Data Driven Stability Analysis of Black-box Switched Linear Systems

Marcus Tullius Cicero [a], Julius Caesar [a], Publius Maro Vergilius [b]

[a] *Department of Electrical Engineering at University of California, Los Angeles (UCLA)*

[b] *UC Louvain/Raphael's affiliation*

Abstract

How do we change the abstract? We address the problem of deciding stability of a "black-box" dynamical system (i.e., a system whose model is not known) from a set of observations. The only assumption we make on the black-box system is that it can be described by a switched linear system. We show that, for a given (randomly generated) set of observations, one can give a stability guarantee, for some level of confidence, with a trade-off between the quality of the guarantee and the level of confidence. We provide an explicit way of computing the best stability guarantee, as a function of both the number of observations and the required level of confidence. Our results rely on geometrical analysis and combining chance-constrained optimization theory with stability analysis techniques for switched systems.

Key words: 5 to 10 keywords to pick from the list: http://www.autsubmit.com/documents/keywords.html

1 Introduction

Today's complex cyber-physical systems are characterized by the interaction of a large number of heterogeneous components. Consequently, the models used to analyze these systems are equally complex and consist of heterogeneous sub-models relying on different assumptions and based on principles from different scientific disciplines. It is not uncommon to encounter a patchwork of differential equations, difference equations, hybrid automata, lookup tables, custom switching logic, low-level legacy code, etc. To further compound the difficulty in analyzing these systems, different components of a complex engineered system are typically designed by different suppliers. Although a high-level specification for these components may be known, detailed models are not available for intellectual property reasons. We are thus faced with a tremendous gap between the existing analysis techniques that rely on closed-form models and the models available in industry. It is, therefore, not surprising the emphasis that industry places on simulation since despite the complexity of models, it is always possible to simulate them. This raises the question of whether we can provide formal guarantees about certain properties of these complex systems based solely on the information obtained via their simulations. In this paper, we focus on one of the most important of such properties in the context of control theory: stability.

More formally, we consider a dynamical system as in:

$$x_{k+1} = f(k, x_k), \tag{1}$$

where, $x_k \in X$ is the state and $k \in \mathbb{N}$ is the time index. If (1) is linear, its identification and stability analysis have been extensively studied. In this work, we take a first step into more complex systems by considering the class of switched linear systems. Although we restrict ourselves to such systems, we believe that the presented results can be extended to more general classes of dynamical systems. We start with the following question to serve as a stepping stone: For some $l \in \mathbb{N}_{>0}$, given N traces of length l, $(x_0^i, x_1^i, \ldots, x_l^i)$, $1 \le i \le N$, belonging to the behavior of the system (1), (i.e., $x_{k+1}^i = f(k, x_k^i)$ for any $0 \le k \le l-1$ and any $1 \le i \le N$), what can we say about

⋆ A. Balkan, J. Kenanian, and P. Tabuada had their work supported by NSF Grants 1645824, 1521617, and NSF Expeditions in Computer Augmented Program Engineering. R. Jungers is a Fulbright Fellow and a F.N.R.S. Fellow currently visiting the Dept. of Electrical Engineering at UCLA. His work is supported by the French Community of Belgium and the IAP network Dysco.

Email addresses: `cicero@senate.ir` (Marcus Tullius Cicero), `julius@caesar.ir` (Julius Caesar), `vergilius@culture.ir` (Publius Maro Vergilius).

the stability of the system (1)? For the rest of the paper, we use the term *black-box* to refer to systems where we do not have access to the model, i.e., to f, yet we can indirectly learn information about f by observing traces of length l (in the particular case of $l = 1$, these traces become couples of points (x_k, y_k) as defined in (1)).

A potential approach to this problem is to first identify the dynamics, i.e., the function f, and then apply existing techniques from the model-based stability analysis literature. However, unless f is a linear function, there are two main reasons behind our quest to directly work on system behaviors and bypass the identification phase:

- Even when the function f is known, in general, stability analysis is a very difficult problem [3];
- Identification can potentially introduce approximation errors, and can be algorithmically hard as well. Again, this is the case for switched systems [14].

A fortiori, the combination of these two steps in an efficient and robust way seems far from obvious.

