrm{suffix}L = \{u \in Q^* \mid \exists w \in Q^* : wu \in L\} \tag{3.10}$$
$$\widetilde{L} = \{u_1^{i_1} \dots u_k^{i_k} \in Q^* \mid u_1 \dots u_k \in \mathrm{suffix}L,\ u_j \in Q,\ i_j \geq 0,\ j = 1, \dots, k,\ i_k > 0,\ k > 0\}. \tag{3.11}$$

Here we used the notation of Section 2.1, *i.e.* $u_j^{i_j}$ stands for the word which is the repetition of the letter u_j precisely i_j times, $j = 1, \dots, k$. The intuition behind the definitions above is the following. The set suffix$L$ is the collection of all *suffixes* of sequences from L. The set $\widetilde{L}$ contains all those sequences which can be obtained from an element of L, or alternatively from an element of suffix$L$, by repeating every letter several times or erasing it, with the restriction that the last letter cannot be erased. The motivation for suffix$L$ and $\widetilde{L}$ is the following. If we know the input-output behavior of a linear switched system for sequences in TL, then we can reconstruct its input-output behavior for all the switching sequences from $T\widetilde{L} = \{(q_1, t_1) \dots (q_k, t_k) \mid k > 0,$ $t_1, \dots, t_k \in T,\ q_1, \dots, q_k \in Q,\ q_1 \dots q_k \in \widetilde{L}\}$.

Definition 3.4 (generalized kernel-representation)**.** The family Φ has a *generalized kernel representation with constraint* L, or simply *generalized kernel representation*, if for all input-output maps $f \in \Phi$ and for all words $w = q_1 q_2 \dots q_k \in \widetilde{L}$, $q_1, q_2 \dots, q_k \in Q$, $k > 0$, there exist functions

$$K_w^{f,\Phi} : \mathbb{R}^k \to \mathbb{R}^p \ \text{ and } G_w^\Phi : \mathbb{R}^k \to \mathbb{R}^{p\times m}$$

such that the following holds:

(1) For each word $w \in \widetilde{L}$ and map $f \in \Phi$, the functions $K^{f,\Phi}_w$ and G^{Φ}_w are analytic.
(2) For each map $f \in \Phi$, for all words $w, v \in Q^*$ and any $q \in Q$ such that $wqqv, wqv \in \widetilde{L}$, and for all $t_1, t_2, \ldots, t_{|w|+|v|}, t, \hat{t} \in T$,

$$\begin{aligned} K^{f,\Phi}_{wqqv}(t_1, t_2, \ldots, t_{|w|}, t, \hat{t}, t_{|w|+1}, \ldots, t_{|w|+|v|}) &= K^{f,\Phi}_{wqv}(t_1, t_2, \ldots, t_{|w|}, t+\hat{t}, t_{|w|+1}, \ldots, t_{|w|+|v|}) \\ G^{\Phi}_{wqqv}(t_1, t_2, \ldots, t_{|w|}, t, \hat{t}, t_{|w|+1}, \ldots, t_{|w|+|v|}) &= G^{\Phi}_{wqv}(t_1, t_2, \ldots, t_{|w|}, t+\hat{t}, t_{|w|+1}, \ldots, t_{|w|+|v|}). \end{aligned}$$

(3) For each words $v, w \in Q^*$ and for each $q \in Q$ such that $vw \in \widetilde{L}$, and $vqw \in \widetilde{L}$, and $|w| > 0$, for each map $f \in \Phi$, and time instances $t_1, t_2, \ldots, t_{|v|+|w|} \in T$,

$$K^{f,\Phi}_{vqw}(t_1, t_2, \ldots, t_{|v|}, 0, t_{|v|+1}, \ldots, t_{|w|+|v|}) = K^{f,\Phi}_{vw}(t_1, t_2, \ldots, t_{|v|+|w|}).$$

For each pair of words $v, w \in Q^+$ and for each discrete mode $q \in Q$ such that $vw \in \widetilde{L}$ and $vqw \in \widetilde{L}$, the following holds. For each $t_1, t_2, \ldots, t_{|v|+|w|} \in T$,

$$G^{\Phi}_{vqw}(t_1, t_2, \ldots, t_{|v|}, 0, t_{|v|+1}, \ldots, t_{|w|+|v|}) = G^{\Phi}_{vw}(t_1, t_2, \ldots, t_{|v|+|w|}).$$

(4) For each map $f \in \Phi$, each switching sequence $s = (q_1, t_1)(q_2, t_2)\ldots(q_k, t_k) \in TL$, where $q_1, q_2, \ldots, q_k \in Q$, and $t_1, t_2, \ldots, t_k \in T$, $k > 0$, and each input $u \in PC(T, \mathcal{U})$, the following holds.

$$f(u, s) = K^{f,\Phi}_{q_1 q_2 \ldots q_k}(t_1, t_2, \ldots, t_k) + \sum_{i=1}^{k} \int_0^{t_i} G^{\Phi}_{q_i q_{i+1} \ldots q_k}(t_i - s, t_{i+1}, \ldots, t_k) u\left(s + \sum_{j=1}^{i-1} t_j\right) \mathrm{d}s. \quad (3.12)$$

The reader may view the functions $K^{f,\Phi}_w$ as the parts of the output which depend on the initial condition and the functions G^{Φ}_w as functions determining the dependence of the output on the continuous inputs. The intuition behind the various conditions of Definition 3.4 are the following. Condition 1 ensures that the response of the elements of Φ to constant continuous-valued inputs is analytic in the switching times. Conditions 2 and 3 make sure that the elements of Φ satisfy certain conditions which are satisfied by any input-output map which is realized by a (not necessarily linear) switched system. More precisely, Condition 2 ensures that staying in a discrete mode q for $t + \hat{t}$ time has the same effect on the output as staying in q for time t and then switching to the very same q and staying there for time $\hat{t}$. Condition 3 ensures that staying in a discrete mode for zero time does not affect the output.

Alternatively, a good intuition can be derived by analogy with input-output maps of linear systems. Recall from [2] that an input-output map $y : PC(T, \mathcal{U}) \times T \to \mathcal{Y}$ can be realized by a linear system (A, B, C) from the initial state x_0, if only if there exists $K : T \to \mathbb{R}^p$ and $G : T \to \mathbb{R}^{p \times m}$ such that

$$y(u, t) = K(t) + \int_0^t G(t - s) u(s) \mathrm{d}s \quad \text{and} \quad K(t) = C\mathrm{e}^{At} x_0 \quad \text{and} \quad G(t) = C\mathrm{e}^{At} B. \quad (3.13)$$

A similar relationship holds for the functions $K^{f,\Phi}_w$ and G^{Φ}_w of a generalized kernel representation of Φ. In order to present the relationship precisely, we need additional terminology.

Definition 3.5 (zero-response of Φ)**.** Let $y^{\Phi} : PC(T, \mathcal{U}) \times TL \to \mathcal{Y}$ be such that for each input $u \in PC(T, \mathcal{U})$ and switching sequence $w = (q_1, t_1)(q_2, t_2)\ldots(q_k, t_k) \in TL$, $q_1, q_2, \ldots, q_k \in Q$ and $t_1, t_2, \ldots, t_k \in T$, $k > 0$

$$y^{\Phi}(u, (q_1, t_1)(q_2, t_2)\ldots(q_k, t_k)) = \sum_{i=1}^{k} \int_0^{t_i} G^{\Phi}_{q_i q_{i+1} \ldots q_k}(t_i - s, t_{i+1}, \ldots, t_k) u\left(s + \sum_{j=1}^{i-1} t_j\right) \mathrm{d}s. \quad (3.14)$$

Remark 3.5. It is easy to see that $y^{\Phi}(u,w) = f(u,w) - f(0,w)$ for all $u \in PC(T,\mathcal{U})$, $w \in TL$ and $f \in \Phi$.

The intuition behind the definition of the function y^{Φ} is the following. If Φ has a realization by a linear switched system Σ, then y^{Φ} is the input-output map induced by Σ from the zero initial state.

Theorem 3.5. *For any linear switched system Σ of the form (3.1) and any map $\mu : \Phi \to \mathcal{X}$, the pair (Σ, μ) is a realization of Φ with constraint L if and only if Φ has a generalized kernel representation defined by*

$$\begin{aligned} G^{\Phi}_{q_1 q_2 \ldots q_k}(t_1, t_2, \ldots, t_k) =& C_{q_k} \mathrm{e}^{A_{q_k} t_k} \mathrm{e}^{A_{q_{k-1}} t_{k-1}} \ldots \mathrm{e}^{A_{q_1} t_1} B_{q_1} \\ K^{f,\Phi}_{q_1 q_2 \ldots q_k}(t_1, t_2, \ldots, t_k) =& C_{q_k} \mathrm{e}^{A_{q_k} t_k} \mathrm{e}^{A_{q_{k-1}} t_{k-1}} \ldots \mathrm{e}^{A_{q_1} t_1} \mu(f), \end{aligned} \tag{3.15}$$

where $q_1 q_2 \ldots q_k \in \widetilde{L}$, $q_1, q_2, \ldots, q_k \in Q$, $k \geq 1$. Moreover, if (Σ, μ) is a realization of Φ, then $y^{\Phi}(u,w) = y_{\Sigma}(0,u,w)$ for each continuous-valued input $u \in PC(T,\mathcal{U})$ and admissible switching sequence $w \in TL$.

Proof of Theorem 3.5. (Σ,μ) is a realization of Φ if and only if for each $f \in \Phi$, $u \in PC(T,\mathcal{U})$, $w \in TL$, $f(u,w) = y_{\Sigma}(\mu(f),u,w) = C_{q_k} x_{\Sigma}(\mu(f),u,w)$ where (q_k,t_k) is the last element of w, *i.e.* $w = \hat{w}(q_k,t_k)$ for some $\hat{w} \in (Q \times T)^*$. The statement of the theorem follows now directly from Theorem 3.1. □

We conclude the section by introducing notation which will be used in the subsequent sections.

Notation 3.1 (input-output maps with fixed switching sequence and input)**.** Consider an input-output map f with the switching constraint L. For a sequence $w = q_1 q_2 \ldots q_k \in L$ where $q_1, q_2, \ldots, q_k \in Q$ and an input $u \in PC(T,\mathcal{U})$, define the map $f_{u,w} : T^k \to \mathcal{Y}$ as follows

$$f_{u,w}(t_1, \ldots, t_k) = f(u, (q_1,t_1)(q_2,t_2), \ldots, (q_k,t_k)). \tag{3.16}$$

I.e. $f_{u,w}$ is obtained by fixing the input u and a sequence of discrete modes w and varying the switching times.

Remark 3.6 (derivatives of input-output maps)**.** Assume that Φ has a generalized kernel representation. For any input value $u \in \mathcal{U}$ identify u with the *constant* input function which takes the value u. Then it follows from Part 4 of Definition 3.4 that for any $f \in \Phi$ and any sequence $w \in L$, the map $f_{u,w}$, defined in (3.16), is analytic. Indeed, if w is of the form $w = q_1 q_2 \ldots q_k$ for $q_1, q_2, \ldots, q_k \in Q$, $k > 0$, then by Part 4 of Definition 3.4 $f_{u,w}(t_1, t_2, \ldots, t_k)$ equals the right-hand side of (3.12). Since on the right-hand side of (3.12) each summand is analytic in $t_1, t_2, \ldots, t_k$, it follows that $f_{u,w}(t_1, t_2, \ldots, t_k)$ is analytic in $t_1, t_2, \ldots, t_k$. Recall the definition of the input-output map y^{Φ} and notice that the notation of Remark 3.1 can be applied to y^{Φ}. In addition, by Remark 3.5, $y^{\Phi}_{u,w} = f_{u,w} - f_{0,w}$ and hence $y^{\Phi}_{u,w}$ is also analytic. Hence, for any $u \in \mathcal{U}$, any sequence $w \in L$ and tuple $\alpha \in \mathbb{N}^k$ where $k = |w|$, the derivatives $D^{\alpha} f_{u,w}$ and $D^{\alpha} y^{\Phi}_{u,w}$ are well-defined. In particular, $D^{\alpha} y^{\Phi}_{e_j,w}$ and $D^{\alpha} f_{0,w}$ are well-defined, where e_j, $j = 1, \ldots, m$ are the jth unit vector of $\mathcal{U} = \mathbb{R}^m$, *i.e.* $e_j = (\underbrace{0,0,\ldots,0}_{j-1\text{-}times},1,0,\ldots,0)$.

3.3.2. *Existence of a realization: arbitrary switching*

Throughout this section, $\Phi \subseteq F(PC(T,\mathcal{U}) \times (Q \times T)^+, \mathcal{Y})$ denotes a family of input-output maps defined for arbitrary switching, and we assume that Φ admits a generalized kernel representation.

We begin with the definition of the Hankel-matrix H_{Φ} of Φ. The entries of H_{Φ} are high-order derivatives of the elements of Φ with respect to the switching times. We collect the derivatives in intermediary vectors $S_{q_1,q_2,j}$ and $S_{f,q}$, as follows. Using Notation 3.1 and Remark 3.6, for each (possibly empty) sequence $w \in Q^*$, map $f \in \Phi$, modes $q, q_0 \in Q$, and indices $j = 1, \ldots, m$,

$$S_{f,q}(w) = D^{(1,1,\ldots,1,0)} f_{0,wq} \text{ and } S_{q,q_0,j}(w) = D^{(1,1\ldots,1,0)} y^{\Phi}_{e_j,q_0 wq}. \tag{3.17}$$

Notice that for $w = \epsilon$, (3.17) yields $S_{f,q}(\epsilon) = D^{(0)} f_{0,q} = f_{0,q}(0)$ and $S_{q,q_0,j}(\epsilon) = D^{(1,0)} y^{\Phi}_{e_j,q_0 q}$. That is, for each word $w = q_1 q_2 \ldots q_k \in Q^*$, the vector $S_{f,q}(w)$ is the derivative of f with respect to the first k switching times evaluate at zero, if the continuous input is zero, the sequence of discrete modes is $q_1 q_2 \dots q_k q$ and the last switching time is zero. Similarly, $S_{q,q_0,j}(w)$ is obtained from y^{Φ} by taking the derivatives at zero with respect to the first $k+1$ switching times, if the continuous input is constant and it equals the jth unit vector in $\mathcal{U}$, the sequence of discrete modes is $q_0 q_1 q_2 \dots q_k q \in Q^+$, and the last switching time is zero. As we have indicated earlier, $S_{f,q}(w)$ and $S_{q,q_0,j}(w)$ collect the high-order derivatives we need for realization theory.

Definition 3.6 (Hankel-matrix)**.** Assume that the cardinality of Q is N, and fix the enumeration

$$Q = \{\sigma_1, \sigma_2, \dots, \sigma_N\}. \tag{3.18}$$

Define the *Hankel-matrix* of Φ as the infinite matrix, the rows of which are indexed by pairs (v, i) where $v \in Q^*$ and $i \in I = \{1, 2, \dots, pN\}$, and the columns of which are indexed by (w, j), where $w \in Q^*$ and $j \in J_{\Phi} = \Phi \cup (Q \times \{1, \dots, m\})$, *i.e.* $H_{\Phi} \in \mathbb{R}^{(Q^* \times I) \times (Q^* \times J_{\Phi})}$. For any $w, v \in Q^*$, $j \in J_{\Phi}$, and any $i \in I$ of the form $i = pK + r$ where $K = 0, 1, \dots, N-1$ and $r = 1, \dots, p$, the entry $(H_{\Phi})_{(v,i),(w,j)}$ with row index (v, i) and column index (w, j) is defined as follows:

$$(H_{\Phi})_{(v,i),(w,j)} = \begin{cases} (S_{\sigma_{K+1},q,z}(wv))_r & \text{if } j = (q,z) \in Q \times \{1, \dots, m\} \\ (S_{f,\sigma_{K+1}}(wv))_r & \text{if } j = f \in \Phi. \end{cases} \tag{3.19}$$

Here $(S_{\sigma_{K+1},q,z}(ww))_r$ and $(S_{f,\sigma_{K+1}}(wv))_r$ denote the rth element of the vectors $S_{\sigma_{K+1},q,z}(wv) \in \mathbb{R}^p$ and $S_{f,\sigma_{K+1}}(wv)$ from (3.17) respectively. Following the convention of Section 2.1, we define the *rank of* H_{Φ}, denoted by rank H_{Φ}, as the dimension of the linear space spanned by the columns of H_{Φ}.

I.e., the block $((H_{\Phi})_{(v,i),(w,j)})_{i=1,\dots,pN} = \begin{bmatrix} (H_{\Phi})_{(v,1),(w,j)} & (H_{\Phi})_{(v,2),(w,j)} & \dots & (H_{\Phi})_{(v,pN),(w,j)} \end{bmatrix}^T \in \mathbb{R}^{pN \times 1}$ of H_{Φ} formed by the entries indexed by the column index (w, j) and row indices (v, i), $i = 1, 2, \dots, Np$ equals

$$((H_{\Phi})_{(v,i),(w,j)})_{i=1,\dots,pN} = \begin{cases} \begin{bmatrix} (S_{\sigma_1,q,z}(wv))^T & \dots & (S_{\sigma_N,q,z}(wv))^T \end{bmatrix}^T & \text{for } j = (q,z) \in Q \times \{1, \dots, m\} \\ \begin{bmatrix} (S_{f,\sigma_1}(wv))^T & \dots & (S_{f,\sigma_N}(wv))^T \end{bmatrix}^T & \text{if } j = f \in \Phi. \end{cases}$$

The main theorem on the existence of a realization for arbitrary switching is as follows.

Theorem 3.6 (existence)**.** *The family of input-output maps Φ has a realization by a linear switched system if and only if Φ has a generalized kernel representation and the rank of H_{Φ} is finite,* i.e., rank $H_{\Phi} < +\infty$.

The proof of Theorem 3.6 will be presented in Section 5.2.

Remark 3.7 (relationship of Hankel-matrix with the functions G^{Φ}_w and $K^{f,\Phi}_w$)**.** The high-order derivatives $S_{q,q_0,z}(w)$ and $S_{f,q}(w)$ can also be expressed through the derivatives of the functions $K^{f,\Phi}_{wq}$ and $G^{\Phi}_{q_0 wq}$.

Remark 3.8 (relationship with the classical Hankel matrix)**.** If we apply the framework above to the classical linear realization problem, *i.e.* if we assume $Q = \{q\}$, $\Phi = \{f\}$, $y^{\Phi} = f$, then the columns of H_{Φ} indexed by $Q^* \times \Phi$ are all zero. The classical Hankel-matrix corresponds to the columns of H_{Φ} indexed by $(w, (q, j))$, $w \in Q^*$ and $q \in Q$, $j = 1, \dots, m$. Hence, the rank H_{Φ} coincides with the rank of the classical Hankel matrix, and thus Theorem 3.6 yields the classical results, if applied to the linear case.

3.3.3. *Existence of a realization: constraint switching*

In this section, $L \subseteq Q^+$ *denotes the set of all admissible sequences of discrete modes and* $\Phi \subseteq F(PC(T, \mathcal{U}) \times TL, \mathcal{Y})$ *is the family of input-output maps with the switching constraint* L.

