Model Reduction of Linear Switched Systems and LPV State-Space Models

Bastug, Mert

DOI (link to publication from Publisher):
10.5278/vbn.phd.engsci.00045

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

MODEL REDUCTION OF LINEAR SWITCHED SYSTEMS AND LPV STATE-SPACE MODELS

BY
MERT BAŞTUĞ

DISSERTATION SUBMITTED 2016

AALBORG UNIVERSITY
DENMARK

Model Reduction of Linear Switched Systems and LPV State-Space Models

Ph.D. Thesis

Mert Baştuğ

Submitted December 23, 2015

Dissertation submitted:	January 2016
PhD supervisors:	Prof. Rafael Wisniewski Aalborg University
	Assoc. Prof. Mihály Petreczky École des Mines de Douai
Assistant PhD supervisor:	Assoc. Prof. John-Josef Leth Aalborg University
PhD committee:	Associate Professor Zhenyu Yang (chairman) Aalborg University, Denmark
	Professor Athanasios C. Antoulas Rice University, USA
	Professor Zhendong Sun Chinese Academy of Sciences, China
PhD Series:	Faculty of Engineering and Science, Aalborg University

ISSN (online): 2246-1248
ISBN (online): 978-87-7112-481-1

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2016

Abstract

A linear switched system (LSS) is a concatenation of more than one linear subsystems, for which the operating subsystem at each time instant is chosen by a function called the "switching signal", whose range is the set of discrete modes. The allowed set of switching sequences for an LSS can be arbitrary (user defined) or constrained due to the modeled physical process itself or due to the physical constraints on choosing the input. In this work, some methods are presented to approximate the input output behavior of an LSS with arbitrary or restricted switching, with another LSS of smaller complexity. Smaller system complexity in this context refers to "smaller continuous state-space dimension".

The methods are based on a non-trivial generalization of Krylov subspace-based moment matching methods, to the linear switched systems. The newly developed methods are numerically much more efficient than some naive approaches appeared in the literature previously for the same problem. The numerical advantage of the given methods stems from the fact that they do not rely on computing the finite Hankel matrices of an LSS, whose size increases exponentially with the number of discrete modes (linear subsystems) of an LSS.

The work consists of five major parts. The first four parts state model order reduction methods for LSSs. The first method can be interpreted as the complete analogue of the solution to the moment matching problem in the linear case, for linear switched systems. The second method is more general and it is based on the so-called "nice selections" of some basis vectors, for some subspaces of reachability/unobservability spaces of an LSS; and it allows for choosing the order of the reduced LSS a priori. In the third part, the problem of model reduction of LSSs with constrained switching is considered. The proposed method (whenever possible) computes a reduced order LSS from a given LSS whose input-output behavior is *exactly* the same with the one of the original LSS. The fourth part further discusses this method, constructing its ties with system theoretical properties like reachability and observability. Namely, the definitions of reachability and observability of LSSs with respect to a constrained set of switching sequences are proposed, and a method to reduce an unreachable and/or unobservable LSS to a reduced order reachable and/or observable LSS (with respect to the same set of constrained switching) is given. The method again preserves the complete input-output behavior. In the last part of the work, a similar approach based on moment matching is taken for the purpose of model reduction of linear parameter varying (LPV) state-space (SS) representations with affine dependence on the scheduling variable. This jump from linear switched systems to LPV-SS representations is possible by observing the Markov parameters (moments) uniquely defining

the input-output behavior of an LSS with arbitrary switching and an LPV-SS representation can be defined in an equivalent way for both cases. For each method, the results are illustrated with numerical examples. The proposed methods are expected to be useful for control synthesis for LSSs and LPV-SS models, since reduced order approximated plants can be used for control synthesis instead of the original ones, yielding the resulting controllers to be of reduced order as well.

Synopsis

Et lineært skifte system (LSS) er en sammenkædning af lineær delsystemer, hvor det aktive delsystemet bestemmes af en tidsafhængig funktion, kaldet et "skifte signal", som tager værdier i en mængde af diskrete tilstande. Den tilladte mængde af skifte signaler hørende til et LSS kan være vilkårlig (brugerdefineret) eller begrænset f.eks. på grund af den modellerede fysiske proces. I dette arbejde præsenteres nogle metoder til tilnærmelse af input-output opførsel af et LSS, med vilkårlig eller begrænset skifte signaler, med et andet LSS af mindre kompleksitet. Mindre kompleksitet henviser i denne sammenhæng til mindre tilstandsrumsdimension.

Metoderne er baseret på en ikke-triviel generalisering af "Krylov subspace-based moment matching" metoder for lineære systemer, til lineære skifte systemer. De nyudviklede metoder er numerisk set mere applikations orienteret end tilsvarende metoder fra litteraturen. De numeriske fordele stammer fra det faktum, at de ikke beror på beregninger af endelige Hankel matricer hørende til et LSS, hvis størrelse stiger eksponentielt med antallet af diskrete tilstande (lineære delsystemer) hørende til et LSS.

Dette arbejde består af fem hoveddele. De første fire dele omhandler "model orden reduktions" metoder for LSSer. Første metode kan fortolkes som værende lineær skifte systemers analog til "moment matching" metoden for lineære systemer. Den anden metode er mere generel og er baseret på de såkaldte "nice selections" af basisvektorer, for underrum af det kontrollerbare/ikke-observerbare rum hørende til et LSS. Metoden giver a prior mulighed for at vælge ordnen af den reducerede LSS. I tredje del behandles problemet omhandlende modelreduktion af LSSer med begrænset skifte signaler. Den foreslåede metode beregner (når det er muligt) et reducerede ordens LSS ud fra et given LSS, således at input-output adfærd stemmer overens med det oprindelige LSS. Den fjerde del diskuterer system teoretiske egenskaber, såsom kontrollerbarhed og observerbarhed, ved metoderne beskrevet i de tre første dele. Mere præcist fremlægges der definitioner af kontrollerbarhed og observerbarhed af LSS med begrænset skifte signaler og der præsenteres en fremgangsmåde til at reducere et ikke-kontrollerbar og/eller ikke-observerbar LSS til et reducerede ordens kontrollerbar og/eller observerbare LSS (med samme mængde af begrænsede skifte signaler). Fremgangsmåden bevarer input-output adfærd. I den sidste del af arbejdet, fokuseres der på "moment matching" med henblik på model reduktion af lineære parameter varierende (LPV) systemer med affin afhængighed af skeduleringsvariablen. Dette spring fra lineære skifte systemer til LPV systemer er mulig da Markov parametrene entydigt definerer input-output opførsel af både et LSS med vilkårlig skift og et LPV system. For hver metode er resultaterne illustreret med numeriske eksempler. Det forventes at de foreslåede metoder er anvendelige ved kontrol syntese af LSSer og LPV systemer, da de reduceret systemer kan anvendes til kontrol syntese i stedet for de oprindelige, hvilket medføre en reduceret ordens regulator.

Contents

Thesis Details

Thesis Title: Model Reduction of Linear Switched Systems and LPV State-Space Models
Ph.D. Student: Mert Baştuğ
Supervisors: Prof. Rafael Wisniewski, Aalborg University
Assoc. Prof. Mihály Petreczky, École des Mines de Douai
Assoc. Prof. John-Josef Leth, Aalborg University

The main body of this thesis consist of the following papers.

[A] Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth, "Model Reduction by Moment Matching for Linear Switched Systems" *Proceedings of the American Control Conference*, pp. 3942–3947, Portland, OR, 2014.

[B] Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth, "Model Reduction by Nice Selections for Linear Switched Systems" cond. accepted for publication in *IEEE Transactions on Automatic Control*, 2015.

[C] Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth, "Model Reduction of Linear Switched Systems by Restricting Discrete Dynamics" *Proceedings of the 53rd IEEE Conference on Decision and Control*, pp. 4422–4427, Los Angeles, CA, 2014.

[D] Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth, "Reachability and Observability Reduction for Linear Switched Systems with Constrained Switching" submitted for publication in *Automatica*, 2015.

[E] Mert Baştuğ, Mihály Petreczky, Roland Tóth, Rafael Wisniewski, John Leth and Denis Efimov, "Moment Matching Based Model Reduction for LPV State-Space Models" accepted for publication in *Proceedings of the 54th IEEE Conference on Decision and Control*, Osaka, JAPAN, 2015.

This thesis has been submitted for assessment in partial fulfillment of the PhD degree. The thesis is based on the submitted or published scientific papers which are listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

Preface and Acknowledgements

This thesis is submitted as a collection of papers in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Section of Automation and Control, Department of Electronic Systems, Aalborg University, Denmark. The work has been carried out partially in the Department of Electronic Systems, Aalborg in the period from April 2013 to October 2014 and in the department of Automatic Control and Computer Science, École des Mines de Douai, from October 2014 to today, under the joint supervision of Professor Rafael Wisniewski, Associate Professor Mihály Petreczky and Associate Professor John-Josef Leth. This work has been partially supported by ESTIREZ project of Région Nord-Pas de Calais, France; the Danish Council for Strategic Research (contract no. 11-116843) within the 'Programme Sustainable Energy and Environment' under the "EDGE" (Efficient Distribution of Green Energy) research project; and the Netherlands Organization for Scientific Research (NWO, grant no: 639.021.127).

I would like to thank everyone at the Section of Automation and Control in Aalborg and Department of Automatic Control and Computer Science in Douai, for providing an inspiring, resourceful, diverse and yet calm enough environment, ideal for research. From both places, I have acquired unique insights on research and life in general. I had the chance to assist a theoretical course on Matrix Computations in Aalborg, and teach a more practical Automatic Control laboratory class in Douai. I had the privilege of being in the organizing committee of a workshop on Hybrid Systems. I have taken a variety of courses from the broad fields of computer science, control theory and mathematics, including but not limited to Classical Logics and Model Theory, Automata Theory, Optimization, Nonlinear Systems and Functional Analysis. In the process, surely I have developed myself in the field of research and teaching, and yet, each new topic I managed to familiarize myself partially, provided to the fact of realizing how little I actually know. Hence, if I will be lucky enough in the future to be able to stay in the field of research, I will always keep in mind that one has to be, first of all, a modest student, motivated to learn and always skeptic of what (s)he thinks (s)he knows.

I would especially like to thank my supervisors Mihály Petreczky, Rafael Wisniewski and John-Josef Leth for their constant positive support, limitless patience and for being also excellent friends to me. I need to thank Mihály Petreczky also specifically, since I have been learning a lot from him since the end of my masters studies, and since he is the reason that I could meet the wonderful researchers (and people) like Rafael Wisniewski and John-Josef Leth, who were always there when I needed help, regarding my PhD studies or life in general. I feel extremely lucky for hav-

ing three supervisors throughout my studies, especially when all three were excellent supervisors in every sense.

Last but not least, I would like to express my gratitude to my family and friends in Turkey, and the great friends I have met in Denmark and France, during the last three years. They were a source of constant encouragement, good advice and fun in times of need.

Mert Baştuğ
Aalborg University, December 23, 2015

Part I

Introduction

Introduction

This chapter presents the context, the problem, the approach and the general outline of the thesis.

1 Context and Motivation

Hybrid systems are a class of nonlinear systems which result from the interaction of continuous time dynamical sub-systems with discrete events [1]. More precisely, a hybrid system is a collection of continuous time dynamical systems. The state of each dynamical system is governed by a set of differential equations. Each of the separate continuous time systems are labeled as a discrete state (mode). Hence, the operating continuous-time subsystem at a specific time instant is called the discrete mode and its state is called the continuous state at a specific time instant. The operating discrete mode in any time instant can be chosen arbitrarily, or it may depend on the value of the continuous state or other constraints. The transitions between the discrete states may result in a jump in the continuous state. If there is a jump, the function which maps the continuous state just before the mode change to the continuous state just after the the mode change is called the reset map. The continuous state evolves continuously in time, whenever there is no transition to another discrete mode. Linear switched systems (LSSs) constitute a subclass of these systems where the discrete events interacting with the sub-systems are governed by a piecewise continuous function called the switching signal. The switching signal may either be considered as an additional external input, or an input satisfying certain constraints, or it may depend on the state of the system. Linear switched systems are called "linear" because all of their subsystems operating with respect to the value of the switching signal in a particular time interval, are individually linear systems. All of the contributions related to linear switched systems in this thesis, consider specifically the case when the reset maps (the maps specifying the next continuous state whenever a jump from one discrete mode to another takes place) are taken as the identity map, i.e., the evolution of the continuous state is really a *continuous* function of time. Such systems are used in modeling, analysis and design of supervisory control systems, mechanical systems with impact, circuits with relays or ideal diodes for instance. These examples and some system theoretic properties of switched systems have been studied in detail in [1], [2], [3] and the references therein. Although there is a remarkable amount of literature about switched systems, the subject is relatively recent and it is still open for research.

Linear parameter varying (LPV) systems can be thought as a collection of linear time invariant (LTI) systems, each created by linearizing a non-linear system in different operating points. A time varying signal (which is called the scheduling signal) is introduced to describe the changes in the operating points, from which the title "parameter varying" follows [4]. In turn, an affine LPV system is an LPV system where the dependence on this scheduling signal is affine. In [5], it is shown that each affine LPV system has an *associated* LSS for which the coefficients describing the input-output behavior of the system are exactly the same as those of the corresponding affine LPV system. In addition, the realization theory of LPV systems and LSSs are closely related, [5].

Model reduction is the other important concept about this thesis. For linear time-invariant (LTI) systems the problem of model reduction can be formulated as follows: Given an LTI discrete or continuous time input/output system represented by the convolution integral (or sum), transfer function or state space representation; approximate this system with a simpler system [6]. For the state space representation, the complexity (sometimes called the order) of the system is defined as the dimension of its state space. Thus, by "approximating the system by a simpler one", it is meant discarding some of the states and finding a reduced order approximation for the system. The trade-off between model complexity and accuracy is a central issue. In order to estimate how accurate is the reduced order system compared to the original one, certain norms must be defined and used for measuring the distance between linear systems. For instance, the $\mathcal{L}_2$ norm of an LTI system is defined to be the norm of its impulse response, and the frequency domain equivalent of this norm is called the $\mathcal{H}_\infty$ norm (both turn out to be equal in time and frequency domains respectively [6]).

The model reduction methods can be divided into two main subtopics: Singular Value Decomposition (SVD)-based methods and Krylov-based methods for LTI systems. One of the most significant methods belonging to the former class is called balanced truncation. The idea of the balanced truncation techniques for LTI systems can be summarized as follows [6]: The states which have bigger components in the space spanned by the eigenvectors corresponding to the smaller eigenvalues of *reachability gramian* $\mathcal{P}$ of the system need more energy to be reached, i.e., they are *difficult to reach*. Whereas, the states which have bigger components in the space spanned by the eigenvectors corresponding to the smaller eigenvalues of *observability gramian* $\mathcal{Q}$ of the system yields less observation energy, i.e., they are *difficult to observe*. A state space transformation is called a balancing transformation, if in the resulting basis for the state space, the reachability and observability gramians are equal, i.e., $\mathcal{P} = \mathcal{Q}$, i.e., the states which are difficult to reach are simultaneously the states which are difficult to observe. The reduced order system is then acquired by truncating these components of the states. In contrast, Krylov methods, make use of the partial realization theory and they are based on finding an approximation of the original system whose first certain number of Markov parameters are equal to the original one's. This approach is called "moment matching" where the word "moment" stands for "Markov parameters" which are the parameters defining the input-output behavior of a linear system uniquely. An extensive reference for model reduction of linear systems is [6].

The methods stated in this thesis for the model reduction of linear switched systems (either with arbitrary or restricted switching) or LPV systems with affine dependence on the scheduling parameter can be considered as a non-trivial extension of Krlyov-based moment matching methods used for model reduction of linear systems. Similar to the linear case, all of the methods are applicable to LSSs with unstable modes and LPV systems with unstable state-space parameters. Krylov based methods have also the property to be computationally less cumbersome than the SVD based methods and since there is no matrix factorization involved, the possible ill-conditioning of the balancing transformation matrix in SVD-based methods is avoided [6]. The motivation for extending moment matching techniques to LSSs is that these techniques can be applied to unstable systems and the corresponding model reduction algorithms are computationally efficient. In this thesis, the underlying theory for applying Krylov type model reduction methods to LSSs and LPV systems are built. However, the numerical aspects and computational efficiency of the presented methods should be investigated further. This remains as a topic of future research. Numerical challenges similar to those encountered in classical Krylov methods are likely to appear for LSSs and LPV systems.