In recent years, increasing number of researchers started addressing various verification and design problems in control of black-box systems [1, 2, 9, 10], Do we add the HSCC '16 paper [13] on machine learning for abstractions here?. In particular, the initial idea behind this paper was influenced by the recent efforts in [12,20], and [4] on using simulation traces to find Lyapunov functions for systems with known dynamics. In these works, the main idea is that if one can construct a Lyapunov function candidate decreasing along several finite trajectories starting from different initial conditions, it should also decrease along every other trajectory. Then, once a Lyapunov function candidate is constructed, this intuition is put to test by verifying the candidate function either via off-the-shelf tools as in [20] and [12], or via sampling-based techniques as in [4]. This also relates to almost-Lyapunov functions introduced in [15], which presents a relaxed notion of stability proved via Lyapunov functions decreasing everywhere except on a small set. Note that, since we do not have access to the dynamics, these approaches cannot be directly applied to black-box systems. However, these ideas raise the following problem that we address in this paper: By observing that a candidate Lyapunov function decreases on a large number of observations, we empirically build a certain confidence that such candidate Lyapunov function is a bona-fide Lyapunov function. *Can we translate this confidence into a confidence that this Lyapunov function decreases at most of the points in the state space?*

Note that, even in the case of a linear system, the connection between these two beliefs is nontrivial. In fact, one can easily construct an example where a candidate Lyapunov function decreases everywhere on its levels sets, except for an arbitrarily small subset, yet, almost all trajectories diverge to infinity (see example in technical report). In this paper, we take the first steps to infer stability from observations of switched linear systems. In addition to the preceding example, there are other reasons to temper our expectations for proving stability from data. First, identifying an arbitrary switched linear system is NP-hard [11]. Second, the stability of switched linear systems is closely related to a quantity on the matrices modeling the dynamics in each mode, whose computation is itself known to be NP hard: the *joint spectral radius* (JSR). Indeed, deciding stability amounts to deciding whether the JSR is less than 1 [11]. In this paper, we present an algorithm to bound the JSR of a switched linear system from a finite number N of observations. This algorithm partly relies on tools from the random convex optimization literature (also known as chance-constrained optimization, see [6, 8, 16]), and provides an upper bound on the JSR with a user-defined confidence level. As N increases, this bound gets tighter. Moreover, with a closed form expression, we characterize what is the exact trade-off between the tightness of this bound and the number of samples. In order to understand the quality of our upper bound, the algorithm also provides a deterministic lower bound. Finally, we provide an asymptotic guarantee on the gap between the upper and the lower bound, for large N.

The organization of the paper is as follows: In Section 2, we introduce the problem studied and provide the necessary background in stability of switched linear systems. Then, based on finite observations for a given switched linear system, we give in Section 3 a deterministic lower bound for the JSR, before presenting in Section 4 the main contribution of this paper, which consists in a probabilistic stability guarantee. We illustrate the performance of the presented techniques with some experiments in Section 5, then we conclude in Section 6, while hinting at our related future work.

2 Preliminaries

2.1 Notations

We consider the usual finite normed vector space $(\mathbb{R}^n, \ell_2)$, $n \in \mathbb{N}_{>0}$, with ℓ_2 the classical Euclidean norm. We denote by $\|x\|$ the ℓ_2-norm of $x \in \mathbb{R}^n$. We also denote the set of linear functions in $\mathbb{R}^n$ by $\mathcal{L}(\mathbb{R}^n)$, and the set of real symmetric matrices of size n by $\mathcal{S}^n$. In particular, the set of positive definite matrices is denoted by $\mathcal{S}^n_{++}$. We write $P \succ 0$ to state that P is positive definite, and $P \succeq 0$ to state that P is positive semi-definite. Given a set $X \subset \mathbb{R}^n$, and $r \in \mathbb{R}_{>0}$ we write $rX := \{x \in X : rx\}$ to denote the scaling of ratio r of this set. We denote by $\mathbb{B}$ (respectively $\mathbb{S}$) the ball (respectively sphere) of unit radius centered at the origin. We denote the ellipsoid described by the matrix $P \in \mathcal{S}^n_{++}$ as E_P, i.e., $E_P := \{x \in \mathbb{R}^n : x^T P x = 1\}$. Finally, we denote the spherical projector on $\mathbb{S}$ by $\Pi_{\mathbb{S}} := x/\|x\|$.