We start with defining the Hankel-matrix H_{Φ} of Φ. We try to extend the definition of a Hankel-matrix for arbitrary switching. More precisely, we will define the Hankel matrix H_{Φ} in terms of vectors $T_{q,q_0,j}(w) \in \mathbb{R}^p$ and $T_{f,q}(w) \in \mathbb{R}^p$ respectively, defined as certain high-order derivatives of the input-output maps. The role of $T_{q,q_0,j}(w)$ and $T_{f,q}(w)$ is similar to that of $S_{q,q_0,j}(w)$ and $S_{f,q}(w)$ for arbitrary switching. More precisely, we collect the set of all those sequence w, for which it holds that we can derive some information on the behavior of the system under w from the behavior of the system under an admissible sequence in L. Obviously, every sequence of discrete modes in L will have this property. Then, for sequences of this class we define $T_{q,q_0,j}(w)$ and $T_{f,q}(w)$ as a certain high-order derivative of y^Φ and $f \in \Phi$. For the sequence which are not in this class we set the values of $T_{f,q}(w)$ and $T_{q,q_0,j}(w)$ to zero. Although this is a rather crude approach, it yields necessary and sufficient conditions for existence of a realization in the case when L is regular. The assumption that L is regular is not very restrictive, as it contains the case when the sequences of discrete modes have to be traces of a known finite-state machine. The details of the approach outlined above go as follows. For each word $w \in Q^*$, and discrete modes $q, q_0 \in Q$ define the sets

$$\begin{aligned} F_{q,q_0}(w) = \{(v,(\alpha,z)) \in Q^* \times (\mathbb{N}^* \times Q^*) \mid vz \in L,\, z = z_1 \dots z_k, z_1,\dots,z_k \in Q,\, k > 0, \\ \alpha = (\alpha_1,\alpha_2,\dots,\alpha_k) \in \mathbb{N}^k,\, q_0 w q = z_1 z_1^{\alpha_1} z_2^{\alpha_2} \dots z_k^{\alpha_k} z_k\} \\ F_q(w) = \{(v,(\alpha,z)) \in Q^* \times (\mathbb{N}^* \times Q^*) \mid vz \in L,\, z = z_1 \dots z_k,\, z_1,\dots,z_k \in Q,\, k > 0, \\ \alpha = (\alpha_1,\alpha_2,\dots,\alpha_k) \in \mathbb{N}^k,\, wq = z_1^{\alpha_1} z_2^{\alpha_2} \dots z_k^{\alpha_k} z_k\}. \end{aligned} \tag{3.20}$$

In words, the triple $(v,(\alpha,z))$ belongs to $F_{q,q_0}(w)$ if and only if $v, z \in Q^*$ are such that $vz \in L$ and the tuple $\alpha \in \mathbb{N}^{|z|}$ has the following property. If $z_1,\dots,z_k \in Q$ are the letters of z, *i.e.* $|z| = k$ and $z = z_1 \dots z_k$, then $\alpha = (\alpha_1,\dots,\alpha_k)$, and the word w can be obtained from z by repeating α_i times the ith letter of z, for all $i = 1,\dots,k$. Notice that repeating a letter zero times amounts to erasing it. Hence, w can be obtained from z by repeating each letter of z several times or erasing it , and α_i, $i = 1,\dots,k$ specifies the number of repetitions or deletion (if $\alpha_i = 0$) of the ith letter of z. In addition, we require that the first letter z equals q_0 and the last letter of z equals q. These requirements are encoded by $q_0 w q = z_1 z_1^{\alpha_1} \dots z_k^{\alpha_k} z_k$. Similarly, the triple $(v,(\alpha,z))$ belongs to $F_q(w)$ if $v, z \in Q^*$ are such that $vz \in L$ and the tuple $\alpha \in \mathbb{N}^{|z|}$ has the following property. Let $z_1,\dots,z_k \in Q$, $k \geq 0$ be the letters of z, that is, $z = z_1 \dots z_k$ and $|z| = k$. Then $\alpha = (\alpha_1,\dots,\alpha_k)$ and the word w can be obtained from z by repeating α_i times the ith letter of z, for each $i = 1,\dots,k$. In addition, the last letter of z is required to be q. The conditions above is encoded as $wq = z_1^{\alpha_1} \dots z_k^{\alpha_k} z_k$. In order to present the intuition behind the above definition, we need the following notation.

Notation 3.2. Denote by $\mathbb{O}_l$ the tuple $(0,0,\dots,0) \in \mathbb{N}^l$, $l > 0$, each entry of which is zero.

Notation 3.3. Let $\alpha = (\alpha_1,\dots,\alpha_k) \in \mathbb{N}^k$. Denote by α^+ the tuple $\alpha^+ = (\alpha_1 + 1, \alpha_2,\dots,\alpha_k) \in \mathbb{N}^k$, $k > 0$.

Notation 3.4. If $\alpha \in \mathbb{N}^k$ and $\beta \in \mathbb{N}^l$ are two tuples of natural numbers, then denote by (α,β) the $k + l$ tuple defined as follows; $(\alpha,\beta) = (\alpha_1,\alpha_2,\dots,\alpha_k,\beta_1,\beta_2,\dots,\beta_l)$.

The intuition behind the definitions of the sets in (3.20) is the following. It can be shown that if (Σ,μ) is a linear switched system realization of Φ, then

$$\begin{aligned} &\forall (v,(\alpha,z)) \in F_{q,q_0}(w) : D^{(\mathbb{O}_{|v|},\alpha^+)} y^\Phi_{e_j,vz} = D^{(1,1,\dots,1,0)} (y_\Sigma(0,.,.))_{e_j,q_0 w q} \quad \text{and} \\ &\forall (v,(\alpha,z)) \in F_q(w) : D^{(\mathbb{O}_{|v|},\alpha)} f_{0,vz} = D^{(1,1,\dots,1,0)} (y_\Sigma(\mu(f),.,.))_{0,wq} \end{aligned}$$

for each $j = 1,\dots,m$, $f \in \Phi$, $w \in Q^*$, $q, q_0 \in Q$. That is, $F_{q,q_0}(w)$ (resp. $F_q(w)$) is non-empty, if we can deduce from Φ some information on the output of Σ when the initial condition is 0 (resp. $\mu(f)$) and the switching sequence is $q_0 w q$ (resp. wq). We define $T_{f,q}(w) \in \mathbb{R}^p$ and $T_{q,q_0,j}(w) \in \mathbb{R}^p$ as follows:

$$T_{q,q_0,j}(w) = \begin{cases} D^{(\mathbb{O}_{|v|},\alpha^+)} y^\Phi_{e_j,vz} & \text{if } F_{q,q_0}(w) \neq \emptyset, \text{ and} \\ & (v,(\alpha,z)) \in F_{q,q_0}(w) \\ 0 & \text{otherwise} \end{cases} \quad \text{and } T_{f,q}(w) = \begin{cases} D^{(\mathbb{O}_{|v|},\alpha)} f_{0,vz} & \text{if } F_q(w) \neq \emptyset, \text{ and} \\ & (v,(\alpha,z)) \in F_q(w) \\ 0 & \text{otherwise.} \end{cases} \tag{3.21}$$

Similarly to the case of arbitrary switching, $T_{q,q_0,j}(w)$ and $T_{f,q}(w)$ can be expressed *via* the functions of the generalized kernel representation of Φ. Notice that it is not entirely trivial that $T_{q,q_0,j}(w)$ and $T_{f,q}(w)$ are well-defined; it will be shown in Section 5.3. Using the definitions above we will define the Hankel-matrix H_Φ of the family of input-output maps Φ in exactly the same way as for the case of arbitrary switching, but one uses the vectors $T_{q,q_0,j}(w)$ and $T_{f,q}(w)$ instead of $S_{q,q_0,j}(w)$ and $S_{f,q}(w)$. The formal definition goes as follows.

Definition 3.7 (Hankel-matrix: constrained switching)**.** As in (3.18), fix an enumeration $Q = \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$ of the set of discrete modes Q. The columns of the Hankel-matrix H_Φ are indexed by all the pairs (w, j) where $w \in Q^*$ and $j \in J_\Phi = \Phi \cup Q \times \{1, \ldots, m\}$. The rows of H_Φ are indexed by pairs (v, i) where $i \in I = \{1, \ldots, Np\}$ and $v \in Q^*$, *i.e.* $H_\Phi \in \mathbb{R}^{(Q^* \times I) \times (Q^* \times J_\Phi)}$. For any $w, v \in Q^*$, $j \in J_\Phi$, and index $i \in I = \{1, \ldots, Np\}$ of the form $i = pK + r$, where $K = 0, 1, \ldots, N-1$ and $r = 1, \ldots, p$, the entry $(H_\Phi)_{(v,i),(w,j)}$ is defined as follows:

$$(H_\Phi)_{(v,i),(w,j)} = \begin{cases} (T_{\sigma_{K+1},q,z}(wv))_r & \text{if} \quad j = (q, z) \in Q \times \{1, \ldots, m\} \\ (T_{f,\sigma_{K+1}}(wv))_r & \text{if} \quad j = f \in \Phi. \end{cases} \tag{3.22}$$

Here $T_{\sigma_i,q,z}(w)$ and $T_{f,\sigma_i}(w)$ are defined as in (3.21), and $(T_{\sigma_{K+1},q,z}(wv))_r$, $(T_{f,\sigma_{K+1}}(wv))_r$ denote the rth entry of the vectors $T_{\sigma_{K+1},q,z}(wv) \in \mathbb{R}^p$ and $T_{f,\sigma_{K+1}}(wv) \in \mathbb{R}^p$ respectively. Following the convention of Section 2.1, the rank of H_Φ, denoted rank H_Φ, is the dimension of the linear space spanned by columns of H_Φ.

Theorem 3.7 (realization of input-output maps: constrained switching)**.** *If Φ admits a generalized kernel representation with constraint L and the rank of the Hankel matrix of Φ is finite,* i.e. rank $H_\Phi < +\infty$, *then Φ has a realization by a linear switched system. Assume that L is regular. Then Φ has a realization by a linear switched system with constraint L if and only if Φ has a generalized kernel representation with constraint L and the rank of the Hankel-matrix H_Φ is finite,* i.e. rank $H_\Phi < +\infty$.

The proof of Theorem 3.7 will be presented in Section 5.3.

Remark 3.9 (algorithms and partial realization theory)**.** The proofs of Theorems 3.6 and 3.7 yield procedures for constructing a linear switched system realization of a family of input-output maps Φ from the columns of the Hankel-matrix H_Φ, both for the case of arbitrary and constrained switching. For the case of arbitrary switching, the thus constructed realization will be minimal, for the case of constrained switching the thus constructed realization will be semi-reachable and observable and quasi-minimal. The details of the construction will be presented in Section 5. In addition, it is possible to formulate a partial realization theory for linear switched systems both for arbitrary and constrained switching. For the details see [20].

4. 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 formal power series, see [1,14]. In order to keep the exposition self-contained, the proofs of those theorems which are not part of the classical theory, will briefly be sketched.

4.1. Formal power series: definition and basic concepts

Let X be a finite set, which we will refer to as alphabet. Recall from Section 2.1 the notion of a word over an alphabet and the related concepts. A *formal power series* S with coefficients in $\mathbb{R}^p$ is a map

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

There are many ways to give an intuition for the definition of a formal power series. For the purposes of this paper the most suitable one is to think of a formal power series as the output of a machine which reads symbols of X from its input tape and writes elements of $\mathbb{R}^p$ to its output tape. We denote by $\mathbb{R}^p\langle\langle X^* \rangle\rangle$ the set of all formal power series with coefficients in $\mathbb{R}^p$. The set $\mathbb{R}^p\langle\langle X^* \rangle\rangle$ forms a vector space with respect to point-wise addition and multiplication. That is, if α, $\beta \in \mathbb{R}$ and $S, T \in \mathbb{R}^p\langle\langle X^* \rangle\rangle$, then the linear combination

$\alpha S+\beta T$ is defined by $\forall w \in X^*$, $\alpha S(w)+\beta T(w)$. Recall from [1], *Hadamard product* on formal power series; if $S, T \in \mathbb{R}^p\langle\langle X^*\rangle\rangle$, then the *Hadamard product* $S \odot T \in \mathbb{R}^p\langle\langle X^*\rangle\rangle$ is defined by

$$(S \odot T)(w) = \begin{bmatrix} S_1(w)T_1(w), & S_2(w)T_2(w), & \dots, & S_p(w)T_p(w) \end{bmatrix}^T \in \mathbb{R}^p \tag{4.1}$$

for all $w \in X^*$, where for each $i = 1, \dots, p$, we denote by $S_i(w)$ and $T_i(w)$ the ith entry of the vector $S(w) \in \mathbb{R}^p$ and $T(w) \in \mathbb{R}^p$ respectively. That is, the ith entry of $(S \odot T)(w)$ is the product of the ith entry of $S(w)$ and the ith entry of $T(w)$ for $i = 1, \dots, p$. In the sequel we will be interested in *families of formal power series*.

Definition 4.1 (family of formal power series)**.** Let J be an arbitrary (possibly infinite) set. A *family of formal power series in* $\mathbb{R}^p\langle\langle X^*\rangle\rangle$ *indexed by* J is simply a collection $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}$ of formal power series from $\mathbb{R}^p\langle\langle X^*\rangle\rangle$ indexed by elements of J.

Notice that we do not require S_j, $j \in J$ to be all distinct, *i.e.* $S_l = S_j$ for some indices $j, l \in J$, $j \neq l$ is allowed. One can think of a family of formal power series as a family of input-output maps of the machine described above, realized from a set of initial states indexed by elements of J.

4.2. Rational representations and rational formal power series

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

$$R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in X}, B, C) \tag{4.2}$$

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 Euclidean bases, and we call them the *state-transition matrices* and the *readout matrix* respectively. If $\mathrm{card}(J) = 1$, then the above definition of a rational representation is essentially the same as the classical definitions of [1,14,24]. In fact, a rational representation can be viewed as a Moore-automaton [4,8] with the state-space $\mathcal{X}$, with input space X^*, with output space $\mathbb{R}^p$. The state transition function $\delta : \mathcal{X} \times X \to \mathcal{X}$ is given by the linear map $\delta(x, \sigma) = A_\sigma x$. The output map $\mu : \mathcal{X} \to \mathbb{R}^p$ is given by $\mu(x) := Cx$. The set of initial states is given by $\{B_j \mid j \in J\}$. The point of view described above was followed in [4,23]. Formal power series represent input-output maps of exactly this kind of systems.

More precisely, let $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}$ be a family of formal power series indexed by J. The representation R from (4.2) is said to be a *representation of* Ψ, if for each index $j \in J$,

$$S_j(\epsilon) = CB_j \text{ and } S_j(\sigma_1\sigma_2\dots\sigma_k) = CA_{\sigma_k}A_{\sigma_{k-1}}\dots A_{\sigma_1}B_j \tag{4.3}$$

for any sequence $\sigma_1, \sigma_2, \dots, \sigma_k \in X$, $k > 0$. We say that the family Ψ is *rational*, if there exists a representation R such that R is a representation of Ψ. A formal power series $S \in \mathbb{R}^p\langle\langle X^*\rangle\rangle$ which is called rational in [1,14,24] is a formal power series such that the family $\{S\}$ is rational according to the definition above.

Notation 4.1. Let $A_\sigma : \mathcal{X} \to \mathcal{X}, \sigma \in X$ be linear maps and let $w \in X^*$ be a word over X. If $w = \epsilon$, then let A_ϵ be the identity map. If $w = \sigma_1\sigma_2\dots\sigma_k \in X^*$, $\sigma_1, \dots \sigma_k \in X$, $k > 0$, then A_w denotes the following composition

$$A_w = A_{\sigma_k}A_{\sigma_{k-1}}\dots A_{\sigma_1}. \tag{4.4}$$

That is, $A_\epsilon(x) = x$ for all $x \in \mathcal{X}$, and $A_{w\sigma} = A_\sigma A_w$ holds for all $w \in X^*$, $\sigma \in X$. With the notation above, (4.3) can be rewritten as $S_j(w) = CA_wB_j$ for all $w \in X^*$, $j \in J$.

A representation $R_{\min}$ of Ψ is called *minimal* if for each representation R of Ψ, $\dim R_{\min} \le \dim R$, *i.e.* $R_{\min}$ is a rational representation of Ψ with the smallest possible state-space dimension.

Next, we define the notions of *observability* and *reachability* for rational representations. Define the subspaces

$$W_R = \mathrm{Span}\{A_w B_j \in \mathcal{X} \mid w \in X^*,\ j \in J\} \text{ and } O_R = \bigcap_{w \in X^*} \ker CA_w. \tag{4.5}$$

The subspace W_R is referred to as the *reachability subspace of* R and the subspace O_R is referred to as the *observability subspace of* R. The subspace above have the following automaton-theoretic interpretation. W_R is the span of states reachable by a word $w \in X^*$ from an initial state B_j, and two states x_1, x_2 are indistinguishable, *i.e.* $CA_w x_1 = CA_w x_2$ for all $w \in X^*$ if and only if $x_1 - x_2 \in O_R$. We will say that the representation R is *reachable* if $\dim W_R = \dim R$, and we will say that R is *observable* if $O_R = \{0\}$.

Remark 4.1 (computability)**.** If J is finite, then the observability and reachability of R can be checked effectively, and the corresponding observability and reachability subspaces can be computed.

Next, we define the notion of morphism between rational representations. This is analogous to algebraic similarity for linear systems. 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 map $T : \mathcal{X} \to \widetilde{\mathcal{X}}$ is called a *representation morphism* from R to $\widetilde{R}$ and is denoted by $T : R \to \widetilde{R}$ if T commutes with A_σ, B_j and C for all $j \in J$, $\sigma \in X$, that is, if the following equalities hold

$$TA_\sigma = \widetilde{A}_\sigma T,\ \forall \sigma \in X, \quad TB_j = \widetilde{B}_j,\ \forall j \in J, \quad C = \widetilde{C}T. \tag{4.6}$$

The representation morphism T is called *surjective, injective, isomorphism* if T is a surjective, injective or isomorphism respectively if viewed as a linear map.

Lemma 4.1. *R is a representation of the family Ψ if and only if $\widetilde{R}$ is a representation of Ψ. If T is an isomorphism, then* $\dim R = \dim \widetilde{R}$ *and R is observable (reachable) if and only if $\widetilde{R}$ is observable (reachable).*

Remark 4.2. Let R be representation of Ψ of the form (4.2), 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 by taking the matrix representation of $TA_\sigma T^{-1}$, CT^{-1} and the column vector representation of TB_j with respect to the natural Euclidean basis of $\mathbb{R}^n$. 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, as opposed to linear maps and elements of abstract vector spaces.

4.3. **Existence and minimality of rational representations: main results**

Below we state the main results on existence and minimality of representations of families of rational formal power series. We start with the definition of the concept of *Hankel matrix* of a family of formal power series. Let $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^* \rangle\rangle \mid j \in J\}$ be a family of formal power series.

Construction 4.1 (Hankel-matrix)**.** Define the *Hankel-matrix* of Ψ as the following infinite matrix H_Ψ. The rows of H_Ψ are indexed by pairs (v, i) where $v \in X^*$ is an arbitrary word and $i = 1, \ldots, p$. The columns of H_Ψ are indexed by pairs (w, j) where $w \in X^*$ and $j \in J$. That is, H_Ψ is a matrix $H_\Psi \in \mathbb{R}^{(X^* \times I) \times (X^* \times J)}$, $I = \{1, \ldots, p\}$. 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 defined as

$$(H_\Psi)_{(v,i)(w,j)} = (S_j(wv))_i \tag{4.7}$$

where $(S_j(wv))_i$ denotes the ith entry of the vector $S_j(wv) \in \mathbb{R}^p$.

Following the convention from Section 2.1, 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 rank H_Ψ.

Theorem 4.1 (existence of a representation)**.** *The family* Ψ *is rational,* i.e. Ψ *admits a rational representation, if and only if* $\operatorname{rank} H_\Psi < +\infty$, i.e. *the rank of the Hankel-matrix* H_Ψ *is finite.*

The proof of Theorem 4.1 is presented in Appendix B.

Remark 4.3. The proof of Theorem 4.1 is constructive and it provides a construction of a rational representation of Ψ from the columns of the Hankel-matrix H_Ψ. The details of the construction will be explained in Procedure B.1, Appendix B. For more details we refer the reader to [20].

Theorem 4.2 (minimal representation)**.** *Assume that* $R_{\min}$ *is a representation of* Ψ*. The following are equivalent:*

(i) $R_{\min}$ *is a minimal representation of* Ψ.
(ii) $R_{\min}$ *is reachable and observable.*
(iii) *If* R *is a reachable representation of* Ψ*, then there exists a surjective morphism* $T : R \to R_{\min}$.
(iv) $\operatorname{rank} H_\Psi = \dim R_{\min}$.

In addition, all minimal representations of Ψ *are isomorphic.*

The proof of Theorem 4.2 is presented in Appendix B.

Remark 4.4. In Appendix B we will present procedures for converting a representation of Ψ to a minimal one. This procedure can be implemented numerically.

4.4. Technical results on rational formal power series

Below we present a number of technical results which will be used in the proof of the realization theorems for switched systems. Let $L \subseteq X^*$ be a language over X. Define the formal power series $\bar{L} \in \mathbb{R}\langle\langle X^*\rangle\rangle$ by

$$\bar{L}(w) = \begin{cases} 1 & \text{if } w \in L \\ 0 & \text{otherwise.} \end{cases} \tag{4.8}$$

Lemma 4.2 ([1])**.** *If* L *is a regular language, then* $\bar{L}$ *is a rational formal power series.*

Recall from (4.1) the definition of the Hadamard product of two rational representations. We can extend the definition to families of formal power series; let $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}$ and $\Theta = \{T_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}$ be two families of formal power series, indexed by the same set J. Define the *Hadamard product* $\Psi \odot \Theta$ as the family of formal power series formed by the Hadamard products $S_j \odot T_j$ of elements of Ψ and Θ, *i.e.*

$$\Psi \odot \Theta := \{S_j \odot T_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}. \tag{4.9}$$

Lemma 4.3. *If* Ψ *and* Θ *are rational, then* $\Psi \odot \Theta$ *is rational. Moreover,* $\operatorname{rank} H_{\Psi\odot\Theta} \le \operatorname{rank} H_\Psi \cdot \operatorname{rank} H_\Theta$.