The remaining of this section is organized as follows: First, we describe the motivation behind the model reduction of hybrid and linear switched systems. Then we will briefly describe the problem formulation and motivation behind each papers A to D, which deal with LSSs, in the contributions part of the thesis. We will conclude with presenting the motivation for model reduction of LPV systems which is the subject of Paper E. A more detailed outline of these papers and the related work will be presented later on in this chapter.

1.1 Motivation for Model Reduction of LSSs

If the problem of control synthesis for hybrid/switched systems are considered, it is evident that the order of the controller and the computation complexity of controller synthesis usually increase as the number of continuous states of the plant model increases. In particular for robust control, H_∞ control synthesis has a tendency to provide controllers with a very high order, which from the implementation point of view should be reduced [7]. Hence, the smaller the plant model is, the easier it is to synthesize the control law and to implement it. This becomes especially relevant for hybrid systems, as many of the existing control synthesis methods are computationally demanding and result in large scale controllers.

For example, many of the existing control synthesis methods rely on computing a finite-state abstraction of the plant model, see [8] and the references therein. After having a finite-state abstraction for the physical phenomenon or the plant which is desired to be modeled, one then applies discrete-event control synthesis techniques to find a discrete controller for the finite-state abstraction of the plant. When building a finite state abstraction, one proceeds as follows: The continuous state space is partitioned into different regions, which have similar characteristics (similar characteristics are defined according to the problem at hand). Then, each of these regions are considered as a discrete state, hence each *region* is considered as a *point* in the discrete mode set. Usually, the discrete states of this abstraction are not directly measurable, only the events (transition labels) are. This means that the controller has to contain a copy of the abstracted plant model, in order to be able to estimate the discrete state of the finite-state abstraction of the plant, [8–10]. In addition, the complexity of the control synthesis algorithm is at best polynomial in the number of continuous states of the finite-state abstraction [8, 9, 11]. The situation becomes even worse when one considers the case of partial observations, i.e., when not all events (transition labels) of the finite-state abstraction are observable. This can be caused by the nature of the problem [12] or by the non-determinism of the abstraction. In this case, the control synthesis algorithm can even have exponential complexity, [9, 11, 13], and the number of states of the controller can be exponential in the number of the continuous states of the abstraction. Even when one considers the behavior of the open loop plant only, depending on the method used and on the application at hand, the size of the finite-state abstraction i.e., the number of discrete modes of the finite-state abstraction, can be very large. It could even be again exponential in the number of continuous states of the original hybrid model, [8]. In such cases as defined above, synthesis or implementation of controller might become very difficult, even for hybrid systems of moderate size. Clearly, model reduction algorithms could be useful for such cases.

1.2 Motivation for Papers A and B

Paper A presents a non-trivial extension of Krylov type moment matching methods for LSSs. The motivation for considering Krylov based model reduction methods is that they are applicable to unstable systems as well as stable systems and they are numerically attractive. Note that the results of paper A represent a first attempt to extend Krylov based methods to LSSs. However, the approach of Paper A has its shortcomings which prompted us to consider a more general approach in paper B. In Paper B we introduce the framework of nice selections. This framework is flexible, it allows to choose the model order of the reduced system, and it allows to ensure that for designated switching signals, the input-output behavior of the reduced-order and of the original system coincide. Moreover, by using nice selections, we can not only select which switching signals we would like to preserve, but we can also aim at preserving the response on some output channels only, or preserving the response only with respect to certain input channels. That is, nice selections give us both flexibility and allow for the following system theoretic interpretation of the reduced system: *the input-output behavior of the reduced system coincides with that of the original one for some inputs, outputs and switching signals*. The latter interpretation is consistent with the usual interpretation of moment matching for LTI systems [14].

1.3 Motivation for Papers C and D

Papers C and D deal with discrete time LSSs. They both provide algorithms for reducing the order of a discrete time LSS while preserving its input-output behavior for a set of switching sequences. The difference between papers C and D are as follows: The method in paper D preserves the input-output behavior *along* all the switching sequences from a certain set. Whereas the method in paper C, computes possibly a smaller order LSS whose input - output behavior is the same as the original LSS for the time instances corresponding to the last element of each allowed switching sequence. For instance, suppose the switching sequence 122 is the only element of the constrained switching set. Then the method in paper D gives a reduced order LSS whose output is the same as the output of the original LSS for the switching sequence 122 for all inputs and time instances $0, 1, 2$; whereas paper C returns a possibly smaller order LSS whose output is the same with the original LSS for the switching sequence

122 for all inputs at only the time instance 2. In addition, the algorithm in Paper D admits an interpretation as reachability and / or observability reduction of LSSs with constrained switching.

The particular model reduction and reachability/observability reduction problems formulated in papers C and D were motivated by the observation that in many instances, we are interested in the input-output behavior of the model only for certain switching sequences. Restrictions on switching sequences could be imposed by certain physical constraints, which prevent generating all switching sequence in practice, or by existing controllers which generate the switching sequences. The method given in Paper D can be useful when one wishes to use the reduced order model to verify the safety properties of the original model, or to synthesize a controller using the reduced model which ensures the safety properties of the original LSS. The method in Paper C would be useful when one wishes to use the reduced order model to verify the liveness properties of the original model, or to synthesize a controller using the reduced model which ensures the liveness properties of the original LSS. A detailed discussion will be given in Section 4.

1.4 Motivation for Paper E

Finally we turn our attention to Paper E, which proposes a similar model reduction method for linear parameter varying - state space (LPV-SS) representations. Such representations are used in a wide variety of applications, see for instance [15–19]. Their popularity is due to their ability to capture nonlinear dynamics, while remaining simple enough to allow effective control synthesis, for example, by using optimal $\mathcal{H}_2/\mathcal{H}_\infty$ control, Model Predictive Control or PID approaches. LPV-SS representations arising in practice, especially which arise from first-principles based modeling methods, often have a large number of states. This is due to the inherent complexity of the physical process whose behavior the LPV-SS representations are supposed to capture. Unfortunately, due to memory limitations and numerical issues, the existing LPV controller synthesis tools are not always capable of handling large state-space representations [20]. Moreover, even if the control synthesis is successful, large plant models lead to large controllers. In turn, large controllers are more difficult and costly to implement, and they often require application of reduction techniques. For this reason, model reduction of LPV-SS representations is extremely relevant for improving the applicability of LPV systems.

2 Methodology

In this section, a brief review of the (partial) realization theory for linear and linear switched systems will be presented. The main idea of moment matching, used throughout the thesis stems from this theory. Hence, we review the relevant concepts from the theory. In particular, we include a simple version of the Silverman realization algorithm for the single-input single-output (SISO) case at the end of the subsection. This algorithm is included to state the result about approximations to a linear system by partial realizations. This in turn, serves as the main idea of all theory of moment matching for LSSs and LPV-SS representations, constructed in this thesis.

In the following, we will use $\mathbb{N}$ and $\mathbb{R}_+$ to denote respectively the set of natural numbers including 0 and the set $[0, +\infty)$ of nonnegative real numbers. Let S be a topological subspace of an Euclidean space $\mathbb{R}^n$. The set of *piecewise-continuous and left-continous maps* of the form $\mathbb{R}_+ \to S$ is denoted by $PC(\mathbb{R}_+, S)$. That is, $f \in PC(\mathbb{R}_+, S)$ if it has finitely many points of discontinuity on any compact subinterval of $\mathbb{R}_+$, and at any point of discontinuity both the left-hand and right-hand side limits exist, and f is continuous from the left. Moreover, when S is a discrete set it will always be endowed with the discrete topology.

In addition, we denote by $AC(\mathbb{R}_+, \mathbb{R}^n)$ the set of *absolutely continuous maps*, and $L_{loc}(\mathbb{R}_+, \mathbb{R}^n)$ the set of *Lebesgue measurable maps* which are integrable on any compact interval.

2.1 Review of Realization Theory for Linear Systems

A linear time invariant (LTI) system Σ is a tuple (A, B, C) with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$. The state $x(t) \in \mathbb{R}^n$ and the output $y(t) \in \mathbb{R}^p$ of the LTI system Σ at time $t \geq 0$ is defined by

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t), \quad x(0) = x_0 \tag{1a}$$

$$y(t) = Cx(t) \tag{1b}$$

where the input $u \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m)$, $x \in AC(\mathbb{R}_+, \mathbb{R}^n)$ and $y \in PC(\mathbb{R}_+, \mathbb{R}^p)$. We denote the fact that the state space dimension of Σ is n by $\dim(\Sigma) = n$. The number n is also called the *order* of Σ.

Since the following discussion will be on realization theory of LTI systems, we take $x(0) = x_0 = 0$ unless stated otherwise, by following the usual convention. Note that such results can easily be extended to the case of non-zero initial states.

We define the *input-to-state map* X_Σ and *input-to-output map* Y_Σ as the maps

$$X_\Sigma : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \to AC(\mathbb{R}_+, \mathbb{R}^n);\ u \mapsto X_\Sigma(u),$$
$$Y_\Sigma : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \to PC(\mathbb{R}_+, \mathbb{R}^p);\ u \mapsto Y_\Sigma(u).$$

defined by letting $t \mapsto X_\Sigma(u)(t)$ be the solution to (1a) with $x(0) = 0$, and letting $Y_\Sigma(u)(t) = CX_\Sigma(u)(t)$ as in (1b).

Next we discuss what kind of input-output maps can be *realized* (described) by an LTI system. Let $f : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \to PC(\mathbb{R}_+, \mathbb{R}^p)$. The necessary condition for f to be realized by an LTI system Σ of the form (1) is that f is of the form

$$f(u)(t) = \int_0^t G(t-\tau)u(\tau)d\tau \tag{2}$$

where $G : \mathbb{R}_+ \to \mathbb{R}^{p\times m}$ is an analytic map.

Definition 1 ((Minimal) Realization). *An LTI system Σ of the form* (1) *is a realization of a map f of the form* (2) *if for all $u \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m)$*

$$f(u) = Y_\Sigma(u).$$

Moreover, we say that Σ is a minimal realization of f, if for any other realization $\hat{\Sigma}$ of f, $\dim(\Sigma) \leq \dim(\hat{\Sigma})$.

Theorem 1. *An LTI system Σ of the form* (1) *is a realization of a map f of the form* (2) *if*

$$G(t) = Ce^{At}B.$$

Since the map G is analytic, it is uniquely described by the coefficients of its Taylor series expansion around $t = 0$. In the following, we will denote this coefficients by

$$h_k = \left.\frac{d^k}{dt^k}G(t)\right|_{t=0}, \forall k \in \mathbb{N}$$

where $\mathbb{N} = \{0, 1, \dots\}$. If f is realized by an LTI system Σ, these coefficients are called the *Markov parameters* of Σ and they are defined as

$$h_k = \left.\frac{d^k}{dt^k}Ce^{At}B\right|_{t=0} = CA^kB, \forall k \in \mathbb{N}$$

where A^0 denotes the $n \times n$ identity matrix I_n. Hence an equivalent formulation for an LTI realization Σ of a map f can be presented as follows:

Theorem 2. *An LTI system Σ of the form* (1) *is a realization of a map of the form* (2) *if*

$$h_k = \left.\frac{d^k}{dt^k}G(t)\right|_{t=0} = CA^kB, \forall k \in \mathbb{N}.$$

A state $x_f \in \mathbb{R}^n$ of an LTI realization Σ is called reachable from the zero initial state if there exists a time instant $T \geq 0$ and an input $u \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m)$ such that $X_\Sigma(u)(T) = x_f$, i.e., if

$$x_f = X_\Sigma(u)(T) = \int_0^T e^{A(T-\tau)}Bu(\tau)d\tau. \tag{3}$$

Let $\mathbb{X}_{\text{reach}}$ denote the set of all reachable states of an LTI realization Σ. The LTI realization Σ is called reachable if $\mathbb{X}_{\text{reach}}$ constitutes the whole state space $\mathbb{R}^n$, i.e., if $\mathbb{X}_{\text{reach}} = \mathbb{R}^n$.

Let the *zero-input state map* and *zero-input response map* of Σ be defined by the maps

$$X_\Sigma^{x_0} : \mathbb{R}^n \to AC(\mathbb{R}_+, \mathbb{R}^n) Y_\Sigma^{x_0} : \mathbb{R}^n \to PC(\mathbb{R}_+, \mathbb{R}^p)$$

by letting $t \mapsto X_\Sigma^{x_0}(t)$ be the solution to (1a) with $x(0) = x_0$ and $u = 0$, and letting $Y_\Sigma^{x_0}(t) = CX_\Sigma^{x_0}(t)$ as in (1b). Then, a state $x_0 \in \mathbb{R}^n$ of Σ is called unobservable if $Y_\Sigma^{x_0} = 0$, i.e.,

$$Ce^{At}x_0 = 0, \ \forall t \geq 0. \tag{4}$$

Note that the zero initial state is vacuously unobservable. Let $\mathbb{X}_{\text{unobs}}$ denote the set of all unobservable states of an LTI realization Σ. The LTI realization Σ is called observable if $\mathbb{X}_{\text{unobs}}$ consists only the zero initial state, i.e., if $\mathbb{X}_{\text{unobs}} = \{0\}$.

The following well-known Theorems 3 and 4 are given without proof for further discussion. Note that they can be proven by replacing the Taylor series expansion of e^{At} around $t = 0$ in (3), (4), and using the Cayley-Hamilton Theorem [21].

Theorem 3. *The reachable space* $\mathbb{X}_{\text{reach}}$ *of* Σ *is given by*

$$\mathbb{X}_{\text{reach}} = \operatorname{im}(\mathscr{R}) = \operatorname{im}(\begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix}) = \operatorname{im}(\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}).$$

Hence Σ *is reachable if*

$$\operatorname{rank}(\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}) = n.$$

Theorem 4. *The unobservable space* $\mathbb{X}_{\text{unobs}}$ *of* Σ *is given by*

$$\mathbb{X}_{\text{unobs}} = \ker(\mathscr{O}) = \ker\left(\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix}\right) = \ker\left(\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}\right).$$

Hence Σ *is observable if*

$$\operatorname{rank}\left(\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n.$$

The following definition will be given to state a necessary and sufficient condition on a map being realizable.

Definition 2. *The Hankel matrix* $\mathscr{H}^f$ *of a map* f *of the form* (2) *is the following infinite matrix, given in* $p \times m$ *real blocks:*

$$\mathscr{H}^f = \begin{bmatrix} h_0 & h_1 & h_2 & \cdots \\ h_1 & h_2 & h_3 & \cdots \\ h_2 & h_3 & h_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Theorem 5. *If* f *is realized by a* Σ *of the form* (1)*, then the Hankel matrix* $\mathscr{H}^\Sigma$ *of* Σ *is given by*

$$\mathscr{H}^\Sigma = \mathscr{H}^f = \begin{bmatrix} CB & CAB & CA^2B & \cdots \\ CAB & CA^2B & CA^3B & \cdots \\ CA^2B & CA^3B & CA^4B & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Theorem 6 ([6]).
1. *A map* f *of the form* (2) *is realizable by an LTI system* Σ *if and only if* $\operatorname{rank}(\mathscr{H}^f) = n < \infty$.
2. *A realization* Σ *of* f *is minimal if and only if it is reachable and observable.*
3. *All minimal realizations have the order* $n = \operatorname{rank}(\mathscr{H}^f)$.
4. *If* $\Sigma_1 = (A_1, B_1, C_1)$ *and* $\Sigma_2 = (A_2, B_2, C_2)$ *are two minimal realizations, then there exists a non-singular matrix (isomorphism map)* $S \in \mathbb{R}^{n \times n}$ *such that*

   $$SA_1 = A_2S, \quad SB_1 = B_2 \quad C_1 = C_2S.$$

   In this case, Σ_1 *and* Σ_2 *are called* isomorphic *systems or realizations.*

One proof of Theorem 6 can be found in [6] (here, the statement of the theorem is slightly simpler than the one in [6] to be consistent with our purposes). There, it is shown how to construct a realization from the finite rank Hankel matrix. By Kalman decomposition, one can reduce a non-reachable and/or a non-observable system to an equivalent reachable and/or observable one. Thus, once any realization of f is constructed, Kalman decomposition can be used to acquire a minimal one. Hence, in the rest, whenever we talk about a realization Σ of f, we may assume that Σ is minimal without loss of generality. In the next, we state a procedure on how to construct a minimal realization for a realizable map f, from its Hankel matrix $\mathscr{H}^f$ for the single-input single-output (SISO) case (this choice is made for simplicity, since the aim of this subsection is only to present the idea of approximations by partial realizations). For this purpose, we define a map f as *SISO-realizable* if f is realizable by a $\Sigma = (A, B, C)$ of the form (1) with $p = m = 1$.