We consider in this work the classical uniform spherical measure on $\mathbb{S}$, denoted by σ^{n-1}, and derived from the Lebesgue measure λ (see the appendix [technical report] for precise definitions). For $m \in \mathbb{N}_{>0}$, we denote by M the set $M = \{1, 2, \ldots, m\}$ and we provide it with the uniform measure μ_M. For any $l \in \mathbb{N}_{>0}$, we denote by M^l the l-Cartesian product of M: $M^l = \Pi_1^l M$, and similarly we provide it with the uniform measure $\mu_{M^l} = \otimes^l \mu_M$. We can then define $Z_l = \mathbb{S} \times M^l$ as the Cartesian product of the unit sphere $\mathbb{S}$ and M^l. On the set Z_l, we define the product measure $\mu_l = \sigma^{n-1} \otimes \mu_{M^l}$. Note that, μ_l is a uniform measure on Z_l and has total mass 1. Finally, we denote by $\pi_{\mathbb{S}}$ the classical projection from Z_l to $\mathbb{S}$

2.2 Stability of Switched Linear Systems

A *switched linear system* with a set of modes $\mathcal{M} = \{A_i, i \in M\}$ is of the form:

$$x_{k+1} = f(k, x_k), \tag{2}$$

with $f(k, x_k) = A_{\tau(k)} x_k$ and switching sequence $\tau \in M^{\mathbb{N}}$. Note that such systems are homogeneous, i.e., for any $\gamma > 0$, $f(k, \gamma x_k) = \gamma f(k, x_k)$.

In this paper, we are interested in the worst-case global stability of this system, that is, we want to guarantee the following property:

$$\forall \tau \in M^{\mathbb{N}},\ \forall x_0 \in \mathbb{R}^n,\ \|x_k\| \xrightarrow[k \to \infty]{} 0.$$

It is well-known that the joint spectral radius of a set of matrices $\mathcal{M}$ closely relates to the stability of the underlying switched linear systems (2) defined by $\mathcal{M}$. This quantity is an extension to switched linear systems of the classic spectral radius for linear systems. It is the maximum asymptotic growth rate of the norm of the state under the dynamics (2), over all possible initial conditions and sequences of matrices of $\mathcal{M}$.

Definition 1 (from [11]) *Given a finite set of matrices $\mathcal{M} \subset \mathbb{R}^{n \times n}$, its* joint spectral radius *(JSR) is given by*

$$\rho(\mathcal{M}) = \lim_{k \to \infty} \max_{i_1, \ldots, i_k} \left\{ ||A_{i_1} \ldots A_{i_k}||^{1/k} : A_{i_j} \in \mathcal{M} \right\}.$$

Property 2.1 (Cor. 1.1, [11]) *Given a finite set of matrices $\mathcal{M}$, the corresponding switched dynamical system is stable if and only if $\rho(\mathcal{M}) < 1$.*

Theorem 2 (Prop. 1.3, [11]) *Should we keep the reference with "Proposition", while we are stating it is a theorem?Given a finite set of matrices $\mathcal{M}$, and any invertible matrix T, $\rho(\mathcal{M}) = \rho(T\mathcal{M}T^{-1})$, i.e., the JSR is invariant under similarity transformations (and is a fortiori a homogeneous function: $\forall \gamma > 0$, $\rho\left(\mathcal{M}/\gamma\right) = \mathcal{M}/\gamma$).*

Definition 3 *Consider a finite set of matrices $\mathcal{M} \subset \mathbb{R}^{n \times n}$. A* common quadratic Lyapunov function (CQLF) *for a system (2) with set of matrices $\mathcal{M}$, is a positive definite matrix $P \in \mathcal{S}_{++}^n$ such that for some $\gamma \geq 0$,*

$$\forall A \in \mathcal{M}, A^T P A \preceq \gamma^2 P. \tag{3}$$

CQLFs are useful because they can be computed, if they exist, with semidefinite programming (see [5]), and they constitute a stability guarantee for switched systems as we formalize next.