The proof of the lemma can be found in Appendix B. The lemmas below state some elementary properties of rational families of formal power series. The proof of the lemmas is routine and they are left to the reader.

Lemma 4.4. *The family of formal power series* $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}$ *is rational if and only if the family* $\Xi = \{S_{(i,j)} \in \mathbb{R}\langle\langle X^*\rangle\rangle \mid (i,j) \in \{1,\dots,p\} \times J\}$ *is rational, where for each* $j \in J$, $i = 1,\dots,p$, *for each word* $w \in X^*$, $S_{(i,j)}(w) \in \mathbb{R}$ *is the ith entry of the vector* $S_j(w) \in \mathbb{R}^p$.

Lemma 4.5. *Let* $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j \in J\}$ *and* $\Psi' = \{T_{j'} \in \mathbb{R}^p\langle\langle X^*\rangle\rangle \mid j' \in J'\}$ *be two families of formal power series with index sets* J *and* J' *respectively. Assume that there exists a map* $f : J' \to J$, *such that* $\forall j' \in J' : S_{f(j')} = T_{j'}$. *If* Ψ *is rational, then* Ψ' *is also rational and* $\operatorname{rank} H_{\Psi'} \le \operatorname{rank} H_\Psi$. *If* f *is surjective, then* $\operatorname{rank} H_{\Psi'} = \operatorname{rank} H_\Psi$.

Lemma 4.6. *If* J *is finite, then* Ψ *is rational if and only if* $S_j \in \mathbb{R}^p\langle\langle X^*\rangle\rangle$ *is rational for each* $j \in J$.

5. Proof of the main results

Below we present the proof of the main results presented in Section 3. Section 5.1 deals with the structure of input-output maps realizable by linear switched systems. Section 5.2 presents the proofs for the case of arbitrary switching. Section 5.3 deals with the case of constrained switching.

5.1. Input-output maps of linear switched systems

In this section we will present a number of technical results related to generalized kernel representations. The main technical result is Lemma 5.1. Let $L \subseteq Q^+$ the set of admissible sequences of discrete modes. Let Φ be a family of input-output maps with the switching constraint L. Let Σ be a linear switched system of the form (3.1). Let $\mu : \Phi \to \mathcal{X}$ be a map. Recall the notation from Notations 2.1–3.4.

Lemma 5.1. *The following are equivalent:*

(i) *(Σ, μ) is a realization of Φ with constraint L;*

(ii) *Φ has a generalized kernel representation with constraint L and for each word $w = q_1 \dots q_k \in L$; $q_1, \dots, q_k \in Q$, $k > 0$, input-output map $f \in \Phi$, multi-index $\alpha \in \mathbb{N}^k$, and integer $j = 1, \dots, m$*

$$\begin{aligned} D^\alpha y^\Phi_{e_j,w} &= D^\beta G^\Phi_{q_l q_{l+1}\dots q_k} e_j = C_{q_k} A^{\alpha_k}_{q_k} A^{\alpha_{k-1}}_{q_{k-1}} \dots A^{\alpha_l - 1}_{q_l} B_{q_l} e_j \text{ if } \alpha \neq (0,0,\dots,0) \\ D^\alpha f_{0,w} &= D^\alpha K^{f,\Phi}_w = C_{q_k} A^{\alpha_k}_{q_k} A^{\alpha_{k-1}}_{q_{k-1}} \dots A^{\alpha_1}_{q_1} \mu(f) \end{aligned} \tag{5.1}$$

where $l \in \{1, \dots, k\}$ is such that $\alpha_1 = \ldots = \alpha_{l-1} = 0$ and $\alpha_l > 0$, e_j is the jth unit vector of $\mathcal{U} = \mathbb{R}^m$, and the tuple β is of the form $\beta = (\alpha_l - 1, \alpha_{l+1}, \dots, \alpha_k)$;

(iii) *Φ has a generalized kernel representation with constraint L and for each word $w \in Q^*$, the following holds. For all discrete modes $q, q_0 \in Q$, if $F_{q,q_0}(w)$ is not empty, then for any element $(v, (\alpha, z)) \in F_{q,q_0}(w)$ such that $z = z_1 \dots z_k$, $z_1 \dots, z_k \in Q$, and for all $j = 1, \dots, m$,*

$$D^{(\mathbb{O}_{|v|}, \alpha^+)} y^\Phi_{e_j, vz} = D^{(0,\alpha,0)} G^\Phi_{q_0 z q} e_j = C_q A^{\alpha_k}_{z_k} A^{\alpha_{k-1}}_{z_{k-1}} \dots A^{\alpha_1}_{z_1} B_{q_0} e_j, \tag{5.2}$$

where $\alpha^+ = (\alpha_1 + 1, \alpha_2, \dots, \alpha_k)$. Similarly, for all discrete modes $q \in Q$, if $F_q(w)$ is not empty, then for all input-output maps $f \in \Phi$, for any $(v, (\alpha, z)) \in F_q(w)$, such that $z = z_1 \dots z_k$, $z_1, \dots, z_k \in Q$,

$$D^{(\mathbb{O}_{|v|}, \alpha)} f_{0,vz} = D^{(\alpha,0)} K^{f,\Phi}_{zq} = C_q A^{\alpha_k}_{z_k} A^{\alpha_{k-1}}_{z_{k-1}} \dots A^{\alpha_1}_{z_1} \mu(f). \tag{5.3}$$

The proof of Lemma 5.1 will be presented in Appendix A. The statement (ii) of Lemma 5.1 is used for realization theory with arbitrary switching, statement (iii) is used for realization theory with constrained switching.

Corollary 5.1. *Assume that $L = Q^+$,* i.e. *arbitrary switching is allowed. Then Σ is a realization of Φ if and only if Φ has a generalized kernel representation and there exists $\mu : \Phi \to \mathcal{X}$ such that for any sequence $w = q_1 \dots q_k \in Q^*$, $k \geq 0$, $q_1, \dots, q_k \in Q$, for any discrete modes $q, q_0 \in Q$, for any $f \in \Phi$ and $j = 1, \dots, m$,*

$$D^{(1,\mathbb{I}_k,0)} y^\Phi_{e_j, q_0 w q} = D^{(0,\mathbb{I}_k,0)} G^\Phi_{q_0 w q} e_j = C_q A_{q_k} A_{q_{k-1}} \dots A_{q_1} B_{q_0} e_j \tag{5.4}$$

$$D^{(\mathbb{I}_k,0)} f_{0,wq} = D^{(\mathbb{I}_k,0)} K^{f,\Phi}_{wq} = C_q A_{q_k} A_{q_{k-1}} \dots A_{q_1} \mu(f) \tag{5.5}$$

where $\mathbb{I}_k = (1,1,\dots,1) \in \mathbb{N}^k$ and e_j denotes the jth unit vector of $\mathcal{U} = \mathbb{R}^m$. Moreover, if $k = 0$, i.e. *$w = \epsilon$, then $(\mathbb{I}_k, 0) = 0$, $(1, \mathbb{I}_k, 0) = (1, 0)$ and $A_{q_k} A_{q_{k-1}} \dots A_{q_1}$ is interpreted as the identity matrix.*

The proof of Corollary 5.1 is presented in Appendix A.1. Notice that in contrast to Lemma 5.1, in Corollary 5.1 only first- and zero-order derivatives are considered. The reason that this can be done is that we can express a high-order derivative of the input-output map for a switching sequence by a zero- and first-order derivative of the same input-output map but for another switching sequence. However, the input-output map must be then defined for the other switching sequence, which may fail if $L \neq Q^+$.

5.2. **Arbitrary switching**

5.2.1. *Existence of a realization: proof of Theorem 3.6*

Consider a family of input-output maps Φ defined for arbitrary switching and assume that Φ has a generalized kernel representation. Below we prove Theorem 3.6 by defining the family of formal power series Ψ_Φ associated with Φ and by showing that existence of a linear switched system realization of Φ is equivalent to rationality of Ψ_Φ. To this end, recall from Section 3.3, (3.17) the definition of the vectors $S_{q,q_0,j}(w)$, $S_{f,q}(w)$, $q, q_0 \in Q$, $f \in \Phi$, $w \in Q^*$, $j = 1, \ldots, m$. The maps $S_{q,q_0,j} : Q^* \ni w \to S_{q,q_0,j}(w) \in \mathbb{R}^p$ and $S_{f,q} : Q^* \ni w \to S_{f,q}(w) \in \mathbb{R}^p$ define formal power series $S_{q,q_0,j}$ and $S_{f,q}$ in $\mathbb{R}^p\langle\langle Q^*\rangle\rangle$. Recall the enumeration of the set of discrete modes Q defined in (3.18), *i.e.* $Q = \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$. For each discrete mode $q \in Q$, index $j = 1, \ldots, m$, and input-output map $f \in \Phi$, define the formal power series $S_{q,j}, S_f \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle$ as follows; for each word $w \in Q^*$ let

$$\begin{aligned} S_{q,j}(w) &= \begin{bmatrix}(S_{\sigma_1,q,j}(w))^T & (S_{\sigma_2,q,j}(w))^T & \ldots & (S_{\sigma_N,q,j}(w))^T\end{bmatrix}^T, \\ S_f(w) &= \begin{bmatrix}(S_{f,\sigma_1}(w))^T & (S_{f,\sigma_2}(w))^T & \ldots & (S_{f,\sigma_N}(w))^T\end{bmatrix}^T. \end{aligned} \tag{5.6}$$

That is, the values of the formal power series $S_{q,j}$ are obtained by stacking up the values of $S_{\sigma_i,q,j}$ for $i = 1, \ldots, N$. Similarly, the values of S_f are obtained by stacking up the values of S_{f,σ_i} for $i = 1, \ldots, N$. Define the set $J_\Phi = \Phi \cup \{(q, z) \mid q \in Q,\ z = 1, \ldots, m\}$. Define the *family of formal power series associated with* Φ by

$$\Psi_\Phi = \{S_j \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle \mid j \in J_\Phi\}. \tag{5.7}$$

Notice that the only information needed to construct Ψ_Φ is the high-order derivatives at zero of the maps from Φ, *the knowledge of the functions* $K^{f,\Phi}_w$, G^Φ_w *is not required in order to construct* Ψ_Φ.

Remark 5.1 (equivalence of Hankel-matrices)**.** The Hankel-matrix H_{Ψ_Φ} of the family of formal power series Ψ_Φ is identical to the Hankel-matrix H_Φ of Φ as defined in Definition 3.6, and hence their ranks coincide.

Below we present the definition of the representation $R_{\Sigma,\mu}$ associated with (Σ, μ) such that (Σ, μ) is a realization of Φ if and only if $R_{\Sigma,\mu}$ is a representation of Ψ_Φ. To this end, we will need the following result.

Lemma 5.2. *Let Σ be a of the form (3.1), and let $\mu : \Phi \to \mathcal{X}$. If (Σ, μ) is a realization of Φ, then for all discrete modes $q_0 \in Q$, for all indices $j = 1, \ldots, m$, for all input-output maps $f \in \Phi$, and for any sequence $w \in Q^*$,*

$$S_{q_0,j}(w) = \begin{bmatrix}C^T_{\sigma_1} & C^T_{\sigma_2} & \ldots & C^T_{\sigma_N}\end{bmatrix}^T A_w B_{q_0} e_j \ \text{ and } \ S_f(w) = \begin{bmatrix}C^T_{\sigma_1} & C^T_{\sigma_2} & \ldots & C^T_{\sigma_N}\end{bmatrix}^T A_w \mu(f). \tag{5.8}$$

Here, Notation 4.1 is used, applied to the the matrices A_q, $q \in Q$ viewed as linear maps, i.e. *$A_\epsilon = I_n$, where I_n is the $n \times n$ identity matrix, and $A_w = A_{q_k} \ldots A_{q_1}$ if $w = q_1 \ldots q_k$, $q_1, \ldots, q_k \in Q$, $k > 0$.*

Proof. By Corollary 5.1, (Σ, μ) is a realization of Φ if and only if for all discrete modes $q, q_0 \in Q$, words, $w = q_1q_2 \ldots q_k \in Q^*$, $q_1, q_2, \ldots, q_k \in Q$, $k \geq 0$, indices $j = 1, \ldots, m$ and elements f of Φ, (5.4–5.5) hold. From the definition of $S_{q,q_0}(w)$ and $S_{f,q}(w)$ it follows that the left-hand side $D^{(1,\mathbb{I}_k,0)}y^\Phi_{e_j,q_0wq}$ of (5.4) equals $S_{q,q_0,j}(w)$ and the left-hand side $D^{(\mathbb{I}_k,0)}f_{0,wq}$ of (5.5) equals $S_{f,q}(w)$. On the other hand, if we apply the convention of Notation 4.1 to the right-hand side of (5.4) we get $C_qA_wB_{q_0}e_j$. Similarly, applying Notation 4.1 to the right-hand side of (5.5) yields $C_qA_w\mu(f)$. Combining the observations stated above, we get

$$D^{(1,\mathbb{I}_k,0)}y^\Phi_{e_j,q_0wq} = S_{q,q_0,j}(w) = C_qA_wB_{q_0}e_j \ \text{ and } \ D^{(\mathbb{I}_k,0)}f_{0,wq} = S_{f,q}(w) = C_qA_w\mu(f). \tag{5.9}$$

By "stacking up" the right-hand sides of equalities in (5.9) and using (5.6), we get the statement of the lemma. □

Since the representation $R_{\Sigma,\mu}$ below will also be used for the case of constrained switching, we will assume that Φ is a family of input-output maps with some switching constraint $L \subseteq Q^+$, *i.e.* $\Phi \subseteq F(PC(T,\mathcal{U}) \times TL, \mathcal{Y})$.

Construction 5.1 (representation associated with (Σ,μ))**.** Assume that $\Phi \subseteq F(PC(T,\mathcal{U}) \times TL, \mathcal{Y})$ for $L \subseteq Q^+$, and assume that Σ is of the form (3.1) and $\mu : \Phi \to \mathcal{X}$. Define the *representation associated with* (Σ,μ) by

$$R_{\Sigma,\mu} = (\mathcal{X}, \{A_q\}_{q\in Q}, \widetilde{B}, \widetilde{C}).$$

The state-space of $R_{\Sigma,\mu}$ is the same as the state-space of Σ, *i.e.* it is $\mathbb{R}^n = \mathcal{X}$. The alphabet of $R_{\Sigma,\mu}$ is set of discrete modes Q. For each discrete mode $q \in Q$, the corresponding state-transition matrix A_q of $R_{\Sigma,\mu}$ is identical to the matrix A_q of Σ. The readout matrix $\widetilde{C}$ is obtained by vertically "stacking up" the matrices $C_{\sigma_1}, \ldots, C_{\sigma_N}$ from top to bottom. That is, $\widetilde{C} = \begin{bmatrix} C_{\sigma_1}^T & C_{\sigma_2}^T, & \ldots, & C_{\sigma_N}^T \end{bmatrix}^T \in \mathbb{R}^{pN\times n}$. Here, $\sigma_1, \ldots, \sigma_N$ is the enumeration of Q defined in (3.18). The set of the initial states of $R_{\Sigma,\mu}$ is of the form $\widetilde{B} = \{\widetilde{B}_j \in \mathcal{X} \mid j \in J_\Phi\}$, where $\widetilde{B}_f = \mu(f)$ for $f \in \Phi$, and $\widetilde{B}_{(q,l)} = B_q e_l$ for $q \in Q$ and $l = 1, \ldots, m$, *i.e.* $\mathbb{B}_{(q,l)}$ is the lth column of B_q.

The intuition behind the definition of $R_{\Sigma,\mu}$ is that if $L = Q^+$, then we would like $R_{\Sigma,\mu}$ to be a representation of Ψ_Φ if and only if (5.8) holds. It then follows that the A_q matrices of the representation $R_{\Sigma,\mu}$ should coincide with the A_q matrices of Σ. The initial states of the representation should be formed by the vectors B_f (in order to generate S_f), and $B_q e_j$ (in order to generate $S_{q,j}$). Finally, the readout map $\widetilde{C}$ should be formed by just "stacking up" the matrices C_q. Next, we construct a linear switched system realization (Σ_R, μ_R) from a representation R.

Construction 5.2 (linear switched system realization associated with a representation)**.** Let Φ be a family of input-output maps with the switching constraint $L \subseteq Q^+$. Consider a representation R of the following form

$$R = (\mathcal{X}, \{A_q\}_{q\in Q}, \widetilde{B}, \widetilde{C}). \tag{5.10}$$

We assume that the range of $\widetilde{C}$ is a subset of $\mathbb{R}^{Np}$, and that $\widetilde{B} = \{\widetilde{B}_j \in \mathcal{X} \mid j \in J_\Phi\}$, where $J_\Phi = \Phi \cup \{(q,z) \mid q \in Q,\, z = 1, \ldots, m\}$, *i.e.* R is a pN-J_Φ representation. If $\mathcal{X} = \mathbb{R}^n$ does not hold, then replace R with the isomorphic copy TR defined in Remark 4.2 whose state-space is $\mathbb{R}^n$. In the rest of the construction, we assume that $\mathcal{X} = \mathbb{R}^n$ for $n = \dim \mathcal{X}$ holds. Hence, we can assume that A_q, $q \in Q$ are $n \times n$ matrices, and $\widetilde{C}$ is a $pN \times n$ matrix. Define the *linear switched system realization* (Σ_R, μ_R) *associated with* R as follows. Let Σ_R be of the form (3.1) that is, the state-space of Σ_R is the same as that of R and for each discrete mode $q \in Q$, the matrix A_q of Σ_R is identical to the state-transition matrix A_q of R. The definition of the $p \times n$ matrices C_q, $q \in Q$, goes as follows. Let $\sigma_1, \sigma_2, \ldots, \sigma_N$ be the enumeration of Q defined in (3.18). Then the $p \times n$ matrix C_{σ_1} is formed by the first p rows of $\widetilde{C}$, the matrix C_{σ_2} is formed by the second block of p rows of $\widetilde{C}$, and so on, up to C_{σ_N} which is formed by the last p rows of $\widetilde{C}$. That is, $\widetilde{C}$ can be expressed *via* the matrices C_q, $q \in Q$ as $\widetilde{C} = \begin{bmatrix} C_{\sigma_1}^T, & C_{\sigma_2}^T, & \ldots, & C_{\sigma_N}^T \end{bmatrix}^T$. For each discrete mode $q \in Q$, the $n \times m$ matrix B_q is obtained as follows; the lth column of B_q equals the initial state $\widetilde{B}_{(q,l)}$ for all $l = 1, \ldots, m$, *i.e.* $B_q e_l = \widetilde{B}_{(q,l)}$ for each $l = 1, \ldots, m$. The map $\mu_R : \Phi \to \mathcal{X}$ assigns to each element f of Φ the initial state of R indexed by f, *i.e.* $\mu_R(f) = \widetilde{B}_f$ for all $f \in \Phi$.

The intuition behind the definition is the following. We would like (Σ_R, μ_R) to be such that if we apply Construction 5.1 to it, then the resulting representation R_{Σ_R,μ_R} coincides with R. Hence, the matrices A_q of Σ_R should be the same as those of R, the matrices B_q should have as columns the vectors $\widetilde{B}_{q,l}$, the matrices C_q should be such that by stacking them up we get the map $\widetilde{C}$. Finally, μ_R should assign each f the initial state $\widetilde{B}_f$. It is easy to see that the above requirement holds, *i.e.* $\Sigma_{R_{\Sigma,\mu}} = \Sigma$, $\mu_{R_{\Sigma,\mu}} = \mu$ and the representation R_{Σ_R,μ_R} is isomorphic to R. In fact, if the state-space of R is of the form $\mathbb{R}^n$, then $R_{\Sigma_R,\mu_R} = R$.

