Stability of discrete-time switching systems with constrained switching sequences.

Matthew Philippe\textsuperscript{a,1}, Ray Essick\textsuperscript{b,2}, Geir Dullerud\textsuperscript{b,2}, Raphaël M. Jungers \textsuperscript{a,1}.

\textsuperscript{a}Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Department of Mathematical Engineering (ICTEAM/INMA) at the Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

\textsuperscript{b}Department of Mechanical Science and Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

Abstract

We present a framework for the stability analysis of a general class of discrete-time linear switching systems and algorithms for the approximation of their exponential growth rate. The framework is applicable to any system for which the switching sequence is constrained to be in a regular language. These systems are referred to as constrained switching systems. Arbitrary switching systems are a special case of constrained switching systems.

We introduce the algebraic concept of multinorms, which characterizes the stability of a constrained switching system. Then, building upon this tool, we develop the first arbitrary accurate approximation schemes for estimating the constrained joint spectral radius $\hat{\rho}$ - that is, the exponential growth rate - of a system. For a given bound $r > 0$, the algorithms provide an estimate of $\hat{\rho}$ within the range $[\hat{\rho}, (1 + r)\hat{\rho}]$. They boil down to solving a convex optimization program involving LMIs, for which the time-complexity depend on the desired relative error $r > 0$.

Key words: Switching systems, Stability

1 Introduction

A switching system is a dynamical system whose state dynamics vary between different operating modes. These modes can be represented by a collection of differential or difference equations. A discrete-time linear switching system is a switching system where each mode is represented by a linear difference equation. It takes the form

$$x_{t+1} = A_{\sigma(t)}x_t,$$

with $x_0 \in \mathbb{R}^n$, $A_{\sigma(t)} \in \Sigma \subset \mathbb{R}^{n \times n}$, where $\Sigma$ is a finite set of $N$ matrices and the mode at time $t$, $\sigma(t) \in \{1, \ldots, N\}$, corresponds to a particular matrix in $\Sigma$.

These systems find application in many theoretical and engineering related domains. Examples include the study of viral mutations [1], of congestion control in computer networks [2], the modelling and analysis of networked and delayed control systems [3], and theoretical implications in the broad field of theoretical computer sciences [4]. More generally, classical methods for simulating more complex systems (such as Hybrid or Cyber-Physical Systems) often boil down to the analysis of such a switching system.

The sequence of modes $\sigma(0), \sigma(1), \ldots$, is called the switching sequence of the system. These sequences may be subject to a set logical rules/constraints (e.g. "mode $a$ cannot be followed by mode $b$"). In the most general case, these rules can be represented using concepts of language theory. A convenient way to represent the constraints is by using a finite automaton. It takes the form of a (strongly) connected, directed, labelled graph $G(V, E)$, on a set of nodes $V$ and edges $E$. Each edge
e = (s, d, ς) ∈ E between two nodes s and d carries a label ς ∈ {1, . . . , N}. A sequence ς(0), ς(1), . . . then satisfies the constraints of the graph G if there exists a path in G such that, ∀ t ≥ 0, the label on the (t + 1)th edge in the path is ς(t). Such sequence is said to be admissible.

The system with a set of matrices Σ having its switching sequences constrained by the graph G will be denoted S(G, Σ).

A system S(G, Σ) is said to be (absolutely) stable if for any admissible switching sequence and any x0 ∈ ℝn,

\[ \lim_{t→∞} A_{ς(t)} \cdots A_{ς(0)} x_0 = 0, \]

or equivalently any admissible infinite product of matrices converges to the zero matrix. The questions related to stability of switching systems (theoretical and algorithmic) carry many be challenges. Surveys on the subject can be found in [5–7].

Many studies have had a particular focus on arbitrary switching between modes (e.g., [4,8–10]), in which ς(t) can take any value in \{1, . . . , N\} at any time t. These systems are a special case of constrained switching systems. For example, an arbitrary switching system can be formed with, the graph G of Figure 1, having one node and one self-loop for each mode.

![Fig. 1. Automaton generating arbitrary switching sequences.](image)

In the arbitrary switching case, a quantity known as the joint spectral radius (JSR), introduced by Rota and Strang [11], has been shown to be of particular interest as it corresponds to the worst case growth rate of the system. The concept was later generalized to systems with constrained switching sequences by Dai [12].

**Definition 1 (Dai [12])** The constrained joint spectral radius (CJSR) of a system S(G(V, E), Σ) is defined as

\[ \hat{ρ}(S) = \lim_{k→∞} \hat{ρ}_k(S), \]

where, letting Ap denote the matrix product associated with a path p,

\[ \hat{ρ}_k(S) = \max \left\{ \|A_p\|^{1/k} : p \text{ is a path of length } k \text{ in } G \right\} \]

is the maximum growth rate up to time k.

**Theorem 1.1 (Dai [12], Corollary 2.8)** A constrained switching system S(G, Σ) is stable if and only if its constrained joint spectral radius satisfies

\[ \hat{ρ}(S) < 1. \]

Any stable system is exponentially stable. For any \( ε > 0 \), there exists a constant C > 0 such that for all trajectories of S(G, Σ),

\[ |x_t| ≤ C(\hat{ρ}(S) + ε)^t|x_0|. \]

If one is able to compute the joint-spectral radius of a given system, then he is able to decide if it is stable (which will also imply exponential stability) or not. There has been a lot of research effort towards that end in the case of arbitrary switching (see for example [4,9,13,14] and references therein.). Due to negative results such that the undecidability of \( \hat{ρ}(S) ≤ 1 \), the methods often focus on approximating the JSR of a system. One common way to do so is to approximate the invariant norm [4] (see Proposition 2.1 below) of the arbitrary switching system. This is done, for example, by computing the quadratic or sum-of-square norm having the highest contractivity factor in time [10,13,15]. These techniques can be used to provide an arbitrary accurate relative approximation of the JSR of a system.

Regarding constrained switching systems, the methods proposed so far focus on Lyapunov arguments without reasoning about joint-spectral radiiuses. Such methods can be found in works including those of Bliman and Ferrari-Trecate [16], Lee and Dullerud [17,18]; Lee and Khargonekar [19,20]; Essick et al. [21], and many others. With the exception of [16] these works tackle several control problems for switching systems (stabilizing controller design, disturbance attenuation, switching sequence design, etc.), showing the importance of these systems in engineering. These approaches rely on sufficient conditions for stability established by the existence of multiple quadratic Lyapunov functions.