Theorem 7 (Silverman realization algorithm (SISO case) [22], [6]). *Let f be a SISO-realizable map of the form* (2) *and* $\operatorname{rank}(\mathscr{H}^f) = n$. *Define the following sub-matrices of $\mathscr{H}^f$:*

$\Phi \in \mathbb{R}^{n\times n}$ is the $n \times n$ principal minor matrix of $\mathscr{H}^f$, i.e.,

$$\Phi = \begin{bmatrix} h_0 & \cdots & h_{n-1} \\ \vdots & \cdots & \vdots \\ h_{n-1} & \cdots & h_{2n-2} \end{bmatrix}.$$

$\sigma\Phi \in \mathbb{R}^{n\times n}$ is the sub-matrix of $\mathscr{H}^f$ having the rows with the same index as those of Φ and the columns one column to the right of the ones of Φ, i.e.,

$$\sigma\Phi = \begin{bmatrix} h_1 & \cdots & h_n \\ \vdots & \cdots & \vdots \\ h_n & \cdots & h_{2n-1} \end{bmatrix}.$$

$\Gamma \in \mathbb{R}^{n\times 1}$ is the sub-matrix of $\mathscr{H}$ composed of the first n rows of the first column of $\mathscr{H}^f$, i.e.,

$$\Gamma = \begin{bmatrix} h_0 \\ \vdots \\ h_{n-1} \end{bmatrix}.$$

$\Lambda \in \mathbb{R}^{1\times n}$ is the sub-matrix of $\mathscr{H}$ composed of the first n columns of the first row of $\mathscr{H}^f$, i.e.,

$$\Lambda = \begin{bmatrix} h_0 & \cdots & h_{n-1} \end{bmatrix}.$$

Then the LTI system $\Sigma = (A, B, C)$ where

$$A = \Phi^{-1}\sigma\Phi \quad B = \Phi^{-1}\Gamma \quad C = \Lambda$$

is a minimal realization of f.

Now we will define the concept of a partial realization, which lies in the heart of the discussion for finding an *approximation* to an LTI realization.

Definition 3 (Partial realization). *Let a map f of the form* (2) *be realized by a* $\Sigma = (A, B, C)$ *of the form* (1). *Another LTI realization $\bar{\Sigma} = (\bar{A}, \bar{B}, \bar{C})$ is called an N-partial realization of Σ (equivalently, of f) if*

$$h_k = CA^kB = \bar{C}\bar{A}^k\bar{B} = \bar{h}_k, \; k = 0, 1, \ldots, N.$$

Let the map realized by $\bar{\Sigma}$ be called $\bar{f}$. Intuitively, an N-partial realization $\bar{\Sigma}$ of Σ is an approximation for the system Σ because first $N+1$ (not all) Markov parameters of both systems are equal. This means that the Taylor series coefficients for some lower order derivatives (namely, up to order N) of the maps comprising G and $\bar{G}$ are equal. This in turn, implies that $\bar{\Sigma}$ can be considered as an approximation for Σ. However, one has to say when exactly a partial realization becomes a full realization of the system in question. More precisely, what is the minimum number N, such that any N-partial realization of Σ is also a complete realization of Σ? The following corollary to Theorem 7 gives the answer to this question.

Corollary 1 (Partial and full realizations). *Let Σ of the form* (1) *be a realization of a map f of the form* (2). *Any N-partial realization of Σ (equivalently, of f) is also a complete realization if $N \geq 2n - 1$.*

The justification of this corollary follows from noticing that the Markov parameter with the highest index number used in constructing a full realization for f in Theorem 7 is $2n-1$, i.e., the element of $\sigma\Phi$ in its nth row and nth column is h_{2n-1}.

Corollary 1 completes the idea of approximation by partial realizations. It states that as long as $N < 2n - 1$, an N-partial realization of an LTI system Σ is a better approximation of the system as N increases. As soon as $N \geq 2n - 1$, all N-partial realizations become full realizations of Σ. This idea of approximating a large scale LTI system with N-partial realizations of smaller order is called *moment matching* in the literature [6]. In this work, this idea is extended to linear switched systems and LPV-SS representations, to approximate such large-scale systems. In the next section, we will review the analogous results to the ones given in this section, in realization theory of linear switched systems.

2.2 Review of Realization Theory for Linear Switched Systems

Analogous results to the ones given for LTI systems in the previous subsection exist in the recent literature like [23], [3], and this section is devoted to a brief review of these results. The procedures for model or reachability / observability reduction of LSSs with arbitrary or constrained switching given in this work (more precisely, from papers A to D) are built upon the realization theory of LSSs.

Definition 4 (Linear Switched System). *A continuous time linear switched system (LSS) Σ is a tuple $\Sigma = (\{(A_q, B_q, C_q) | q \in Q\}, x_0)$ [1] with $Q = \{1, \ldots, D\}$, $D > 0$, $A_q \in \mathbb{R}^{n \times n}$,*

[1] It is unconventional in the classical linear system theory to include the initial state x_0 as a system parameter, but this choice is taken in the linear switched case for the sake of precision, since most of the following discussion in this work will be about equivalence or approximation of input-output behaviors of linear switched systems resulting from a specific initial state, not necessarily zero. Having stated that, we also remark that the assumption $x_0 = 0$ will be taken for some of the following discussion to keep the notation simpler, whenever the arguments can be easily extended to the nonzero initial state case, or a nonzero initial state is not in the core of the discussion. The feed-forward matrix D, present in the classical theory, is taken to be zero for the same reason.

$B_q \in \mathbb{R}^{n\times m}$, $C_q \in \mathbb{R}^{p\times n}$ for all $q \in Q$ and $x_0 \in \mathbb{R}^n$. The state $x(t)$ and the output $y(t)$ of the LSS Σ at time $t \geq 0$ is defined by

$$\frac{d}{dt}x(t) = A_{\sigma(t)}x(t) + B_{\sigma(t)}u(t), \quad x(0) = x_0 \tag{5a}$$

$$y(t) = C_{\sigma(t)}x(t) \tag{5b}$$

where $\sigma \in PC(\mathbb{R}_+, Q)$ is the switching signal*, $u \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m)$ is the input, $x \in AC(\mathbb{R}_+, \mathbb{R}^n))$ is the state, and $y \in PC(\mathbb{R}_+, \mathbb{R}^p)$ is the output and $Q = \{1, \ldots, D\}$, $D > 0$ is the set of discrete modes. We denote the fact that the state space dimension of Σ is n by* $\dim(\Sigma) = n$. *The number n is also called the* order *of Σ.*

The *input-to-state* map $X_{\Sigma,x}$ and *input-to-output* map $Y_{\Sigma,x}$ of Σ are the maps

$$X_{\Sigma,x_0} : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q) \to AC(\mathbb{R}_+, \mathbb{R}^n); (u, \sigma) \mapsto X_{\Sigma,x}(u, \sigma),$$
$$Y_{\Sigma,x_0} : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q) \to PC(\mathbb{R}_+, \mathbb{R}^p); (u, \sigma) \mapsto Y_{\Sigma,x}(u, \sigma).$$

defined by letting $t \mapsto X_{\Sigma,x_0}(u, \sigma)(t)$ be the solution to the Cauchy problem (5a) with $x(0) = x_0$, and letting $Y_{\Sigma,x_0}(u, \sigma)(t) = C_{\sigma(t)}X_{\Sigma,x}(u, \sigma)(t)$ as in (5b).

Notice that the restriction to a finite interval $[0, t]$ of any switching signal $\sigma \in PC(\mathbb{R}_+, Q)$ can be interpreted as a finite sequence of pairs

$$\mu = (q_1, t_1)(q_2, t_2) \cdots (q_k, t_k) \tag{6}$$

where $q_1, \ldots, q_k \in Q$ and $t_1, \ldots, t_k \in \mathbb{R}_+ \setminus \{0\}$, $t_1 + \cdots + t_k = t$, such that for all $s \in \left[\sum_{l=1}^{i-1} t_l, \sum_{l=1}^{i} t_l\right)$, $i \in \{1, \ldots, k\}$,

$$\sigma(s) = q_i \tag{7}$$

and $\sigma(t) = q_k$. Hence the first element of each pair in (6) represents a discrete mode, and the second element represents the time this mode is active (the time periods $t_1, \ldots, t_k$ are called the *dwell times* of the modes $q_1, \ldots, q_k$ respectively). Note that this encoding is not one-to-one, since if $q_{i-1} = q_i$ for any $i \in \{2, \ldots, k\}$ and $\mu = (q_1, t_1)(q_2, t_2) \cdots (q_k, t_k)$ corresponds to $\sigma|_{[0,t]}$, then

$$(q_1, t_1)(q_2, t_2) \cdots (q_{i-1}, t_{i-1} + t_i)(q_{i+1}, t_{i+1}) \cdots (q_k, t_k)$$

also corresponds to $\sigma|_{[0,t]}$.

From [23], it follows that a necessary condition for an input-output map of the form $f : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q) \to PC(\mathbb{R}_+, \mathbb{R}^p)$ to be realizable by an LSS is that f has a *generalized kernel representation*. For a detailed definition of a generalized kernel representation of f, we refer the reader to [23, Definition 19]. If f has a generalized kernel representation, then there exists a unique family of analytic functions $K^f_{q_1,\ldots,q_k} : \mathbb{R}^k_+ \to \mathbb{R}^p$ and $G^f_{q_1,\ldots,q_k} : \mathbb{R}^k_+ \to \mathbb{R}^{p\times m}$, $q_1, \ldots, q_k \in Q$, $k \geq 1$, such that for all $(u, \sigma) \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q)$, $t > 0$ and for any $\mu = (q_1, t_1)(q_2, t_2) \cdots (q_k, t_k)$ which corresponds to σ on $[0, t]$,

$$\begin{aligned} f(u, \sigma)(t) =& K^f_{q_1 q_2 \cdots q_k}(t_1, t_2, \ldots, t_k) + \\ & \sum_{i=1}^{k} \int_0^{t_i} G^f_{q_i q_{i+1} \cdots q_k}(t_i - s, t_{i+1}, \ldots, t_k) u\left(s + \sum_{j=1}^{i-1} t_j\right) ds, \end{aligned} \tag{8}$$

and the functions $\{K^f_{q_1\cdots q_k}, G^f_{q_1\cdots q_k} \mid q_1,\ldots,q_k \in Q, k \geq 1\}$ satisfy a number of technical conditions, see [23, Definition 19] for details.

Definition 5 ((Minimal) Realization). *An LSS Σ of the form* (5) *is a realization of a map f of the form* (8) *if for all $u \in L_{loc}(\mathbb{R}_+,\mathbb{R}^m)$ and $\sigma \in PC(\mathbb{R}_+, Q)$*

$$f(u,\sigma) = Y_{\Sigma,x_0}(u,\sigma).$$

Moreover, we say that Σ is a minimal realization of f, if for any other realization $\hat{\Sigma}$ of f, $\dim(\Sigma) \leq \dim(\hat{\Sigma})$.

Theorem 8. *An LSS Σ of the form* (5) *is a realization of a map f of the form* (8) *if*

$$K^f_{q_1q_2\cdots q_k}(t_1,t_2,\ldots,t_k) = C_{q_k}e^{A_{q_k}t_k}e^{A_{q_{k-1}}t_{k-1}}\cdots e^{A_{q_1}t_1}x_0,$$

and for all $i \in \{1,\ldots,k\}$

$$G^f_{q_iq_{i+1}\cdots q_{k-1}q_k}(t_i,t_{i+1},\ldots,t_{k-1},t_k) = C_{q_k}e^{A_{q_k}t_k}e^{A_{q_{k-1}}t_{k-1}}\cdots e^{A_{q_{i+1}}t_{i+1}}e^{A_{q_i}t_i}B_{q_i}.$$

Since the maps K^f and G^f are analytic, they are again uniquely described by the coefficients of their Taylor series expansions around $t_1 = \cdots = t_k = 0$. Similar to the LTI case, these coefficients will be called *sub-Markov parameters* of the map f. Some of these coefficients can be collected in a matrix to create a totally analogous notion of Markov parameters in the LTI case, hence these matrices will be called the *Markov parameters* of f. To denote these high-order derivatives of the maps K^f and G^f, we will need the following notation, which is standard in automata theory [24].

Notation 1. *The notation $|S|$ is used to denote the cardinality of a set S. Consider a finite non-empty set Q which will be called the* alphabet. *Denote by Q^* the set of finite sequences of elements of Q. The elements of Q^* are called* words *over Q and any set $L \subseteq Q^*$ is called a* language *over Q. Each non-empty word w is of the form $w = q_1q_2\cdots q_k$ for some $q_1,q_2,\ldots,q_k \in Q$. The element q_i is called the i*th letter *of w, for $i = 1,2,\ldots,k$, and k is called the* length *of w. The* empty sequence (word) *is denoted by ε. The length of word w is denoted by $|w|$; we define $|\varepsilon| = 0$. The set of non-empty words is denoted by Q^+, i.e., $Q^+ = Q^*\setminus\{\varepsilon\}$. The subset of Q^* containing all the words of length at most (resp. at least) $N \in \mathbb{N}$ will be denoted by $Q^{\leq N}$ (resp. $Q^{\geq N}$). The* concatenation *of word $v \in Q^*$ with $w \in Q^*$ is denoted by vw: If $v = v_1v_2\cdots v_k$, and $w = w_1w_2\cdots w_m$, $k > 0, m > 0$, then $vw = v_1v_2\cdots v_kw_1w_2\cdots w_m$. If $v = \epsilon$, then $vw = w$; if $w = \epsilon$, then $vw = v$. The notation $(q)^\omega$ will be used to denote the word $\overbrace{qq\cdots q}^{\omega\ times}$ where $q \in Q$, $\omega \in \mathbb{N}$. We define $(q)^0 = \varepsilon$. For simplicity, the finite set Q will be identified with its index set, that is $Q = \{1,2,\ldots,D\}$ if $|Q| = D$.*

We can now define the sub-Markov parameters of f as follows:

Definition 6 (Sub-Markov parameters of f). *The sub-Markov parameters of f are the values of the maps*

$$S^f : Q^* \to \mathbb{R}^{p\times m}, \quad S^f_0 : Q^* \to \mathbb{R}^p$$

where for all $q_0, q \in Q$,

$$S_0^f(q) = K_q^f(0) \text{ and } S^f(q_0 q) = G_{q_0 q}^f(0,0).$$

and for all $q_0, q \in Q$, $v \in Q^*$, $v \neq \varepsilon$ *by*

$$S_0^f(vq) = \frac{d}{dt_1} \cdots \frac{d}{dt_k} K^f_{q_1 \cdots q_k q}(t_1, \ldots, t_k, 0)\Big|_{t_1 = t_2 = \cdots = t_k = 0}$$
$$S^f(q_0 v q) = \frac{d}{dt_1} \cdots \frac{d}{dt_k} G^f_{q_0 q_1 \cdots q_k q}(0, t_1, \ldots, t_k, 0)\Big|_{t_1 = t_2 = \cdots = t_k = 0}$$

where $v = q_1 q_2 \cdots q_k$, $k \geq 1$, $q_1, q_2, \ldots, q_k \in Q$.

From [23], it follows that the values S^f, S_0^f are the coefficients of the Taylor series of the analytic maps $\{K^f_{q_1 \cdots q_k}, G^f_{q_1 \cdots q_k} \mid q_1, \ldots, q_k \in Q, k \geq 1\}$ around zero, and hence they determine those maps uniquely. More precisely, in [23] it was shown that for any $\omega_1, \ldots, \omega_k \in \mathbb{N}$, $q_1, \ldots, q_k \in Q$, $k > 0$,

$$\begin{aligned} S_0^f(q_1^{\omega_1} \cdots q_k^{\omega_k} q_k) &= \frac{d^{\omega_1}}{dt_1^{\omega_1}} \cdots \frac{d^{\omega_k}}{dt_k^{\omega_k}} K^f_{q_1 \cdots q_k}(t_1, \ldots, t_k)|_{t_1 = \cdots = t_k = 0}, \\ S^f(q_1 q_1^{\omega_1} q_2^{\omega_2} \cdots q_{k-1}^{\omega_{k-1}} q_k^{\omega_k} q_k) &= \frac{d^{\omega_1}}{dt_1^{\omega_1}} \cdots \frac{d^{\omega_k}}{dt_k^{\omega_k}} G^f_{q_1 \cdots q_k}(t_1, \ldots, t_k)|_{t_1 = \cdots = t_k = 0}. \end{aligned} \tag{9}$$

Note that Definition 6 already includes the statement in (9). Nevertheless, (9) is given as an alternative, to emphasize that the Markov parameters in the linear switched case corresponds to *all* combinations of high order derivatives of the maps K^f and G^f with respect to the individual dwell times of each operating mode evaluated at zero, i.e., at $t_1 = \cdots = t_k = 0$.

