

Learning stability guarantees for data-driven constrained switching linear systems

Adrien Banse* Zheming Wang* Raphaël Jungers*

* *The ICTEAM Institute, UCLouvain, Louvain-la-Neuve, 1348, Belgium (email: adrien.banse@student.uclouvain.be, zheming.wang@uclouvain.be, raphael.jungers@uclouvain.be).*

Abstract: We consider stability analysis of constrained switching linear systems in which the dynamics is unknown and whose switching signal is constrained by an automaton. We propose a data-driven Lyapunov framework for providing probabilistic stability guarantees based on data harvested from observations of the system. By generalizing previous results on arbitrary switching linear systems, we show that, by sampling a finite number of observations, we are able to construct an approximate Lyapunov function for the underlying system. Moreover, we show that the entropy of the language accepted by the automaton allows to bound the number of samples needed in order to reach some pre-specified accuracy.

Keywords: Stability analysis, Constrained switching linear systems, Data-driven optimization, Scenario approach

1. INTRODUCTION

In this paper we address the problem of finding probabilistic guarantees for the stability of *constrained switching linear systems* whose dynamics is unknown.

Switching systems. We consider discrete-time *switching linear systems (SLS)* defined by a set $\mathcal{A} = \{A_i\}_{i \in \{1, \dots, m\}}$ of m matrices. Their dynamics is given by the following equation:

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

for any $t \in \mathbb{N}$, where $x_t \in \mathbb{R}^n$ and $\sigma(t) \in \{1, \dots, m\}$ are respectively the *state* and the *mode* at time t . The sequence $(\sigma(0), \sigma(1), \dots) \subseteq \{1, \dots, m\}^{\mathbb{N}}$ is the *switching sequence*.

Switching linear systems are an important family of hybrid systems which often arise in Cyber-Physical systems (see Tabuada (2009)). Indeed, the interaction between continuous and discrete dynamics causes hybrid behaviors which makes the stability analysis challenging. In recent years, many model-based stability analysis techniques have been proposed (see Lin and Antsaklis (2009) and references therein, or Jungers (2009)).

A *constrained switching linear system (CSLS)* is a switching linear system with logical rules on its switching signal. We represent these rules by an *automaton*. The stability of CSLS has also been studied extensively (see e.g. Dai (2011), Philippe et al. (2016) and Xu and Acikmese (2020)). In particular, we are interested in asymptotic stability of CSLS, whose definition is given as follows. A CSLS whose dynamics is given by (1) is said to be *asymptotically stable* (or *stable*, for short) if for all $x_0 \in \mathbb{R}^n$,

$$\lim_{t \rightarrow \infty} x_t = 0. \quad (2)$$

Data-driven approach. In many practical applications, the engineer cannot rely on having a model, but rather has to analyze stability in a *data-driven* fashion. Most classical data-driven methods (see e.g. Karimi and Kammer (2017), Hjalmarsson et al. (1998) and Campi et al. (2003)) are limited to linear systems and based on classical identification and frequency-domain approaches. These methods may not be well suited for complex systems such as constrained switching linear systems.

In order to tackle hybrid behaviors in switching systems, novel data-driven stability analysis methods have been recently developed based on *scenario optimization* (see Kenanian et al. (2019), Berger et al. (2021) and Rubbens et al. (2021)). In this paper we seek to take one more step towards complexity. To do that, we develop a data-driven method for providing probabilistic guarantees on the stability of noise-free constrained switching linear systems.

Outline. The rest of this paper is organized in two parts. We introduce the problem that we tackle in Section 2. All concepts needed to this end are introduced in Section 2.1, and the problem is formulated in Section 2.2. In Section 3, we propose a lifting result allowing us to reduce the computation of the *constrained joint spectral radius* to the *joint spectral radius* of a certain set of matrices. Moreover, we state the main theorem of this paper, which extends data-driven results from Berger et al. (2021) to constrained switching linear systems. Finally, we investigate further the obtained generalization. We show that the notion of *entropy* can be used to characterize the number of samples needed to reach a specified guarantee on the stability. We will show that, under some assumptions, a smaller entropy allows for a better probabilistic guarantee.

2. PROBLEM SETTING

2.1 Preliminaries

In this subsection, we introduce the notions necessary to formally present the problem that we solve in this paper.