Multiple Lyapunov functions, in general, have proven to be very effective stability analysis tools for switching and hybrid systems [6,22]. Multiple quadratic Lyapunov functions, however, usually only provide sufficient conditions for stability [23,24]. In some cases, they can be embedded within a hierarchy of multiple quadratic Lyapunov functions [16,17], to asymptotically provide a necessary condition for stability as well.

However, to the best of our knowledge, tools for the approximation of the constrained joint spectral radius have not been investigate yet.

One of the main inspiration in our work can be found in [9]. The focus of the paper is on systems S(G, Σ)
with $G$ allowing for arbitrary switching. Such graphs are called path-complete graphs. The results show a link between the JSR of an arbitrary switching system and the performances of the multiple Lyapunov function-based methods presented above, allowing to see these methods as JSR approximation schemes. The author left the generalization of these results to general constrained switching systems as an open question, which we tackle in our work.

The plan of the paper is as follows. Section 2 introduces the algebraic concept of multinorm as a characterization of stability for constrained switching systems. Their structure is linked to the marginal stability of a system.

In Section 3, we focus on the more algorithmic question of the approximation - in finite time and with arbitrary accuracy - of the constrained joint spectral radius of a system $S(G, \Sigma)$.

In Section 4, we present a system inspired from an application, which is a controlled system experiencing dropouts in its control feedback. We apply and compare the different methods introduced in section 2 in terms of time-efficiency and approximation accuracy to this system.

**Notations**

By $|·|$ we denote a vector norm and by $∥·∥$ an induced matrix norm.

An automaton $G(V, E)$ is seen as a labeled, directed graph; $N_V$ and $N_E$ denote, respectively, the number of nodes and edges of a graph $G(V, E)$. Paths in $G$ are written $p \subseteq E$. A cycle $c$ on node $v \in V$ is simple if there is no partitions $c = [c_1, c_2] \subseteq E$ of the cycle with $c_1$ and $c_2$ being cycles on $v$.

The matrix $A^T \in \mathbb{R}^{m \times n}$ is the transpose of $A \in \mathbb{R}^{n \times m}$. For $Q \in \mathbb{R}^{n \times n}$ symmetric, we write $Q > 0$ if it is positive definite, i.e., $\forall x \in \mathbb{R}^n$, $x \neq 0$, $x^T Q x > 0$. Given a set of matrices $\Sigma$ and a constant $\alpha \in \mathbb{R}$, $\alpha \Sigma$ is the set of scaled matrices $\{\alpha A, A \in \Sigma\}$. Given a system $S(G, \Sigma)$ and a path $p \subseteq E$ of length $t$ carrying a sequence of successive labels $\sigma(0), \ldots, \sigma(t - 1)$, the matrix product associated with $p$ is

$$A_p = A_{\sigma(t-1)} \cdots A_{\sigma(0)},$$

with $A_{\sigma(i)} \in \Sigma$, $0 \leq i \leq t - 1$. For a path $p$ of length 0 (empty path), we define $A_p = I$, the identify matrix of appropriate dimensions.

2 Lyapunov functions for constrained switching systems

The stability of arbitrary switching systems is linked to the existence of a contractive norm serving as a Lyapunov function. The result is a direct corollary of the following proposition:

**Proposition 2.1** (e.g. [4], Proposition 1.4) The joint spectral radius of a set of matrices $\Sigma$ is given by

$$\hat{\rho}(\Sigma) = \inf_{\|A\|} \max_{A \in \Sigma} \|A\|,$$

where the infimum is taken over all induced matrix norms in $\mathbb{R}^{n \times n}$.

Equation (2) can be rewritten as follows, developing the induced matrix norm,

$$\hat{\rho}(\Sigma) = \inf_{\|A\|} \left( \min_{\gamma \in \mathbb{R}} \left\{ \gamma : |Ax| \leq \gamma |x|, \forall x \in \mathbb{R}^n, \forall A \in \Sigma \right\} \right).$$

As a consequence, when $\hat{\rho}(\Sigma) < 1$, there is always a norm $|·|^{*}$ such that $|Ax|^{*} < |x|^{*}$. This norm is said to be contractive for the system and serves as a Lyapunov function.

It is however straightforward to find examples of stable constrained switching systems for which there are no contractive norms.

**Example 1** We consider a scalar system with two modes:

$$A_1 = 2, A_2 = 1/4.$$

An arbitrary switching system built on $\Sigma = \{A_1, A_2\}$ is unstable with a joint spectral radius of $\hat{\rho}(\Sigma) = 2$, and consequently has no contractive norm. Consider now the switching graph $G$ of Figure 2. It becomes clear that the system $S(G, \Sigma)$ is stable with $\hat{\rho}(S) = 1/\sqrt{2}$. Still, there cannot be any contractive norm for the system...

The switching automaton $G$ clearly plays an important role in the (un)stability of $S(G, \Sigma)$. In what follows, we introduce an algebraic concept called multinorm, integrating the structure of the automaton to form a multiple Lyapunov function. Similar ideas have appeared in the literature (see [9, 17, 22, 24, 25], for previous works making use of multiple Lyapunov functions), but often with underlying algorithmically motivated assumptions about the nature - often quadratic - of the functions.

**Definition 2** (Multinorm) A multinorm $H$ for a system $S(G(V, E), \Sigma)$ is a set of $N_V$ norms, $H = \{|·|_1, \ldots, |·|_{N_V}\}$. The value $\gamma^*$ of a multinorm is
defined as
\[
\gamma^*(\mathcal{H}) = \min_{\gamma \in \mathbb{R}} \{ \gamma : |A_{\sigma}x|_j \leq \gamma|x|, \forall x \in \mathbb{R}^n, \forall (v_i, v_j, \sigma) \in E \}.
\]  
(5)

The value of a multinorm plays a role similar to $\gamma$ in Equation (3), that is a contraction factor applied at each time step. This observation leads to the following generalization of Proposition 2.1.

**Proposition 2.2** The constrained joint spectral radius (Definition 1) of a system $S(G, \Sigma)$ satisfies
\[
\hat{\rho}(S) = \inf \{ \gamma^*(\mathcal{H}) : \mathcal{H} \text{ is multinorm for } S(G, \Sigma) \}. 
\]  
(6)

**PROOF.** We first show that the value of any multinorm for system is an upper bound of its CJSR. Let $p \in E$ be a path of length $k \geq 1$, with origin and destination nodes $v_i$ and $v_j$ (resp.).