Theorem 4 (Prop. 2.8 and Thm. 2.11, [11]) *Consider a finite set of matrices $\mathcal{M}$.*

- *If there exist $\gamma \geq 0$ and $P \succ 0$ such that the Lyapunov equation (3) holds, then $\rho(\mathcal{M}) \leq \gamma$.*
- *If $\rho(\mathcal{M}) < \frac{\gamma}{\sqrt{n}}$, there exists a CQLF, P, such that $\forall A \in \mathcal{M}$, $A^T P A \preceq \gamma^2 P$.*

This theorem provides us with a *converse Lyapunov result*: if there exists a CQLF, then our system is stable. If, on the contrary, there is no such stability guarantee, one may conclude a lower bound on the JSR. We obtain then an approximation algorithm for the JSR. It turns out that one can still refine this technique, in order to improve the error factor $1/\sqrt{n}$, and asymptotically get rid of it. This is a well-known technique for the "white-box" computation of the JSR, which we summarize in the following corollary.

Corollary 5 *For any finite set of matrices such that $\rho(\mathcal{M}) < \frac{\gamma^l}{\sqrt[2l]{n}}$ with $\gamma \geq 0$, there exists a CQLF for $\mathcal{M}^l := \{\Pi_{j=1}^l A_{i_j} : A_{i_j} \in \mathcal{M}\}$, that is, a $P \succ 0$ such that:*

$$\forall\ \mathbf{A} \in \mathcal{M}^l,\ \mathbf{A}^T P \mathbf{A} \preceq \gamma^{2l} P.$$

Proof. It is easy to see from the definition of the JSR that $\rho(\mathcal{M}^l) = \rho(\mathcal{M})^l$. Thus, applying Theorem 4 to the finite set $\mathcal{M}^l$, one directly obtains the corollary.

Note that, the smaller γ is in Theorem 4 and Corollary 5, the tighter is the upper bound we get on $\rho(\mathcal{M})$. Therefore, we could consider in particular, for any $l \in \mathbb{N}_{>0}$, the optimal solution γ^* of the following optimization problem:

$$\begin{aligned} \min_{\gamma, P} \quad & \gamma \\ \text{s.t.} \quad & (\mathbf{A}x)^T P \mathbf{A} x \leq \gamma^{2l} x^T P x, \\ & \qquad \forall \mathbf{A} \in \mathcal{M}^l, \forall x \in \mathbb{S} \\ & P \succ 0. \end{aligned} \tag{4}$$

Notice that we restrict the set of constraints x to $\mathbb{S}$ due to the homogeneity of the system, which implies that it is sufficient to show the decrease of a CQLF on an arbitrary set enclosing the origin.

2.3 Problem Formulation

We now formally present the problem we will be considering from now. We recall that our observations are traces of the form $(x_k, x_{k+1}, \ldots, x_{k+l})$ for some arbitrary $l \in \mathbb{N}_{>0}$, and that we do not have access to the mode applied to the system at each time step. To generate these traces, we assume that we can randomly draw a finite number of initial conditions $x_0^i \in \mathbb{S}$, and that a random sequence of l modes is applied to each of these points. Hence, the probability event corresponding to a given observed trace $(x_k, x_{k+1}, \ldots, x_{k+l})$ is another $(l+1)$-tuple $(x_k, j_1, \ldots, j_l)$. More precisely, we assume that we can uniformly sample N such $(l+1)$-tuples $(x, \mathbf{j})$ in $Z_l = \mathbb{S} \times M^l$, giving us a sample denoted by

$$\omega_N := \{(x_0^i, j_{i,1}, \ldots, j_{i,l}), 1 \leq i \leq N\} \subset Z_l$$

and a corresponding set of N available observations $\{(x_0^i, x_1^i, \ldots, x_l^i), 1 \leq i \leq N\}$ which satisfy for all $1 \leq i \leq N$ and $1 \leq k \leq l$, $x_k^i = A_{j_{k,i}} \ldots A_{j_{1,i}} x_0^i$.