Joint spectral radius. For arbitrary SLS, given a set of matrices $\mathcal{A} = \{A_1, \dots, A_m\} \subseteq \mathbb{R}^{n \times n}$, the quantity

$$\rho(\mathcal{A}) = \lim_{t \rightarrow \infty} \max_{\sigma(\cdot) \in \{1, \dots, m\}} \|A_{\sigma(t-1)} \dots A_{\sigma(0)}\|^{1/t} \quad (3)$$

is known as the *joint spectral radius (JSR)* of a switching linear system defined on \mathcal{A} . The JSR of an switching system rules the stability of the latter:

Proposition 1. (Jungers (2009), Corollary 1.1.). Given a set of matrices \mathcal{A} , the switching linear system defined by \mathcal{A} is asymptotically stable if and only if $\rho(\mathcal{A}) < 1$.

It is a well known fact that for any stable arbitrary switching linear system, there is a norm acting as a *common Lyapunov function* (see Jungers (2009), Proposition 1.4.). The following proposition gives a sufficient condition for stability, by restricting the search to *common quadratic Lyapunov functions (CQLF)*.

Proposition 2. (Jungers (2009), Proposition 2.8). Consider a finite set of matrices \mathcal{A} . If there exists $\gamma \geq 0$ and a symmetric matrix $P \succ 0$ such that $A^T P A \preceq \gamma^2 P$ holds for any matrix $A \in \mathcal{A}$, then $\rho(\mathcal{A}) \leq \gamma$.

Constrained joint spectral radius. First, we give the definition of an *automaton*. An automaton is a strongly connected, directed and labelled graph $\mathbf{G}(V, E)$ with V the set of nodes and E the set of edges. Note that we drop the explicit writing of V and E when it is clear from the context. The edge $(v, w, \sigma) \in E$ between two nodes $v, w \in V$ carries the *label* $\sigma \in \{1, \dots, m\}$, which maps to a mode of the switching system. A sequence of labels $(\sigma(0), \sigma(1), \dots)$ is a *word* in the language *accepted* by the automaton \mathbf{G} if there is a path in \mathbf{G} carrying the sequence as the succession of the labels on its edges. A CSLS defined on the set of matrices Σ and constrained by the automaton \mathbf{G} is noted $S(\mathbf{G}, \Sigma)$. We define the set of all possible products of matrices in Σ of length l given an automaton \mathbf{G} as

$$\Pi_l = \{A_{\sigma(l-1)} A_{\sigma(l-2)} \dots A_{\sigma(0)} : (\sigma(0), \sigma(1), \dots, \sigma(l-1)) \text{ is a word of } \mathbf{G}\}. \quad (4)$$

The constrained joint spectral radius (*CJSR*), which is a generalization of the JSR to CSLS, was first introduced in Dai (2011). Given a set of matrices Σ and an automaton \mathbf{G} , the CJSR of the constrained switching linear system $S(\mathbf{G}, \Sigma)$ is defined as

$$\rho(\mathbf{G}, \Sigma) = \lim_{t \rightarrow \infty} \max \left\{ \|\mathbf{A}\|^{1/t} : \mathbf{A} \in \Pi_t \right\}. \quad (5)$$

In the same way, the stability of a constrained switching linear system is characterized by its CJSR:

Proposition 3. (Dai (2011), Corollary 2.8.). Given a set of matrices Σ and an automaton \mathbf{G} , the constrained switching linear system $S(\mathbf{G}, \Sigma)$ is asymptotically stable if and only if $\rho(\mathbf{G}, \Sigma) < 1$.

2.2 Problem formulation

We will now formally present the problem that we solve in this paper.

Model-based setting. Consider a given constrained switching linear system $S(\mathbf{G}, \Sigma)$ with $\Sigma \subseteq \mathbb{R}^{n \times n}$. Let $\Delta = \mathbb{S} \times \Pi_l$ with $\mathbb{S} \subseteq \mathbb{R}^n$ the unit sphere and Π_l the set of all admissible products of length l . We introduce the following optimization problem¹ (see Berger et al. (2021)):

$$\begin{aligned} \mathcal{P}(\Delta) : \min_{\substack{P \in \mathbb{R}^{n \times n} \\ \gamma \geq 0}} & (\gamma, \|P\|_F^2) \\ \text{s.t. } & P \in \mathcal{X} := \{P : I \preceq P \preceq CI, P = P^T\}, \\ & (\mathbf{A}x)^T P (\mathbf{A}x) \leq \gamma^{2l} x^T P x \quad \forall (x, \mathbf{A}) \in \Delta, \end{aligned} \quad (6)$$

for a large $C \in \mathbb{R}_{\geq 0}$, where $\|\cdot\|_F$ is the Frobenius norm. We denote $(\gamma^*(\Delta), P^*(\Delta))$ as the solution of optimization problem (6).