If $\mathcal{H}$ is a multinorm of value $\gamma$, then by definition
\[
|A_p x|_j \leq \gamma^k |x|.
\]  
(7)

For any norm $|\cdot|$, by equivalence of norms in $\mathbb{R}^n$, there exists $0 < \alpha < \beta$ such that the inequalities
\[
\forall x \in \mathbb{R}^n, \alpha|x| \leq |x| \leq \beta|x|
\]
holds for all norms $|\cdot|_i$ in $\mathcal{H}$. Therefore, we have
\[
\|A_p\| = \max_{|x|=1} \frac{|A_p x|}{|x|} \leq \frac{\beta}{\alpha} \max_{|x|=1} \frac{|A_p x|_j}{|x|} \leq \frac{\beta}{\alpha} \gamma^k,
\]
which allows us to conclude. Indeed, taking the limit $k \to \infty$, we obtain that the kth root of any products corresponding to any path is lower than $\gamma$, and by Definition 1, $\hat{\rho} \leq \gamma$ as well.

In the second part of the proof we show that for any $\epsilon > 0$ there exists a multinorm of value at most $(\hat{\rho}(S) + \epsilon)$.

Consider the scaled set of matrices
\[
\Sigma' = \{ A'_i = A_i/(\hat{\rho} + \epsilon), i = 1, \ldots, N \}.
\]

The CJSR of $S(G, \Sigma')$ is an homogeneous function of $\Sigma$. Thus the system $S(G, \Sigma')$ is stable $(\hat{\rho}(\Sigma') < 1)$.

We now define, at each node $v_i \in V$, the following norm:
\[
|x|_i := \sup \{|A'_p x| : p \text{ is a path with origin } v_i\},
\]  
(8)

where $|\cdot|$ is the euclidean norm. These functions are well-defined (by stability of $S(G, \Sigma')$), sub-additive, homogeneous (due to the use of the euclidean norm $|\cdot|$ in their definition), and positive definite (for paths of length $t = 0$, we have $|x|_i \geq |x|$).

Moreover, for any edge $e = (v_i, v_j, \sigma) \in E$, and all $x \in \mathbb{R}^n$, we have
\[
|x|_i = \sup \{|A'_p x| : p \text{ is a path with origin } v_i\},
\]
\[
\leq \sup \{|A'_q A'_e x| : q \text{ is a path with origin } v_j\},
\]
\[
= |A'_e x|_j.
\]

Since $A'_e = A_e/(\hat{\rho} + \epsilon)$, we conclude that for any edge $(v_i, v_j, \sigma)$
\[
|A_{\sigma} x|_j \leq (\hat{\rho} + \epsilon)|x|_i,
\]  
(10)
and therefore the value of the multinorm is lower than $\hat{\rho}(S) + \epsilon$.

As a direct consequence, multinorms characterize the stability of constrained switching systems through their link with the CJSR.

**Theorem 2.3** A constrained switching system $S(G, \Sigma)$ is stable if and only if it admits a multinorm $\mathcal{H}$ with value $\gamma(\mathcal{H}) < 1$.

Before illustrating the result, we introduce these definitions for later use.

**Definition 3** Given a system $S(G, \Sigma)$,
\begin{enumerate}
\item a Lyapunov multinorm is a multinorm $\mathcal{H}$ of value $\gamma(\mathcal{H}) < 1$;
\item an extremal multinorm is a multinorm $\mathcal{H}$ of value $\gamma(\mathcal{H}) = \hat{\rho}(S)$.
\end{enumerate}

**Example 2** Let $v_1$ and $v_2$ be respectively the nodes on the left and right of the automaton on Figure 2. The reader can then verify that the multinorm
\[
\mathcal{H} = \{|x|_1, |x|_2\} = \left\{ 2 \sqrt{2} \text{ abs}(x), \text{ abs}(x) \right\},
\]
where abs$(x)$ is the absolute value of $x$, is an extremal Lyapunov multinorm for the scalar switching system of example 1.

### 2.1 Extremal multinorms and marginal stability

In this subsection we give conditions for the existence of extremal multinorm. The notion of extremality will become more relevant in Section 3, where we approximate the CJSR by approximating these multinorms.
In the following, we will focus on systems having \( \rho(S) = 1 \). Such a system can always be obtained by scaling all the matrices of a given system by the inverse of its CJSR. While a unitary CJSR implies instability, it does not allow to decide the (un)boundness of the trajectories of a system, which is its marginal (un)stability.

**Theorem 2.4** A system \( S(G, \Sigma) \) with \( \rho(S) = 1 \) admits an extremal multinorm if and only if it is marginally stable, i.e. there exists \( K \geq 0 \) such that

\[
|A_p x| \leq K |x|,
\]

for all \( x \in \mathbb{R}^n \) and all paths \( p \) of any length in \( G \).

**PROOF.** We start with the only-if part. Let \( \mathcal{H} = \{ |.|, 1 \leq i \leq N_V \} \) be an extremal multinorm for the system. The boundedness of the system’s trajectories is proved by an application of Definition 2. Indeed, we have the following relationship between the norms: for any any pair of nodes \( v_i, v_j \) and any paths \( p \subset E \) of any length from \( v_i \) to \( v_j \),

\[
|A_p x|_j \leq |x|_i.
\]

We now show that boundedness implies the existence of an extremal multinorm. Given any vector norm \( |.| \), we define, for each node \( v_i \in V \), the following norm:

\[
|x|^* = \sup\{|A_p x| : p \subset E, p \text{ starts at } v_i \}.
\]

One can verify that \( |x|^* \) is indeed a norm since its definition implies that \( |x|^* \geq |x| \). Also, one can easily check that for any edges \( (v_i, v_j, \sigma) \in E \), \( |A_{\sigma} x|^* \leq |x|^* \), which points out that the multinorm \( \{ |.|^*, 1 \leq i \leq N_V \} \) is extremal by construction.

Deciding whether a system is marginally stable remains an open question [4]. In [26], the authors give a sufficient, decidable condition for marginal stability. The condition generalizes the notion of the irreducibility of a matrix set to the context of constrained switching systems.