Theorem 5.1. *Let Φ be a family of input-output maps defined for arbitrary switching. Assume that Φ has a generalized kernel representation. Then the following holds:*

(a) *The (Σ,μ) is realization of Φ if and only if the associated representation $R_{\Sigma,\mu}$ from Construction 5.1 is a rational representation of Ψ_Φ.*

(b) *The representation R is a representation of Ψ_Φ if and only if the associated linear switched system realization (Σ_R, μ_R) from Construction 5.2 is a realization of Φ.*

Proof. Notice that if R is a representation of Ψ_Φ, then R satisfies the assumptions of Construction 5.2. Part (a) follows from Lemma 5.2. Since R is isomorphic to R_{Σ_R,μ_R}, part (b) follows from part (a). □

Corollary 5.2. *If (Σ, μ) is a minimal realization of Φ, then $R_{\Sigma,\mu}$ is a minimal representation of Ψ_Φ. Conversely, if R is a minimal representation of Ψ_Φ, then (Σ_R, μ_R) is a minimal realization of Φ.*

Proof. Notice that $\dim \Sigma = \dim R_{\Sigma,\mu}$ and $\dim \Sigma_R = \dim R$. The statement follows now from Theorem 5.1. □

Proof of Theorem 3.6. If Φ has a realization, then Φ has a generalized kernel representation, moreover, by Theorem 5.1, Ψ_Φ has a representation, *i.e.* Ψ_Φ is rational. If Φ has a generalized kernel representation and Ψ_Φ is rational, *i.e.* it has a representation, then by Theorem 5.1 Φ has a realization. By Theorem 4.1 and Remark 5.1, rank $H_\Phi < +\infty$ is equivalent to Ψ_Φ being rational. □

5.2.2. *Minimality: proof of Theorem 3.3*

First we will formulate results establishing the relationship between observability and reachability and morphism for representations and observability, semi-reachability and system morphism for linear switched systems.

Lemma 5.3. *Assume that Φ is a family of input-output maps with the switching constraint $L \subseteq Q^+$. Let Σ be of the form (3.1), and $\mu : \Phi \to \mathcal{X}$. Then Σ is observable if and only if $R_{\Sigma,\mu}$ is observable, and (Σ, μ) is semi-reachable if and only if $R_{\Sigma,\mu}$ is reachable. Assume that R is a pN-J_Φ representation. Then R is reachable if and only if (Σ_R, μ_R) is semi-reachable and R is observable if and only if Σ_R is observable.*

Proof. Since R_{Σ_R,μ_R} is isomorphic to R, R is reachable or observable if and only if R_{Σ_R,μ_R} is reachable, respectively observable. Hence, it is enough to prove the first part of the lemma. Notice that $W_{R_{\Sigma,\mu}} = WR(\mathrm{Im}\mu)$, where $WR(\mathrm{Im}\mu)$ is the space $WR(\mathcal{X}_0)$ for $\mathcal{X}_0 = \mathrm{Im}\mu$ as defined in Proposition 3.1. Similarly, $O_{R_{\Sigma,\mu}} = O_\Sigma$ where O_Σ is the observability kernel of Σ, defined in Theorem 3.2. Here $W_{R_{\Sigma,\mu}}$ is the reachability subspace of $R_{\Sigma,\mu}$ as defined for $R = R_{\Sigma,\mu}$ in (4.5), and $O_{R_{\Sigma,\mu}}$ is the observability subspace of the representation $R_{\Sigma,\mu}$ as defined for $R = R_{\Sigma,\mu}$ in (4.5). Now the statement follows easily from Theorem 3.2 and Proposition 3.1 and the definitions of observability and reachability for representations. □

Lemma 5.4. *Assume that Φ is a family of input-output maps with switching constraint $L \subseteq Q^+$, and (Σ, μ) and $(\Sigma^{'}, \mu^{'})$ are linear switched system realizations such that the domains of μ and $\mu^{'}$ equal Φ. Then $T : (\Sigma, \mu) \to (\Sigma^{'}, \mu^{'})$ is a linear switched system morphism if and only if $T : R_{\Sigma,\mu} \to R_{\Sigma^{'},\mu^{'}}$ is a representation morphism.*

Recall that $T : (\Sigma, \mu) \to (\Sigma^{'}, \mu^{'})$ is a linear switched system morphism if T is a linear map from the state-space of Σ to the state-space of $\Sigma^{'}$ satisfying certain properties. Recall that a representation morphism between two representations is a linear map between the state-spaces of the representations which satisfies certain properties. Since the state space of $R_{\Sigma,\mu}$ and of $R_{\Sigma^{'},\mu^{'}}$ coincide with the state-space of Σ and $\Sigma^{'}$ respectively, it is justified to denote both the linear switched system morphism and the representation morphism by the same symbol, indicating that the underlying linear map is the same.

Proof of Lemma 5.4. Assume that Σ is of the form (3.1) and that $\Sigma^{'}$ is of the form $\Sigma^{'} = (\mathcal{X}^{'}, \mathcal{U}, \mathcal{Y}, \{(A^{'}_q, B^{'}_q, C^{'}_q) \mid q \in Q\})$. Recall from Construction 5.1 the definition of the representations $R_{\Sigma,\mu}$ and $R_{\Sigma^{'},\mu^{'}}$. Assume that $R_{\Sigma,\mu}$ is of the form $R_{\Sigma,\mu} = (\mathcal{X}, \{A_q\}_{q\in Q}, \widetilde{B}, \widetilde{C})$ and $R_{\Sigma^{'},\mu^{'}}$ is of the form $R_{\Sigma^{'},\mu^{'}} = (\mathcal{X}^{'}, \{A^{'}_q\}_{q\in Q}, \widetilde{B}^{'}, \widetilde{C}^{'})$ where $\widetilde{B} = \{\widetilde{B}_j \mid j \in J_\Phi\}$ and $\widetilde{B}^{'} = \{\widetilde{B}^{'}_j \mid j \in J_\Phi\}$ and $J_\Phi = \Phi \cup (Q \times \{1, \ldots, m\})$. Then T is a switched linear system morphism if and only if $TA_q = A^{'}_q T$, $C_q = C^{'}_q T$, $TB_q = B^{'}_q$ and $T\mu(f) = \mu^{'}(f)$ for each discrete mode $q \in Q$, and input-output map $f \in \Phi$. But this is equivalent to requiring that

(1) For each discrete mode $q \in Q$, and for each element $x \in \mathcal{X}$, $TA_q x = A^{'}_q Tx$;
(2) For each index $j \in \Phi \cup (Q \times \{1, \ldots, m\})$, $T\widetilde{B}_j = \widetilde{B}^{'}_j$; and

(3) For each $x \in \mathcal{X}$, $\widetilde{C}x = \left[(C_{\sigma_1}x)^T \quad \dots \quad (C_{\sigma_N}x)^T\right]^T = \widetilde{C}'Tx$.

In turn, Conditions (1)–(3) are equivalent to saying that T is a representation morphism. □

Proof Theorem 3.3. We will proof the following equivalences; (i) $\iff$ (ii), (i) $\iff$ (iii) and (i) $\iff$ (iv). Finally, we will prove that all minimal linear switched system realizations of Φ are equivalent.

(i) $\iff$ (ii). By Corollary 5.2 system (Σ, μ) is minimal if and only if $R = R_{\Sigma,\mu}$ from Construction 5.1 is minimal. By Theorem 4.2, R is minimal if and only if R is reachable and observable. By Lemma 5.3 the latter is equivalent to Σ being semi-reachable from $\mathrm{Im}\mu$ and observable.

(i) $\iff$ (iii). By Corollary 5.2 (Σ, μ) is minimal if and only if $R_{\Sigma,\mu}$ is minimal. By Theorem 4.2, $R_{\Sigma,\mu}$ is minimal if and only if $\dim R_{\Sigma,\mu} = \dim \Sigma = \mathrm{rank}\ H_{\Psi_\Phi} = \mathrm{rank}\ H_\Phi$.

(i) $\iff$ (iv). Again (Σ, μ) is minimal if and only if $R_{\Sigma,\mu}$ is minimal. Hence, by Theorem 4.2, we get that (Σ, μ) is minimal if and only if for any reachable representation R of Ψ_Φ there exists a surjective representation morphism $T : R \to R_{\Sigma,\mu}$. But any reachable representation R of Ψ_Φ can arise as an associated representation of a semi-reachable linear switched system realization of Φ. Indeed, by possibly replacing R with an isomorphic copy, we can construct the associated linear switched system realization (Σ_R, μ_R), which by Theorem 5.1 is a realization of Φ. By Lemma 5.3, if R is reachable, then (Σ_R, μ_R) is semi-reachable. In addition, the representation associated with (Σ_R, μ_R) is isomorphic to R. That is, we get that (Σ, μ) is minimal if and only if for any semi-reachable realization $(\hat{\Sigma}, \hat{\mu})$ of Φ there exists a surjective representation morphism $T : R_{\hat{\Sigma},\hat{\mu}} \to R_{\Sigma,\mu}$. By Lemma 5.4 we get that the latter is equivalent to $T : (\hat{\Sigma}, \hat{\mu}) \to (\Sigma, \mu)$ being a surjective linear switched system morphism.

Finally, we will show that minimal linear switched system realizations of Φ are algebraically similar. Let (Σ, μ) and $(\hat{\Sigma}, \hat{\mu})$ be two minimal linear switched system realizations of Φ. By Corollary 5.2, $R_{\Sigma,\mu}$ and $R_{\hat{\Sigma},\hat{\mu}}$ are minimal representations of Ψ_Φ. Then from Theorem 4.2 it follows that there exists a representation isomorphism $T : R_{\hat{\Sigma},\hat{\mu}} \to R_{\Sigma,\mu}$. The latter means that $T : (\hat{\Sigma}, \hat{\mu}) \to (\Sigma, \mu)$ is a linear switched system isomorphism. □

Procedure 5.1 (minimal realization from the Hankel-matrix)**.** Assume that Φ is a family of input-output maps with arbitrary switching. Using Procedure B.1, Appendix B we construct a minimal representation R of Ψ_Φ from $H_\Phi = H_{\Psi_\Phi}$. Then, we construct the linear switched system realization (Σ_R, μ_R). By Corollary 5.2, (Σ_R, μ_R) will be a minimal realization of Φ.

Procedure 5.2 (minimization)**.** Assume that Φ is a family of input-output maps with arbitrary switching. Let (Σ, μ) be a realization of Φ and compute the representation $R_{\Sigma,\mu}$ from Construction 5.1. By Theorem 5.1, $R_{\Sigma,\mu}$ is a representation of Ψ_Φ. Use Procedure B.4, Appendix B to transform $R_{\Sigma,\mu}$ into a minimal representation R of Ψ_Φ. Construct the realization (Σ_R, μ_R). By Corollary 5.2 (Σ_R, μ_R) is a minimal realization of Φ.

5.3. **Constrained switching**

5.3.1. *Existence of a realization: proof of Theorem* 3.7

Let $L \subseteq Q^+$ be the set of admissible sequences of discrete modes. Let Φ be a family of input-output maps with the switching constraint L. Assume that Φ has a generalized kernel representation with constraint L. We start with introducing the notion of a family of formal power series Ψ_Φ associated with Φ. Recall from Section 3.3, (3.20) the definition of the sets $F_{q,q_0}(w)$ and $F_q(w)$, $q, q_0 \in Q$. Define the languages $\widetilde{L}_{q,q_0}$, $\widetilde{L}_q$

$$\widetilde{L}_{q,q_0} = \{w \in Q^* \mid F_{q,q_0}(w) \neq \emptyset\} \text{ and } \widetilde{L}_q = \{w \in Q^* \mid F_q(w) \neq \emptyset\}. \tag{5.11}$$

That is, $\widetilde{L}_{q,q_0}$ (resp. $\widetilde{L}_q$) consists of all those words $w \in Q^*$ for which $F_{q,q_0}(w)$ (resp. $F_q(w)$) is not empty. Recall from (3.21) the definition of the vectors $T_{q,q_0,j}(w)$ and $T_{f,q}(w)$ for each word $w \in Q^*$, discrete modes $q, q_0 \in Q$, indices $j = 1, \dots, m$, and input-output maps $f \in \Phi$. It is easy to see that the maps $T_{q,q_0,j} : Q^* \ni w \mapsto T_{q,q_0,j}(w) \in \mathbb{R}^p$ and $T_{f,q} : Q^* \ni w \mapsto T_{f,q}(w) \in \mathbb{R}^p$ can be viewed as formal power series.

Lemma 5.5. *The formal power series $T_{q,q_0,j}$ and $T_{f,q}$ are well-defined.*

Proof. The lemma follows from Lemma 5.6 presented below. □

Lemma 5.6. *With the notation above, the formal power series $T_{q,q_0,j}$ and $T_{f,q}$ admit the following representation:*

$$T_{q,q_0,j}(w) = \begin{cases} D^{\alpha}G^{\Phi}_{z}e_j = D^{(0,\alpha,0)}G^{\Phi}_{q_0zq}e_j & \text{if } w \in \widetilde{L}_{q,q_0} \text{ and } (v,(\alpha,z)) \in F_{q,q_0}(w) \\ 0 & \text{otherwise} \end{cases}$$
$$T_{f,q}(w) = \begin{cases} D^{(\mathbb{O}_{|v|},\alpha)}K^{f,\Phi}_{vz} = D^{\alpha}K^{f,\Phi}_{z} = D^{(\alpha,0)}K^{f,\Phi}_{zq} & \text{if } w \in \widetilde{L}_{q} \text{ and } (v,(\alpha,z)) \in F_{q}(w) \\ 0 & \text{otherwise.} \end{cases} \tag{5.12}$$

Moreover, in (3.21) and (5.12) the values of $T_{f,q}(w)$ and $T_{q,q_0,j}(w)$ are independent from the particular choice of the elements $(v,(\alpha,z)) \in F_q(w)$ and $(v,(\alpha,z)) \in F_{q,q_0}(w)$ respectively.

The proof of Lemma 5.6 can be found in Appendix A.

Fix the enumeration $Q = \{\sigma_1,\sigma_2,\ldots,\sigma_N\}$ as in (3.18). Define the formal power series $T_{q,l}$, $T_f \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle$, $l \in \{1,\ldots,m\}$, $q \in Q$ and $f \in \Phi$ by requiring that for all $w \in Q^*$,

$$\begin{aligned} T_{q,l}(w) &= \left[(T_{\sigma_1,q,l}(w))^T, \quad (T_{\sigma_2,q,l}(w))^T, \quad \ldots, \quad (T_{\sigma_N,q,l}(w))^T\right]^T \\ T_f(w) &= \left[(T_{f,\sigma_1}(w))^T, \quad (T_{f,\sigma_2}(w))^T, \quad \ldots, \quad (T_{f,\sigma_N}(w))^T\right]^T. \end{aligned} \tag{5.13}$$

That is, the formal power series $T_{q,l}$ and T_f are simply formed by stacking up the values of $T_{\sigma_i,q,l}$ and T_{f,σ_i} respectively, $i = 1,\ldots,N$. Define the *family of formal power series associated with* Φ as

$$\Psi_{\Phi} = \{T_j \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle \mid j \in J_{\Phi}\} \tag{5.14}$$

where the index set J_{Φ} is defined as $J_{\Phi} = \Phi \cup (Q \times \{1,2,\ldots,m\})$.

Remark 5.2 (equivalence of Hankel-matrices)**.** The Hankel-matrix $H_{\Psi_{\Phi}}$ of the family of formal power series Ψ_{Φ} and the Hankel-matrix defined in Definition 3.7 are identical, and hence their respective ranks are identical.

In order to prove Theorem 3.7, we need the following two theorems.

Theorem 5.2. *We can construct a family of formal power series Ω_{Φ}, elements of which depend only on L, and for which the following holds. Assume that Σ is a linear switched system of the form (3.1) and let $\mu : \Phi \to \mathcal{X}$. If (Σ,μ) is a realization of Φ, then there exists a family of formal power series $K_{\Sigma,\mu}$ such that:*

- *$K_{\Sigma,\mu}$ is rational, and in addition all the elements of $K_{\Sigma,\mu}$ depend only on the parameters of (Σ,μ).*
- *The family of formal power series Ψ_{Φ} associated with Φ can be expressed as the Hadamard-product of $K_{\Sigma,\mu}$ and Ω_{Φ},* i.e.

$$\Psi_{\Phi} = \Omega_{\Phi} \odot K_{\Sigma,\mu}. \tag{5.15}$$

In addition, if L is a regular language, the family of formal power series Ω_{Φ} is rational.

The proof of Theorem 5.2 will be presented at the end of this section. The next theorem relates rational representations of Ψ_{Φ} and realizations of Φ. Recall the definition of $\mathrm{comp}(L)$ from (3.7).

Theorem 5.3. *If $R = (\mathcal{X},\{A_q\}_{q\in Q},B,C)$ is a representation of Ψ_{Φ}, then the associated linear switched system realization (Σ_R,μ_R) (defined in Construction 5.2) is a realization of Φ. Moreover, for each input-output map $f \in \Phi$, for each input $u \in PC(T,\mathcal{U})$ and for any switching sequence $w \in T(\mathrm{comp}(L))$,*

$$y_{\Sigma_R}(\mu_R(f),u,w) = 0. \tag{5.16}$$

The theorem above states the input-output map f and the input-output map induced by the initial state $\mu_R(f)$ of Σ_R coincide on admissible switching sequences. The output of Σ_R for those switching sequences which are not related to any admissible switching sequence is assumed to be zero.

Proof of Theorem 3.7. Recall that the Hankel-matrix of Φ as defined in Construction 3.7 coincides with the Hankel-matrix of the family of formal power series Ψ_Φ. If Φ has a generalized kernel representation with constraint L and rank $H_\Phi < +\infty$, then by Theorem 4.1 and Remark 5.2 the family of formal power series Ψ_Φ is rational. If Ψ_Φ is rational, then there exists a representation R of Ψ_Φ and by Theorem 5.3, (Σ_R, μ_R) is a realization of Φ with constraint L. That is, we have proved the first statement of the theorem.

Assume that L is regular and Φ is realized by (Σ, μ). Then by Theorem 5.2 Φ has a generalized kernel representation and with the notation of Theorem 5.2 it holds that $\Psi_\Phi = \Omega_\Phi \odot K_{\Sigma,\mu}$. Moreover, by Theorem 5.2, both $K_{\Sigma,\mu}$ and Ω_Φ is rational. Hence, by Lemma 4.3 their Hadamard-product is rational, and hence Ψ_Φ is rational. Again, by Theorem 4.1 and Remark 5.2 the rationality of Ψ_Φ means that rank $H_\Phi < +\infty$. □

We devote the rest of the section to proving Theorems 5.2 and 5.3. We start by presenting the construction of the families of formal power series Ω_Φ and $K_{\Sigma,\mu}$ defined in Theorem 5.2. To this end, recall from (5.11) the definition of the languages $\widetilde{L}_{q,q_0}$ and $\widetilde{L}_q$ for discrete modes $q, q_0 \in Q$.

Construction 5.3 (definition of Ω_Φ)**.** Define the power series $Z_{q,q_0}, Z_q \in \mathbb{R}^p\langle\langle Q^*\rangle\rangle$ by

$$Z_{q,q_0}(w) = \begin{cases} (1,1,\dots,1)^T \in \mathbb{R}^p & \text{if } w \in \widetilde{L}_{q,q_0} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad Z_q(w) = \begin{cases} (1,1,\dots,1)^T \in \mathbb{R}^p & \text{if } w \in \widetilde{L}_q \\ 0 & \text{otherwise} \end{cases}$$

for all $w \in Q^*$. That is, Z_{q,q_0} (resp. Z_q) is just the p tuple of the indicator functions, each of which returns one if a word belongs to $\widetilde{L}_{q,q_0}$ (resp. $\widetilde{L}_q$) and zero otherwise. We will define the power series $\Gamma_q, \Gamma \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle$ by stacking up the power series $Z_{\sigma_1,q}, \dots, Z_{\sigma_N,q}$, respectively $Z_{\sigma_1}, \dots, Z_{\sigma_N}$ in this order, that is, for all $w \in Q^*$,

$$\Gamma_q(w) = \left[(Z_{\sigma_1,q}(w))^T, \ (Z_{\sigma_2,q}(w))^T, \ \dots, \ (Z_{\sigma_N,q}(w))^T\right]^T \tag{5.17}$$

$$\Gamma(w) = \left[(Z_{\sigma_1}(w))^T, \ (Z_{\sigma_2}(w))^T, \ \dots, \ (Z_{\sigma_N}(w))^T\right]^T. \tag{5.18}$$

The family of formal power series Ω_Φ is indexed by $J_\Phi = \Phi \cup (Q \times \{1, \dots, m\})$ and it is of the form

$$\Omega_\Phi = \{\Xi_j \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle \mid j \in J_\Phi\} \text{ where } \forall j \in J : \Xi_j = \begin{cases} \Gamma & j = f \in \Phi \\ \Gamma_q & j = (q,l) \in Q \times \{1, \dots, m\}. \end{cases} \tag{5.19}$$

Lemma 5.7. *If L regular, then Ω_Φ is rational, and the rank of the Hankel-matrix H_{Ω_Φ} depends only on L.*

Lemma 5.7 is a corollary of the following lemma.