One way to encode the maps S^f, S_0^f is by defining the following matrix valued map $M^f : Q^* \to \mathbb{R}^{pD \times (mD+1)}$, where for all $v \in Q^*$,

$$M^f(v) = \begin{bmatrix} S_0^f(v1) & S^f(1v1) & \cdots & S^f(Dv1) \\ S_0^f(v2) & S^f(1v2) & \cdots & S^f(Dv2) \\ \vdots & \vdots & \cdots & \vdots \\ S_0^f(vD) & S^f(1vD) & \cdots & S^f(DvD) \end{bmatrix}. \tag{10}$$

The matrix representation of the map $M^f(v)$ as in (10) will be called the *Markov parameter* of f, related to the word v. This notion is the total analogue of the notion of a Markov parameter for the linear case.

If f has a realization by an LSS Σ of the form (5), then the Markov-parameters of f can be expressed as products of the matrices of Σ. In order to present the corresponding formula, we will use the following notation.

Notation 2. *Let* $w = q_1 q_2 \cdots q_k \in Q^*$, $q_1, \ldots, q_k \in Q$, $k > 0$ *and* $A_{q_i} \in \mathbb{R}^{n \times n}$, $i = 1, \ldots, k$. *Then the matrix* A_w *is defined as*

$$A_w = A_{q_k} A_{q_{k-1}} \cdots A_{q_1}. \tag{11}$$

If $w = \varepsilon$, *then* A_ε *is the identity matrix.*

Example 1. *This example illustrates Notation 2. Consider the bimodal ($D = 2$) LSS Σ. Since the system has two modes, the* alphabet *set is: $Q = \{1, 2\}$. If we consider the two words of the set Q^* defined by $w = 112$ and $v = 212$, A_w denotes the matrix $A_w = A_2A_1A_1$ and A_v denotes the matrix $A_v = A_2A_1A_2$. The concatenation of these two words is $wv = 112212$, thus A_{wv} denotes the matrix $A_{wv} = A_vA_w = A_2A_1A_2A_2A_1A_1$. If $N = 2$ the set $Q^{\leq N}$ is defined as $Q^{\leq N} = \{\varepsilon, 1, 2, 11, 12, 21, 22\}$.*

Definition 7 ([23]). *An LSS of the form* (5) *is a realization of a map f of the form* (8) *if*

$$S_0^f(vq) = C_qA_vx_0 \text{ and } S^f(q_0vq) = C_qA_vB_{q_0}, \quad \forall v \in Q^*, \tag{12}$$

i.e.,

$$M^f(v) = \widetilde{C}A_v\widetilde{B},\ \forall v \in Q^* \tag{13}$$

with $\widetilde{C} = \begin{bmatrix} C_1^{\mathrm{T}}, & \cdots, & C_D^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ and $\widetilde{B} = \begin{bmatrix} x_0, & B_1, & B_2, & \cdots, & B_D \end{bmatrix}$.

Note that Definition 7 corresponds to the analogous definition in the previous subsection (Definition 2) of a realization in the LTI case, for the linear switched case. This relation can be explained as follows: Consider an LSS realization Σ of the form (5) with $x_0 = 0$ and $Q = \{1\}$, i.e, the LSS only has one discrete mode, in other words, the LSS is actually just an LTI system. In this case the set Q^* would be defined with $Q^* = \{\varepsilon, 1, 11, 111, \dots\}$, hence a word $v \in Q^*$ could just be defined by the number of the letter 1 it contains. Hence $M^f(v) = C_1A_1^{|v|}B_1$. It means that if $\Sigma = (A, B, C)$ is an LTI system of the form (1) with $A = A_1$, $B = B_1$ and $C = C_1$ then the corresponding Markov parameters of both systems would be equal, i.e., $M^f(v) = C_1A_1^{|v|}B_1 = CA^kB$ for all words $v \in Q^*$ and $k \in \mathbb{N}$ such that $|v| = k$.

A state $x_f \in \mathbb{R}^n$ of an LSS realization Σ is called reachable from the initial state $x(0) = x_0$ if there exists a time instant $T \geq 0$, an input $u \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m)$ and a switching signal $\sigma \in PC(\mathbb{R}_+, Q)$ associated with the switching sequence $\mu = (q_1, t_1) \cdots (q_k, t_k)$, $t_1 + \cdots + t_k = T$, $k \in \{1, 2, \dots\}$ on the time interval $[0, T]$, such that $X_{\Sigma,x_0}(u, \sigma)(T) = x_f$, i.e., if

$$\begin{aligned} x_f =& X_{\Sigma,x_0}(u,\sigma)(T) \\ =& e^{A_{q_k}t_k} \cdots e^{A_{q_1}t_1}x_0 + \sum_{i=1}^{k} \int_0^{t_i} e^{A_{q_k}t_k} \cdots e^{A_{q_i}(t_i-\tau)}B_{q_i}u\left(\tau + \sum_{j=1}^{i-1} t_j\right) d\tau. \end{aligned} \tag{14}$$

where $\mu = (q_1, t_1) \cdots (q_k, t_k)$, $t_1 + \cdots + t_k = T$, $k \in \{1, 2, \dots\}$ is the switching sequence associated with σ on the time interval $[0, T]$. Let $\mathbb{X}_{\text{reach}}$ denote the set of all reachable states of Σ. The LSS realization Σ is called span-reachable if the linear span of the elements of $\mathbb{X}_{\text{reach}}$ constitutes the whole state space $\mathbb{R}^n$, i.e., if $\text{span}(\mathbb{X}_{\text{reach}}) = \mathbb{R}^n$. Note that there is a difference between the analogous definition for the LTI case. The reason for it is that the reachable set from a nonzero initial state in the linear switched case (both in continuous and discrete time) is not necessarily a subspace of the whole state space [3]. For continuous time LSSs, the reachable set from the *zero* initial state constitutes a vector space (for discrete time LSSs this is not necessarily true), but the proof is highly non-trivial [3]. This is why the definition is made in terms of the linear span of the reachable set and the term "span-reachability" is adopted in the LSS realization theory literature [23], [25]. In the rest of this work, whenever the term "reachability" is used for LSSs, the concept of "span-reachability" is meant, as defined here.

A non-zero initial state $x_0 \in \mathbb{R}^n$ of Σ is called unobservable if $Y_{\Sigma,x_0}(0,\sigma) = 0$ for all switching signals $\sigma \in PC(\mathbb{R}_+, Q)$, i.e., if

$$C_{q_k} e^{A_{q_k} t_k} \cdots e^{A_{q_1} t_1} x_0 = 0 \tag{15}$$

for all switching sequences $\mu = (q_1, t_1) \cdots (q_k, t_k)$ of any length $k \in \{1, 2, \dots\}$. Note that the zero initial state is again vacuously unobservable. Let $\mathbb{X}_{\mathrm{unobs}}$ denote the set of all unobservable states of Σ. The LSS realization Σ is called observable if $\mathbb{X}_{\mathrm{unobs}}$ consists only the zero initial state, i.e., if $\mathbb{X}_{\mathrm{unobs}} = \{0\}$.

To state the counterparts of Theorems 3 and 4 we will define the following ordering on Q^* as follows:

Definition 8. (Ordering on Q^*). *Suppose that $Q = \{1, \dots, D\}$. Let the map $\phi : Q^* \to \mathbb{N}$ be defined as follows:*

$$\begin{aligned} &\phi(\varepsilon) = 0 \\ &\phi(v) = q_1(D+1)^{k-1} + q_2(D+1)^{k-2} + \cdots + q_k. \end{aligned} \tag{16}$$

where $v = q_1 q_2 \cdots q_k$ with $q_1, \dots, q_k \in Q$, $k \geq 1$. Then an ordering *$\prec$ on the elements of Q^* can be defined as follows: For any two words $v, w \in Q^*$, if $\phi(v) < \phi(w)$, then $v \prec w$.*

Intuitively, this ordering states that $v \prec w$ if w is bigger than v when the words v, w are interpreted as integer numbers in the basis $D + 1$. Note that for any $v, w \in Q^*$, $v \prec w$ implies $|v| \leq |w|$, and $|v| < |w|$ implies $v \prec w$. In the following, we will assume that the elements of the set Q^* are ordered with this ordering, i.e., if $Q^* = \{v_1, v_2, \dots\}$ then $v_1 \prec v_2 \prec \cdots$.

Let Σ be an LSS realization. We will define the *N-reachability matrix of* Σ $\mathscr{R}_N$ with $N \in \mathbb{N}$ as

$$\mathscr{R}_N = \begin{bmatrix} A_{v_1} \widetilde{B} & A_{v_2} \widetilde{B} & \cdots & A_{v_{|Q^{\leq N}|}} \widetilde{B} \end{bmatrix}. \tag{17}$$

where $|Q^{\leq N}|$ denotes the cardinality of the set $Q^{\leq N}$; $v_1, v_2, \dots, v_{|Q^{\leq N}|} \in Q^{\leq N}$ and $v_1 \prec v_2 \prec \cdots \prec v_{|Q^{\leq N}|}$ with respect to the ordering in Definition 8, and $\widetilde{B} = \begin{bmatrix} x_0 & B_1 & \cdots & B_D \end{bmatrix}$.

In addition, we will define the *N-observability matrix of* Σ $\mathscr{O}_N$ with $N \in \mathbb{N}$ as

$$\mathscr{O}_N = \begin{bmatrix} \widetilde{C} A_{v_1} \\ \widetilde{C} A_{v_2} \\ \cdots \\ \widetilde{C} A_{v_{|Q^{\leq N}|}} \end{bmatrix}. \tag{18}$$

where $v_1, v_2, \dots, v_{|Q^{\leq N}|} \in Q^{\leq N}$ and $v_1 \prec v_2 \prec \cdots \prec v_{|Q^{\leq N}|}$ with respect to the ordering in Definition 8, and $\widetilde{C} = \begin{bmatrix} C_1^{\mathrm{T}} & \cdots & C_D^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$.

Theorem 9. *The linear span of the reachable set* $\mathrm{span}(\mathbb{X}_{\mathrm{reach}})$ *of an LSS Σ is given by*

$$\mathrm{span}(\mathbb{X}_{\mathrm{reach}}) = \mathrm{im}(\mathscr{R}) = \mathrm{im}(\begin{bmatrix} A_{v_1} \widetilde{B} & A_{v_2} \widetilde{B} & A_{v_3} \widetilde{B} & \cdots \end{bmatrix}) = \mathrm{im}(\mathscr{R}_{n-1}).$$

where n is the order of Σ, $v_1, v_2, \dots \in Q^*$ and $v_1 \prec v_2 \prec \cdots$. Hence Σ is reachable if

$$\mathrm{rank}(\mathscr{R}_{n-1}) = n.$$

Theorem 10. *The unobservable space $\mathbb{X}_{\text{unobs}}$ of an LSS Σ is given by*

$$\mathbb{X}_{\text{unobs}} = \ker(\mathscr{O}) = \ker\left(\begin{bmatrix} \widetilde{C}A_{v_1} \\ \widetilde{C}A_{v_2} \\ \widetilde{C}A_{v_3} \\ \vdots \end{bmatrix}\right) = \ker(\mathscr{O}_{n-1}).$$

where n is the order of Σ, $v_1, v_2, \dots \in Q^*$ and $v_1 \prec v_2 \prec \cdots$. Hence Σ is observable if

$$\text{rank}(\mathscr{O}_{n-1}) = n.$$

Theorems 9 and 10 can be proven in a similar fashion to the LTI case, by (14) and repeated use of Cayley-Hamilton Theorem [3], [23].

The following definition of Hankel matrix of a map f of the form (8) will be given to state a necessary and sufficient condition of a map being realizable by an LSS Σ.

Definition 9. *The Hankel matrix $\mathscr{H}^f$ of a map f of the form* (8) *is the following infinite matrix, given in $(pD) \times (mD+1)$ real blocks:*

$$\mathscr{H}^f = \begin{bmatrix} M(v_1v_1) & M(v_2v_1) & \dots & M(v_kv_1) & \dots \\ M(v_1v_2) & M(v_2v_2) & \dots & M(v_kv_2) & \dots \\ M(v_1v_3) & M(v_2v_3) & \dots & M(v_kv_3) & \dots \\ \vdots & \vdots & \dots & \vdots & \ddots \end{bmatrix}.$$

where $v_1, v_2, \dots \in Q^*$ and $v_1 \prec v_2 \prec \cdots$.

If f is realized by a Σ of the form (5), then the Hankel matrix $\mathscr{H}^\Sigma$ of Σ is defined by

$$\mathscr{H}^\Sigma =$$

$$\begin{bmatrix}
C_1x_0 & C_1B_1 & \dots & C_1B_D & C_1A_1x_0 & C_1A_1B_1 & \dots \\
\vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots \\
C_Dx_0 & C_DB_1 & \dots & C_DB_D & C_DA_1x_0 & C_DA_1B_1 & \dots \\
C_1A_1x_0 & C_1A_1B_1 & \dots & C_1A_1B_D & C_1A_1A_1x_0 & C_1A_1A_1B_1 & \dots \\
\vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots \\
C_DA_1x_0 & C_DA_1B_1 & \dots & C_DA_1B_D & C_DA_1A_1x_0 & C_DA_1A_1B_1 & \dots \\
C_1A_2x_0 & C_1A_2B_1 & \dots & C_1A_2B_D & C_1A_2A_1x_0 & C_1A_2A_1B_1 & \dots \\
\vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots \\
C_DA_2x_0 & C_DA_2B_1 & \dots & C_DA_2B_D & C_DA_2A_1x_0 & C_DA_2A_1B_1 & \dots \\
C_1A_1A_1x_0 & C_1A_1A_1B_1 & \dots & C_1A_1A_1B_D & C_1A_1A_1A_1x_0 & C_1A_1A_1A_1B_1 & \dots \\
\vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots \\
C_DA_1A_1x_0 & C_DA_1A_1B_1 & \dots & C_DA_1A_1B_D & C_DA_1A_1A_1x_0 & C_DA_1A_1A_1B_1 & \dots \\
C_1A_2A_1x_0 & C_1A_2A_1B_1 & \dots & C_1A_2A_1B_D & C_1A_2A_1A_1x_0 & C_1A_2A_1A_1B_1 & \dots \\
\vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots \\
C_DA_2A_1x_0 & C_DA_2A_1B_1 & \dots & C_DA_2A_1B_D & C_DA_2A_1A_1x_0 & C_DA_2A_1A_1B_1 & \dots \\
\vdots & \vdots & \dots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.$$

A better intuition of the Hankel matrix $\mathscr{H}^{\Sigma}$ of an LSS realization can be acquired by noticing that it can be written as the product of the infinite observability and reachability matrices of Σ, which are implicitly defined respectively in Theorems 9 and 10:

$$\mathscr{H}^{\Sigma} = \mathscr{O}\mathscr{R} = \begin{bmatrix} \widetilde{C}A_{v_1} \\ \widetilde{C}A_{v_2} \\ \widetilde{C}A_{v_3} \\ \vdots \end{bmatrix} \begin{bmatrix} A_{v_1}\widetilde{B} & A_{v_2}\widetilde{B} & A_{v_3}\widetilde{B} & \cdots \end{bmatrix}.$$

The Hankel matrix of a map f plays a crucial role in the realization theory of LSSs (similar to the LTI case) as it can be used to get a realization Σ for f. Below we present the principal theorem of realizability with LSSs which can be considered as the counterpart of the Theorem 6 given for the LTI case in the previous subsection.

Theorem 11 ([23]). *1. A map f of the form* (8) *is realizable by an LTI system Σ of the form* (5) *if and only if* $\operatorname{rank}(\mathscr{H}^{f}) = n < \infty$.