Following Proposition 2, Program (6) allows us to study stability in a model-based setting i.e., when Δ is known. Indeed if $\gamma < 1$, then the ellipsoidal norm $\|\cdot\|_{P^*(\Delta)}$ is a CQLF for the considered CSLS (Jungers, 2009). Observe that, in addition to the problem of Proposition 2, a tie-breaking rule is defined in Program (6). This tie-breaking rule allows for improving the probabilistic guarantees we obtain in Theorem 5 (see Kenanian et al. (2019) for details). A constraint $P \preceq CI$ is also added to ensure that the set of feasible P is compact, so that the existence of a solution is guaranteed².

Data-driven setting. In this work, we analyze the same problem in a data-driven framework: we assume that the system is not known (i.e., \mathbf{A} is not known in Program (6)), but that we sample N trajectories of length l of a system $S(\mathbf{G}, \Sigma)$. The i -th trajectory is noted $(x_{i,0}, \dots, x_{i,l})$ for $i \in \{1, \dots, N\}$. The trajectories are assumed to be generated from initial states $x_{i,0}$ drawn randomly, uniformly and independently from \mathbb{S} , the unit sphere.

For each trajectory $i \in \{1, \dots, N\}$, the l matrices are generated from the automaton $\mathbf{G}(V, E)$ in the following way. An initial state u_0 is drawn randomly and uniformly from V . Then a random walk of length l is performed on \mathbf{G} , where, from $u_j \in V$, the next state u_{j+1} is drawn randomly, uniformly and independently from the set of its out-neighbours $\{u_{j+1} \in V : (u_j, u_{j+1}, \sigma_i(j)) \in E\}$ where $\sigma_i(j)$ is the label corresponding to the edge linking u_j and u_{j+1} . The sequence of nodes $(u_0, \dots, u_j, u_{j+1}, \dots, u_l)$ form a switching sequence $\sigma_i(0), \dots, \sigma_i(l-1)$, which maps to the matrices $A_{\sigma_i(0)}, \dots, A_{\sigma_i(l-1)}$.

We define the set of N observations ω_N as

$$\omega_N = \{(x_{i,0}, \mathbf{A}_i), i = 1, \dots, N\} \quad (7)$$

where $\mathbf{A}_i = A_{\sigma_i(l-1)} \dots A_{\sigma_i(0)} \in \Pi_l$. Note that the observations in ω_N are assumed to be noise-free.

¹ We note $\min(f(x), g(x))$ the multiobjective optimization problem where $g(x)$ is used as a *tie-breaking rule*. That is, the objective is to minimize the function $f(x)$, and, in case there are several optimizers, the solution is the one which minimizes $g(x)$. Observe that the latter is unique because the problem is quasi-convex, and because $\|\cdot\|$ is a strongly convex function.

² For more details about these additions, see Berger et al. (2021).

We define $\mathbb{P} = \mathbb{P}_x \times \mathbb{P}_\sigma$ the probability measure on Δ with \mathbb{P}_x the uniform distribution on \mathbb{S} and \mathbb{P}_σ the probability distribution describing the distribution of paths in $\mathbf{\Pi}_l$ as explained above. Note that \mathbb{P}_σ is not necessarily a uniform measure.

Now, for a given set ω_N , let us define the *sampled optimization problem* $\mathcal{P}(\omega_N)$ associated to \mathcal{P} :

$$\begin{aligned} \mathcal{P}(\omega_N) : \min_{\substack{P \in \mathbb{R}^{n \times n} \\ \gamma \geq 0}} (\gamma, \|P\|_F^2) \\ \text{s.t. } P \in \mathcal{X} := \{P : I \preceq P \preceq CI, P = P^T\}, \\ (\mathbf{A}x)^T P (\mathbf{A}x) \leq \gamma^{2l} x^T P x \quad \forall (x, \mathbf{A}) \in \omega_N, \end{aligned} \quad (8)$$

We denote $(\gamma^*(\omega_N), P^*(\omega_N))$ as the solution of optimization problem (8), and $\text{Cost}(\omega_N)$ its optimal cost. The problem $\mathcal{P}(\omega_N)$ defined in Program (6) is the *data-driven* version of the optimization problem $\mathcal{P}(\Delta)$ defined in Program (8). The issue that we tackle in this paper is the inference of $\gamma^*(\Delta)$, the solution of optimization problem (6) from $(\gamma^*(\omega_N), P^*(\omega_N))$ with a certain user-defined level of confidence.