Lemma 5.8. *If $L \subseteq Q^+$ is regular, then $\widetilde{L}$, $\widetilde{L}_{q,q_0}$ and $\widetilde{L}_q$ are regular languages for each q, $q_0 \in Q$.*

The proof of the Lemma 5.8 can be found in the Appendix A.3.

Proof of Lemma 5.7. If L is regular, then by Lemma 5.8, $\widetilde{L}_{q,q_0}$ and $\widetilde{L}_q$ are regular languages for all $q, q_0 \in Q$. Recall from (5.18) the definition of Γ and recall from (5.17) the definition of Γ_q. Then it is easy to see that for each $l = 1, \dots, pN$, such that $l = p(z-1) + i$ for some $z = 1, \dots, N$, $i = 1, \dots p$, the lth coordinate of the vector $\Gamma(w)$ is of the form $(\Gamma(w))_l = \begin{cases} 1 & \text{if } w \in \widetilde{L}_{\sigma_z} \\ 0 & \text{otherwise} \end{cases}$ and the lth coordinate of $\Gamma_q(w)$ is of the form $(\Gamma_q(w))_l = \begin{cases} 1 & \text{if } w \in \widetilde{L}_{\sigma_z,q} \\ 0 & \text{otherwise} \end{cases}$. For each $l = 1, \dots, Np$, denote by $(\Gamma_q)_l$ and respectively by Γ_l the scalar valued formal power series formed by the lth coordinate of Γ_q and respectively Γ. From the regularity of $\widetilde{L}_{q_1,q_2}$ and $\widetilde{L}_{q_1}$, $q_1, q_2 \in Q$ and Lemma 4.3 it follows that $(\Gamma_q)_l$ and Γ_l are rational for all $l = 1, \dots, Np$ and $q \in Q$. Consider the family of formal power series $\Theta = \{\Gamma_j \mid j \in \{\emptyset\} \cup Q\}$, where $\Gamma_\emptyset = \Gamma$. Hence, since $(\Gamma_q)_i$ and Γ_i, $i = 1, \dots, Np$ are rational, by Lemmas 4.6 and 4.4 we get that Θ is rational. Notice that Θ depends only on L, hence the rank of the Hankel matrix of Θ depends only on L.

It is left to show that Ω_Φ is rational and the rank of its Hankel-matrix depends only on L. To this end, let $R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in\Sigma}, B, C)$ be a minimal representation of Θ. From Theorem 4.2 it follows then that R is reachable and observable, and $\dim R = \operatorname{rank} H_\Theta$. Define the indexed set $\widetilde{B} = \{\widetilde{B}_j \in \mathcal{X} \mid j \in \Phi \cup Q \times \{1,\dots,m\}\}$ as follows. For each $f \in \Phi$ let $\widetilde{B}_f = B_\emptyset$, and for each $i = 1,\dots,m$, $q \in Q$ let $\widetilde{B}_{(q,i)} = B_q$. Then it is easy to see that $\widetilde{R} = (\mathcal{X}, \{A_\sigma\}_{\sigma\in\Sigma}, \widetilde{B}, C)$ is a well-defined rational representation of the family Ω_Φ. Hence, Ω_Φ is rational. Moreover, it is easy to see that $\dim \widetilde{R} = \dim R$ and $\widetilde{R}$ is reachable and observable as well. Hence, by Theorem 4.2, $\dim \widetilde{R} = \operatorname{rank} H_{\Omega_\Phi}$. The latter implies that $\operatorname{rank} H_{\Omega_\Phi} = \operatorname{rank} H_\Theta$ depends only on L. □

Next, we proceed with defining the formal power series $K_{\Sigma,\mu}$ from (5.15).

Construction 5.4 (definition of $K_{\Sigma,\mu}$)**.** Let (Σ,μ) be a linear switched system realization of Φ. Define the family of input-output maps $\Theta_{\Sigma,\mu} = \{y_\Sigma(\mu(f),.,.) \mid f \in \Phi\}$. The elements of $\Theta_{\Sigma,\mu}$ are simply those input-output maps of Σ (defined for arbitrary switching) which are induced by an initial state of the form $\mu(f)$ for some $f \in \Phi$. Define $U(\mu) : \Theta_{\Sigma,\mu} \to \Phi$ by $U(\mu)(y_\Sigma(\mu(f),.,.)) = f$. The map $U(\mu)$ is well defined. Indeed, if $y_\Sigma(\mu(f_1),.,.) = y_\Sigma(\mu(f_2),.,.)$, then for all $u \in PC(T,\mathcal{U})$ and $w \in TL$, $f_1(u,w) = y_\Sigma(\mu(f_1),u,w) = y_\Sigma(\mu(f_2),u,w) = f_2(u,w)$. Notice that $(\Sigma, \mu \circ U(\mu))$ is a realization of $\Theta_{\Sigma,\mu}$ and for each $g = y_\Sigma(\mu(f),.,.) \in \Theta_{\Sigma,\mu}$, $\mu \circ U(\mu)(g) = \mu(f)$. Assume that the family of formal power series associated with $\Theta_{\Sigma,\mu}$ as defined in Section 5.2, (5.7), is of the form

$$\Psi_{\Theta_{\Sigma,\mu}} = \{S_z \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle \mid z \in \Theta_{\Sigma,\mu} \cup (Q \times \{1,2,\dots,m\})\}. \tag{5.20}$$

The family $K_{\Sigma,\mu}$ is indexed by $J_\Phi = \Phi \cup (Q \times \{1,\dots,m\})$ and it represents the following re-indexing of $\Psi_{\Theta_{\Sigma,\mu}}$,

$$K_{\Sigma,\mu} = \{V_j \in \mathbb{R}^{pN}\langle\langle Q^*\rangle\rangle \mid j \in J_\Phi\},\ \ \forall j \in J_\Phi : V_j = \begin{cases} S_{y_\Sigma(\mu(f),.,.)} & \text{if } j = f \in \Phi \\ S_{(q,l)} & \text{if } l = (q,j) \in Q \times \{1,\dots,m\}. \end{cases} \tag{5.21}$$

Lemma 5.9. *The family of formal power series $K_{\Sigma,\mu}$ is rational, moreover* $\operatorname{rank} H_{K_{\Sigma,\mu}} = H_{\Psi_{\Theta_{\Sigma,\mu}}} \le \dim\Sigma$.

Proof. From Theorem 3.6 it follows that $\Psi_{\Theta_{\Sigma,\mu}}$ is rational. Define the map $\phi : \Theta_{\Sigma,\mu} \cup (Q \times \{1,\dots,m\}) \to \Phi \cup (Q \times \{1,\dots,m\})$ by $\phi(g) = U(\mu)(g)$ for $g \in \Theta_{\Sigma,\mu}$ and $\phi((q,j)) = (q,j)$ for $q \in Q$, $j = 1,\dots,m$. The map ϕ is surjective and $V_{\phi(j)} = S_j$ for all $j \in \Theta_{\Sigma,\mu} \cup (Q \times \{1,\dots,m\})$. By Lemma 4.5, rationality of $\Psi_{\Theta_{\Sigma,\mu}}$ implies the rationality of $K_{\Sigma,\mu}$, and since ϕ is surjective we get that $\operatorname{rank} H_{K_{\Sigma,\mu}} = \operatorname{rank} H_{\Psi_{\Theta_{\Sigma,\mu}}}$. Finally, $\operatorname{rank} H_{\Psi_{\Theta_{\Sigma,\mu}}} = \operatorname{rank} H_{\Theta_{\Sigma,\mu}}$ by Remark 5.1, and by Theorem 3.3 we get that $\operatorname{rank} H_{\Psi_{\Theta_{\Sigma,\mu}}} \le \dim\Sigma$. □

Proof of Theorem 5.2. We show that Theorem 5.2 is satisfied by choosing Ω_Φ and $K_{\Sigma,\mu}$ as defined in (5.19) and (5.21). Note that the elements Ω_Φ depend only L and from Lemma 5.7 it follows that Ω_Φ is rational if L is regular. The elements of $K_{\Sigma,\mu}$ depend on the parameters of (Σ,μ) only and by Lemma 5.9 $K_{\Sigma,\mu}$ is rational. Hence, it is left to show that (Σ,μ) is a realization of Φ if and only if (5.15) holds. To this end, notice (5.15) is equivalent to

$$\forall f \in \Phi,\ q, q_0 \in Q,\ j = 1,2,\dots,m : T_{f,q} = S_{y_\Sigma(\mu(f),.,.),q} \odot Z_q \quad \text{and} \quad T_{q,q_0,j} = S_{q,q_0,j} \odot Z_{q,q_0}. \tag{5.22}$$

Here we used the notation of (3.17) applied to $\Theta_{\Sigma,\mu}$ and (5.6). That is, for each word $w \in Q^*$, $S_{y_\Sigma(\mu(f),.,.),\sigma_i}(w)$ and respectively $S_{\sigma_i,q_0,j}(w)$ are formed by the block of rows of $S_{y_\Sigma(\mu(f),.,.)}(w)$ and respectively $S_{q_0,j}(w)$ indexed by indices in the range $[p(i-1)+1, pi]$. Hence, it is enough to show that (Σ,μ) is a realization of Φ if and only if (5.22) holds for all $q \in Q$, $f \in \Phi$ and $j = 1,\dots,m$.

By Lemma 5.1, (Σ,μ) is a realization of Φ, if and only if Φ has a generalized kernel representation with constraint L, and (5.2) and (5.3) hold. Notice the following facts about the expressions in (5.2) and (5.3). If $(v,(\alpha,z)) \in F_{q,q_0}(w)$, and $z = z_1z_2\dots z_k$ for some $z_1,z_2,\dots,z_k \in Q$, then $w = z_1^{\alpha_1}z_2^{\alpha_2}\dots z_k^{\alpha_k}$. Similarly, if $(v,(\alpha,z)) \in F_q(w)$ and $z = z_1z_2\dots z_k$ for some $z_1,z_2,\dots,z_k \in Q$, then $w = z_1^{\alpha_1}z_2^{\alpha_2}\dots z_k^{\alpha_k}$. Recall the notation of Notation 4.1, (4.4). Then we get that $A_{z_k}^{\alpha_k}A_{z_{k-1}}^{\alpha_{k-1}}\dots A_{z_1}^{\alpha_1} = A_w$ if $w = z_1^{\alpha_1}z_2^{\alpha_2}\dots z_k^{\alpha_k}$. Combining (5.2), (5.3) and the definition of $T_{f,q}(w)$ and $T_{q,q_0,j}(w)$ presented in (5.12) we get that

$$T_{f,q}(w) = C_qA_w\mu(f) \ \text{ if } \ w \in \widetilde{L}_q, \qquad \text{and} \qquad T_{q,q_0,j}(w) = C_qA_wB_{q_0}e_j \ \text{ if } \ w \in \widetilde{L}_{q,q_0}. \tag{5.23}$$

Notice that $(\Sigma, \mu \circ U(\mu))$ is also a realization of $\Theta = \Theta_{\Sigma,\mu}$. Recall the definition of $R_{\Sigma,\mu\circ U(\mu)}$ from Construction 5.1. By Theorem 5.1, $R_{\Sigma,\mu\circ U(\mu)}$ is a representation of Ψ_Θ. Hence, from (5.9)

$$\forall q, q_0 \in Q,\ j = 1, \ldots, m, w \in Q^* : C_q A_w B_{q_0} e_j = S_{q,q_0,j}(w) \ \text{ and } \ C_q A_w \mu(f) = S_{y_\Sigma(\mu(f),.,.),q}(w). \tag{5.24}$$

Notice that if $w \notin \widetilde{L}_{q,q_0}$, then $T_{q,q_0,j}(w) = 0$; and if $w \notin \widetilde{L}_q$, then $T_{f,q}(w) = 0$. Combining this with (5.24),

$$T_{f,q}(w) = \begin{cases} S_{y_\Sigma(\mu(f),.,.),q}(w) & \text{if } w \in \widetilde{L}_q \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad T_{q,q_0,j}(w) = \begin{cases} S_{q,q_0,j}(w) & \text{if } w \in \widetilde{L}_{q,q_0} \\ 0 & \text{otherwise.} \end{cases} \tag{5.25}$$

From the definition of $Z_q(w)$ and $Z_{q,q_0}(w)$ and from the definition of the Hadamard-product we get that the right-hand sides of (5.25) equal $(S_{y_\Sigma(\mu(f),.,.)} \odot Z_q)(w)$ and $(S_{q,q_0,j} \odot Z_{q,q_0})(w)$ respectively, *i.e.* (5.22) holds. □

Proof of Theorem 5.3. Let $(\Sigma, \mu) = (\Sigma_R, \mu_R)$ and assume that Σ is of the form (3.1).

First, we show that (Σ, μ) is a realization of Φ. To this end, notice that R is a representation of Ψ_Φ, and hence for all $q_0 \in Q$, $f \in \Phi$, $j = 1, \ldots, m$, $w \in Q^*$, $T_{q_0,j}(w) = CA_wB_{(q_0,j)}$ and $T_f(w) = CA_wB_f$. From this, the definition of $T_{q,q_0,j}(w)$, and that of the matrices C_q, B_{q_0} of Σ, we get that $T_{q,q_0,j}(w) = C_qA_wB_{q_0}e_j$. Let $w \in \widetilde{L}_{q,q_0}$ and let $(v, (\alpha, z)) \in F_{q,q_0}(w)$. It then follows that $w = z_1^{\alpha_1} \ldots z_k^{\alpha_k}$ and hence $A_w = A_{z_k}^{\alpha_1} A_{z_{k-1}}^{\alpha_{k-1}} \ldots A_{z_1}^{\alpha_1}$, where $z = z_1 \ldots z_k$ $z_1 \ldots, z_k \in Q$. Using (3.21) we get

$$D^{(\mathbb{O}_{|v|},\alpha^+)} y^{\Phi}_{e_j,vz} = T_{q,q_0,j}(w) = C_qA_wB_{q_0}e_j = C_qA_{z_k}^{\alpha_k} A_{z_{k-1}}^{\alpha_{k-1}} \ldots A_{z_1}^{\alpha_1} B_{q_0}e_j. \tag{5.26}$$

From $T_f(w) = CA_wB_f$, the definition of $T_{q,f}(w)$, and the definition of the parameters C_q and $\mu(f) = B_f$ of Σ, we get that $T_{f,q}(w) = C_qA_w\mu(f)$. Let $w \in \widetilde{L}_q$ and $(v, (\alpha, z)) \in F_q(w)$. Then we get $w = z_1^{\alpha_1} z_2^{\alpha_2} \ldots z_k^{\alpha_k}$, where $z = z_1 \ldots z_k$, $z_1, \ldots z_k \in Q$, and $A_w = A_{z_k}^{\alpha_k} A_{z_{k-1}}^{\alpha_{k-1}} \ldots A_{z_1}^{\alpha_1}$. Combining this with (3.21) we get that

$$D^{(\mathbb{O}_{|v|},\alpha)} f_{0,vz} = T_{q,f}(w) = C_qA_w\mu(f) = C_qA_{z_k}^{\alpha_k} A_{z_{k-1}}^{\alpha_{k-1}} \ldots A_{z_1}^{\alpha_1} \mu(f). \tag{5.27}$$

By Lemma 5.1 the equalities (5.26) and (5.27) are equivalent to (Σ, μ) being a realization of Φ with constraint L.

Next, we show that (5.16) holds. Apply Construction 5.4 to (Σ, μ) and recall from Construction 5.4 the definition of the family of input-output maps $\Theta_{\Sigma,\mu}$. Recall from Construction 5.4 the definition of the map $U(\mu)$ and recall that $(\Sigma, \mu \circ U(\mu))$ is a realization of $\Theta_{\Sigma,\mu}$. Hence, $\Theta_{\Sigma,\mu}$ admits a generalized kernel representation. Using the notation of Definition 3.4, the functions of the generalized kernel representation of $\Theta_{\Sigma,\mu}$ are denoted $K_w^{g,\Theta_{\Sigma,\mu}}$ and $G_w^{\Theta_{\Sigma,\mu}}$ for all $w \in Q^+$, $g \in \Theta_{\Sigma,\mu}$. We are going to show that

$$\forall w \in \mathrm{comp}(L),\ g \in \Theta_{\Sigma,\mu}\colon \quad G_w^{\Theta_{\Sigma,\mu}} = 0 \ \text{ and } \ K_w^{g,\Theta_{\Sigma,\mu}} = 0. \tag{5.28}$$

It is easy to see that if a word belongs to $\mathrm{comp}(L)$ then any of its suffixes belongs to $\mathrm{comp}(L)$. Then from Definition 3.4, part 4 it follows that (5.28) implies (5.16). Since $G_w^{\Theta_{\Sigma,\mu}}$ and $K_w^{g,\Theta_{\Sigma,\mu}}$ are analytic entire functions, we obtain (5.28), if we show that the high-order derivatives of $G_w^{\Theta_{\Sigma,\mu}}$, $K_w^{g,\Theta_{\Sigma,\mu}}$ at zero are zero, *i.e.* if we show that

$$\forall g \in \Theta_{\Sigma,\mu},\ w \in \mathrm{comp}(L),\ \beta \in \mathbb{N}^{|w|}\colon \quad D^\beta G_w^{\Theta_{\Sigma,\mu}} = 0 \ \text{ and } \ D^\beta K_w^{g,\Theta_{\Sigma,\mu}} = 0. \tag{5.29}$$

Fix a word $w = q_1 \ldots q_k \in \mathrm{comp}(L)$, $q_1, \ldots, q_k \in Q$, and a tuple $\beta = (\beta_1, \ldots, \beta_k)$. Apply the first equality of (5.1) of Lemma 5.1 with w and $\alpha = (\beta_1 + 1, \beta_2, \ldots, \beta_k)$, and the second equality of (5.1) with w and $\alpha = \beta$, to the family $\Theta_{\Sigma,\mu}$ and to the realization $(\Sigma, \mu \circ U(\mu))$. Then for all $g = y_\Sigma(\mu(f), ., .) \in \Theta_{\Sigma,\mu}$, $f \in \Phi$,

$$D^\beta G_w^{\Theta_{\Sigma,\mu}} = C_{q_k}A_vB_{q_1} \ \text{ and } \ D^\beta K_w^{g,\Theta_{\Sigma,\mu}} = C_{q_k}A_v(\mu \circ U(\mu))(g) = C_{q_k}A_v\mu(f) \tag{5.30}$$

where $v = q_1^{\beta_1} q_2^{\beta_2} \ldots q_k^{\beta_k}$. But $w \in \mathrm{comp}(L)$ implies $\widetilde{L}_{q_k} = \emptyset$. Hence, v cannot be an element of $\widetilde{L}_{q_k}$. If $v \in Q^*$ is such that $v \notin \widetilde{L}_{q_k}$, then $v \notin \widetilde{L}_{q_k,q}$ for all $q \in Q$. From the definition of $T_{f,q_k}(v)$ and $T_{g_k,q,j}(v)$ we get that

$T_{f,q_k}(v) = 0$ and $T_{q_k,q,j}(v) = 0$ for all $f \in \Phi$, $j = 1,\dots,m$ and $q \in Q$. Hence from (5.26) and (5.27) it follows that $C_{q_k} A_v B_q e_j = T_{q_k,q,j}(v) = 0$, $j = 1,\dots,m$ and $C_{q_k} A_v \mu(f) = T_{q_k,f}(v) = 0$. Using (5.30), (5.29) then follows. □

5.3.2. *Quasi-minimality: proof of Theorem* 3.4

Proof of Theorem 3.4. From Theorem 3.7, if L is a regular language and Φ has a realization with constraint L, then rank $H_\Phi < +\infty$. Since the Hankel-matrix of Φ and Ψ_Φ coincide, by Theorem 4.1 we get that Ψ_Φ is rational. Let R be a minimal representation of Ψ_Φ. Consider $(\Sigma, \mu) = (\Sigma_R, \mu_R)$, *i.e.* the linear switched system realization associated with R. Then by Theorem 5.3 (Σ, μ) is a realization of Φ with constraint L such that (3.8) holds. Since R is reachable and observable, by Lemma 5.3 we get that (Σ, μ) is semi-reachable and observable. If $(\widetilde{\Sigma}, \widetilde{\mu})$ is a realization of Φ, then by Theorem 5.2, $\Psi_\Phi = K_{\widetilde{\Sigma},\widetilde{\mu}} \odot \Omega_\Phi$. From Lemma 4.3, rank $H_{\Psi_\Phi} \le$ rank $H_{K_{\widetilde{\Sigma},\widetilde{\mu}}} \cdot$rank H_{Ω_Φ}. By Lemma 5.9, rank $H_{K_{\widetilde{\Sigma},\widetilde{\mu}}} =$ rank $H_{\Psi_{\Theta_{\widetilde{\Sigma},\widetilde{\mu}}}} \le \dim \widetilde{\Sigma}$. Since R is a minimal, by Theorem 4.2, $\dim \Sigma = \dim R =$ rank H_Φ. Combining these observations we get $\dim \Sigma \le$ rank $H_{\Omega_\Phi} \cdot \dim \widetilde{\Sigma}$. By Lemma 5.7 rank H_{Ω_Φ} depends only on L. Hence, for $M =$ rank H_{Ω_Φ} we get (3.9). □

Procedure 5.3 (construction of a realization from the Hankel-matrix)**.** Construct a minimal representation R from $H_\Phi = H_{\Psi_\Phi}$ using Procedure B.1, Appendix B. Construct the linear switched system realization (Σ_R, μ_R) associated with R as described in Construction 5.2. By Theorem 5.3 and the proof of Theorem 3.4, (Σ_R, μ_R) is a quasi-minimal realization of Φ. For the corresponding algorithm see [20].