**Definition 4** A system \( S(G(V, E), \Sigma) \) is said to be irreducible if

1. **between any two nodes** \( v_i \) and \( v_j \) in \( V \), there exists at least one path \( p \subset E \) such that
   \[
   A_p \neq 0.
   \]

2. **for any node** \( v_i \) in \( G \) and any non-trivial linear subspace \( X \subset \mathbb{R}^n \) there exists at least one cycle \( c \subset E \), cycling on \( v_i \), such that
   \[
   A, X \not\subset X.
   \]

With the following theorem, we show that an irreducible system has an extremal multinorm. Moreover, this multinorm satisfies a stronger property than extremality alone:

**Theorem 2.5** If the system \( S(G, \Sigma) \) with \( \rho(S) = 1 \) is irreducible, then it admits a multinorm \( H = \{ |.|, i = 1 \ldots N_V \} \) such that for all \( x \in \mathbb{R}^n \), for any connected pair of node \( v_1, v_2 \in V \), there is an edge \( (v_1, v_2, \sigma(x)) \in E \) such that

\[
|A_{\sigma(x)} x|_2 = \rho(S) |x|_1.
\]

**PROOF.** We provide such a multinorm in a constructive way. Take any vector norm \( |.| \) and define, for each node \( v_i \in V \), the following norm:

\[
|x|^*_i = \limsup_{t \to \infty} \max\{|A_p x| : p \subset E, |p| = t, p \text{ starts at } v_i \}.
\]

A major part of the proof is to show that the above functions are indeed norms. The most crucial part is to prove that they are definite: \( |x|^*_i = 0 \Rightarrow x = 0 \). Assume it is not the case, that there is a node \( v_i \), such that there exists \( x \neq 0 \) with \( |x|^*_i = 0 \).

Define the following linear subspaces at all nodes \( v_j \in V \)

\[
X_j = \{ x : |x|^*_j = 0 \}.
\]

Clearly, these subspaces are invariant under all products associated with cycles on the relevant node. Moreover, for the particular node \( v_i \), this subspace is non-empty. Therefore, if \( X_i \neq \mathbb{R}^n \), the second part of Definition 4 does not hold.

We now assume that \( X_i = \mathbb{R}^n \) and begin reasoning on the other nodes.

If there is a node \( v_j \) with \( X_j = 0 \), then the first part of Definition 4 fails between \( v_i \) and \( v_j \), leading to a contradiction.

If all nodes \( v_j \) are such that \( X_j = \mathbb{R}^n \), then the system is stable, so \( \rho(S) < 1 \), which again leads to a contradiction. Last, if there is at least one node \( v_j \) with \( 0 < \dim(X_j) < n \), then the second part of Definition 4 does not hold, leading once again to a contradiction.

It is interesting to notice that the above result implies that constrained version of irreducibility is sufficient to have marginal stability. This shows the usefulness of multinorms in solving more theoretical questions.

### 3 Approximation algorithms for stability analysis

Given a system \( S(G, \Sigma) \) and a relative error tolerance \( r > 0 \), we wish to obtain an estimate \( \hat{\rho} \) of the CJSR \( \rho(S) \), such that

\[
\hat{\rho}(S) \leq \hat{\rho} \leq (1 + r) \rho(S).
\]
This section, we will provide several approximation algorithms solving the above problem. Moreover, for any value of \( r \), we are able to give algorithmic complexity bounds for obtaining the estimation.

All the methods share the same core mechanism: the approximation of an extremal multinorm by a quadratic multinorm (for which all norms are quadratic norms). This approximation can be expressed as a quasi-convex optimization program, which is solved using a bisection procedure, iteratively checking the feasibility of a set of LMIs.

Theorem 3.1 Consider a system \( S(\mathbf{G}(V,E), \Sigma) \).

The value \( \gamma^* \) such that

\[
\gamma^* = \inf_{Q} \gamma
\]

s.t. \( \forall (v_i, v_j, \sigma) \in E \),

\[ -A^T_\sigma Q_j A_\sigma + \gamma^2 Q_i \preceq 0; \]

\[ \forall i \in \{1, \ldots, |V|\}, Q_i > 0. \]  

satisfies the following inequalities :

\[
\hat{\rho}(\mathbf{G}, \Sigma) \leq \gamma \leq \sqrt{n}(\hat{\rho}(\mathbf{G}, \Sigma)).
\]  

Moreover, the feasibility sub-problem in the program (11) is solved in

\[ O\left(n^{13/2} \cdot N_V^{7/2} \cdot N_E^{3/2}\right) \]

operations.

The rest of this subsection is devoted to the proof of Theorem 3.1.

PROOF.

We will start with proving the inequalities (12), and will discuss the complexity afterwards.

First, we show that for any system \( S(\mathbf{G}, \Sigma) \) and any \( \epsilon > 0 \), there is a quadratic multinorm with a value \( \gamma \) satisfying

\[
\hat{\rho}(S) \leq \gamma \leq \sqrt{n}(\hat{\rho}(S) + \epsilon).
\]

The result is obtained by an application the so-called John’s Ellipsoid Theorem [27] to every norm in a multinorm. Such approach is very close to the ones presented in [13, 15] for the arbitrary switching case. More formally, John’s ellipsoid theorem states that for any norm \( |\cdot| \) of \( \mathbb{R}^n \), there exists a quadratic norm \(|\cdot|_Q : x \rightarrow (x^T Q x)^{1/2} \), with \( Q > 0 \), such that

\[ \forall x \in \mathbb{R}^n : |x|_Q \leq |x| \leq \sqrt{n}|x|_Q. \]

Consider now a multinorm \( \mathcal{H}_Q = \{|\cdot|, i = 1, \ldots, N_V\} \) with value (lower than) \( (\hat{\rho}(S) + \epsilon) \). Such a multinorm exists (Proposition 2.2). By John’s ellipsoid theorem, there exists quadratic norms, forming a quadratic multinorm \( \mathcal{H}_Q = \{|\cdot|_i, i = 1, \ldots, N_V\} \), such that for any edge \((v_i, v_j, \sigma) \in E\), \( \forall x \in \mathbb{R}^n : |A_v x|_Q \leq |A_\sigma x|_j \leq (\hat{\rho}(S) + \epsilon)|x| \leq \sqrt{n}(\hat{\rho}(S) + \epsilon)|x|_Q. \)

Since the above holds for any edge, we can state, at this point, that for any \( \epsilon \), there is a quadratic multinorm of value lesser than \( \sqrt{n}(\hat{\rho}(S) + \epsilon) \).

In order to obtain the inequalities (12), it then suffice to take \( \epsilon \rightarrow 0 \).

The complexity computations are based on a classic result of semi-definite programming, presented in the reference book [28], p.424. The number of variables of the problem is \( n N_V \). The LMIs constraints can be represented by a \( n (N_V + N_E) \) bloc diagonal symmetric matrix with diagonal blocs of size \( n \times n \).

Before going further, it is worth noticing that 1) the above result gives a sufficient condition for a given system to possess a quadratic Lyapunov multinorm (if \( \hat{\rho}(S) < 1/\sqrt{n} \)) and that 2) it gives an algorithm to solve the relative approximation problem with maximum error \( r \geq \sqrt{n} - 1 \).

The methods we present in Subsection 3.2 through Subsection 3.5 present ways to increase the precision of the approximation method by performing an algebraic lifting of the structures (graph and/or matrix set) of a system \( S(\mathbf{G}, \Sigma) \).