Remark 6 *Let us motivate the choice of considering a uniform sampling for the modes. We assumed that we only have access to random observations of the state of the system, and ignore the process that generates them. In particular, we ignore the process that picks the modes at each time step, and model it as a random process. We suppose that with nonzero probability, each mode is active: the problem would indeed not make a lot of sense otherwise, since in a such case, the system would be unidentifiable with probability 1 and would prevent to ever observe some of its behaviors. We take this distribution uniform here since we cannot say that some modes are privileged a priori. We can still take any other distribution satisfying our assumption: if we have a nonzero lower bound on the probability of each mode, our guarantees naturally extend to them.*

In this work, we aim at understanding what type of guarantees one can obtain on the stability of System (2) (that is, on the JSR of $\mathcal{M}$) from such data. More precisely, we answer the following problem:

Problem 7 *Consider a finite set of matrices, $\mathcal{M}$, describing a switched system (2), and suppose that one has a set of N observations relative to ω_N, finite sample of Z_l drawn according to the uniform measure μ_l.*

- *For a given number $\beta \in [0, 1)$, provide an upper bound on $\rho(\mathcal{M})$, together with a level of confidence β, that is, a number $\overline{\rho(\omega_N)}$ such that*

$$\mu_l \left(\{\omega_N : \; \rho(\mathcal{M}) \leq \overline{\rho(\omega_N)}\} \right) \geq \beta.$$

- *For the same given level of confidence β, provide a lower bound $\underline{\rho(\omega_N)}$ on $\rho(\mathcal{M})$.*

Remark 8 *We will see in Section 3 that we can even derive a deterministic lower bound for a given (sufficiently high) number of observations.*

The idea from now will be to leverage the fact that conditions for the existence of a CQLF for (2) can be obtained by considering a finite number of traces in $\mathbb{R}^n$ of the form $(x_k, x_{k+1}, \ldots, x_l)$. It will lead us to the following algorithm, that is the main result of our work and that answers Problem 7:

Algorithm 1 (Probabilistic upper bound)
Input: *observations produced by a uniform random sample $\omega_N \subset Z$ of size $N \geq \frac{n(n+1)}{2} + 1$;*

Input: *β desired level of certainty;*

Compute: *a candidate for the upper bound, $\gamma^*(\omega_N)$, solution of a convex optimization problem;*

Compute: *$\varepsilon(\beta, \omega_N)$ the proportion of points where our inference on the upper bound may be invalid;*

Compute: *$\delta(\varepsilon)$ a correcting factor;*

Output: *$\frac{\gamma^*(\omega_N)}{\sqrt[2l]{n}} \leq \rho \leq \frac{\gamma^*(\omega_N)}{\sqrt[l]{\delta(\varepsilon)}}$, (right-hand inequality valid with probability at least β and $\delta(\beta, \omega_N) \xrightarrow[N \to \infty]{} 1$).*

3 A Deterministic Lower Bound

In Section 2.2, we gave an optimization problem, (4), that provides a stability guarantee. Nevertheless, solving this problem as stated solely from observation of traces (that gives a finite number of constraints) is not possible since (4) involves infinitely many constraints. We consider then the following optimization problem:

$$\begin{array}{ll} \min_P & \gamma \\ \text{s.t.} & (\mathbf{A}_l x)^T P \mathbf{A}_l x \leq \gamma^{2l} x^T P x, \; \forall (x, \mathbf{j}) \in \omega_N \\ & P \succ 0, \; \gamma \geq 0. \end{array} \tag{5}$$

where $\mathbf{A}_l := A_{j_l} A_{j_{l-1}} \ldots A_{j_1}$ and $\mathbf{j} := \{j_1, \ldots, j_l\}$. Note that, (5) can be efficiently solved by semidefinite programming and bisection on the variable γ (see [5]). Let us denote from now by $\gamma^*(\omega_N)$ the optimal solution of this problem, which we will use to compute a deterministic lower bound and a probabilistic upper bound on the JSR. In this section, we provide a theorem for a deterministic lower bound based on the observations given by ω_N, whose accuracy depends on the “horizon” l.