Procedure 5.4 (quasi-minimization)**.** Using the parameters of (Σ, μ) and (5.23) in the proof of Theorem 5.2, construct the Hankel-matrix H_Φ of Φ. Then use Procedure 5.3 to construct a semi-reachable and observable realization of Φ which satisfies (3.8) and (3.9). The procedure outlined above can be made effective, see [20].

6. Conclusions

The current paper is the first part of a series of papers dealing with realization theory of switched systems. In this paper realization theory of linear switched systems was presented. The forthcoming Part II of the series deals with realization theory of bilinear switched systems. The paper uses the theory of formal power series, to derive the results. To this end, the paper also presents an extension of the classical theory of formal power series to *families* of power series.

A. Proof of technical results on linear switched systems

A.1. **Technical results on input-output maps of linear switched systems**

The results of the section are necessary for the proof of Lemma 5.1 and Corollary 5.1, and Lemma 5.6. Assume that Φ has a generalized kernel representation defined in Definition 3.4. In the sequel we will use the notation of Definition 3.4 and the notation introduced in Notations 2.1–3.4.

Lemma A.1. *For each $w = q_1 \dots q_k \in L$, $q_1,\dots,q_k \in Q$, $k > 0$, and for all $j = 1,\dots,m$, $l = 1,\dots,k$,*

$$\forall \alpha \in \mathbb{N}^k\colon D^\alpha K^{f,\Phi}_{q_1 q_2 \dots q_k} = D^\alpha f_{0,q_1 q_2 \dots q_k} \quad \textit{and} \quad \forall \alpha \in \mathbb{N}^{k-l+1}\colon D^\alpha G^\Phi_{q_l q_{l+1} \dots q_k} e_j = D^\beta y^\Phi_{e_j, q_1 q_2 \dots q_k} \tag{A.1}$$

where $\mathbb{N}^k \ni \beta = (0,0,\dots,0,\alpha_1 + 1, \alpha_2, \dots, \alpha_{k-l+1})$. Here e_j is the jth unit vector of $\mathbb{R}^m$.

Proof. It follows from the formula $\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t g(t,\tau)\mathrm{d}\tau = g(t,t) + \int_0^t \frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau)\mathrm{d}\tau$ and Part 4 of Definition 3.4. □

The next lemma together with its numerous corollaries states a number of relationships between the functions $K^{f,\Phi}_w$ and G^Φ_w and functions $K^{f,\Phi}_z$ and G^Φ_z, where w is obtained from z by repeating zero or more times the letters of z. Moreover, these relationships are of great importance for the proof of Lemma 5.1 and for the derivation of results on realization theory with switching constraints. Recall from (3.11) the definition of the language $\widetilde{L}$.

Lemma A.2. *Consider the words $v, w, s \in Q^*$. Assume that $w = z_1^{\alpha_1} z_2^{\alpha_2} \dots z_k^{\alpha_k}$ for some $z_1, \dots, z_k \in Q$ and $(\alpha_1, \alpha_2 \dots, \alpha_k) \in \mathbb{N}^k$. Assume that vws and $vz_1 \dots z_k s$ both belong to $\widetilde{L}$. In addition, assume that either $\alpha_k > 0$ or, if $\alpha_k = 0$, then s is non-empty,* i.e. *$|s| > 0$. Then the following holds:*

(a) *For all $r_1, \dots, r_{|v|}, p_1, \dots, p_{|s|}, t_1, \dots, t_{|w|} \in T$, $f \in \Phi$*

$$K^{f,\Phi}_{vws}(r_1, \dots, r_{|v|}, t_1, \dots, t_{|w|}, p_1, \dots, p_{|s|}) = K^{f,\Phi}_{vz_1 \dots z_k s}(r_1, \dots, r_{|v|}, T_1, \dots, T_k, p_1, \dots, p_{|s|}). \tag{A.2}$$

If in addition either $|v| > 0$ or $\alpha_1 > 0$, then

$$G^{\Phi}_{vws}(r_1, \dots, r_{|v|}, t_1, \dots, t_{|w|}, p_1, \dots, p_{|s|}) = G^{f,\Phi}_{vz_1 \dots z_k s}(r_1, \dots, r_{|v|}, T_1, \dots, T_k, p_1, \dots, p_{|s|}). \tag{A.3}$$

Here $T_i = \sum_{j=1+\alpha_1+\dots+\alpha_{i-1}}^{\alpha_1+\dots+\alpha_i} t_j$, $i = 1, \dots, k$.

(b) *Let $\mathbb{I}_{|w|} = (1, 1, \dots, 1) \in \mathbb{N}^{|w|}$. Then for any $\beta \in \mathbb{N}^{|v|}$, $\gamma \in \mathbb{N}^{|s|}$, for all $f \in \Phi$,*

$$D^{(\beta, \mathbb{I}_{|w|}, \gamma)} K^{f,\Phi}_{vws} = D^{(\beta, \alpha, \gamma)} K^{f,\Phi}_{vz_1 z_2 \dots z_k s} \tag{A.4}$$

$$D^{(\beta, \mathbb{I}_{|w|}, \gamma)} G^{\Phi}_{vws} = D^{(\beta, \alpha, \gamma)} G^{\Phi}_{vz_1 z_2 \dots z_k s}, \text{ if } |v| > 0 \text{ or } \alpha_1 > 0. \tag{A.5}$$

Proof of Lemma A.2. We will show only that (A.2) and (A.4) hold. The proof of (A.3) and (A.5) can be obtained from the proof of (A.2) and (A.4) which will be presented below by simply replacing $K^{f,\Phi}_w$ with G^{Φ}_w and replacing the references to the first equality of Part 2 and Part 3 of Definition 3.4 with references to the *second equality* of Part 2 and Part 3 of Definition 3.4.

The proof goes by induction on k. Let $k = 1$ and denote by α the power α_1 and denote by z the letter z_1. If $\alpha = 0$, then $w = \epsilon$, *i.e.* the word w is the empty word. From Part 3 of Definition 3.4 it follows that $K^{f,\Phi}_{vzs}(r_1, \dots, r_{|v|}, 0, p_1, \dots, p_{|s|}) = K^{f,\Phi}_{vs}(r_1, \dots, r_{|v|}, p_1, \dots, p_{|s|})$. Hence, for $k = 1$ and $\alpha = 0$ (A.2) and (A.4) hold. Notice that $\mathbb{I}_0$ is the empty tuple, hence $(\beta, \mathbb{I}_0, \gamma) = (\beta, \gamma)$. It follows from definition of $\frac{\mathrm{d}^0}{\mathrm{d}t^0}$ and Part 3 of Definition 3.4 that $D^{(\beta,\gamma)} K^{f,\Phi}_{vs} = D^{(\beta,\alpha,\gamma)} K^{f,\Phi}_{vzs}$. Notice that for $k = 1$ and $\alpha = 1$ (A.2) and (A.4) are trivially true. For $k = 1$ and $\alpha > 1$, (A.2) and (A.4) can be shown to be true by using Part 2 of Definition 3.4 and $\frac{\mathrm{d}}{\mathrm{d}t_1} \frac{\mathrm{d}}{\mathrm{d}t_2} g(t_1 + t_2)|_{t_1=t_2=0} = \frac{\mathrm{d}^2}{\mathrm{d}t^2} g(t)|_{t=0}$.

Finally, assume that (A.2) and (A.4) hold for all $k < n$. We will prove that (A.2) and (A.4) hold for $k = n$. Indeed, let's apply the induction hypothesis for $k = n - 1$ to $\hat{w} = z_1^{\alpha_1} \dots z_{n-1}^{\alpha_{n-1}}$. Then we get that

$$\begin{aligned} &K^{f,\Phi}_{vz_1^{\alpha_1} \dots z_n^{\alpha_n} s}(r_1, \dots, r_{|v|}, t_1, \dots, t_{|w|-\alpha_n}, t_{|w|-\alpha_n+1}, \dots, t_{|w|}, p_1, \dots, p_{|s|}) \\ &\qquad = K^{f,\Phi}_{vz_1 \dots z_{n-1} z_n^{\alpha_n} s}(r_1, \dots, r_{|v|}, T_1, \dots, T_{n-1}, t_{|w|-\alpha_n+1}, \dots, t_{|w|}, p_1, \dots, p_{|s|}) \end{aligned}$$

$$D^{(\beta, \mathbb{I}_{|w|}, \gamma)} K^{f,\Phi}_{vws} = D^{(\beta, \mathbb{I}_{|\hat{w}|}, 1, \gamma)} K^{f,\Phi}_{v\hat{w} z_n^{\alpha_n} s} = D^{(\beta, \alpha, 1, \gamma)} K^{f,\Phi}_{vz_1 \dots z_{n-1} z_n^{\alpha_n} s}$$

where $T_i = \sum_{j=1+\alpha_1+\dots\alpha_{i-1}}^{\alpha_1+\dots+\alpha_i} t_j$. Using the equalities above and applying the induction hypothesis for $k = 1$ to $z_n^{\alpha_n}$ we get (A.2) and (A.4). □

Corollary A.1. *Let $w \in \widetilde{L}$ be a word such there exist $z_1, z_2, \dots, z_k \in Q$, $k > 0$ and a k-tuple $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k) \in \mathbb{N}^k$ such that $w = z_1^{\alpha_1} z_2^{\alpha_2} \dots z_k^{\alpha_k}$ and $z_1 z_2 \dots z_k \in \widetilde{L}$, and $\alpha_k > 0$. Then for all $t_1, \dots, t_{|w|} \in T$, $f \in \Phi$,*

$$K^{f,\Phi}_w(t_1, \dots, t_{|w|}) = K^{f,\Phi}_{z_1 z_2 \dots z_k}(T_1, \dots, T_k) \text{ and } G^{\Phi}_w(t_1, \dots, t_{|w|}) = G^{\Phi}_{z_1 z_2 \dots z_k}(T_1, \dots, T_k) \text{ if } \alpha_1 > 0. \tag{A.6}$$

Here it is assumed $T_i = \sum_{j=1+\alpha_1+\ldots+\alpha_{i-1}}^{\alpha_1+\ldots+\alpha_i} t_j$, $i = 1,\ldots,k$. *That is,* $\{K^{f,\Phi}_w, G^{\Phi}_v \mid f \in \Phi, w, v \in \mathrm{suffix}L\}$ *uniquely determines the collection of functions* $\{K^{f,\Phi}_w, G^{\Phi}_v \mid f \in \Phi, w, v \in \widetilde{L}\}$.

Proof. Apply part (a) of Lemma A.2 with $v = \epsilon$ and $s = \epsilon$ and substituting $z_1^{\alpha_1} z_2^{\alpha_2} \ldots z_k^{\alpha_k}$ for w. □

Corollary A.2. *Let* $z_1, z_2, \ldots, z_k, d_1, d_2, \ldots, d_l, q, q_0 \in Q$ *be discrete modes. Assume that* $z_1^{\alpha_1} z_2^{\alpha_2} \ldots z_k^{\alpha_k} = d_1^{\beta_1} d_2^{\beta_2} \ldots d_l^{\beta_l}$ *for some* $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \in \mathbb{N}^k$, $\beta = (\beta_1, \beta_2, \ldots, \beta_l) \in \mathbb{N}^l$. *Then the following holds:*

- *If* $q_0 z_1 z_2 \ldots z_k q \in \widetilde{L}$ *and* $q_0 d_1 d_2 \ldots d_l q \in \widetilde{L}$, *then* $D^{(0,\alpha,0)} G^{\Phi}_{q_0 z_1 z_2 \ldots z_k q} = D^{(0,\beta,0)} G^{\Phi}_{q_0 d_1 d_2 \ldots d_l q}$.
- *If* $z_1 z_2 \ldots z_k q$ *and* $d_1 d_2 \ldots d_l q \in \widetilde{L}$, *then* $D^{(\alpha,0)} K^{f,\Phi}_{z_1 z_2 \ldots z_k q} = D^{(\beta,0)} K^{f,\Phi}_{d_1 d_2 \ldots d_l q}$.

Proof. Denote by w the word $z_1^{\alpha_1} z_2^{\alpha_2} \ldots z_k^{\alpha_k} = d_1^{\beta_1} d_2^{\beta_2} \ldots d_l^{\beta_l}$. Using (A.4) with $v = q_0$ and $s = q$ one gets that $D^{(0,\alpha,0)} G^{\Phi}_{q_0 z q} = D^{(0,\mathbb{I}_{|w|},0)} G^{\Phi}_{q_0 w q} = D^{(0,\beta,0)} G^{\Phi}_{q_0 d q}$. Similarly, from (A.5) with $v = \epsilon$, and $s = q$ we get $D^{(\alpha,0)} K^{f,\Phi}_{z_1 z_2 \ldots z_k q} = D^{(\mathbb{I}_{|w|},0)} K^{f,\Phi}_{wq} = D^{(\beta,0)} K^{f,\Phi}_{d_1 \ldots d_l q_1}$. □

Corollary A.3. *For any* $w \in \widetilde{L}$, *for any* $q_1, q_2, q \in Q$, $j = 1, \ldots, m$,

$$(v, (\alpha, z)) \in F_{q_1,q_2}(w) \implies D^{(\mathbb{O}_{|v|},\alpha^+)} y^{\Phi}_{e_j,vz} = D^{\alpha} G^{\Phi}_z e_j = D^{(0,\alpha,0)} G^{\Phi}_{q_2 z q_1} e_j \tag{A.7}$$

$$(v, (\alpha, z)) \in F_q(w) \implies D^{(\mathbb{O}_{|v|},\alpha)} f_{0,vz} = D^{\alpha} K^{f,\Phi}_z = D^{(\alpha,0)} K^{f,\Phi}_{zq}. \tag{A.8}$$

Proof of Corollary A.3. First we prove (A.8). From (A.1) it follows that $D^{(\mathbb{O}_{|v|},\alpha)} f_{0,vz} = D^{\alpha} K^{f,\Phi}_z$. Indeed, from (A.1), $D^{(\mathbb{O}_{|v|},\alpha)} f_{0,vz} = D^{(\mathbb{O}_{|v|},\alpha)} K^{f,\Phi}_{vz}$. By applying Part 3 of Definition 3.4, $K^{f,\Phi}_{vz}(0, 0, \ldots, 0, t_1, t_2, \ldots, t_{|z|}) = K^{f,\Phi}_z(t_1, \ldots, t_k)$ and hence $D^{(\mathbb{O}_{|v|},\alpha)} K^{f,\Phi}_{vz} = K^{f,\Phi}_z$. Hence, it is enough to show that

$$D^{(\alpha,0)} K^{f,\Phi}_{zq} = D^{\alpha} K^{f,\Phi}_z. \tag{A.9}$$

To this end, notice that $(v, (\alpha, z)) \in F_q(w)$, implies that the last letter of z equals q. Hence, by Part 2 of Definition 3.4 we get that $K^{f,\Phi}_{zq}(t_1, t_2, \ldots, t_{k-1}, t_k, 0) = K^{f,\Phi}_{z_1 z_2 \ldots z_k}(t_1, t_2, \ldots, t_k)$. By applying $\frac{\mathrm{d}^{\alpha_1}}{\mathrm{d}t_1^{\alpha_1}} \cdots \frac{\mathrm{d}^{\alpha_k}}{\mathrm{d}t_k^{\alpha_k}}$ to the left- and right-hand sides of the equation above, and evaluating the result at zero, we get the desired equality.

Next, we will prove (A.7). In order to prove (A.7) it is enough to show that

$$D^{\alpha} G_z = D^{(0,\alpha,0)} G_{q_0 z q}. \tag{A.10}$$

Indeed, from (A.1) we get that for $(v, (\alpha, z)) \in F_{q,q_0}(w)$, $D^{(\mathbb{O}_{|v|},\alpha^+)} y^{\Phi}_{e_j,vz} = D^{\alpha} G^{\Phi}_z e_j$, and hence (A.10) implies (A.7). Before proving (A.10) we will show that for any $q_0 w \in \widetilde{L}$, $w \in \widetilde{L}$, and any $q_0 \in Q$

$$D^{(0,\alpha,\beta)} G^{\Phi}_{q_0 q_0 w} = D^{(\alpha,\beta)} G^{\Phi}_{q_0 w} \tag{A.11}$$

for any $\alpha \geq 0$ and multi-index $\beta \in \mathbb{N}^{|w|}$. Assume that $|w| = k$ and $\beta = (\beta_1, \ldots, \beta_k)$. From Part 2 of Definition 3.4 it follows that $G^{\Phi}_{q_0 q_0 w}(\tau_1, \tau_2, t_1, \ldots, t_k) = G^{\Phi}_{q_0 w}(\tau_1 + \tau_2, t_1, \ldots, t_k)$. Hence, we get that

$$\frac{\mathrm{d}^0}{\mathrm{d}\tau_1^0} \frac{\mathrm{d}^{\alpha}}{\mathrm{d}\tau_2^{\alpha}} \frac{\mathrm{d}^{\beta_1}}{\mathrm{d}t_1^{\beta_1}} \cdots \frac{\mathrm{d}^{\beta_k}}{\mathrm{d}t_k^{\beta_k}} G^{\Phi}_{q_0 q_0 w}(\tau_1, \tau_2, t_1, \ldots, t_k) = \frac{\mathrm{d}^{\alpha}}{\mathrm{d}\tau^{\alpha}} \frac{\mathrm{d}^{\beta_1}}{\mathrm{d}t_1^{\beta_1}} \cdots \frac{\mathrm{d}^{\beta_k}}{\mathrm{d}t_k^{\beta_k}} G^{\Phi}_{q_0 w}(\tau, t_1, \ldots, t_k)|_{\tau = \tau_1 + \tau_2}.$$

By evaluating the equation above at $\tau_1 = \tau_2 = 0$, $t_1 = \ldots = t_k = 0$, we get (A.11).

But $(v, (\alpha, z)) \in F_{q_0,q}(w)$ implies that if $z = z_1 \ldots z_k$ for some $z_1, \ldots, z_k \in Q$, $k > 0$, then $z_1 = q_0$ and $z_k = q$. But applying (A.11) to $w = z_2 \ldots z_k q$ yields $D^{(0,\alpha,0)} G^{\Phi}_{q_0 z q} = D^{(\alpha,0)} G^{\Phi}_{zq}$. Therefore, in order to prove (A.10) it is left to show that $D^{(\alpha,0)} G^{\Phi}_{zq} = D^{\alpha} G^{\Phi}_z$. The latter can be shown in exactly the same way as (A.9) by simply replacing $K^{f,\Phi}$ with G^{Φ} and using the second equality of Part 2 of Definition 3.4. □

Proof of Lemma 5.1. First, we show the following equivalence:

`(i) if and only if (ii)`. First, notice that the left-most equalities in (5.1) follow immediately from (A.1). By Theorem 3.5 (Σ, μ) is a realization of Φ if and only if Φ has a generalized kernel representation of the form (3.15). Assume that `(i)` holds. Then (3.15) holds. Consider $q_1 \dots q_k \in L$ with $q_1, \dots, q_k \in Q$. By taking derivatives of $G^{\Phi}_{q_l q_{l+1} \dots q_k}$ for any $l = 1, \dots, k$, from (3.15) we get

$$D^{\beta} G^{\Phi}_{q_l q_{l+1} \dots q_k} = C_{q_k} A^{\alpha_k}_{q_k} A^{\alpha_{k-1}}_{q_{k-1}} \dots A^{\alpha_l - 1}_{q_l} B_{q_l} \text{ where } \beta = (\alpha_l - 1, \alpha_{l+1}, \dots, \alpha_k). \tag{A.12}$$

Notice that if $q_1 q_2 \dots q_k \in L$, then $q_l q_{l+1} \dots q_k \in \widetilde{L}$, and hence $G^{\Phi}_{q_l q_{l+1} \dots q_k}$ is well-defined. Lemma A.1 implies that $D^{\alpha} y^{\Phi}_{e_j, q_1 q_2 \dots q_k} = D^{\beta} G_{q_l q_{l+1} \dots q_k} e_j$, where $\alpha = (0, 0, \dots, 0, \alpha_l, \alpha_{l+1}, \dots, \alpha_k) \in \mathbb{N}^k$ and $j = 1, \dots, m$. Combining the equation above with (A.12), we get the first equation of (5.1). The second equation of (5.1) can be proved analogously. That is, `(i)` implies `(ii)`.