3.1 T-Product lift.

This method increases the accuracy of the multinorm approximation scheme at the cost of adding edges of the graph.

Definition 5 (T-product lift) Given a system \( S(\mathbf{G}, \Sigma) \), the T-product lift of \( S \), denoted \( S^T(\mathbf{G}, \Sigma) \), is a constrained switching system on an automaton \( \mathbf{G}'(V', E') \) and a matrix set \( \Sigma' \) defined as follows:

1. The automata have the same sets of nodes \( (V' = V) \). To each path \( p \in \mathbf{G}, |p| = T \), from a node \( v_i \in V \) to a node \( v_j \in V \), is associated an edge \( e = (v_i, v_j, \{\sigma_1, \ldots, \sigma(T)\}) \in E' \). The label of this edge is a concatenation of all labels across the path.
2. The set of matrices \( \Sigma' \) is the set of all product of size \( T \) of matrices in \( \Sigma \).

Given a system \( S(\mathbf{G}, \Sigma) \) describing the evolution of a state \( x, t = \{1, 2, \ldots\} \), the trajectories of \( S^T \) describe the evolution of the state at times \( kT, k = \{1, 2, \ldots\} \).

Example 3 Consider the automaton of Figure 3. The 2-product lift of the graph is presented on Figure 4.
Theorem 3.1 to the system \( \gamma \). The CJSR estimation error bound \( r > 0 \) as a function of \( T \). The maximum relative error, it means that one can pick \( T \) such that \( \rho(S) = \hat{\rho}(S) \). By Theorem 3.1 and Proposition 3.2, we have \( \hat{\rho}(S)^T = \hat{\rho}(S^T) \). The result can be obtained from Definition 1. First, observe that for all \( k \geq 1 \), \( \rho_k(S^T) = \hat{\rho}_k(T)^T \), where \( \hat{\rho}_k(S) \) is the maximum growth rate of \( S \) after \( k \) time-steps. Applying the limit as \( k \) grows to infinity, we obtain the desired result.

\textbf{Theorem 3.3} The optimal value \( \gamma_* \) obtained with the bisection procedure of Theorem 3.1 applied to the system \( S^T(G, \Sigma) \) is such that \( \hat{\rho}(S) \leq \gamma_*^{1/T} \leq n^{1/2T} \hat{\rho}(S) \).

\textbf{PROOF.} By Theorem 3.1 and Proposition 3.2, we have \( \hat{\rho}(S)^T = \hat{\rho}(S^T) \leq \gamma_* \leq n^{1/2} \hat{\rho}(S^T) = n^{1/2} \hat{\rho}(S)^T \).

As the parameter \( T \) increases, Theorem 3.3 guarantees that the estimation of the CJSR becomes tighter and tighter. Regarding the problem of approximation with maximum relative error, it means that one can pick \( T \) as a function of \( r \) to obtain the estimation.

\textbf{Corollary 3.4} For any system \( S(G, \Sigma) \) and relative error bound \( r > 0 \), for any integer \( T \) satisfying

\[ T \geq \log(n)/(2\log(1 + r)) \]

the CJSR estimation \( \gamma_* \) obtained by applying the procedure of Theorem 3.1 to the system \( S^T(G, \Sigma) \) satisfies

\[ \hat{\rho}(S) \leq \gamma_*^{1/T} \leq (1 + r) \hat{\rho}(S) \].

Ending this section, notice that the \( T \)-product lift conserves both the number of nodes \( N \) and the dimension \( n \) of the matrix set. The number of edges, however, grows quite fast with \( T \) (one per path of length \( T \)).

\textbf{3.2 Symmetric algebra and sum-of-squares multinorms.}

In this section, we generalize the tools presented in [15] to constrained switching systems. The main idea lies in refining the approximation of Theorem 3.1 by allowing oneself to use a more flexible approximation tool than the quadratic approximation. Amongst these tools are homogeneous sum-of-squares polynomials. A sum-of-squares polynomial of degree \( 2d \) is, by name, the sum of the squares homogeneous polynomials of degree \( d \).

Before being able to use these tools in our context, the key concept of symmetric algebras needs to be introduced. This theoretical tool defined the \( d \)-lift, a transformation \( x \rightarrow x[d] \in \mathbb{R}^{C(n + d - 1, d)} \), where \( x[d] \) is a vector of monomials of degree \( d \) and \( C(n + d - 1, d) \) is the combinatorial number. For more information, we refer the reader to Annex 5.

\textbf{Lemma 3.5} Given a system \( S(G, \Sigma) \) and an integer \( d \geq 1 \), let \( S[d](G, \Sigma) \) denote the constrained switching system on the graph \( G \) and the set of matrices \( \Sigma[d] \). Then,

\[ \hat{\rho}(S[d]) = \hat{\rho}(S)^d \].

\textbf{PROOF.} The \( d \)-lift has the property \( |x[d]| = |x|^d \) for the euclidean norm. The result follows from the definition of the CJSR.

\textbf{Theorem 3.6} The optimal value \( \gamma_* \) obtained with the bisection procedure of Theorem 3.1 applied to the system \( S[d](G, \Sigma) \) is such that

\[ \hat{\rho}(S) \leq \gamma_*^{1/d} \leq (C(n, d))^{1/2d} \hat{\rho}(S) \].

\textbf{PROOF.} The dimension of the matrix set \( \Sigma[d] \) is \( C(n, d) \).

By Theorem 3.1, we have \( \hat{\rho}(S[d]) \leq \gamma_* \leq (C(n, d))^{1/2} \hat{\rho}(S[d]) \). The result then follows from Lemma 3.5.

The relationship between the above and sum-of-square polynomials is that \( p(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R} \) is a sum-of-squares polynomial of degree \( 2d \) if and only if it can be written as

\[ p(x) = (x[d]^T Q x[d])^{1/2} \]
for \( Q \) positive semi-definite. By approximating an extremal multinorm on \( S^{[d]} \) with a quadratic multinorm, we actually approximate the extremal multinorm of \( S \) with a sum-of-squares multinorm of degree \( 2d \).

In regarding the CJSR approximation problem, notice with a sum-of-squares multinorm of degree 2

\[
S
\]

Ending this section, note that the transformation \( S \) does not affect the graph of the system. It affects the dimension of the matrix set, raising it from \( n \) to \( C(n, d) \).