2. *A realization Σ of f is minimal if and only if it is span-reachable and observable.*

3. *All minimal realizations have the order $n = \operatorname{rank}(\mathscr{H}^{f})$.*

4. *If $\Sigma = (\{(A_q, B_q, C_q) | q \in Q\}, x_0)$ and $\hat{\Sigma} = (\{(\hat{A}_q, \hat{B}_q, \hat{C}_q) | q \in Q\}, \hat{x}_0)$ are two minimal realizations, then there exists a non-singular matrix (isomorphism map) $S \in \mathbb{R}^{n\times n}$ such that*

$$SA_q = \hat{A}_q S, \quad SB_q = \hat{B}_q \quad C_q = \hat{C}_q S \ \forall q \in Q \text{ and } Sx_0 = \hat{x}_0.$$

In this case, Σ and $\hat{\Sigma}$ are called isomorphic *LSS realizations.*

In the same reference [23] where the proofs for the statements in the above theorem can be found, a procedure for reducing a non-reachable and/or a non-observable LSS to an equivalent reachable and/or observable one is given. Thus, again in the linear switched case, whenever we talk about a realization Σ of f, we may assume that Σ is minimal without loss of generality.

In [23], a realization algorithm for LSSs in a similar spirit to the Silverman realization algorithm (Theorem 7 in Subsection 2.1) given. The extension to the case of linear switched systems is far from being trivial, hence the procedure itself will not be explicitly stated here.

Now we will define the concept of a partial realization for LSSs.

Definition 10 (Partial realization (of an LSS)). *Let a map f of the form* (8) *be realized by a $\Sigma = (\{(A_q, B_q, C_q) | q \in Q\}, x_0)$ of the form* (5). *Another LSS realization $\bar{\Sigma} = (\{(\bar{A}_q, \bar{B}_q, \bar{C}_q) | q \in Q\}, \bar{x}_0)$ is called an N-partial realization of Σ (equivalently, of f) if*

$$M^{f}(v) = \widetilde{C}A_v\widetilde{B} = \widetilde{\bar{C}}\widetilde{\bar{A}}_v\widetilde{\bar{B}} = M^{\bar{f}}(v), \ \forall v \in Q^{\leq N}$$

where the map $\bar{f}$ is realized by $\bar{\Sigma}$.

The main idea of the works presented in Papers A to D, concerning LSSs stems from the following: Similar to the LTI case, an N-partial realization $\bar{\Sigma}$ of Σ can be considered as an approximation for the system Σ because Markov parameters indexed by words up to length N of both systems are equal, i.e., $M^f(v) = M^{\bar{f}}(v), \forall v \in Q^{\leq N}$. This means again that the Taylor series coefficients for some lower order derivatives (namely, up to order N) of the maps comprising f and $\bar{f}$ are equal. This, in turn, implies that $\bar{\Sigma}$ is an approximation for Σ. The following natural question arises also in the linear switched case: Is there a number k as in the linear case, such that all N-partial realizations of a Σ with $N \geq k$ are also complete realizations? If yes, what is this number k? From [23] it follows that the answer is yes, and the following corollary to Theorem 4 in [23] presents this number.

Corollary 2 (Partial and full realizations of LSSs). *Let Σ of the form* (5) *be a realization of a map f of the form* (8). *Any N-partial realization of Σ (equivalently, of f) is also a complete realization if $N \geq 2n-1$ (where n is the order of Σ).*

It is worthwhile to note the analogy between this corollary and the corresponding one in the LTI case. In the LTI case, one has to match all derivatives of f of order up to $2n-1$ (evaluated at $t=0$) and it turns out this number is exactly the same for the linear switched case. One important difference is that in the linear case, there is only one coefficient to be matched for a derivative of order $k \leq 2n-1$ whereas in the linear switched case this number increases exponentially with the number of discrete modes. Namely there are D^k coefficients to be matched, corresponding to the derivatives of order $k \leq 2n-1$, where D is the number of discrete modes of Σ (note that this also follows from the fact that there are exactly D^k words of length k in the set Q^* related to Σ). The justification of this corollary follows also from the Silverman-like realization algorithm given in [23] for LSSs, relying on the Hankel matrices.

Corollary 2 completes the circle of ideas lying in the basis of approximation by partial realizations for LSSs. It states that as long as $N < 2n-1$, an N-partial realization of an LSS Σ is a better approximation of the system as N increases. As soon as $N \geq 2n-1$, all N-partial realizations become full realizations of Σ. This idea of approximating a large scale LSS with N-partial realizations of smaller order is called also *moment matching* throughout this work, evoking the LTI case. In the next section, we will review some of the analogous results for discrete time LPV-SS representations (which is the subject of Paper E).

2.3 Review of LPV-SS Realizations

In this section, we will review the Markov parameters and related concepts in realization theory of LPV-SS representations analogous to the two previous subsections. The notation used in this section follows the convention of the related literature on LPV-SS representations [4], [26], hence it is slightly different than the previous two subsections.

A discrete time linear parameter-varying state-space representation with affine dependence on parameters (abbreviated as *LPV-SS representations* in the sequel) is a tuple $\Sigma = (\{(A_i, B_i, C_i)\}_{i=0}^{n_p})$ [2] with $A_i \in \mathbb{R}^{n_x \times n_x}$, $B_i \in \mathbb{R}^{n_x \times n_u}$, $C_i \in \mathbb{R}^{n_y \times n_x}$ for all $i \in \{0, 1, \ldots, n_p\}$, $n_p \geq 1$. The state $x(t) \in \mathbb{R}^{n_x}$ and the output $y(t) \in \mathbb{R}^{n_y}$ of the

[2]For all discussion about LPV-SS representations in this work, the initial state $x(0) = x_0$ is taken to be zero to keep the notation less cumbersome, unless stated otherwise. Note that all the related results can be generalized to the case of nonzero initial states, in a similar fashion shown in the previous subsection.

LPV-SS representation at time $t \in \mathbb{N}$ is defined by

$$\begin{aligned} x(t+1) &= A(p(t))x(t) + B(p(t))u(t) \\ y(t) &= C(p(t))x(t), \end{aligned} \tag{19}$$

where $u(t) \in \mathbb{R}^{n_u}$ is the input, and $p(t) = \begin{bmatrix} p_1(t) & \cdots & p_{n_p}(t) \end{bmatrix}^{\mathrm{T}} \in \mathbb{P} \subseteq \mathbb{R}^{n_p}$ is the scheduling signal at time $t \in \mathbb{N}$. Here $\mathbb{P}$ is an arbitrary but fixed subset of $\mathbb{R}^{n_p}$ with a non-empty interior. The matrices $A(p(t))$, $B(p(t))$, $C(p(t))$ in (19) are assumed to be affine and static functions of $p(t)$, where this dependence is defined explicitly by:

$$\begin{aligned} A(p(t)) &= A_0 + \sum_{i=1}^{n_p} A_i p_i(t), \\ B(p(t)) &= B_0 + \sum_{i=1}^{n_p} B_i p_i(t), \\ C(p(t)) &= C_0 + \sum_{i=1}^{n_p} C_i p_i(t), \end{aligned} \tag{20}$$

where $A_i \in \mathbb{R}^{n_x \times n_x}$, $B_i \in \mathbb{R}^{n_x \times n_u}$, $C_i \in \mathbb{R}^{n_y \times n_x}$ are constant matrices for all $i \in \{0, 1, \ldots, n_p\}$ [3]. The fact that the state space dimension of Σ is n_x, is denoted by $\dim(\Sigma) = n_x$. The number n_x is called again the *order* of Σ.

In this subsection, we will skip the formalities regarding the input-output maps which can be potentially realized by an LPV-SS representation Σ of the form (19). Such formal discussion can be found in [5]. We will keep the discussion for already realized maps, and focus on the similarities of Markov parameters of LPV-SS representations and LSSs. We will end the subsection by stating a similar result as in the last parts of the two previous subsections, which justifies the interpretation of partial LPV-SS realizations for a Σ of the form (19) as an approximation for Σ.

Now we recall the concepts of an *infinite impulse response (IIR)* representation of an input-output map [26] and the concept of sub-Markov parameters.

Consider an LPV-SS representation Σ of the form (19), and consider its input-output map $f : \mathbb{R}^{n_u \mathbb{N}} \times \mathbb{R}^{n_p \mathbb{N}} \to \mathbb{R}^{n_y \mathbb{N}}$ where for any set S, $S^{\mathbb{N}}$ denotes the set of all functions g of the form $g : \mathbb{N} \to S$. Recall from [26] that for any input sequence $\mathbf{u} = \{u(k)\}_{k=0}^{\infty}$ and scheduling sequence $\mathbf{p} = \{p(k)\}_{k=0}^{\infty}$,

$$f(\mathbf{u}, \mathbf{p})(t) = \sum_{m=0}^{t} (h_m \diamond p)(t) u(t-m) \tag{21}$$

for all $t \in \mathbb{N}$ where

$$\begin{aligned} (h_0 \diamond p)(t) &= 0, \\ (h_1 \diamond p)(t) &= C(p(t))B(p(t-1)), \\ &\vdots \\ (h_m \diamond p)(t) &= C(p(t))A(p(t-1)) \cdots A(p(t-m+1))B(p(t-m)), \ \forall m > 1. \end{aligned} \tag{22}$$

[3]In this subsection, we use the notations n_x, n_u, n_y for the corresponding analogous numbers n, m, p in the previous two subsections. We prefer this notation to be consistent with the related literature on LPV-SS representations, see [4], [26] for instance.

The representation above is called the IIR of f. From (22) and (20), it can be seen that the terms $(h_m \diamond p)(t)$, $m \geq 0$ can be written as follows:

$$
\begin{aligned}
(h_0 \diamond p)(t) &= 0, \\
(h_1 \diamond p)(t) &= \sum_{q=0}^{n_p} \sum_{q_0=0}^{n_p} C_q B_{q_0} p_q(t) p_{q_0}(t-1) \\
&\vdots \\
(h_m \diamond p)(t) &= \sum_{q=0}^{n_p} \sum_{j_1=0}^{n_p} \cdots \sum_{j_{m-1}=0}^{n_p} \sum_{q_0=0}^{n_p} C_q A_{j_1} \cdots A_{j_{m-1}} B_{q_0} \hat{p}_{q j_1 \cdots j_{m-1} q_0}
\end{aligned}
\tag{23}
$$

where $p_0(k) = 1$ for all $k \in \mathbb{I}_0^t$ and $\hat{p}_{q j_1 \cdots j_{m-1} q_0} = p_q(t) p_{j_1}(t-1) \cdots p_{j_{m-1}}(t-m+1) p_{q_0}(t-m)$.

Now we are ready to define the sub-Markov parameters of Σ. To this end, we recall that the symbol ϵ denote the empty sequence of integers, i.e. ϵ will stand for a sequence of length zero and we denote by $\mathcal{S}(\mathbb{I}_0^{n_\mathrm{p}})$ the set $\{\epsilon\} \cup \{j_1 \cdots j_m \mid m \geq 1, j_1, \ldots, j_m \in \mathbb{I}_0^{n_\mathrm{p}}\}$ of all sequence of integers from $\mathbb{I}_0^{n_\mathrm{p}}$, including the empty sequence. If $s \in \mathcal{S}(\mathbb{I}_0^{n_\mathrm{p}})$, then $|s|$ denotes the length of the sequence s. By convention, if $s = \epsilon$, then $|s| = 0$.

The coefficients

$$
\begin{aligned}
\eta^{\Sigma}_{q,q_0}(\epsilon) &= C_q B_{q_0}, \\
&\vdots \\
\eta^{\Sigma}_{q,q_0}(j_1 \cdots j_m) &= C_q A_{j_1} \cdots A_{j_m} B_{q_0},
\end{aligned}
\tag{24}
$$

$m \geq 1$; $q, j_1, \ldots, j_m, q_0 \in \mathbb{I}_0^{n_p}$ appearing in (23) are called the *sub-Markov parameters* of the LPV-SS representation Σ. In the sequel, the sub-Markov parameters $\eta^{\Sigma}_{q,q_0}(s)$, $q, q_0 \in \mathbb{I}_0^{n_p}$, $s \in \mathcal{S}(\mathbb{I}_0^{n_\mathrm{p}})$, $|s| = m$ will be called *sub-Markov parameters of* Σ *of length* m. The intuition behind this terminology is totally analogous with the linear switched case: The length of a sub-Markov parameter is determined by the number of A_j matrices which appear in (24) as factors.

Example 2 (Output of an LPV-SS realization). *Let* $\Sigma = (\{(A_i, B_i, C_i)\}_{i=0}^{2})$ *be an LPV-SS realization. Then the output of* Σ *due to the input* $\boldsymbol{u} = \{u(k)\}_{k=0}^{\infty}$ *and scheduling sequence* $\boldsymbol{p} = \{p(k)\}_{k=0}^{\infty}$ *at time* $t = 2$ *will be*

$$
\begin{aligned}
y(2) &= \sum_{i=0}^{2} (h_i \diamond p)(2) \cdot u(2-i) \\
&= 0 + (h_1 \diamond p)(2) \cdot u(2-1) + (h_2 \diamond p)(2) \cdot u(2-2) \\
&= C(p) B(p(t-1)) u(1) + C(p) A(p(t-1)) B(p(t-2)) u(0) \\
&= \sum_{q=0}^{2} \sum_{q_0=0}^{2} C_q B_{q_0} p_q(2) p_{q_0}(1) u(1) + \sum_{q=0}^{2} \sum_{j_1=0}^{2} \sum_{q_0=0}^{2} C_q A_{j_1} B_{q_0} p_q(2) p_{j_1}(1) p_{q_0}(0) u(0).
\end{aligned}
$$

Now we can define the Markov parameters of an ALPV realization in an analogous way to the linear switched case. Intuitively, a Markov parameter related to a word (sequence) $j_1 \cdots j_m$ will be again the collection of all $\eta^{\Sigma}_{q,q_0}(j_1 \cdots j_m) = C_q A_{j_1} \cdots A_{j_m} B_{q_0}$ such that $q, q_0 \in \mathbb{I}_0^{n_p}$, in a block-matrix.

For an LPV-SS representation

$$\Sigma = (\{(A_i, B_i, C_i)\}_{i=0}^{n_p})$$

with the scheduling space dimension n_p, consider an *associated* LSS

$$\Sigma^{\mathrm{LSS}} = (\{(A_q^{\mathrm{LSS}}, B_q^{\mathrm{LSS}}, C_q^{\mathrm{LSS}}) | q \in Q\}, 0)$$

with $D = |Q| = n_p + 1$ discrete modes. Note that the state space parameters of such an LSS and LPV-SS representation satisfy $A_q^{\mathrm{LSS}} = A_{q-1}$, $B_q^{\mathrm{LSS}} = B_{q-1}$, $C_q^{\mathrm{LSS}} = C_{q-1}$ for all $q \in Q = \{1, \ldots, D\}$. For a detailed formal discussion see [5]. Using the correspondence above between LPV-SS and LSS realizations we can see that the sub-Markov parameter $\eta^{\Sigma}_{q,q_0}(j_1 \cdots j_m) = C_q A_{j_1} \cdots A_{j_m} B_{q_0}$ of Σ, equals the sub-Markov parameter $S^f(q_0 v q)$ of Σ^{LSS} where $f = Y_{\Sigma^{\mathrm{LSS}},0}$ is the input-output map of Σ^{LSS} from the zero initial state, and $v = (j_m + 1) \cdots (j_1 + 1)$. Note that in turn, reachable and observable (hence, minimal) realizations can be equivalently formulated for LPV-SS representations [5], as it is done in the linear switched case. These steps are omitted here to avoid the repetition of analogous concepts in different contexts (namely LPV and linear switched) and the focus is given more to the final model reduction idea. To state once again this idea for the LPV case, we present the definition of partial realizations for LPV-SS representations as follows.

Definition 11. *Let $\Sigma = (\{(A_i, B_i, C_i)\}_{i=0}^{n_p})$ be an LPV-SS representation of the form* (19). *Another LPV-SS representation $\bar{\Sigma} = (\{(\bar{A}_i, \bar{B}_i, \bar{C}_i)\}_{i=0}^{n_p})$ is called an N-partial realization of Σ, for some $N \in \mathbb{N}$, if*

$$\forall s \in \mathcal{S}(\mathbb{I}_0^{n_\mathrm{p}}), q, q_0 \in \mathbb{I}_0^{n_\mathrm{p}}, |s| \le N : \eta^{\Sigma}_{q,q_0}(s) = \eta^{\bar{\Sigma}}_{q,q_0}(s)$$

That is, $\bar{\Sigma}$ is an *N-partial realization* of Σ, if sub-Markov parameters of Σ and $\bar{\Sigma}$ up to length N are equal. In other words, $\bar{\Sigma}$ is an N-partial realization of Σ, if

$$\begin{aligned} C_q B_{q_0} &= \bar{C}_q \bar{B}_{q_0}, \ \forall q, q_0 \in \mathbb{I}_0^{n_p}, \\ C_q A_{j_1} \cdots A_{j_k} B_{q_0} &= \bar{C}_q \bar{A}_{j_1} \cdots \bar{A}_{j_k} \bar{B}_{q_0}, \ \forall k \in \mathbb{I}_1^N, \forall q, q_0, j_1, \ldots, j_k \in \mathbb{I}_0^{n_p}. \end{aligned}$$

Now we can build the relationship with partial and full realizations with the following theorem. In a similar way to the linear and linear switched case, this theorem constructs the idea of approximations by N-partial realizations for LPV-SS representations.