Assume that `(ii)` holds. We show that then (3.15) holds, which by Theorem 3.5 implies `(i)`. From Corollary A.1 it follows that it is enough to consider $\{K^{f,\Phi}_v, G^{\Phi}_w \mid w \in \text{suffix}L,\ v \in L,\ f \in \Phi\}$. Using (A.1) and (5.1) we obtain that the derivatives of $G^{\Phi}_{q_l q_{l+1} \dots q_k}$ at zero are equal to the derivatives of the corresponding right-hand side of the first equation of (3.15). From analicity of $G^{\Phi}_{q_l q_{l+1} \dots q_k}$ we get that $G^{\Phi}_{q_l q_{l+1} \dots q_k}$, $q_1 q_2 \dots q_k \in L$, is as in (3.15). We can show that $K^{f,\Phi}_{q_1 q_2 \dots q_k}$ is of the form (3.15) analogously.

`(ii) if and only if (iii)`. First, notice that the left-most equalities in (5.2) and (5.3) follow from (A.7) and (A.8) of Corollary A.3. We show that `(iii)` implies `(ii)`. Consider any word $w \in L$ of the form $w = q_1 q_2 \dots q_k$, $q_1, q_2, \dots, q_k \in Q$ and any tuple $\alpha \in \mathbb{N}^k$. Let $0 < l \le k$ be such that $\alpha_l > 0$ and $\alpha_1 = \dots \alpha_{l-1} = 0$ and assume that such l exists. Define $v = q_1 q_2 \dots q_{l-1}$, $z = q_l q_{l+1} \dots q_k$ and $x = q_l^{\alpha_l - 1} q_{l+1}^{\alpha_{l+1}} \dots q_k^{\alpha_k}$. Then $(v, (\beta, z)) \in F_{q_l, q_k}(x)$, where $\beta = (\alpha_l - 1, \dots, \alpha_{|w|})$. Notice that $(\mathbb{O}_{l-1}, \beta^+) = \alpha$ and $vz = w$. Hence from (5.2) we get the first equation of (5.1). Similarly, let $y = q_1^{\alpha_1} \dots q_k^{\alpha_k}$. Then $(\epsilon, (\alpha, w)) \in F_{q_k}(y)$ and (5.3) implies the second equation of (5.1).

Conversely, assume that `(ii)` holds and we will show that `(iii)` holds too. Indeed, for any $s \in \widetilde{L}$, q, $q_0 \in Q$, $(v, (\beta, z)) \in F_{q, q_0}(s)$ it holds that $vz \in L$, $z = z_1 z_2 \dots z_k$, $z_1, z_2, \dots, z_k \in Q$, $z_1 = q_0$, $z_k = q$. Applying (5.1) to $w = vz$ and $\alpha = (\mathbb{O}_{|v|}, \beta^+) \in \mathbb{N}^{|w|}$ yields that (5.2) holds for $(v, (\beta, z))$. Similarly, for any $(v, (\beta, z)) \in F_q(s)$ it holds that $z = z_1 z_2 \dots z_k$, $z_1, z_2, \dots, z_k \in Q$, $z_k = q$ and $vz \in L$. Then the application of (5.1) to $w = vz$ and $\alpha = (\mathbb{O}_{|v|}, \beta)$ yields (5.3). □

Proof of Corollary 5.1. The pair (Σ, μ) is a realization of Φ if and only if part (ii) of Lemma 5.1 holds. Hence, in order to prove the corollary, it is enough to show that (5.1) holds for all $w \in L = Q^+$ and $\alpha \in \mathbb{N}^{|w|}$ if and only if (5.4) and (5.5) holds for all $w \in Q^+$. To this end, it is enough to show that:

(1) The first equation of (5.1) holds for all $w \in Q^+$ and $\alpha \in \mathbb{N}^{|w|}$ if and only if (5.4) holds for any $w \in Q^*$.

(2) The second equation of (5.1) holds for all $w \in Q^+$ and $\alpha \in \mathbb{N}^{|w|}$ if and only if (5.5) holds for any $w \in Q^*$.

We will present only the proof of (1), the proof of (2) is completely analogous to that of (1).

Assume that the first equality of (5.1) holds. Apply the first equality of (5.1) for $\alpha = (1, \mathbb{I}_{|w|}, 0)$ and $w = q_0 q_1 q_2 \dots q_k q$. By noticing that $A^0_{q_0} = A^0_q = I_n$ is the identity matrix, we get precisely (5.4). Conversely, assume that (5.4) holds for all $w \in Q^+$, q, $q_0 \in Q$, $j = 1, \dots, m$. We will show that then the first equality (5.1) holds for $\alpha \in \mathbb{N}$, $w \in Q^+$. The left-most equalities of (5.1) follow from Lemma A.1. The rest follows from (5.4) if it is applied with $w = q_l^{\alpha_l - 1} q_{l+1}^{\alpha_{l+1}} \dots q_k^{\alpha_k}$ with $q_0 = q_l$ and $q = q_k$ and from the following remark. Applying (A.5) from Lemma A.2, and (A.10) from the proof of Corollary A.3, we get that for $\beta = (\alpha_l - 1, \alpha_{l+1}, \dots, \alpha_k) \in \mathbb{N}^{k-l+1}$, $s = q_l q_{l+1} \dots q_k \in Q^+$, $D^{\beta} G^{\Phi}_s = D^{(0, \mathbb{I}_{|w|}, 0)} G^{\Phi}_{q_l w q_k}$, where $w = q_l^{\alpha_l - 1} q_{l+1}^{\alpha_{l+1}} \dots q_k^{\alpha_k}$ as above. □

A.2. **Proof of the characterization of semi-reachability**

Proof of Proposition 3.1. The second statement of the proposition follows from the first by taking $\text{Im}\mu$ as $\mathcal{X}_0$. Hence, it is enough to prove that first statement of the proposition. We will show that $WR(\mathcal{X}_0)$ is the smallest vector space containing the set $\text{Reach}(\Sigma, \mathcal{X}_0)$ of states reachable from the set of initial states $\mathcal{X}_0$. First, we show that $\mathrm{Reach}(\Sigma,\mathcal{X}_0)$ is contained in $WR(\mathcal{X}_0)$. Then we show that if W is a linear space containing $\mathrm{Reach}(\Sigma,\mathcal{X}_0)$, then $WR(\mathcal{X}_0)\subseteq W$.

Reach $(\Sigma,\mathcal{X}_0)\subseteq WR(\mathcal{X}_0)$

From Theorem 3.2, (3.5) it follows that the set $\mathrm{Reach}(\Sigma,\{0\})$ of states reachable from 0 is contained in $WR(\mathcal{X}_0)$. From Theorem 3.1, (3.3) it follows that each element of $\mathrm{Reach}(\Sigma,\mathcal{X}_0)$ is the sum of the controlled and autonomous parts. More precisely, each reachable element is of the form

$$\mathrm{e}^{A_{q_k}t_k}\mathrm{e}^{A_{q_{k-1}}t_{k-1}}\dots\mathrm{e}^{A_{q_1}t_1}x_0+x_\Sigma(0,u,(q_1,t_1)(q_2,t_2)\dots(q_k,t_k))$$

for some piecewise-continuous input $u\in PC(T,\mathcal{U})$, discrete modes $q_1,\dots,q_k\in Q$, $k\geq 0$, switching times $t_1,\dots,t_k\in T$, and some initial state $x_0\in\mathcal{X}_0$. Since $x_\Sigma(0,u,(q_1,t_1)(q_2,t_2)\dots(q_k,t_k))$ belongs to the set $\mathrm{Reach}(\Sigma,\{0\})$, we get that $x_\Sigma(0,u,(q_1,t_1)(q_2,t_2)\dots(q_k,t_k))$ belongs to $WR(\mathcal{X}_0)$. If we can show that any vector of the form $\mathrm{e}^{A_{q_k}t_k}\mathrm{e}^{A_{q_k-1}t_{k-1}}\dots\mathrm{e}^{A_{q_1}t_1}x_0$ belongs to $WR(\mathcal{X}_0)$ for all initial states $x\in\mathcal{X}_0$ and switching sequences $(q_1,t_1)(q_2,t_2)\dots(q_k,t_k)$, then it follows that $\mathrm{Reach}(\Sigma,\mathcal{X}_0)$ is a subset of $WR(\mathcal{X}_0)$. To this end, notice that for each discrete state $q\in Q$, $\mathrm{e}^{A_qt}x=\sum_{k=0}^{+\infty}\frac{t^k}{t!}A_q^kx$, hence $\mathrm{e}^{A_qt}x$ belongs to the linear span of the vectors A_q^jx, $j\in\mathbb{N}$. This implies that $\mathrm{e}^{A_{q_k}t_k}\mathrm{e}^{A_{q_{k-1}}t_{k-1}}\dots\mathrm{e}^{A_{q_1}t_1}x_0$ belongs to the linear span of the vectors of the form $A_{q_1}^{j_1}A_{q_2}^{j_2}\dots A_{q_k}^{j_k}x_0\in WR(\mathcal{X}_0)$ for integers $j_1,j_2,\dots,j_k\in\mathbb{N}$, and hence it belongs to $WR(\mathcal{X}_0)$.

$WR(\mathcal{X}_0)\subseteq W$

First, notice that for any initial state $x_0\in\mathcal{X}_0$, for any switching sequence $s=(q_1,t_1)(q_2,t_2)\dots(q_k,t_k)\in(Q\times T)^*$, $k\geq 0$, and for any input $u\in PC(T,\mathcal{U})$, $x_\Sigma(x_0,u,s)=x_\Sigma(x_0,0,s)+x_\Sigma(0,u,s)$, and hence, $x_\Sigma(0,u,s)=x_\Sigma(x_0,u,s)-x_\Sigma(x_0,0,s)$. Since both $x_\Sigma(x_0,u,s)$ and $x_\Sigma(x_0,0,s)$ belong to the vector space W, we get that $x_\Sigma(0,u,s)$ belongs to W as well. Hence, we get that $\mathrm{Reach}(\Sigma,\{0\})$ belongs to W. Notice that

$$WR(\mathcal{X}_0)=\mathrm{Reach}(\Sigma,\{0\})+\mathrm{Span}\{A_{q_k}A_{q_{k-1}}\dots A_{q_1}x_0\mid q_1,q_2,\dots,q_k\in Q,\ k\geq 0,\ x_0\in\mathcal{X}_0\}. \tag{A.13}$$

Hence, if we can show that vectors of the form $A_{q_k}A_{q_{k-1}}\dots A_{q_1}x_0$ for $q_1,\dots,q_k\in Q$, $k\geq 0$, $x_0\in\mathcal{X}_0$ belong to W, then we obtain that $WR(\mathcal{X}_0)\subseteq W$. In order to show that $A_{q_k}A_{q_{k-1}}\dots A_{q_1}x_0\in W$, define for the sequence $w=q_1q_2\dots q_k\in Q^+$ and $x_0\in\mathcal{X}_0$, the map $\exp_{w,x_0}:T^k\to\mathcal{X}$ by

$$\exp_{w,x_0}(t_1,t_2,\dots,t_k)=\mathrm{e}^{A_{q_k}t_k}\mathrm{e}^{A_{q_{k-1}}t_{k-1}}\dots\mathrm{e}^{A_{q_1}t_1}x_0=x_\Sigma(x_0,0,(q_1,t_1)(q_2,t_2)\dots(q_k,t_k)). \tag{A.14}$$

It follows that the values $\exp_{w,x_0}$ belong to $\mathrm{Reach}(\Sigma,\mathcal{X}_0)$ and hence to W. Since W is a vector space, it follows that all the high-order derivatives at zero must also belong to W. It is easy to see that $D^{(1,1,\dots,1)}\exp_{w,x_0}=A_{q_k}A_{q_{k-1}}\dots A_{q_1}x_0$. Hence, $A_{q_k}A_{q_{k-1}}\dots A_{q_1}x_0$ belongs to W as well. $\square$

A.3. Technical proofs for Section 5.3

Proof of Lemma 5.6. Formula (5.12) follows by using formulas (A.7), (A.8) from Section 5.1.

From the right-hand sides of (5.12) it follows that the values $T_{q,q_0,j}(w)$ and $T_{q,f}(w)$ do not depend on the choice of v in $(v,(\alpha,z))\in F_{q,q_0}(w)$ or $(v,(\alpha,z))\in F_q(w)$ respectively. We will argue that the value of $T_{q,q_0,j}(w)$, $j=1,\dots,m$ and $T_{q,f}(w)$ do not depend on the choice of (α,z), *i.e.* the right-hand sides of (5.12) are the same for any $(v,(\alpha,z))$ as long as $(v,(\alpha,z))$ belongs to $F_{q,q_0}(w)$ or $F_q(w)$ respectively.

If $(v,(\alpha,z)),(u,(\beta,x))\in F_{q,q_0}(w)$ are two elements of $F_{q,q_0}(w)$, then from the definition of the set $F_{q,q_0}(w)$ it follows that $x_1^{\beta_1}\dots x_{|x|}^{\beta_{|x|}}=z_1^{\alpha_1}\dots z_{|z|}^{\alpha_{|z|}}=w$, and $z_1=x_1=q_0$, $z_{|z|}=x_{|x|}=q$, and $q_0zq,q_0xq\in\widetilde{L}$. Hence, by Corollary A.2, $D^{(0,\alpha,0)}G^{\Phi}_{q_0zq}=D^{(0,\beta,0)}G^{\Phi}_{q_0xq}$. Similarly, if $(v,(\alpha,z)),(u,(\beta,x))$ are two elements of $F_q(w)$, then $x_1^{\beta_1}\dots x_{|x|}^{\beta_{|x|}}=z_1^{\alpha_1}\dots z_{|z|}^{\alpha_{|z|}}=w$ and $zq,xq\in\widetilde{L}$. Hence, by Corollary A.2, $D^{(\alpha,0)}K^{f,\Phi}_{zq}=D^{(\beta,0)}K^{f,\Phi}_{xq}$. $\square$

Proof of Lemma 5.8. Notice that the languages $\widetilde{L}_{q,q_0}$ and $\widetilde{L}_q$ can be written as $\widetilde{L}_{q,q_0}=\{w\in Q^*\mid q_0wq\in\widetilde{L}\}$ and $\widetilde{L}_q=\{w\in Q^*\mid wq\in\widetilde{L}\}$. That is, $\widetilde{L}_{q,q_0}$ consists of all those words w for which the word q_0wq belongs to $\widetilde{L}$, and $\widetilde{L}_q$ consists of all those words w for which the word wq belongs to $\widetilde{L}$. Then it is easy to see that if $\widetilde{L}$ is regular, then so are $\widetilde{L}_{q,q_0}$ and $\widetilde{L}_q$. Hence, it is enough to show that $\widetilde{L}$ is regular, if L is regular.

To this end, notice that if L is regular then suffix$L$ is regular. Let $A = (S, Q, \delta, F, s_0)$ be a deterministic automaton (see [4,8]) accepting suffix$L$. Here S is the state-space, Q is the alphabet of the automaton, F is the set of accepting states, $\delta : S \times Q \to S$ is the state-transition function, s_0 is the set of initial states. Recall from [4,8] that the extended state-transition function is defined as follows. For each $s_0 \in S$, $w \in Q^*$, $\delta(s_0, w) = s$ if there exists a sequence of states $s_1, s_2 \ldots, s_k = s \in S$, such that if w is of the form $w = q_1q_2 \ldots q_k \in Q^*$, $q_1, q_2, \ldots, q_k \in Q$, $k \geq 0$, then $s_i = \delta(s_{i-1}, q_i)$ for each $i = 1, \ldots, k$.

Define the non-deterministic automaton[2] $B = ((S \times Q) \cup \{\hat{s}_0\}, Q, \delta_B, F \times Q, \hat{s}_0)$ as follows. The set of accepting states of B is $F \times Q$, the set of states of B is $S \times Q \cup \{\hat{s}_0\}$, $\hat{s}_0 \notin S \times Q$. The initial state is $\hat{s}_0$. The state-transition relation δ_B is as follows. For any discrete mode $q \in Q$, $\delta_B(\hat{s}_0, q) \ni (s, q)$ holds, if $\delta(s_0, wq) = s$ for some $w \in Q^*$. For any discrete mode $q, u \in Q$ and state $s \in S$ of A, $(\hat{s}, u) \in \delta_B((s, q), u)$ holds if either (i) $u = q$ and $\hat{s} = s$, or (ii) there exists $wu \in Q^*$, such that $\delta(s, wu) = \hat{s}$.

We argue that B accepts precisely the language $\widetilde{L}$. Denote the fact that $s \in \delta_B(\hat{s}, q)$, for some states $s, \hat{s} \in (S \times Q) \cup \{\hat{s}_0\}$, by $\hat{s} \xrightarrow{q} s$. Then B accepts a word $z = q_1q_2 \ldots q_k \in Q^*$, $q_1, q_2, \ldots, q_k \in Q$, $k \geq 0$, if and only if there exists a run

$$\hat{s}_0 \xrightarrow{q_1} (s_1, q_1) \xrightarrow{q_2} (s_2, q_2) \ldots \xrightarrow{q_k} (s_k, q_k) \tag{A.15}$$

where $s_k \in F$. This is equivalent to the existence of integers $0 < \alpha_1, \ldots, \alpha_l \in \mathbb{N}$ and words $w_1, w_1, \ldots, w_l \in Q^*$ such that $\sum_{j=1}^{l} \alpha_j = k$ and the following holds. The first state (s_1, q_1) in (A.15) satisfies, $\delta(s_0, w_1q_1) = s_1$ and the subsequent states in (A.15) are of the following form. For each $0 \leq d \leq l$ denote by n_d the sum $n_d = \sum_1^d \alpha_j$. Then for each $d = 0, \ldots, l-1$, $(s_i, q_i) = (s_{i+1}, q_{i+1})$ for each $1 + n_d \leq i < n_{d+1}$ and $\delta(s_{n_d}, w_{d+1}q_{n_d+1}) = s_{n_d+1}$. Define $u_d = q_{n_d+1}$ for all $1 \leq d \leq l-1$. Then in the automaton A it holds that $\delta(s_0, w_1u_1w_2u_2 \ldots w_lu_l) = s_k \in F$. That is, $w_1u_1 \ldots w_lu_l \in$ suffix$L$. In addition, the word $z = q_1q_2 \ldots q_k \in Q^*$ from above is then of the form

$$z = w^0_{1,1} \ldots w^0_{1,m_1} u_1^{\alpha_1} w^0_{2,1} \ldots w^0_{2,m_2} u_2^{\alpha_2} \ldots w^0_{l,1} \ldots w^0_{l,m_l} u_l^{\alpha_l} \tag{A.16}$$

where $w_{i,1}, w_{i,2} \ldots, w_{i,m_i} \in Q$ are the letters of w_i for all $i = 1, \ldots, l$, *i.e.* $w_i = w_{i,1}w_{i,2} \ldots w_{i,m_i}$. But (A.16) means exactly that z belongs to $\widetilde{L}$. Hence, we get that B accepts exactly the elements of $\widetilde{L}$. □

B. Proof of the main results of Section 4 on formal power series

In this section we will present the proof of the main results of Section 4 on rationality of families of formal power series. We start with the proof of Theorem 4.1. To this end, we need the following notation and terminology. Let $w \in X^*$ be a word over X^* and let $S \in \mathbb{R}^p\langle\langle X^* \rangle\rangle$ be a formal power series. Define the formal power series $w \circ S \in \mathbb{R}^p\langle\langle X^* \rangle\rangle$, called the *left shift of S by w*, as follows; we require that for all $v \in X^*$ the value of $w \circ S$ at v is as follows

$$(w \circ S)(v) = S(wv). \tag{B.1}$$

Notice that for any word $w, v \in X^*$, $wv \circ S = v \circ (w \circ S)$ and $\epsilon \circ S = S$. Moreover, notice that the shift operation is linear, that is, for any $T, S \in \mathbb{R}^p\langle\langle X^* \rangle\rangle$, and for any scalars $\alpha, \beta \in \mathbb{R}$, and for any word $w \in X^*$, $w \circ (\alpha S + \beta T) = \alpha(w \circ S) + \beta(w \circ T)$. In the rest of the subsection, let $\Psi = \{S_j \in \mathbb{R}^p\langle\langle X^* \rangle\rangle \mid j \in J\}$ be a family of formal power series.