3.3 Improving accuracy by adding memory to the graph.

Path-dependent Lyapunov functions have been introduced by Lee and Dullerud [17] as tools for the stability analysis and control of discrete-time switching systems. The concept, which follows a similar idea to that of Bliman and Ferrari-Trecate [16], is to build a multiple Lyapunov function that associates a different quadratic form to every switching sequences of a certain length \( M \geq 0 \), with \( M \) being an integer parameter called the memory of the Lyapunov function.

The authors showed that, for any stable switching system, there is a finite \( M \) such as the system admits a Path-dependent Lyapunov function with memory \( M \).

Remark that the CJSR approximation tools presented in the above subsections can also be expressed in a similar manner. For any stable system \( S \), there is an integer \( d \) (or \( T \)) such that the system \( S^{[d]} \) (or \( S^T \)) admits a quadratic Lyapunov multinorm. These statements are direct corollaries of Theorems 3.6 and 3.3: if \( \rho(S) \leq 1 \) as \( d \) (or \( T \)) increases, the estimate \( \gamma_s \) will at some point become lower than 1 (and the corresponding multinorm is then Lyapunov). Following this line of reasoning, it is natural to ask ourselves whether the Path-Dependent Lyapunov functions present similar bounds approximation bounds to that of Theorems 3.6 and 3.3.

In [17], Lee and Dullerud concentrated on systems where \( G \) is a Markov Chain. However, Path-Dependent Lyapunov functions can of course be expressed for general graphs. In this general case, they can be understood as quadratic multinorms on a system \( S_M \) obtained by transforming an original system \( S \).

**Definition 6 (M-Path-Dependent Lifting)** Given a system \( S(G, \Sigma) \), the M-Path-Dependent lift of \( S \), denoted \( S_M(G, \Sigma) \), is a constrained switching system on an automaton \( G' \) and matrix set \( \Sigma' \) defined as follows:

1. The graph \( G'(V', E') \) is constructed from \( G(V, E) \).
   For all nodes \( v_i \in V \), for all paths \( p \in E \) of length \( M \) ending at \( v_i \), there is a node \( v_{i,p} \in V' \).
   To any path \( p \in E \) of length \( M+1 \) is associated an edge in \( e' \in E' \). For \( p \) of the form
   \[
p = [e_1, p', e_2] \subset E,
   \]
   with \( |p'| = M-1 \) and \( e_2 = (v_i, v_j, \sigma) \in E \), the edge in \( E' \) is
   \[
e' = (v_{i,[e_1,p']}, v_{j,[p',e_2]}, \sigma).
   \]
2. The set of matrices remains unchanged, \( \Sigma' = \Sigma \).

**Example 4** Figure 5 presents the 1-Path-Dependent lift of the automaton of Figure 3. Each node in the original system will give birth to 2 nodes in the lifted automaton, since they have 2 ingoing edges (paths of length 1).

If is worth noticing that the path-dependent lifting preserves the constrained joint spectral radius of a system. Indeed, by construction, any switching sequence is accepted by a system if and only if it is accepted by its path-dependent lifting with memory \( M \), for all \( M \geq 1 \).

**Theorem 3.7** Given a system \( S(G, \Sigma) \) and an integer \( T \geq 1 \), \( \gamma_s(S^T) \) and \( \gamma_s(S_{T-1}) \) be the values obtained by applying Theorem 3.1 to \( S^T(G, A) \) and \( S_{T-1}(G, A) \) respectively. Then,

\[
\gamma_s(S_{T-1}) \leq (\gamma_s(S^T))^{1/T}.
\]

**PROOF.** For \( T = 1 \), \( S = S_{T-1} = S^T \), so the claim holds. Assume now \( T \geq 2 \).

To ease the reading we will adapt the notations defining a path by referring to it by its succession of nodes and labels,

\[
p = \{v_0, \sigma(1)v_1, \cdots, \sigma(T)v_T\} \subset E,
\]
with \( T \) the length of the path. Of course, for the notation to be consistent, we assume that

\[
(v_i, v_{i+1}, \sigma(i+1)) \in E.
\]
A quadratic multinorm with value $\gamma_1$ for the T-product lifted system must be such that, for all paths $p \in G$ of length $T$,
\[
A_p^{\text{Tr}} Q_{v_T} A_p - \gamma^{2T} Q_{v_t} \leq 0,
\] (13)
where $v_t$ and $v_T$ are the source and destination nodes of the path, and $Q_i > 0$, $i = 1, \ldots, N_T$.

We will construct, from a solution to the set of LMIs (13), a quadratic multinorm for the $(T - 1)$-Path-Dependent lift and eliminating terms two by two. Doing so, we can verify this by injecting the solution (17) in the corresponding lift, and using the Schur complement formula, the LMIs (13) are equivalent to
\[
A_p R_{v_T} A_p^{\text{Tr}} - \gamma^{2T} R_{v_T} \leq 0,
\] (15)
and the LMIs (14) to
\[
A_{\sigma(T)} R_{(v_{T-1}, (v_0, \sigma(1)v_1, \ldots, \sigma(T-1)v_{T-1}))} A_{\sigma(T)}^{\text{Tr}}
- \gamma^{2} R_{(v_T, (v_1, \sigma(2)v_2, \ldots, \sigma(T)v_T))} \leq 0.
\] (16)

From a solution $\{R_v, v \in V\}$, to the LMIs (15), the following provides a solution to the set of LMIs (16):
\[
R_{(v_T, (v_0, \sigma(1)v_1, \ldots, \sigma(T-1)v_{T-1}))} =
R_{v_{T-1}} + \gamma^{-2} A_{\sigma(T-1)} R_{v_T} A_{\sigma(T-1)}^{\text{Tr}} +
\gamma^{-4} (A_{\sigma(T-1)} A_{\sigma(T-2)}) R_{v_T} (A_{\sigma(T-1)} A_{\sigma(T-2)})^{\text{Tr}} +
\ldots
+ \gamma^{-2(T-1)} (A_{\sigma(T-1)} \cdots A_{\sigma(1)}) R_{v_T} (A_{\sigma(T-1)} \cdots A_{\sigma(1)})^{\text{Tr}}.
\]

One can verify this by injecting the solution (17) in the LMIs (16), and eliminating terms two by two. Doing so, one obtains the LMI (15), which concludes the proof.

**Corollary 3.8** Given a system $S(G, \Sigma)$, $M \geq 1$ integer, the value $\gamma_*$ obtained by applying Theorem 3.1 to the $(M - 1)$-Path-Dependent lift of $S$ satisfies
\[
\hat{\rho}(S) \leq \gamma_* \leq n^{1/2M} \hat{\rho}(S).
\] This transformation preserves the dimension $n$ of the matrix set. The graph, however, can grow quite quickly. By construction, for $M \geq 1$, the graph of $S_M(G, \Sigma)$ there is one node per path of length $M$, and one edge for each path of length $M + 1$.