3. MAIN RESULTS

In this section, we present our main results. First, in Proposition 4 given an automaton \mathbf{G} and a set of matrices $\mathbf{\Sigma}$, we show that the CJSR can be bounded by the classical JSR of the set of all admissible products of a given length $\mathbf{\Pi}_l$. Even though other reductions of the CJSR computation problems to a simpler JSR have already been proposed in the literature (see e.g. Dai (2011) and Philippe et al. (2016)), to the best of our knowledge, Proposition 4 is new, and will be useful for our purposes. Second, we use this result in order to derive a probabilistic guarantee allowing to relate the data-driven problem (8) to the model-based problem (6). This guarantee is given in Theorem 5.

Proposition 4. For all $l > 0$, given an automaton \mathbf{G} and a set of matrices $\mathbf{\Sigma}$, the CJSR of $S(\mathbf{G}, \mathbf{\Sigma})$ and the JSR of the switching linear system defined by $\mathbf{\Pi}_l$ satisfy

$$\rho(\mathbf{G}, \mathbf{\Sigma}) \leq \rho(\mathbf{\Pi}_l)^{1/l}. \quad (9)$$

Moreover, the equality holds asymptotically i.e.,

$$\rho(\mathbf{G}, \mathbf{\Sigma}) = \lim_{l \rightarrow \infty} \rho(\mathbf{\Pi}_l)^{1/l}. \quad (10)$$

Proposition 4 allows us to reduce the problem of approximating the CJSR to the problem of approximating the JSR of another arbitrary switching linear system. Therefore we can generalize previous data-driven works on arbitrary systems. In particular, we draw our results on top of Berger et al. (2021) in order to obtain data-driven stability guarantees for constrained systems.

We remark that the data-driven problem (8) is a quasi-linear optimization problem, as defined in (Berger et al., 2021, Equation 1). Thus, a very similar analysis as in Berger et al. (2021), based on scenario-approach results Calafiore (2010), can be done. First, we recall the definition of a *Barabanov* matrix (see Berger et al. (2021), Definition 7). A matrix $A \in \mathbb{R}^{n \times n}$ is said to be Barabanov if there exists a symmetric matrix $P \succ 0$ and $\gamma \geq 0$ such that $A^T P A = \gamma^2 P$.

Given Proposition 3, the following theorem generalizes Corollary 14 of Berger et al. (2021). It gives probabilistic guarantees for the stability of a constrained switching linear system. In the following theorem, $\Phi(\cdot, a, b)$ denotes the *regularized incomplete beta function* for the two parameters $a, b \in \mathbb{N}$ (see Kenanian et al. (2019), Definition 2).

Theorem 5. Consider an automaton \mathbf{G} , a set of matrices $\mathbf{\Sigma} \subseteq \mathbb{R}^{n \times n}$, samples $\omega_N \subset \Delta$ obtained as explained in Section 2.2, a fixed length $l > 0$ and $N \geq d := n(n+1)/2$. Suppose $\mathbf{\Pi}_l$ contains no Barabanov matrices. Consider problem $\mathcal{P}(\omega_N)$ with solutions $\gamma^*(\omega_N)$ and $P^*(\omega_N)$. Then, for a given level of confidence $\beta \in (0, 1)$,

$$\mathbb{P} \left(\left\{ \omega_N \in \Delta^N : \rho(\mathbf{G}, \mathbf{\Sigma}) \leq \frac{\gamma^*(\omega_N)}{\sqrt{l \delta(\beta, \omega_N)}} \right\} \right) \geq \beta, \quad (11)$$

and the function $\delta(\beta, \omega_N)$ takes the form

$$\sqrt{1 - \Phi^{-1}(\varepsilon(\beta, N) \kappa(P^*(\omega_N)) / p_{l, \min}, (n-1)/2, 1/2)}. \quad (12)$$

where $p_{l, \min}$ is the minimal probability of all matrices in $\mathbf{\Pi}_l$, $\kappa(P) = \sqrt{\det(P) / \lambda_{\min}(P)^n}$, and $\varepsilon(\beta, N)$ takes the closed form

$$\varepsilon(\beta, N) = 1 - \Phi(1 - \beta, d+1, N-d). \quad (13)$$

Theorem 5 provides a general way of obtaining probabilistic stability guarantees. Indeed, for a given confidence level β , if one computes an upper bound (11) strictly less than 1, then following Proposition 3, stability holds with probability at least β .