Definition B.1. The *the smallest shift invariant space of* Ψ, denoted by W_Ψ, is the subspace of $\mathbb{R}^p\langle\langle X^* \rangle\rangle$, spanned by all formal power series $w \circ S_j$, $j \in J$, $w \in X^*$, *i.e.* $W_\Psi = \mathrm{Span}\{w \circ S_j \in \mathbb{R}^p\langle\langle X^* \rangle\rangle \mid j \in J, w \in X^*\}$.

Remark B.1. There is one-to-one correspondence between the formal power series $w \circ S_j$ and the column of H_Ψ indexed by (w, j) for any word $w \in X^*$ and index $j \in J$. In particular, it follows that W_Ψ is isomorphic to the span of columns of H_Ψ and hence $\dim W_\Psi = \mathrm{rank}\ H_\Psi$.

[2]See [4,8] or any other standard textbook on automata theory for the definition of the concept of non-deterministic automaton.

We will need the following two auxiliary statements, proof of which is routine.

Lemma B.1. *Assume that* $\dim W_\Psi < +\infty$ *holds. Consider the p-J-representation*

$$R_\Psi = (W_\Psi, \{A_\sigma\}_{\sigma\in X}, B, C) \tag{B.2}$$

where for each $\sigma \in X$*, the map* $A_\sigma : W_\Psi \to W_\Psi$ *is defined as the shift by* σ, i.e. *for each* $T \in W_\Psi$, $A_\sigma(T) = \sigma \circ T$*; the collection* $B = \{B_j \in W_\Psi \mid j \in J\}$ *is such that* $B_j = S_j$ *for each* $j \in J$*; the linear map* $C : W_\Psi \to \mathbb{R}^p$ *is defined as* $C(T) = T(\epsilon)$ *for all* $T \in W_\Psi$*. Then* R_Ψ *is a representation of* Ψ*. The representation* R_Ψ *is called the* free *representation of* Ψ.

Lemma B.2. *If* Ψ *is rational, then* $\dim W_\Psi < +\infty$ *and for each representation* R *of* Ψ, $\dim W_\Psi \leq \dim R$.

Proof of Lemma B.2. Assume that Ψ is rational and assume that $R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in X}, B, C)$ is a representation of Ψ. Let $\dim \mathcal{X} = n$ and let $e_l \in \mathcal{X}$, $l = 1, 2, \ldots, n$ be a basis of $\mathcal{X}$. Define the formal power series $Z_l \in \mathbb{R}^p\langle\langle X^*\rangle\rangle$, $l = 1, \ldots, n$ by $Z_l(w) = CA_w e_l$ for each word $w \in X^*$. For each index $j \in J$ there exist reals $\alpha_{j,1}, \ldots, \alpha_{j,n} \in \mathbb{R}$ such that $B_j = \sum_{l=1}^n \alpha_{j,l} e_l$. We get that for any word $w \in X^*$ and index $j \in J$, $S_j(w) = CA_w B_j = \sum_{l=1}^n \alpha_{j,l} CA_w e_l = \sum_{l=1} \alpha_{j,l} Z_l(w)$. That is, for any $j \in J$, S_j belongs to the linear span of $Z_1, \ldots, Z_n$. Hence, $w \circ S_j$ belongs to the linear span of $w \circ Z_1, \ldots, w \circ Z_n$. But for any $w, v \in X^*$ and index $l = 1, \ldots, n$, $w \circ Z_l(v) = Z_l(wv) = CA_v A_w e_l = \sum_{k=1}^n \beta_{k,l} CA_v e_k = \sum_{k=1}^n \beta_{k,l} Z_k(v)$ where $\beta_{1,l}, \ldots, \beta_{n,l} \in \mathbb{R}$ satisfy $A_w e_l = \sum_{k=1}^n \beta_{k,l} e_k$. That is, the shift $w \circ Z_l$ belongs to the linear span of $Z_1, \ldots, Z_n$. Hence, we get that for any word $w \in X^*$, $w \circ S_j$ belongs to the linear span of $Z_1, \ldots, Z_n$. Hence, W_Ψ is a subspace of the linear span of $Z_1, \ldots, Z_n$ and therefore $\dim W_\Psi \leq n < +\infty$. $\square$

Proof of Theorem 4.1. From Remark B.1 it follows that $\dim W_\Psi = \operatorname{rank} H_\Psi$. If $\operatorname{rank} H_\Psi < +\infty$, then Lemma B.1 implies that R_Ψ is a well-defined representation of Ψ, hence Ψ is rational. Conversely, if Ψ is rational then Lemma B.2 implies that $\dim W_\Psi = \operatorname{rank} H_\Psi < +\infty$. $\square$

Procedure B.1. For each word $w \in X^*$ and index $j \in J$ denote the column of H_Ψ indexed by (w, j) as $(H_\Psi)_{.,(w,j)}$. Let $\mathrm{Im} H_\Psi = \mathrm{Span}\{(H_\Psi)_{.,(w,j)} \in \mathbb{R}^{X^*\times I} \mid (w, j) \in X^* \times J\}$ be the vector space spanned by the columns of H_Ψ. Then the map $T : W_\Psi \to \mathrm{Im} H_\Psi$ defined by $T(w \circ S_j) = (H_\Psi)_{.,(w,j)}$ is a linear isomorphism. Define the representation $R_{H,\Psi} = (\mathrm{Im} H_\Psi, \{TA_\sigma T^{-1}\}_{\sigma\in X}, TB, CT^{-1})$, where $TB = \{T(B_j) \mid j \in J\}$. It then follows that $T : R_\Psi \to R_{H,\Psi}$ is a representation isomorphism and $R_{H,\Psi}$ is a representation of Ψ.

Next, we present the proof of Theorem 4.2. To this end, we formulate constructions, similar to reachability and observability reductions for linear systems. Let R be a p-J representation of Ψ of the form (4.2).

Procedure B.2 (reachability reduction)**.** Define the p-J representation $R_r = (W_R, \{A^r_\sigma\}_{\sigma\in X}, B^r, C^r)$, where for each $\sigma \in X$, the linear map A^r_σ is the restriction of A_σ to W_R, *i.e.* for all $x \in W_R$, $A^r_\sigma x = A_\sigma x$; $B^r = \{B_j \in \mathcal{X} \mid j \in J\}$, *i.e.* the indexed set B^r coincides with B; finally, the linear map C^r equals the restriction of the map C to W_R, *i.e.* for all $x \in W_R$, $C^r x = Cx$.

Lemma B.3. *The representation* R_r *defined above is well defined, it is a representation of* Ψ *and it is reachable. Moreover,* $\dim R_r \leq \dim R$*, and* $\dim R_r = \dim R$ *holds if and only if* R *is reachable.*

Procedure B.3 (observability reduction)**.** Define the p-J representation $R_o = (\mathcal{X}/O_{R_r}, \{\widetilde{A}_\sigma\}_{\sigma\in X}, \widetilde{B}, \widetilde{C})$. Here $\mathcal{X}/O_R$ denotes the quotient space of $\mathcal{X}$ with respect to O_R, *i.e.* $\mathcal{X}/O_R$ is the linear space formed by equivalence classes $[x]$ with $x \in \mathcal{X}$, where $[x] = [y]$ if and only if $x - y \in O_R$. The map $\widetilde{A}_\sigma : \mathcal{X}/O_R \to \mathcal{X}/O_R$ is defined as $\widetilde{A}_\sigma[x] = [A_\sigma x]$ for all $x \in \mathcal{X}$; the indexed set $\widetilde{B} = \{\widetilde{B}_j \in \mathcal{X}/O_R \mid j \in J\}$ is defined by requiring that for all $j \in J$, $\widetilde{B}_j = [B_j]$, and $\widetilde{C} : \mathcal{X}/O_R \to \mathbb{R}^p$ is defined by $\widetilde{C}[x] = Cx$ for each $x \in \mathcal{X}$.

Lemma B.4. *The representation* R_o *is an observable representation of* Ψ*. If* R *is reachable, then* R_o *is reachable. In addition,* $\dim R_o \leq \dim R$*, and* $\dim R = \dim R_o$ *if and only if* R *is observable.*

Procedure B.4 (transformation to a canonical representation)**.** Use Procedure B.2 to construct the reachable representation R_r. Apply then Procedure B.3 to R_r and obtain the observable representation $R_{\mathrm{can}} = (R_r)_o$.

Lemma B.5. *R_{can} is a well defined representation of Ψ, it is reachable and observable. Moreover,* $\dim R_{\mathrm{can}} \leq \dim R$*, and* $\dim R = \dim R_{\mathrm{can}}$ *holds if and only if R is reachable and observable.*

Remark B.2 (computations)**.** If J is finite, then Procedures B.2, B.3 and B.4 can be implemented, see [20].

Proof of Theorem 4.2. The proof of the theorem will be divided into the proof of the following implications: (i) $\implies$ (ii), (ii) $\implies$ (iii), and (iii) $\implies$ (i). These implications prove that the first three statements are equivalent. In addition, we will show that (i) and (iv) are equivalent. Finally, we will show that any two minimal representations of Ψ are isomorphic.

(i) $\implies$ (ii). Assume that $R_{\min}$ is a minimal representation of Ψ, but $R_{\min}$ is either not reachable or it is not observable. Then by Lemma B.5 we can transform $R_{\min}$ to a reachable and observable representation R_{can} of Ψ, such that $\dim R_{\mathrm{can}} < \dim R_{\min}$. But this contradicts to minimality of $R_{\min}$.

(ii) $\implies$ (iii). Let $R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in X}, B, C)$ be a reachable representation of Ψ. Assume that $R_{\min}$ is of the form $R_{\min} = (\mathcal{X}_m, \{A^m_\sigma\}_{\sigma\in X}, B^m, C^m)$. Define the map $T : \mathcal{X} \to \mathcal{X}_m$ as follows. From reachability of R it follows that for any element of x of $\mathcal{X}$ there exists a finite subset $I \subseteq J$ and a collection of reals α_j, $j \in I$ and words $w_j \in X^*$, $j \in I$ such that $x = \sum_{j\in I} \alpha_j A_{w_j} B_j$. Define then the action of T on x by $T(x) = \sum_{j\in I} \alpha_j A^m_{w_j} B^m_j$. Using observability of R_m and the fact that for all $w \in X^*$, $j \in J$, $CA_w B_j = S_j(w) = C^m A^m_w B^m_j$, it can be shown that T is a well-defined linear map, moreover, it is a representation morphism. In addition, reachability of $R_{\min}$ implies that T is surjective.

(iii) $\implies$ (i). Let R be a representation of Ψ, let R_r be the representation obtained by applying Procedure B.2 to R. It follows then from Lemma B.5 that R_r is a reachable representation of Φ and $\dim R_r \leq \dim R$. By part (iii) there exists a surjective map $T : R_r \to R_{\min}$. But $\dim R \geq \dim R_r = \dim T(W_R) = \dim R_{\min}$, so $R_{\min}$ is indeed a minimal representation of Ψ.

(iv) $\iff$ (i). From Lemma B.2 and Remark B.1 it follows that the dimension of any rational representation of Ψ is at least $\dim W_\Psi = \operatorname{rank} H_\Psi$. From Lemma B.1 it follows that $\dim R_\Psi = \operatorname{rank} H_\Psi$. Hence, R_Ψ is a minimal rational representation of Ψ. Hence, if R_m is another minimal representation of Ψ, then $\operatorname{rank} H_\Psi = \dim R_m$, *i.e.* (i) implies (iv). Conversely, if R_m is a rational representation such that $\dim R_m = \operatorname{rank} H_\Psi$, then for any rational representation R of Ψ $\dim R_m = \operatorname{rank} H_\Psi \leq \dim R$, *i.e.* R_m is minimal.

Finally, it is left to show that all minimal representations of Ψ are isomorphic. To this end, let $R_{\min} = (\mathcal{X}_m, \{A^m_\sigma\}_{\sigma\in X}, B^m, C^m)$ be a minimal representation of Ψ. Let $R = (\mathcal{X}, \{A_\sigma\}_{\sigma\in X}, B, C)$ be another minimal representation of Ψ. Then R is reachable and there exists a surjective representation morphism $T : R \to R_{\min}$. Since R and $R_{\min}$ are both minimal, they must have the same dimension, *i.e.* $\dim R = \dim R_{\min}$. But the latter implies that $\dim \mathcal{X}_m = \dim T(\mathcal{X})$, which implies that T is an isomorphism. □

Corollary B.1. *The free representation R_Ψ from Lemma* B.1 *is a minimal representation of Ψ. In addition, the representation $R_{\Psi,H}$ defined in Procedure* B.1 *is a minimal representation of Ψ.*

Remark B.3. From Remark B.2, Lemma B.5 and Theorem 4.2 it follows that if J is finite, then any representation R of Ψ can be transformed to a minimal representation of Ψ by a numerical algorithm. Again, see [20] for details.

Proof of Lemma 4.3. Notice that it is enough to show that $\operatorname{rank} H_{\Psi\odot\Theta} \leq \operatorname{rank} H_\Psi \cdot \operatorname{rank} H_\Theta$ holds if Ψ and Θ are rational. Indeed, if this is the case then from Theorem 4.1 it follows that $\operatorname{rank} H_{\Psi\odot\Theta} < +\infty$ and hence $\Psi \odot \Theta$ is rational, if Ψ and Θ are rational. Recall from Definition B.1 the shift invariant space $W_{\Psi\odot\Theta}$. Since by Remark B.1 the dimension of $W_{\Psi\odot\Theta}$, W_Ψ and W_Θ are equal to $\operatorname{rank} H_{\Psi\odot\Theta}$, $\operatorname{rank} H_\Psi$ and $\operatorname{rank} H_\Theta$ respectively, it is enough to show that $\dim W_{\Psi\odot\Theta} \leq \dim W_\Psi \cdot \dim W_\Theta$. if Ψ and Θ are rational. To this end, notice that for any two formal power series $T_1, T_2 \in \mathbb{R}^p\langle\langle X^*\rangle\rangle$ and any word $w \in X^*$, it holds that $w \circ (T_1 \odot T_2) = (w \circ T_1) \odot (w \circ T_2)$. Then we get that $W_{\Psi\odot\Theta}$ is spanned by formal power series of the form $(w \circ S_j) \odot (w \odot T_j)$ where $j \in J$, $w \in X^*$. Let $m = \dim W_\Theta$ and $n = \dim W_\Psi$. Fix a basis $w_l \circ T_{z_l}$, $l = 1, \ldots m$, $z_l \in J$, $w_l \in X^*$ of W_Θ. Fix a basis $v_k \circ S_{j_k}$, $v_k \in X^*$, $k = 1, \ldots n$, $j_k \in J$ of W_Ψ. Since the Hadamard product is bilinear, it follows that each formal power series $(w \circ S_j) \odot (w \circ T_j)$, $j \in J$, $w \in X^*$, and hence $W_{\Psi\odot\Theta}$, belongs to the linear space spanned by the formal power series $(v_k \circ S_{j_k}) \odot (w_l \circ T_{z_l})$, $k = 1, \ldots, n$, $l = 1, \ldots, m$. Hence, it follows that $\dim W_{\Psi\odot\Theta} \leq nm$. □

Acknowledgements. The author thanks Jan H. van Schuppen, Pieter Collins and Luc Habets for the useful discussions and suggestions. The author thanks the anonymous referee for the useful comments.

References

[1] J. Berstel and C. Reutenauer, *Rational series and their languages*, *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag (1984).
[2] M.F. Callier and A.C. Desoer, *Linear System Theory*. Springer-Verlag (1991).
[3] P. D'Alessandro, A. Isidori and A. Ruberti, Realization and structure theory of bilinear dynamical systems. *SIAM J. Control* **12** (1974) 517–535.
[4] S. Eilenberg, *Automata, Languages and Machines*. Academic Press, New York-London (1974).
[5] M. Fliess, Matrices de Hankel. *J. Math. Pures Appl.* **53** (1974) 197–222.
[6] M. Fliess, Realizations of nonlinear systems and abstract transitive Lie algebras. *Bull. Amer. Math. Soc.* **2** (1980) 444–446.
[7] M. Fliess, Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bull. Soc. Math. France* **109** (1981) 3–40.
[8] F. Gécseg and I. Peák, *Algebraic theory of automata*. Akadémiai Kiadó, Budapest (1972).
[9] A. Isidori, Direct construction of minimal bilinear realizations from nonlinear input-output maps. *IEEE Trans. Automat. Contr.* **AC-18** (1973) 626–631.
[10] A. Isidori, *Nonlinear Control Systems*. Springer-Verlag (1989).
[11] N. Jacobson, *Lectures in Abstract Algebra*, Vol. **II**: *Linear algebra*. D. van Nostrand Company, Inc., New York (1953).
[12] B. Jakubczyk, Existence and uniqueness of realizations of nonlinear systems. *SIAM J. Control Optim.* **18** (1980) 455–471.
[13] B. Jakubczyk, Realization theory for nonlinear systems, three approaches, in *Algebraic and Geometric Methods in Nonlinear Control Theory*, M. Fliess and M. Hazewinkel Eds., D. Reidel Publishing Company (1986) 3–32.
[14] W. Kuich and A. Salomaa, Semirings, Automata, Languages, in *EATCS Monographs on Theoretical Computer Science*, Springer-Verlag (1986).
[15] D. Liberzon, *Switching in Systems and Control*. Birkhäuser, Boston (2003).
[16] M. Petreczky, Realization theory for linear switched systems, in *Proceedings of the Sixteenth International Symposium on Mathematical Theory of Networks and Systems* (2004). [Draft available at http://www.cwi.nl/~mpetrec.]
[17] M. Petreczky, Realization theory for bilinear hybrid systems, in *11th IEEE Conference on Methods and Models in Automation and Robotics* (2005). [CD-ROM only.]
[18] M. Petreczky, Realization theory for bilinear switched systems, in *Proceedings of 44th IEEE Conference on Decision and Control* (2005). [CD-ROM only.]
[19] M. Petreczky, Hybrid formal power series and their application to realization theory of hybrid systems, in *17th International Symposium on Mathematical Networks and Systems* (2006).
[20] M. Petreczky, *Realization Theory of Hybrid Systems*. Ph.D. Thesis, Vrije Universiteit, Amsterdam (2006). [Available online at: http://www.cwi.nl/~mpetrec.]
[21] M. Petreczky, Realization theory for linear switched systems: Formal power series approach. *Syst. Control Lett.* **56** (2007) 588–595.
[22] C. Reutenauer, The local realization of generating series of finite lie-rank, in *Algebraic and Geometric Methods in Nonlinear Control Theory*, M. Fliess and M. Hazewinkel Eds., D. Reidel Publishing Company (1986) 33–43.
[23] M.-P. Schtzenberger, On the definition of a family of automata. *Inf. Control* **4** (1961) 245–270.
[24] E.D. Sontag, *Polynomial Response Maps*, *Lecture Notes in Control and Information Sciences* **13**. Springer Verlag (1979).
[25] E.D. Sontag, Realization theory of discrete-time nonlinear systems: Part I – The bounded case. *IEEE Trans. Circuits Syst.* **26** (1979) 342–356.
[26] Z. Sun, S.S. Ge and T.H. Lee, Controllability and reachability criteria for switched linear systems. *Automatica* **38** (2002) 115–786.
[27] H. Sussmann, Existence and uniqueness of minimal realizations of nonlinear systems. *Math. Syst. Theory* **10** (1977) 263–284.
[28] Y. Wang and E. Sontag, Algebraic differential equations and rational control systems. *SIAM J. Control Optim.* **30** (1992) 1126–1149.