### 3.4 A sufficient condition for extremality of quadratic multinorms.

In this subsection, we present an easily checkable sufficient condition for the extremality of an optimal quadratic multinorms. By optimal, we denote a quadratic multinorm obtained as a solution of the optimization program of Theorem 3.1.

The technique we present aims at detecting a cyclic product achieving the maximum growth rate in the system. It is based on a further analysis of the output of Theorem 3.1.

We start by the following fact on the constrained joint spectral radius.

**Lemma 3.9** For a system $S(G, \Sigma)$, let $c \subset E$ be a cycle of length $T$ in $G$. Then
\[
\rho(A_c)^{1/T} \leq \hat{\rho}(G, \Sigma),
\]
where $\rho(A_c)$ is the spectral radius of any cyclic product associated to $c$.

**PROOF.** For any induced matrix norm $\| \cdot \|$, the following holds (Gelfand’s formula)
\[
\rho(A_c)^{1/T} = \lim_{k \to \infty} \| A_c^k \|^{1/kT}
\leq \rho(G, \Sigma).
\] (17)

**Theorem 3.10** (Sufficient extremality condition)

Let $H$ be the optimal quadratic multinorm with value $\gamma_*$, obtained by applying Theorem 3.1 to the system $S(G(V, E), \Sigma)$.

If the subset $E' = \{(v_i, v_j, \sigma)\} \subset E$ of edges for which $-A_{\sigma}^{\text{Tr}} Q_{\star} A_{\sigma} + \gamma^{2\sigma} Q_{\star}$ has a zero eigenvalue is a simple cycle $c \subset E$, then $H$ is extremal.

Moreover,
\[
\hat{\rho}(S) = \rho(A_c)^{1/T},
\] where $T$ is the length of $c$.

**PROOF.** The set of edges $E'$ represent all the LMI constraints of Theorem 3.1 that are tight from the optimization. It is a known fact of convex optimization that
removing constraints that are not tight at the optimal point of a given program does not affect neither the optimal point or optimal objective cost of the program.

In our case, this implies the following result:
If $\gamma^*$ is the value of an optimal quadratic multinorm for $S(G(V,E), \Sigma)$, then it is also the value of an optimal quadratic multinorm for $S'(G(V',E'), \Sigma)$. Therefore, the value of the multinorm $\mathcal{H}$ is the one of an optimal multinorm on the cycle.

The next step is to show that the value of an optimal multinorm on the cycle $c$ of length $T$ is to its CJSR, which is $\rho(A_c)^{1/T}$. One (original) way to do so is to use results such as Corollary 3.8. Indeed, it can be seen that $S_M(G(V,E'), \Sigma) = S'(G(V',E'), \Sigma)$ for any $M$ if $G(V,E')$ is a simple cycle. Therefore,

$$\rho(A_c)^{1/T} \leq \gamma_* \leq \lim_{M \to \infty} \left( n^{1/2M} \right) \rho(A_c)^{1/T},$$

and so optimal multinorms on cycles are extremal.

Finally, from Lemma 3.9 and Proposition 2.2, we obtain

$$\rho(A_c)^{1/T} \leq \rho(G, \Sigma) \leq \gamma_* = \rho(A_c)^{1/T},$$

which concludes the proof.

### 3.5 Lifting to an arbitrary switching system

The following lifting procedure, introduced independently by Kozyakin [29] and Wang [30], allows for such a reduction, but at the cost of raising significantly the dimension of the matrices.

**Definition 7** Given a system $S(G(V,E), \Sigma)$, the lifted set $\Sigma_S \subset \mathbb{R}^{N_V \times N_V}$ is defined as follows:

$$\Sigma_S = \{ A_e, e \in E \},$$

with $e = (v_i, v_j, \sigma)$,

$$A_{(v_i,v_j,\sigma)} = (1(j)1(i)^{Tr}) \otimes A_{\sigma},$$

where $1(k) \in \{ 0,1 \}^n$ is a column vector with $1(k)_\ell = 1$ if $k = \ell$, $1(k)_\ell = 0$ if $k \neq \ell$.

This transformation allows to see any constrained switching system as an arbitrary switching system. Since a lot of research effort has been dedicated to developing approximation algorithms for the joint-spectral radius, the idea for reduce a constrained system to an arbitrary one before applying classical JSR estimation method comes naturally.

However, with the following result, we show that approximation scheme based on computing a common Lyapunov function give the same bound than multinorm-based methods. Therefore, there are no incentive of using this lift within an approximation procedure, since it augments the size of the matrices, which is the most penalizing factor in the approximation (see the complexity in Theorem 3.1).

**Proposition 3.11** The system $S(G, \Sigma)$ admits a quadratic multinorm of value $\gamma$ if and only if there is a quadratic form $Q_S > 0$ such that $\forall A \in \Sigma_S$,

$$A^{Tr} Q_S A - \gamma^2 Q_S \preceq 0.$$

**PROOF.** We assume $\gamma = 1$ without loss of generality. We start with the if part. For each edge, we define the negative definite matrix $B_e = A_e^{Tr} Q_G A_e - Q_G$, with $A_e \in \Sigma_S$.

By construction, the matrices $A_e$, $B_e$ and $Q_G$ have a clear block-structure. To each pair of nodes corresponds a square $n \times n$ bloc in the above mentioned matrices. For example, the bloc associated to the pair $v_k, v_l$ in the matrix $A_{(v_k,v_l,\sigma)} \in \Sigma_S$ is given by

$$A_{(v_k,v_l,\sigma)}^{[k,l]} = \delta_{k,l} A_{\sigma},$$

the blocs in $B_e$ and $Q_G$ being defined in a similar way.

Since $B_e \preceq 0$ and $Q_G > 0$, the diagonal blocs of both matrices are respectively negative and positive definite. The blocs in $B_{(v_k,v_l,\sigma)}$ are given by

$$B_{(v_k,v_l,\sigma)}^{[k,l]} = \sum_{i,j} A_{(v_k,v_l,\sigma)}^{[k,l]} A_{(v_j,v_l,\sigma)}^{[i,j]} - Q_G^{[i,j]} = \delta_{k,l} \left( A_{(v_k,v_l,\sigma)}^{[k,l]} A_{(v_j,v_l,\sigma)}^{[i,j]} \right) - Q_G^{[i,j]}.$$

Therefore, the diagonal blocs of $B_e$ are

$$B_{(v_k,v_l,\sigma)}^{[k,k]} = A_{\sigma}^{Tr} Q_G^{[k,l]} A_{\sigma} - Q_G^{[i,i]} \preceq 0,$$

where $Q_G^{[j,j]} > 0$ and $Q_G^{[i,i]} > 0$ because $Q_G > 0$, and concludes the first part of the proof.