Theorem 12 ([5]). *Let $\Sigma = (\{(A_i, B_i, C_i)\}_{i=0}^{n_p})$ be an LPV-SS representation of the form* (19). *Any N-partial realization $\bar{\Sigma}$ is also a complete realization of Σ (i.e., an equivalent realization of Σ) if $N \ge 2n - 1$.*

The main idea of paper E lies, similarly with LSSs, in approximating an LPV-SS Σ representation with N-partial realizations $\bar{\Sigma}$, where $N < 2n_x - 1$ and $\bar{n}_x < n_x$, i.e., the approximated model is of reduced order with respect to the original model ($\bar{n}_x$ denotes the order of $\bar{\Sigma}$).

3 Related Work

In the linear case, model reduction is a mature research area, see [6] and the references therein. The subject of model reduction for hybrid and switched systems was addressed in several papers [27–40]. The paper [29] addresses the question of reducing the number of discrete modes. All the remaining papers address the question of reducing the number of continuous states. Namely, the paper [31] deals with the problem of observability reduction for affine systems whose state spaces are restricted to polytopes and does not study the hybrid case. The rest of the cited papers are given for hybrid systems and they propose methods based on some extensions of balanced truncation techniques for model reduction. These techniques require existence of a solution to an LMI, which implies at least local stability of the subsystems. In contrast, the approach taken in this work functions for LSSs with unstable local modes and LPV-SS representations with unstable state map matrices. However, this comes at a price, since we are not able to propose analytic error bounds, like the ones for balanced truncation [40]. From a practical point of view, the lack of an analytic error bound need not be a very serious disadvantage, since it may often be acceptable to evaluate the accuracy of the the approximation after the reduced model has been computed. For example, one can compute the L_2 distance between the original and reduced order model [40], provided the LSSs in question are quadratically stable. Such analysis on a posteriori error bounds for quadratically stable LSSs is provided in Paper B.

One of the main problems arising in the context of model reduction for switched systems is to define and use a valid norm for comparing how close are the original system and approximated model. For this purpose, metrics and topology for hybrid systems are studied in references as [41] and [42] (though it should be noted that, they can only serve as a posteriori error computations between the original and reduced systems. No clear a priori error bound is provided in this work). Note that each of the proposed methods in the present work does have a system theoretical interpretation, as each method operates on the Markov parameters. Markov parameters of switched systems characterize their input-output behavior uniquely [23]. Moreover, the Euclidean distance between Markov parameters can be used to define a natural distance for LSSs with state-space representations [41, 42]. For this reason we believe that it might be possible to improve the theoretical justification for the proposed algorithms in terms of error bounds. However, this remains a topic of future research. Note that even in the linear case, it is not possible to give clear analytic error bounds for algorithms based on moment matching, [6].

3.1 Work Related to Papers A and B

To the best of our knowledge, the only result provided for moment matching of LSSs has appeared in [43]. In [43], it is pointed out that the partial realization theory can be used for developing theory for model reduction and identification of switched linear systems as well. In the basis of this work lies the following idea: The partial realization theory used for model reduction for switched linear systems. The motivation is that a realization procedure can be interpreted as a model reduction method when it is used for acquiring a *partial* realization for the original system rather than computing a full realization. It can be used to compute a reduced order model in the sense some certain number of Markov parameters (moments) of the reduced order model coincides with the original one. This is similar to the case in Krylov-based methods for approximation of linear systems. However, even though the possibility of model reduction by moment matching for LSSs was hinted in [43], no details were provided, no efficient algorithm was proposed, and no numerical experiments were done. Note that a naive application of the realization algorithm of [43] yields an algorithm whose computational complexity is exponential.

The basis of the model reduction algorithms proposed in these works are similar in spirit to moment matching for linear systems [6, 44] and bilinear systems [45–47], however, the details and the system class considered are entirely different.

3.2 Work Related to Papers C and D

Results on realization theory of linear switched systems with constrained switching appeared in [23]. However, [23] does not yield a model reduction algorithm. There, the continuous time case is considered and it was shown that if Σ is an LSS realization of an input-output map f and M is a number which depends on the cardinality of the state-space of a deterministic finite state automaton accepting L (the restricted switching set), then it is possible to find a $\bar{\Sigma}$ such that $\bar{\Sigma}$ has the same input-output behavior for all the allowed switching sequences in L and

$$\dim \bar{\Sigma} \leq M \dim \Sigma. \tag{25}$$

Let us call such a $\bar{\Sigma}$ as an L-realization of Σ. This result may also be extended for the discrete time case in a similar way. However, as (25) shows, the obtained L-realization can even be of higher dimension than the original system. Whereas the methods given in Paper C and D compute a (possibly) reduced order system, whose order is at most the same as the original one, while preserving the input-output behavior for the switching sequences in L.

In [3], the definition of reachability under a constrained switching set is given, but the counterpart of this definition for observability is not provided. Moreover, there, no procedure for reducing a non-reachable and/or non-observable LSS with constrained switching to a reachable and/or observable one is given. Paper D precisely gives the answers to these questions, by building systematically the related definitions and specifying the conditions under which such a reduction would work. The procedures for reachability / observability reduction for discrete time LSSs with arbitrary switching appeared before in [25], [3]. These procedures and their counterparts given in Paper D for the constrained switching case, actually resemble a Kalman-like decomposition of the reachable / unreachable or observable / unobservable subspaces of a linear system [48] and an LSS [3]. But the question of minimality for constrained switching is still an open problem. Note that it is not clear if the reachability and observability reduction procedures stated in Paper D yields a minimal realization with respect to a constrained switching set.

3.3 Work Related to Paper E

Results on model reduction of LPV-SS representations can be found in several papers such as [49–53]. However, except for [53] they are only applicable to quadratically stable LPV systems. The method of [53] is applicable to quadratically stabilizable and detectable LPV-SS representations. In contrast, paper E does not impose any restrictions on the class of LPV-SS representations. In [54] joint reduction of the number of states and the number of scheduling parameters has been investigated. However, the method of [54] requires constructing the Hankel matrix explicitly. Hence, it suffers from the same curse of dimensionality. The model reduction algorithm for LPV systems described in [26] is related to the method given in Paper E, as it also relies on a realization algorithm and Markov parameters. However, when used for model reduction, the method in [26] does not yield a partial realization. In addition, the computational complexity of the method in [26] increases exponentially with the dimension of the scheduling space, since this method rely on constructing the partial Hankel matrix of an LPV-SS representation explicitly. In turn, the size of these matrices increase exponentially with the dimension of the scheduling space. These issues concerning LPV-SS representations and improvements provided by this work discussed more in detail in Paper E. In contrast, the algorithm proposed in Paper E does not require the explicit computation of Hankel matrices, and its worst-case computational complexity is polynomial. We present an example where the algorithm of [26] is not feasible due to the large size of the Hankel-matrix, while the algorithm given in paper E works without problems.In addition, the system theoretic interpretation of the algorithm is less clear.

4 Content Outline and Comparison of Papers A to E

4.1 Outline of Paper A

Paper A presents a procedure of model order reduction for LSSs with arbitrary switching and it can be considered as the direct analogue of the moment matching problem for linear systems described in Section 2.1. As described in Section 2.2, the Markov parameters of an LSS Σ can be considered as a multiplication of some state space parameters of different local modes, starting with a C matrix, ending with either the initial state x_0 or a B matrix, and factors of individual A matrices in the middle. In this work, more specifically from papers A to D, the initial state $x(0) = x_0$ exciting the autonomous behavior of the LSS is treated as a system parameter, whereas conventionally, the initial state is viewed as a disturbance in the LTI case, and it is taken to be zero to state the results related to the LTI systems theory. There are two main reasons for this rather unconventional choice taken for the LSSs in the current work:

1. LTI models are usually results of linearization of nonlinear systems around the equilibrium point zero. For modeling a nonlinear process by an LSS, this choice may be less clear since LSS modeling can be used when the equilibrium point around which we linearize changes.

2. Consider an LTI system $\Sigma = (A, B, C)$ of the form (1). By Kalman decomposition, one can separate the controllable and uncontrollable parts of Σ as

$$A = \begin{bmatrix} A_c & * \\ 0 & A_{uc} \end{bmatrix}, B = \begin{bmatrix} B_c \\ 0 \end{bmatrix}, C = \begin{bmatrix} C_c & C_{uc} \end{bmatrix}.$$

where $\Sigma_c = (A_c, B_c, C_c)$ and $\Sigma_{uc} = (A_{uc}, 0, C_{uc})$ are respectively controllable and uncontrollable. Hence if we decompose the initial state also as $x_0 = \begin{bmatrix} x_{0_c} \\ x_{0_{uc}} \end{bmatrix}$, then either $x_{0_{uc}} = 0$ or the contribution of $x_{0_{uc}}$ cannot be controlled. In other words, if A_{uc} is stable, in practical cases we can forget about $x_{0_{uc}}$ since its contribution in the output will converge to zero. In addition, if A_{uc} is not stable, we cannot influence the uncontrollable component $e^{A_{uc}t}x_{0_{uc}}$ in the output anyway. Hence, since the set of states reachable from zero is the same as the set of the states which can be steered (controlled) to zero in the continuous time LTI-case, an unreachable state is not of interest in the following sense: Either the unreachable part converges to zero in the output, or there is nothing one can do to control it. On the other hand, if x_0 is reachable in finite time T with the input $u(t)$, $t \in [0, T]$, then we may consider the system being initialized with the zero initial state and some input is applied until the real initial state is reached. Hence in both cases, adding a nonzero initial state to the discussion of the theory can be considered redundant.

In contrast, for LSSs the situation can be more complicated. Since for LSSs, the switching signal σ can also be considered as a control together with the input u, the definition of an uncontrollable state may change depending on the allowed switching sequences. For example, there are examples of LSSs with all local modes unstable but for which a stabilizing switching sequence exists [55]. Consequently, we have no concrete reason to exclude initial states which are not reachable from 0 from the discussion.

Paper A defines the moments of a linear switched state space representation (as explained in Subsection 2.2, it turns out these coefficients called 'moments' uniquely determine the input-output behavior of the LSS, hence they are representation independent parameters of the system itself) analogous to the linear case, and proposes a method of model reduction such that the original and reduced order LSSs have exactly the same moments (Markov parameters) which consists up to a certain number of A matrix factors. In Paper A, these Markov parameters which include up to a certain number of A matrix factors are named as "Markov parameters indexed by sequences up to a certain length N". That approach works for unstable systems, it has a clear system theoretic interpretation (namely, it is the analogue of moment matching methods for linear systems, in the linear switched case) and the corresponding algorithm is computationally efficient. However, despite being easier to understand and apply, that approach has a number of drawbacks. First, it does not contain any rigorous interpretation related to switching sequences, i.e., no precise mathematical statement could be formulated regarding the input-output behavior of the original and reduced order systems for switching sequences of interest. Second, it is quite conservative in the following sense: If the number of discrete modes is sufficiently large, then even for small N, the reduced system tends to have the same dimension as the original one. More precisely, it can happen that even when the number N is chosen to be 0 (the smallest possible value for it), the approximation system computed with the method can already be of the same order with the original system. This was confirmed both by numerical examples, and by analytic results, showing that for any given N, for a large enough number of discrete states the reduced order system will generically be of the same dimension as the original one. The reason for this is intuitively the following: The number of Markov parameters indexed by sequences of discrete modes of length at most N is of the order of magnitude D^N, where D is the number of discrete modes. The Hankel matrix formed by these Markov parameters will have $O(D^{N/2})$ rows and columns, i.e., even for small N, its size maybe larger than $n \times n$ where n is the order of the related LSS realization [23]. Generically, a matrix with more than n rows and n columns will have at least rank n. Consequently, even for small N, we may have an N-partial realization in the sense defined in Subsection 2.2 which is already a complete realization of the original system with the same order. Third, even when the LSS at hand allows for at least one reduced order model with the use of the method, it is not possible to choose this reduced order a priori (before using the method). In generic cases, the reduced order can be *calculated* by using the value N, but in non-generic cases even this prediction on the reduced order is not possible. An obvious solution is to try to preserve not all the Markov parameters, but only those which are of interest for the following reason: Either they occur in those output responses related to some switching sequences which we are really interested in, or we know that in order to preserve them, an LSS of dimension $r < n$ suffices (where n is the order of the original LSS).

4.2 Outline of Paper B and Comparison with Paper A

The considerations mentioned in the previous subsection motivate us to build a less conservative framework for model reduction of linear switched systems and it turns out that the framework of nice selections given in Paper B, allow us to select Markov parameters complying with both criteria discussed above. That is, we can either fix the desired dimension $r < n$ of the reduced LSS and choose a nice selection (essentially, a set of Markov parameters) such that there exists an LSS of dimension r generating those Markov parameters. Alternatively, we can choose a nice selection so that any LSS which generates those Markov parameters will yield the same response as the original one for certain switching signals of interest. These points can be further clarified: Namely, with the approach of nice selections, the user chooses arbitrarily the basis vectors for smaller subspaces of the reachability (respectively observability) space of LSSs, of arbitrary dimensions. With the help of these basis vectors, a projection can be defined and all of the states can be restricted to this smaller subspace. This fact in turn, allows for the a priori choice of the reduced order. In fact, computing the reduced order models in the sense defined in paper A can also be formulated just as a special case of the method based on nice selections, given in paper B. In addition, choosing an arbitrary subspace of the reachability (resp. observability) space and restricting all the states to this subspace has another advantage: It turns out that there are particular switching sequences related to this subspace, where the resulting state trajectories have relatively bigger components in this subspace, when compared to any other. Hence choosing the subspace in such a way results in a cleverly tailored reduced order system, which serves as a *better* approximation for these particular switching sequences. Whenever one is interested not all possible switching sequences but only a subset of them (the restriction of the allowed switching sequences may result from physical constraints if the switching is considered as a control input or it can be directly the result of the modeling or abstraction [8]), the framework of nice selections gives also the tool for a "good approximation". Finally in paper B, for continuous time LSSs, it is shown that whenever the nice selections are chosen such that they contain a basis for the set of states reached by allowed switching sequences, it is possible to get a reduced order system which has *exactly* the same input-output behavior with the original system, for all of the allowed switching sequences. It is also proven in Paper B that if the original LSS considered is quadratically stable, then the resulting reduced order LSS will be quadratically stable as well.

Here is a simple example just to illustrate the basic idea of the method given in paper B (hence also for the method in paper A since it can be considered as a special case of the method in paper B). Consider a continuous time, single input - single output SISO LSS Σ of order 3 with 2 discrete modes of the form:

$$\frac{d}{dt}x(t) = A_{\sigma(t)}x(t) + B_{\sigma(t)}u(t),\, x(0) = x_0 \tag{26}$$

where $\sigma : \mathbb{R}_+ \to Q$ is the switching signal, $Q = \{1, 2\}$ is the set of discrete modes and

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},\, B_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\, A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},\, B_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\, x_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{27}$$

Simple calculation by integrating (26) shows that the LSS Σ is reachable (from zero), i.e., the reachability space of Σ is $\mathbb{R}^3$, i.e., any state $x \in \mathbb{R}^3$ can be reached in a finite time by integrating (26) in time with some switching signal σ and input signal u. Assume we would like to get a reduced order approximation to this LSS of order 2. In this case, both of the methods given in paper A and paper B are applicable. Namely, applying the method in paper A with $N = 0$ or the method in paper B by choosing $\beta = \{(\varepsilon, 1, 1), (\varepsilon, 2, 1)\}$ will be equivalent, both yielding the reduced order approximation LSS $\bar{\Sigma}$ of order 2, whose parameters are given as follows:

$$\bar{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\, \bar{B}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\, \bar{A}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\, \bar{B}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\, x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{28}$$

Figure 1 illustrates the action of both methods in this case. Basically, both methods creates a new basis for the reduced reachability space, namely the columns of the full column rank matrix V whose image is $\mathrm{im}(V) = \mathrm{im}(\begin{bmatrix} B_1 & B_2 \end{bmatrix})$. Then, this matrix is used for defining the orthogonal projection matrix V^{-1} (i.e., the left inverse of V) as $V^{-1} = (V^{\mathrm{T}}V)^{-1}V^{\mathrm{T}}$. So when the projector V^{-1} is applied to a state in $\mathbb{R}^3$ which does not belong to the space $\mathrm{im}(\begin{bmatrix} B_1 & B_2 \end{bmatrix})$, it produces the orthogonal projection of this state onto the space spanned by columns of V. In other words, V^{-1} restricts all the states in $\mathbb{R}^3$ to the space $\mathrm{im}(\begin{bmatrix} B_1 & B_2 \end{bmatrix})$. This action is illustrated on Figure 1. The restricted space of dimension 2 is defined by $R_2 = \mathrm{im}(\begin{bmatrix} B_1 & B_2 \end{bmatrix}) = \mathrm{span}\{e_1, e_3\}$ where e_i, $i = 1, 2$ denotes the ith canonical basis vector in $\mathbb{R}^3$. The state $v \in \mathbb{R}^3$ is projected onto R_2 with V^{-1}, i.e., $V^{-1}v = v_p$ where $v \in \mathbb{R}^3$ and $v_p \in R_2$. The resulting LSS $\bar{\Sigma}$ of order 2, consistent with this state transformation is an approximation to Σ, since with the help of such a projection first some number of Markov parameters (the coefficients of the Taylor series expansion of the input-output map of an LSS around $t = 0$) of Σ and $\bar{\Sigma}$ are equal, hence they have *similar* input-output behaviors.