We now show how one can use it in practice, by deriving a few corollaries. The following corollary holds if the distribution of drawing a product in $\mathbf{\Pi}_l$ is uniform.

Corollary 6. Suppose \mathbb{P}_σ is a uniform measure. Then the function $\delta(\beta, \omega_N)$ in Theorem 5 can be written

$$\sqrt{1 - \Phi^{-1}(\varepsilon(\beta, N) |\mathbf{\Pi}_l| \kappa(P^*(\omega_N)), (n-1)/2, 1/2)}. \quad (14)$$

We now show that we can push further our analysis of the upper bound expressed in Corollary 6 by using the notion of *entropy* (Lind and Marcus, 1995, Definition 4.1.1). Let $|\mathcal{L}_{\mathbf{G}, l}|$ be the language accepted by \mathbf{G} restricted to length l . The entropy $h(\mathbf{G})$ of \mathbf{G} is the growth rate of $|\mathcal{L}_{\mathbf{G}, l}|$ i.e.,

$$h(\mathbf{G}) = \lim_{l \rightarrow \infty} \frac{\log_2 |\mathcal{L}_{\mathbf{G}, l}|}{l}. \quad (15)$$

Since $|\mathbf{\Pi}_l| \leq |\mathcal{L}_{\mathbf{G}, l}|$, the definition of the entropy gives the following corollary.

Corollary 7. For $l \rightarrow \infty$, the function $\delta(\beta, \omega_N)$ in Corollary 6 can be written

$$\lim_{l \rightarrow \infty} \sqrt{1 - \Phi^{-1}(\varepsilon(\beta, N) 2^{lh(\mathbf{G})} \kappa(P^*(\omega_N)), (n-1)/2, 1/2)}. \quad (16)$$

Corollary 7 provides an asymptotic estimate of the probabilistic upper bound in Theorem 5, as a function of the entropy of the automaton \mathbf{G} . One can see that an automaton with small entropy allows for a better estimate of the CJSR, for a fixed number of samples. This is illustrated in Figure 1.

Now we show that we can also derive a practical bound for any finite $l > 0$, unlike Corollary 7 which holds

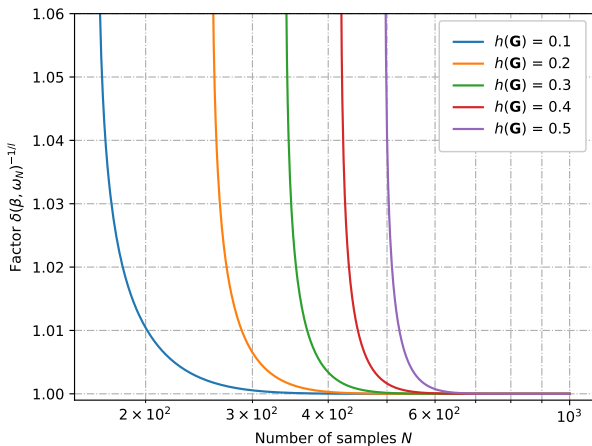


Fig. 1. Shape of the factor $1/\sqrt{l\delta(\beta, \omega_N)}$ in Theorem 5 with respect to the entropy, for a confidence level $\beta = 95\%$, a large l (here $l = 50$) and $n = 2$. One can see that this factor converges to 1 as N increases, and that a smaller entropy allows to converge faster.

asymptotically. For this we use classical results from graph theory.

Proposition 8. Let A be the adjacency matrix of some automaton $\mathbf{G}(V, E)$. Let $\lambda_1 \leq \dots \leq \lambda_{|V|}$ be the eigenvalues of A . Assume A is diagonalizable. Then for any $l \geq 0$, $|\Pi_l| \leq |V|\lambda_n^l$.

Proposition 8 directly gives the following corollary.