For the second part (only if), notice that for all $1 \leq k, \ell \leq N_V$ with $k \neq \ell$, the terms $Q_G^{[i,i]}$ never appears on the diagonal of $B_e$. Therefore, as long as we are able to build the diagonal of $B_e$ and guarantee $B_e \preceq 0$, we do not have to bother with the off-diagonal terms. And since we have a multinorm, we can indeed build the diagonal, which concludes the proof.
4 Numerical example

In this section, we present an example based on a mod-
elization of a system with state-feedback control that might undergo dropouts in its state feedback. More precisely, the system is as follows:

\[ x_{t+1} = (A + BK_{\sigma(t)}) x_t, \]

with
\[ A = \begin{pmatrix} 0.94 & 0.56 \\ 0.14 & 0.46 \end{pmatrix}, \]

and \( K_{\sigma(t)} = (k_{1,\sigma(t)}, k_{2,\sigma(t)}) \) will switch to represent the dropouts. In this example, dropouts corresponds to a failure from the controller to read either the first or the second state variable of the system. In that case, the control feedback dismisses the corresponding variable. There are 4 cases. First, \( \sigma(t) = 1 \), which is no dropouts, and

\[ K_1 = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} -0.49 \\ 0.27 \end{pmatrix}. \]

Second and third, \( K_2 = \begin{pmatrix} 0 \\ k_2 \end{pmatrix} \) and \( K_3 = \begin{pmatrix} k_1 \\ 0 \end{pmatrix} \) correspond respectivelly to a dropout of the first and second inputs. Finally, \( K_4 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \) is the case where both outputs are dropped. Assuming that a single output cannot be dropped twice in a row, we model the system with dropouts as a constraint switching system on the set

\[ \Sigma = \{ A + BK_i, i = 1 \ldots 4 \}, \]

with the automaton \( G(V, E) \) depicted in Figure 6.

![Fig. 6. Dropout automaton. It has a markovian structure, each node being directly linked to a mode.](image)

We will test the methods of Section 3: 1) the T-product lift \( S^{T_k} \) with \( T_k = 1, \ldots, 5 \); 2) the M-Path-Dependent lift \( S^{M_k} \) lift for \( M_k = T_k - 1 \); 3) the T-product lift applied on the equivalent arbitrary switching systems. Then, we compose methods to do 4) the T-product lift on the system \( S^{[2]} \), noted \( (S^{[2]})^{T_k} \); and 5) \( (S^{[2]})^{M_k} \). Remark that the three first methods provide the same guaranteed accuracy, as well as the two last.

The first comparison is about the running times of the approximations schemes using each of these methods. The result is presented on Figure 7.

![Fig. 7. The T-lift product solves the approximation problem faster than the two others.](image)

The second comparison is about the absolute quality of the approximation (Figure 8). Since the arbitrary lifting methods produce the same result that the T-product lift, it is omitted from the comparison. As expected, the methods using the path-dependent lifting outperforms the others. Remark that as \( T_k \) grows, the absolute quality of the approximation obtained by the T-product lift can decrease. The maximum relative error, on the other hand, still decreases. It is easy to show that the upper bounds obtained by path-dependent lift always decrease with the value of \( T_k \), and that if \( T_k = 1 \) is a multiple of \( T_k \), the bound obtained by the T-product lift with \( T_k \) will be better than the one with \( T_k \).

![Fig. 8. Path-dependent Lyapunov functions obtain better bounds than the T-product lift. Stability is detected by the Path-dependent method for \( T_k \geq 2 \), by T-product lift for \( T_k \geq 3 \).](image)

References

[1] E. A. Hernandez-Vargas, R. H. Middleton, and P. Colaneri, “Optimal and mpc switching strategies for mitigating viral mutation and escape,” in Preprints of the 18th IFAC World Congress Milano (Italy) August, 2011.

[2] R. Shorten, F. Wirth, and D. Leith, “A positive systems model of tcp-like congestion control: asymptotic results,” IEEE/ACM Transactions on Networking, vol. 14, no. 3, pp. 616-629, 2006.

[3] R. M. Jungers, A. D’Innocenzo, and M. D. Di Benedetto, “Feedback stabilization of dynamical systems with switched delays,” arXiv preprint arXiv:1207.5119, 2012.
Given a vector \( x = (x_1, \ldots, x_n)^\text{T} \in \mathbb{R}^n \), and \( n \) indexes \( \alpha = \alpha_1, \ldots, \alpha_n \), with \( \sum \alpha_i = d \), we define

\[
x^{[d]}_\alpha = \frac{d!}{\alpha_1! \cdots \alpha_n!} x_1^{\alpha_1} \cdots x_n^{\alpha_n},
\]

which represent a product of monomials, whose degree sum to \( d \), scales by the multinomial coefficient for the indexes \( \alpha \). This scaling as a very practical justification, which we will highlight later. Let \( x^{[d]} \) be the vector containing all \( x^{[d]}_\alpha \) for all indexes \( \alpha \) summing to \( d \).

With these definitions, we have defined a lifting procedure \( x \rightarrow x^{[d]} \), defining a transformation from the space \( \mathbb{R}^n \) to \( \mathbb{R}^{N(n,d)} \), where

\[
N(n,d) = \binom{n + d - 1}{d}.
\]
Example 5  Given a vector \( x = (x_1, x_2)^T \),

\[
x^{[1]} = x, \quad x^{[2]} = \begin{pmatrix}
    x_1^2 \\
    \sqrt{2}x_1x_2 \\
    x_2^2
\end{pmatrix}, \quad x^{[3]} = \begin{pmatrix}
    x_1^3 \\
    \sqrt{3}x_1^2x_2 \\
    \sqrt{3}x_1x_2^2 \\
    x_2^3
\end{pmatrix}, \ldots
\]

The reason behind the scaling in Equation (19) is that defines a norm-preserving transformation (called the Veranese embedding [15]) between \( \mathbb{R}^n \) and \( \mathbb{R}^N \). One can easily verify that \( |x^{[d]}| = |x|^d \), where \( |\cdot| \) is the euclidean norm on the appropriate vector space. Now, for any matrix \( A \in \mathbb{R}^n \), we define the matrix \( A^{[d]} \) associated to the linear map

\( A^{[d]} : x^{[d]} \rightarrow (Ax)^{[d]} \).

We can naturally extend the definition above to sets of matrices \( \Sigma^{[d]} \).