Now assume that a reduced order approximation of order 1 to the same LSS Σ is desired to be computed. Such a task is not possible with the formulation given in paper A, because even choosing the number $N = 0$ will yield an LSS $\bar{\Sigma}$ of order 2 and if $N \geq 1$ the method will always yield a $\bar{\Sigma}$ of order 3, which is the order of Σ already (in fact, in this case the original LSS Σ and the computed LSS $\bar{\Sigma}$ will be isomorphic). However, with the method in paper B the set β can be chosen as $\beta = \{(\varepsilon, 2, 1)\}$, and hence all the states can be restricted to the subspace $R_1 = \mathrm{im}(\begin{bmatrix} B_2 \end{bmatrix}) = \mathrm{span}\{e_3\}$ with the projection V^{-1} where $\mathrm{im}(V) = R_1$. This example illustrates that the method in paper B is more general than the method in paper A, since it allows the user to choose the order of the reduced model a priori. However, it should be noted that paper A is formulated as it is since it has a more intuitive and easier-to-understand system theoretic interpretation.

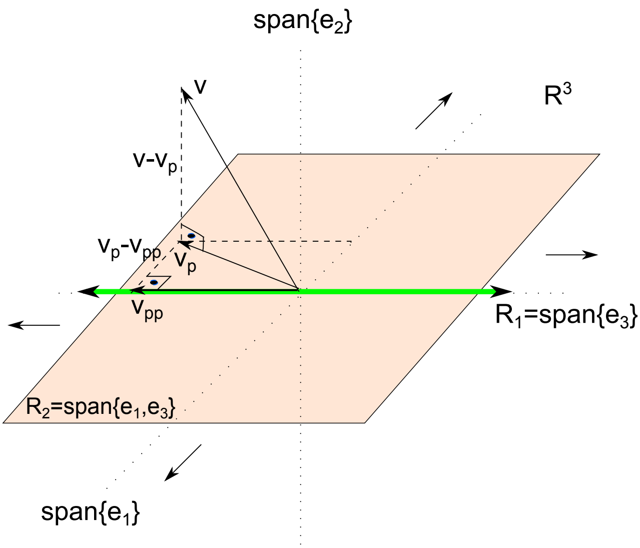

Fig. 1: *Example* 1 *(ACC* 2014 *IEEE TAC* 2015*)*

4.3 Outline and Comparison of Papers C and D

In papers C and D an entirely different problem is considered, but a similar, moment matching based solution is proposed. First of all, these papers deal with the discrete time case rather than the continuous-time. Second, these papers provide a solution to the problem of *exactly* matching the input-output behavior of an LSS with another LSS with possibly reduced order, rather than approximating the input-output behavior. Third, the case of constrained switching is considered. Hence, the LSSs considered in these papers have a specific set of allowed switching sequences (note that *all* possible switching sequences, i.e., the case of arbitrary switching, is also included by simply taking the “restricted” switching sequence set as the whole set Q^*). Hence, the solutions presented in those papers are intended for reachability and / or observability reduction of LSSs with constrained switching, where the definitions of reachability and observability are made in a slightly different manner. The motivation for having these different definitions and formulating two solutions to these two slightly different problem is stated in this subsection.

We will again consider a simple example to illustrate the aim of the methods in papers C and D. Consider this time the discrete time SISO LSS Σ of order 3 with 3 discrete modes

$$x(t+1) = A_{\sigma(t)}x(t) + B_{\sigma(t)}u(t),\ x(0) = x_0 \tag{29}$$

where $\sigma : \mathbb{N} \to Q$ is the switching signal, $Q = \{1, 2\}$ is the set of discrete modes,

$$\begin{aligned} A_1 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ A_2 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ A_3 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned} \tag{30}$$

and $x_0 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$. Let us consider the case of arbitrary switching first, i.e., the set of allowed switching sequences L consists of all possible sequences which can be generated by the alphabet Q, i.e., $L = Q^*$ where Q^* is as defined in Notation 1. Clearly, the set of reachable states for this LSS is not the whole space $\mathbb{R}^3$. Simple calculation by (29) reveals that the reachable set is the union of bold black, blue and green lines on Figure 4, i.e., $R = \operatorname{span}\{e_1\} \cup \operatorname{span}\{e_2\} \cup \operatorname{span}\{e_3\}$ (note that no matter which switching and input signal is chosen, at all time instants $t \in \mathbb{N}$, two components of the state are certainly zero). It can be seen that the reachable set is not a subspace and indeed this can be the case for many discrete or continuous time LSSs (it can also happen that the reachable set is not even a finite union of subspaces [3], unlike even the case in the current example).

The set Q^* of arbitrary switching (non-constrained switching sequence set) can be represented by the language of the non-deterministic finite state automaton (NDFA) shown in Figure 2. The formal definition of an NDFA can be found in the related papers (papers C and D). For now, it is enough to first define a *word* accepted by an NDFA as the concatenation of labels along a path starting from the initial state (named with s_0 in Figure 2) and ending in one of the final states (indicated by double circles in Figure 2); and then define the *language* accepted by the NDFA as the union of all words accepted by it.

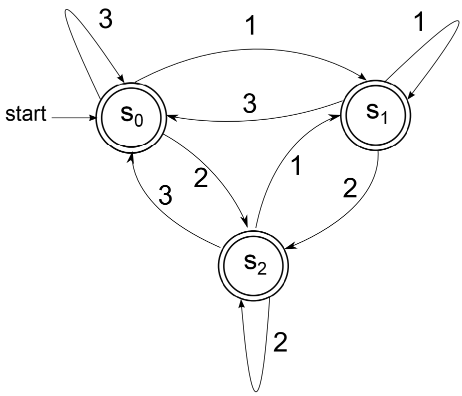

Fig. 2: The NDFA A accepting the arbitrary switching sequence set Q^*

Now consider the case when some of the edges are removed from the NDFA (namely edges labeled with the discrete mode 3) in Figure 2, i.e., the NDFA in Figure 3 (hence, the set of allowed switching sequences is now constrained). Call the language accepted by this NDFA as L. Now the set L of all allowed switching sequences is the set of all sequences in the discrete modes 1 and 2, hence it is a proper subset of Q^*. Let us denote the set of all the states reachable using only the switching sequences in L as $\hat{R}_L$.

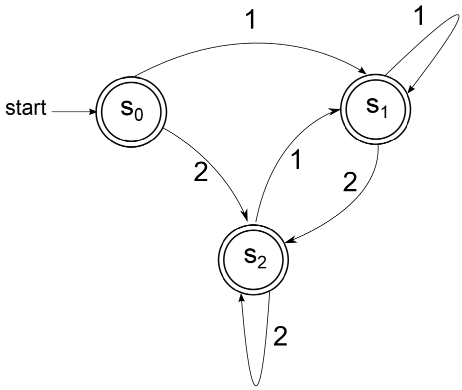

Fig. 3: The NDFA A_L accepting the restricted switching sequence set $L \subset Q^*$

Again, it is clear by simple calculations using (29) that the set $\hat{R}_L$ is the union of bold black and blue lines on Figure 4, i.e., $\hat{R}_L = \text{span}\{e_1\} \cup \text{span}\{e_3\}$ (note that when a jump to discrete mode 3 is not allowed, all of the reachable states rest inside the space $\text{span}\{e_1, e_3\}$, and in this space, one component either in $\text{span}\{e_1\}$ or $\text{span}\{e_3\}$ is always zero). Let us also define the smallest subspace containing the set $\hat{R}_L$ as R_L, i.e., R_L is the linear span of all elements in the set $\hat{R}_L$. For such a case of constrained switching, the methods given in papers C and D basically perform the following: They compute a matrix V for which $\hat{R}_L \subseteq R_L = \text{im}(V)$ and define its left inverse as

$V^{-1} = (V^{\mathrm{T}}V)^{-1}V^{\mathrm{T}}$. With this left inverse, all of the states which are not anymore reachable with the restricted set of switching sequences, are restricted (projected) to the space $R_L = \mathrm{im}(V) = \mathrm{span}\{e_1, e_3\}$. Hence the reduced order LSS consistent with this new basis of the state space is of order 2 and this LSS has clearly the same input-output behavior for all switching sequences in the restricted switching set (language) L. One example state for this procedure can be seen as v on Figure 4, its restricted counterpart to the space R_L being v_p, i.e., $V^{-1}v = v_p$. Note that V projects onto the space R_L, which is larger than $\hat{R}_L$. Hence we keep some states which cannot be reached by the set of allowed switching sequences. One such example state is u in Figure 4: The state u is in the space R_L but not in the set $\hat{R}_L$, and the action of the projection V^{-1} on this state is to leave it unchanged.

In fact the method restricts the system to the subspace R_L which may be larger than the set $\hat{R}_L$. If $\hat{R}_L$ happens to be a vector space then $R_L = \hat{R}_L$. However, in general, the reachable set $\hat{R}_L$ is not a vector space. If we want to represent the restricted system as an LSS, its state space must be a vector space. Therefore, the best one can do is to consider the smallest vector space containing $\hat{R}_L$, which is exactly R_L.

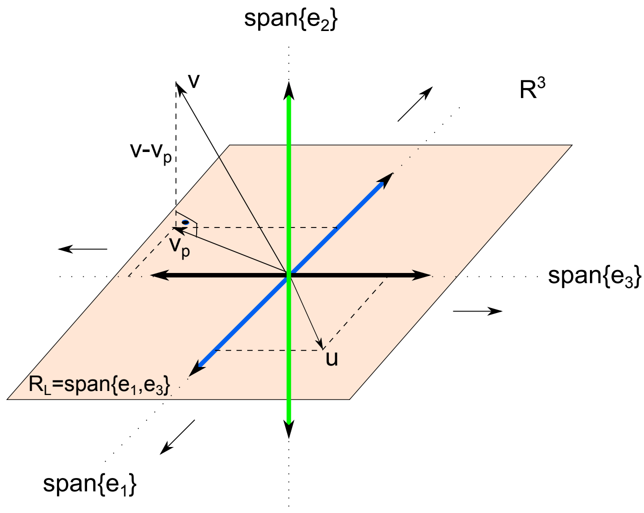

Fig. 4: *Example* 2 *(CDC* 2014 *Automatica* 2015*)*

Note that all the explanations and illustrations so far formulated for the reachability case, can also be formulated dually for the observability case. These dual methods which deal with the issue of observability are also provided in the papers A, B, C and D.

The main difference between the motivations for considering the problems in Paper C and Paper D separately can be explained as follows: The method given in paper C can be useful for verification of a hybrid finite state abstraction models for liveness type of properties (whenever safety concerns are of negligible importance) whereas the method in paper D can be preferred whenever safety and/or liveness related properties have to be verified [56]. In other words, whenever the concern for checking the model can be summed up as "something good *eventually* happens" in the output behavior, one can ignore the output behavior corresponding to the intermediate states of the NDFA which generates the allowed discrete mode (switching) sequences. In such cases, one is only interested in the output behavior corresponding to the final states of the governing NDFA, and it is of little interest *how* this desired behavior is achieved, since the safety concerns are of lesser importance (due to the specific physical problem taken at hand). On the other hand, when one wishes to verify *safety* oriented properties, the concern can be summed up as "something bad *never* happens". In such cases, one has to consider the output behavior *along* each allowed switching sequence (intuitively, it is clear that considering the whole behavior along the switching sequences decreases the amount of possible reduction of the model, with respect to the case of considering the behavior only at the end of the sequences). For the model checking of such properties, it can be useful to reduce the unreachable or unobservable LSS with constrained switching, to a reachable or observable one, by using the method in paper D. This will essentially mean that some *useless* portion of the continuous state space is removed (the state space model is reduced). This in turn may be very beneficial for computational complexity when verifying the model, especially for switched or hybrid systems. Note that since the "behavior along the switching sequence" also includes the behavior at the final states of the automaton, the method in Paper D can be used again for model reduction with the purpose of verifying liveness properties.

The discussion in the previous paragraph can be summarized in the following three items: **(1)** The advantage of Paper C compared with Paper D, could be in the cases when safety issues are not of paramount importance. In such cases, the given method in Paper C results in the possibility of reducing the model order more than the one given in Paper D. **(2)** Whereas the method given in Paper D, even though its reduction capabilities are more limited than the one given in Paper C, is useful also for verifying safety related properties. **(3)** Moreover, analogous with the linear case, the method in Paper D can be interpreted as a *quasi-minimization* procedure for discrete time LSSs with constrained switching since it restricts an unreachable and/or unobservable state space representation of a discrete time LSS to a reachable and/or observable one. Whereas the method in Paper C should only be interpreted as a model reduction method which preserves the output behavior at the end of allowed switching sequences. The reachability or observability properties of the reduced models acquired by the procedure in paper C are not clear.

To illustrate the discussion above, we will briefly sketch a case when the results of paper D could be useful for verification and control design for safety properties. In [57] one considers LSSs with discrete outputs. One then would like to verify the discrete outputs generated by the LSS satisfy a certain temporal logic formula (this results in a restriction in the allowed switching sequence set). In addition, one is also interested in finding a controller, such that the symbolic output generated by the closed loop system satisfies the temporal logic formula. The discrete outputs correspond to a polyhedral partition of the state-space: An output symbol is generated if the current state belongs to the corresponding polyhedron. Such modeling can be represented by the following piecewise-linear hybrid system H [58, 59],

$$H \begin{cases} \Sigma \begin{cases} x(t+1) = A_{q(t)}x(t) + B_{q(t)}u(t) \text{ and } x(0) = x_0 \\ y(t) = C_{q(t)}x(t) \end{cases} \\ \phi \begin{cases} q(t) = \phi(y_{t-1}, q_{t-1}) \text{ and } q(0) = q_0 \end{cases} \end{cases} \tag{31}$$

where $q(t) \in Q = \{1, \ldots, D\}$ is the switching signal, A_q, B_q, C_q are matrices of suitable size and $\phi : Q \times \mathbb{R}^p \to Q$ is a discrete-state transition function which defines the conditions for a change in the discrete mode. The map ϕ is defined by polyhedral sets, i.e. $\{y \mid \phi(y, q_1) = q_2\}$ is a polyhedral set for all $q_1, q_2 \in Q$. Such systems arise either by piecewise-linear modeling of a complex plant, or by explicit modeling of a switching controller. It is easy to see that (31) can be viewed as a feedback interconnection of a discrete time LSS Σ (first two lines of (31)) with the discrete time controller ϕ (last line of (31)) whose next operating discrete state is specified by the past discrete states and past outputs.