Corollary 9. Let A be the adjacency matrix of some automaton $\mathbf{G}(V, E)$. Let $\lambda_1 \leq \dots \leq \lambda_{|V|}$ be the eigenvalues of A . Assume A is diagonalizable. Then for any $l > 0$, the function $\delta(\beta, \omega_N)$ in Corollary 6 can be written

$$\sqrt{1 - \Phi^{-1}(\varepsilon(\beta, N)|V|\lambda_n^l \kappa(P^*(\omega_N)), (n-1)/2, 1/2)}. \quad (17)$$

Corollary 9 provides a probabilistic upper bound in Theorem 5, as a function of the largest eigenvalue of the adjacency matrix of \mathbf{G} . One can see that an automaton with a small largest eigenvalue allows for a better estimate of the CJSR, for a fixed number of samples N and length l .

4. CONCLUSION

In this work, we extended the scope of data-driven stability analysis of hybrid systems by generalizing previous data-driven results to the constrained case. In particular we have built our results on the basis of Berger et al. (2021).

We proceeded as follows. We first proposed a lifting result allowing us to reduce the computation of the CJSR of a given CSLS to the computation of a simpler JSR. We then stated the main theorem of this paper, which provides probabilistic guarantees for the stability of a given noise-free CSLS. Finally, we claimed that in case of uniformity on the distribution of switching sequences, we can investigate further the obtained bound. We showed that a smaller entropy of the automaton allows for a better guarantee about the stability.

In further research, we plan to extend this type of method to noisy observations. We also plan to investigate different approaches. For example, getting rid of the lifting result would allow to reduce the conservatism introduced by the latter, i.e. the gap between the lifted JSR $\rho(\Pi_l)^{1/l}$ and the true CJSR $\rho(\mathbf{G}, \Sigma)$ in (9). In this regard we plan to directly approximate *multiple Lyapunov functions* (Philippe and Jungers, 2015, Definition 2).

REFERENCES

- Berger, G.O., Jungers, R.M., and Wang, Z. (2021). Chance-constrained quasi-convex optimization with application to data-driven switched systems control. *arXiv:2101.01415 [cs, eess, math]*.
- Calafiore, G. (2010). Random convex programs. *SIAM Journal on Optimization*, 20, 3427–3464. doi:10.1137/090773490.
- Campi, M., Lecchini, A., and Savaresi, S. (2003). An application of the virtual reference feedback tuning method to a benchmark problem. *European Journal of Control*, 9(1), 66–76. doi:https://doi.org/10.3166/ejc.9.66-76.
- Dai, X. (2011). A gel'fand-type spectral radius formula and stability of linear constrained switching systems. *arXiv:1107.0124 [cs, math]*.
- Hjalmarsson, H., Gevers, M., Gunnarsson, S., and Lequin, O. (1998). Iterative feedback tuning: theory and applications. *IEEE Control Systems Magazine*, 18(4), 26–41. doi:10.1109/37.710876.
- Jungers, R. (2009). *The Joint Spectral Radius: Theory and Applications*. Springer.
- Karimi, A. and Kammer, C. (2017). A data-driven approach to robust control of multivariable systems by convex optimization. *Automatica*, 85, 227–233. doi:10.1016/j.automatica.2017.07.063.
- Kenanian, J., Balkan, A., Jungers, R.M., and Tabuada, P. (2019). Data driven stability analysis of black-box switched linear systems. *Automatica*, 109.
- Lin, H. and Antsaklis, P. (2009). Stability and stabilizability of switched linear systems: A survey of recent results. *Automatic Control, IEEE Transactions on*, 54, 308 – 322. doi:10.1109/TAC.2008.2012009.
- Lind, D. and Marcus, B. (1995). *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press. doi:10.1017/CBO9780511626302.
- Philippe, M., Essick, R., Dullerud, G.E., and Jungers, R.M. (2016). Stability of discrete-time switching systems with constrained switching sequences. *Automatica*, 72, 242–250. doi:https://doi.org/10.1016/j.automatica.2016.05.015.
- Philippe, M. and Jungers, R.M. (2015). Converse lyapunov theorems for discrete-time linear switching systems with regular switching sequences. In *2015 European Control Conference (ECC)*, 1902–1907. doi:10.1109/ECC.2015.7330816.
- Rubbens, A., Wang, Z., and Jungers, R.M. (2021). Data-driven stability analysis of switched linear systems with sum of squares guarantees.
- Tabuada, P. (2009). *Verification and Control of Hybrid Systems: A Symbolic Approach*. doi:10.1007/978-1-4419-0224-5.
- Xu, X. and Acikmese, B. (2020). Approximation of the constrained joint spectral radius via algebraic lifting.