As a consequence, the solutions of H corresponds to the solutions $\{q_t, x_t, u_t, y_t\}_{t=0}^{\infty}$ of the discrete time LSS Σ with $q_t = \phi(\{y_s, q_s\}_{s=0}^{t-1})$. A simple example of such a system is $q_t = \phi(y_{t-1})$, $t > 0$, and q_0 is fixed, where ϕ is a piecewise affine map. Often, it is desired to verify if the system is *safe*, i.e., that the sequences of discrete modes generated by the system H belong to a certain set of safe sequences L for all (some) continuous input signals. Consider now another piecewise-affine hybrid system $\bar{H}$ obtained by interconnecting the discrete event generator ϕ with a reduced order discrete time LSS $\bar{\Sigma}$, such that the input-output behavior of $\bar{\Sigma}$ coincides with that of Σ for all the switching sequences from L. If L is prefix closed, then H is safe if and only if $\bar{H}$ is safe, and hence it is sufficient to perform safety analysis on $\bar{H}$. Since the number of continuous-states of $\bar{H}$ is smaller than that of H, it is easier to perform verification for $\bar{H}$ than for the original model. Note that verification of piecewise-affine hybrid systems has high (in certain cases exponential) computational complexity, [60, 61]. Likewise, assume that it is desired to design a control law for H which ensures that the switching signal generated by the closed-loop system belongs to a certain prefix closed set L. Such problems arise in various settings for hybrid systems [8]. Since the next discrete state generated by the controller ϕ only depends on the output and the current discrete state, $\bar{H}$ can be used instead of H for synthesizing the controller ϕ. The computational burden for control synthesis can hence decrease, since the output depends on the state and $\bar{H}$ has a smaller number of states than H.

Given the above discussion, still it should be noted that it remains as future work to demonstrate the usefulness of our approach by investigating several practical case studies.

4.4 Outline of Paper E

In paper E, an entirely different class of systems, namely LPV systems are considered. Paper E, provides a solution to the model reduction problem for LPV systems by moment matching. The tools which have been used in paper E stem from the realization theory of LPV-SS representations [5, 26]. As hinted in Subsections 2.2 and 2.3 the Markov parameters and realization theory of LPV-SS representations are closely related with the counterpart concepts for LSSs. In fact, we use the relationship between

LPV-SS representations and linear switched systems derived in [5] to adapt the tools developed for model reduction of LSSs in the Papers A - D, to LPV-SS representations. Namely, from [5] it follows that there exists a linear switched SS representation with $n_p + 1$ local modes associated with an LPV-SS representation with scheduling space dimension of n_p. Moreover, this switched SS representation has exactly the same sub-Markov parameters (the coefficients uniquely determining the input-output behavior of a system) as the corresponding LPV-SS representation. Hence, in paper E, by making use of this connection between linear switched SS and LPV-SS representations, the moment matching method used in Paper A for linear switched systems is adapted to LPV-SS models.

5 Conclusion

Four methods for model reduction of LSSs suitable for different purposes and one method for model reduction of LPV-SS representations are given. One of the proposed model reduction methods given for LSSs is analogous to the moment matching techniques used for LTI systems. The other method which is based on nice selections, allows model reduction specific to certain switching sequences and a priori choice of the order of reduced model. The rest of the two methods given for LSSs, considers the problem of model order reduction, in the case of restricted switching. One of these methods allows for reducing the order possibly more than the other, but lacks some system theoretical interpretations which the other holds. The second method given for model reduction of LSSs with constrained switching can actually be considered as a reachability and / or observability reduction procedure for LSSs with restricted discrete dynamics. Finally, a similar method relying on a moment matching technique is given for model reduction of LPV-SS representations.

The following important aspects of the problems summarized above have not been investigated and remain as topics of future research: The numerical aspects of the algorithms stated in the contributions of the thesis can be improved and detailed complexity analyses can be made. Practical case studies motivated by real world examples can be made for each method. The performance of the model reduction methods can be assessed when they are used for control synthesis. In addition, the reduction methods given for LSSs with arbitrary switching can be extended to hybrid systems with reset maps. Finally, the question of minimality for LSSs with restricted discrete dynamics can be investigated, with the help of the reachability / observability reduction method given in the thesis.

References

[1] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid Dynamical Systems: Modeling, Stability, and Robustness*. New Jersey, NJ: Princeton University Press, 2012.

[2] D. Liberzon, *Switching in Systems and Control*. Boston, MA: Birkhäuser, 2003.

[3] Z. Sun and S. Ge, *Switched Linear Systems: Control and Design*. London: Springer-Verlag, 2005.

[4] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*. Germany: Springer, 2010.

[5] M. Petreczky and G. Mercere, "Affine LPV systems: Realization theory, input-output equations and relationship with linear switched systems," in *Proc. of the IEEE 51st Conference on Decision and Control (CDC)*, Maui, HI, USA, Dec. 2012, pp. 4511 – 4516.

[6] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: SIAM, 2005.

[7] A. V. Savkin and R. J. Evans, "A new approach to robust control of hybrid systems over infinite time," *IEEE Transactions on Automatic Control*, vol. 43, no. 9, pp. 1292 – 1296, 1998.

[8] P. Tabuada, *Verification and Control of Hybrid Systems: A Symbolic Approach*. Springer-Verlag, 2009.

[9] W. M. Wonham, "Supervisory control of discrete-event systems," lecture notes, `http://www.control.utoronto.ca/~wonham`.

[10] O. Kupferman, P. Madhusudan, and P. S. Thiagarajan, "Open systems in reactive environments: Control and synthesis," in *CONCUR'00*, 2000.

[11] E. Grädel, W. Thomas, and T. Wilke, *Automata, Logic and Infinite Games*. Springer, 2002, vol. LNCS 2500.

[12] M. Petreczky, P. Collins, D. A. van Beek, J. H. van Schuppen, and J. E. Rooda, "Sampled-data control of hybrid systems with discrete inputs and outputs," in *Proc. of 3rd IFAC Conference on Analysis and Design of Hybrid Systems (ADHS09)*, 2009, pp. 334–339.

[13] A. Arnold, A. Vincent, and I. Walukiewicz, "Games for synthesis of controllers with partial observation," *Theoretical Computer Science*, vol. 303, no. 1, pp. 7–34, jun 2003. [Online]. Available: http://www.sciencedirect.com/science/article/B6V1G-487DRHS-1/1/de6f26f5c11d7a24bede23fc09baa6df

[14] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 55, no. 10, pp. 2321–2336, 2010. [Online]. Available: http://dx.doi.org/10.1109/TAC.2010.2046044

[15] A. Marcos and G. J. Balas, "Development of Linear-Parameter-Varying models for aircraft," *Journal of Guidance, Control and Dynamics*, vol. 27(2), pp. 218 – 228, 2004.

[16] V. Verdult, M. Lovera, and M. Verhaegen, "Identification of Linear Parameter-Varying state space models with application to helicopter rotor dynamics," *Int. Journal of Control*, vol. 77(13), pp. 1149 – 1159, 2004.

[17] F. D. Bianchi, H. D. Battista, and R. J. Mantz, *Wind Turbine Control Systems; Principles, modeling and gain scheduling design*. Heidelberg: Springer, 2007.

[18] R. Tóth and D. Fodor, "Speed sensorless mixed sensitivity linear parameter variant H-infinity control of the induction motor," *Journal of Electrical Engineering*, vol. 6(4), p. 12/1–6, 2006.

[19] M. Dettori and C. W. Scherer, "LPV design for a CD player: An experimental evaluation of performance," in *Proc. of the 40th IEEE Conference on Decision and Control (CDC)*, 2001, pp. 4711 – 4716.

[20] C. Hoffmann and H. Werner, "Complexity of implementation and synthesis in Linear Parameter-Varying Control," in *19th IFAC World Congress*, Cape Town, RSA, aug 2014, pp. 11 749 – 11 760.

[21] E. D. Sontag, *Mathematical Control Theory Deterministic Finite Dimensional Systems*. New York, NY: Springer, 1990.

[22] L. M. Silverman, "Realization of linear dynamical systems," *IEEE Transactions on Automatic Control*, vol. 16, pp. 554 – 567, Dec. 1971.

[23] M. Petreczky, "Realization theory for linear and bilinear switched systems: formal power series approach - part i: realization theory of linear switched systems," *ESAIM Control, Optimization and Caluculus of Variations*, vol. 17, pp. 410–445, 2011.

[24] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*. Reading, MA: Addison - Wesley, 2006.

[25] M. Petreczky, L. Bako, and J. H. van Schuppen, "Realization theory of discrete-time linear switched systems," *Automatica*, vol. 49, no. 11, pp. 3337 – 3344, 2013.

[26] R. Tóth, H. S. Abbas, and H. Werner, "On the state-space realization of LPV input-output models: Practical approaches," *IEEE Trans. Contr. Syst. Technol.*, vol. 20, pp. 139–153, Jan. 2012.

[27] A. Birouche, J. Guilet, B. Mourillon, and M. Basset, "Gramian based approach to model order-reduction for discrete-time switched linear systems," in *Proc. of the 18th Mediterranean Conference on Control and Automation*, Marrakech, Morocco, jun 2010, pp. 1224 – 1229.

[28] L. Zhang, P. Shi, E. Boukas, and C. Wang, "Model reduction for uncertain switched linear discrete-time systems," *Automatica*, vol. 44, no. 11, pp. 2944 – 2949, 2008.

[29] E. Mazzi, A. S. Vincentelli, A. Balluchi, and A. Bicchi, "Hybrid system model reduction," in *Proc. of the 47th IEEE Conference on Decision and Control*, Cancun, dec 2008, pp. 227 – 232.

[30] Y. Chahlaoui, "Model reduction of hybrid switched systems," in *Proceeding of the 4th Conference on Trends in Applied Mathematics in Tunisia, Algeria and Morocco, May 4-8, Kenitra, Morocco*, 2009.

[31] L. C. G. J. M. Habets and J. H. van Schuppen, "Reduction of affine systems on polytopes," in *Proc. of the 15th International Symposium on Mathematical Theory of Networks and Systems*, 2002.

[32] L. Zhang, E. Boukas, and P. Shi, "Mu-dependent model reduction for uncertain discrete-time switched linear systems with average dwell time," *International Journal of Control*, vol. 82, no. 2, pp. 378– 388, 2009.

[33] L. Zhang and P. Shi, "Model reduction for switched lpv systems with average dwell time," *IEEE Transactions on Automatic Control*, vol. 53, pp. 2443–2448, 2008.

[34] H. Gao, J. Lam, and C. Wang, "Model simplification for switched hybrid systems," *Systems & Control Letters*, vol. 55, pp. 1015–1021, 2006.

[35] G. Kotsalis, A. Megretski, and M. A. Dahleh, "Balanced truncation for a class of stochastic jump linear systems and model reduction of hidden Markov models," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2543 – 2557, 2008.

[36] G. Kotsalis and A. Rantzer, "Balanced truncation for discrete-time Markov jump linear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2606 – 2611, 2010.

[37] N. Monshizadeh, H. Trentelman, and M. Camlibel, "A simultaneous balanced truncation approach to model reduction of switched linear systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3118 – 3131, 2012.

[38] H. R. Shaker and R. Wisniewski, "Generalized gramian framework for model/controller order reduction of switched systems," *International Journal of Systems Science*, vol. 42, no. 8, pp. 1277 – 1291, 2011.

[39] H. Shaker and R. Wisniewski, "Generalized gramian framework for model/controller order reduction of switched systems," *International Journal of System Science*, vol. 42, pp. 1277—-1291, Aug. 2011.

[40] M. Petreczky, R. Wisniewski, and J. Leth, "Balanced truncation for linear switched systems," *Nonlinear Analysis: Hybrid Systems*, vol. 10, pp. 4–20, Nov. 2013.

[41] M. Petreczky and R. Peeters, "Spaces of nonlinear and hybrid systems representable by recognizable formal power series," in *Proc. 19th International Symposium on Mathematical Theory of Networks and Systems*, Budapest, Hungary, Jul. 2010, pp. 1051–1058.

[42] M. Petreczky and R. Vidal, "Metrics and topology for nonlinear and hybrid systems," *Hybrid Systems: Computation and Control (HSCC) 2007, LNCS*, vol. 4416, pp. 459–472, Apr. 2007.

[43] M. Petreczky and J. H. van Schuppen, "Partial-realization theory for linear switched systems - a formal power series approach," *Automatica*, vol. 47, pp. 2177 – 2184, Oct. 2011.

[44] S. Gugercin, "Projection methods for model reduction of large-scale dynamical systems," Ph.D. dissertation, Rice Univ., Houston, TX, May 2003. [Online]. Available: http://scholarship.rice.edu/handle/1911/18536

[45] Y. Lin, L. Bao, and Y. Wei, "A model-order reduction method based on krylov subspaces for mimo bilinear dynamical systems," *Journal of Applied Mathematics and Computing*, vol. 25, no. 1-2, pp. 293–304, 2007.

[46] Z. Bai and D. Skoogh, "A projection method for model reduction of bilinear dynamical systems," *Linear Algebra and its Applications*, vol. 415, no. 2 - 3, pp. 406 – 425, 2006.

[47] G. M. Flagg, "Interpolation methods for the model reduction of bilinear systems," Ph.D. dissertation, Virginia Polytechnic Institute, 2012.

[48] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[49] M. Farhood, C. Beck, and G. Dullerud, "On the model reduction of nonstationary lpv systems," in *Proc. of the American Control Conference (ACC)*, Denver, CO, USA, Jun. 2003, pp. 3869 – 3874.

[50] S. D. Hillerin, G. Scorletti, and V. Fromion, "Reduced-complexity controllers for LPV systems: Towards incremental synthesis," in *Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Orlando, FL, USA, Dec. 2011, pp. 3404 – 3409.

[51] F. Adegas, I. Sonderby, M. Hansen, and J. Stoustrup, "Reduced-order LPV model of flexible wind turbines from high fidelity aeroelastic codes," in *Proc. of the IEEE International Conference on Control Applications (CCA)*, Hyderabad, Aug. 2013, pp. 424 – 429.

[52] G. Wood, P. Goddard, and K. Glover, "Approximation of linear parameter-varying systems," in *Proc. of the 35th IEEE Conference on Decision and Control*, Kobe, Dec. 1996, pp. 406 – 411.

[53] Widowati, R. Bambang, R. Saragih, and S. M. Nababan, "Model reduction for unstable LPV systems based on coprime factorizations and singular perturbation," in *Proc. of the 5th Asian Control Conference*, Melbourne, Jul. 2004, pp. 963 – 970.

[54] M. M. Siraj, R. Tóth, and S. Weiland, "Joint order and dependency reduction for LPV state-space models," in *Proc. of the 51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 6291–6296.

[55] Z. Sun, "Stabilizing switching design for switched linear systems: A state-feedback path-wise switching approach," *Automatica*, vol. 45, no. 7, pp. 1708–1714, 2009.

[56] J. Magee and J. Kramer, *Concurrency: State Models and Java Programs*. Hoboken, NJ: Wiley, 2006.

[57] E. A. Gol, X. C. Ding, M. Lazar, and C. Belta, "Finite bisimulations for switched linear systems," vol. 59, no. 12, pp. 3122 – 3134, 2014.

[58] A. Bemporad, G. Ferrari-Trecate, and M. Morari, "Observability and controllability of piecewise affine and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1864–1876, 2000.

[59] A. van der Schaft and H. Schumacher, *An Introduction to Hybrid Dynamical Systems*. Springer-Verlag London, 2000.

[60] G. Frehse, "Phaver: algorithmic verification of hybrid systems past hytech," *International Journal on Software Tools for Technology Transfer*, vol. 10, no. 3, pp. 263–279, 2008. [Online]. Available: http://dx.doi.org/10.1007/s10009-007-0062-x

[61] B. Yordanov and C. Belta, "Formal analysis of discrete-time piecewise affine systems," *Automatic Control, IEEE Transactions on*, vol. 55, no. 12, pp. 2834–2840, Dec 2010.

Part II

Contributions

Paper A

Model Reduction by Moment Matching for Linear Switched Systems

Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth

The paper has been published in the
Proceedings of the American Control Conference, pp. 3942–3947, 2014.

Paper B

Model Reduction by Nice Selections for Linear Switched Systems

Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth

The paper is conditionally accepted for publication in the
IEEE Transactions on Automatic Control, 2015.

Paper C

Model Reduction of Linear Switched Systems by Restricting Discrete Dynamics

Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth

The paper has been published in the
Proceedings of the 53rd IEEE Conference on Decision and Control, pp. 4422—4427, 2014.

Paper D

Reachability and Observability Reduction for Linear Switched Systems with Constrained Switching

Mert Baştuğ, Mihály Petreczky, Rafael Wisniewski and John Leth

The paper is submitted for publication in
Automatica, 2015.

The layout has been revised, and small editorial changes have been made. Content relavant changes, if any, are marked with explicit footnotes.

Paper E

Moment Matching Based Model Reduction for LPV State-Space Models

Mert Baştuğ, Mihály Petreczky, Roland Tóth, Rafael Wisniewski, John Leth and Denis Efimov

The paper is accepted for publication in the
Proceedings of the 54th IEEE Conference on Decision and Control, 2015.

ISSN (online): 2246-1248
ISBN (online): 978-87-7112-481-1

AALBORG **UNIVERSITY PRESS**