Theorem 9 *For an arbitrary $l \in \mathbb{N}_{>0}$ and a given uniform sample $\omega_N \subset Z_l$, by considering $\gamma^*(\omega_N)$ the optimal solution of the optimization problem (5), we have*

$$\rho(\mathcal{M}) \geq \frac{\gamma^*(\omega_N)}{\sqrt[2l]{n}}.$$

Proof. Let $\nu > 0$. By definition of $\gamma^*(\omega_N)$, there exists no matrix $P \in \mathcal{S}^n_{++}$ such that, $\forall x \in \mathbb{S}$, $\forall \mathbf{A}_l \in \mathcal{M}^l$,

$$(\mathbf{A}_l x)^T P \mathbf{A}_l x \leq (\gamma^*(\omega_N) - \nu)^{2l} x^T P x.$$

By Theorem 2, this means that there exists no CQLF for the scaled set of matrices $\frac{\mathcal{M}^l}{(\gamma^*(\omega_N)-\nu)^l}$. Since this is valid for every $\nu > 0$, using Theorem 4, and the fact that $\rho(\mathcal{M}^l) = \rho(\mathcal{M})^l$, we conclude:

$$\frac{\rho(\mathcal{M})}{\gamma^*(\omega_N)} \geq \frac{1}{\sqrt[2l]{n}}.$$

4 A Probabilistic Stability Guarantee

4.1 A Partial Upper Bound

In this section, we show how to compute an upper bound on ρ, with a user-defined confidence $\beta \in [0, 1)$. We do this by constructing a l-step CQLF which is valid with probability at least β. Note that, the existence of a l-step CQLF implies $\rho \leq \gamma^*$ due to Theorem 4. Even though to find this CQLF in practice, one would solve the feasibility problem on P induced by problem (5) once γ^* is found, for the sake of rigor and clarity of our proofs, we introduce a slighly different optimization problem. We consider the following optimization problem, referred by $\mathrm{Opt}(\omega_N)$ for the rest of the discussion:

$$\begin{aligned} \min_{P} \quad & \lambda_{\max}(P) \\ \text{s.t.} \quad & (\mathbf{A}x)^T P \mathbf{A} x \leq ((1+\eta)\gamma^*(\omega_N))^{2l} x^T P x, \\ & \qquad\qquad\qquad\qquad \forall\ (x, \mathbf{j}) \in \omega_N \\ & P \succeq I, \end{aligned} \tag{6}$$

with $\eta > 0$, and $\gamma^*(\omega_N)$ the optimal solution to the optimization problem (5). Let us analyze the relationship between $\mathrm{Opt}(\omega_N)$ and the optimization problem (5). Firstly, thanks to the homogeneity of the system, we could replace the constraint $P \succ 0$ in the initial problem with the constraint $P \succeq I$. Secondly, for reasons that will become clear later in the discussion, we want to take the objective function as $\lambda_{\max}(P)$ (which is convex), instead of solving a feasibility problem in P. Lastly, we introduced a "regularization parameter", $\eta > 0$, which ensures strict feasibility of $\mathrm{Opt}(\omega_N)$. As the reader will see, we will derive results valid for arbitrarily small values of η. This will then not hamper the practical accuracy of our technique, while allowing us to derive a theoretical asymptotic guarantee (i.e., for large number of observations). From now, we denote the optimal solution of $\mathrm{Opt}(\omega_N)$ by $P(\omega_N)$, and drop the explicit dependence of P on ω_N when it is clear from the context.

The curious question whether the optimal solution of this sampled problem is a feasible solution to (4) has been widely studied in the literature [6]. It turns out that under certain technical assumptions, one can bound the proportion of the constraints of the original problem (4) that are violated by the optimal solution of $\mathrm{Opt}(\omega_N)$, with some probability which is a function of the sample size N. In the following theorem, we adapt a classical result from random convex optimization literature to our problem.

Theorem 10 (adapted from Theorem 3.3 [1] , [6]) *Let d be the dimension of $Opt(\omega_N)$ and $N \geq d+1$. Consider the optimization problem $Opt(\omega_N)$ given in (6), where ω_N is a uniform random sample drawn from the set Z_l. Then, for all $\varepsilon \in (0, 1]$ the following holds:*

$$\mu_l^N\{\omega_N \in Z_l^N : \mu_l\left(V(\omega_N)\right) \leq \varepsilon\} \geq \beta(\varepsilon, N), \tag{7}$$

where μ_l^N denotes the product probability measure on Z_l^N, $\beta(\varepsilon, N) = 1 - \sum_{j=0}^{d} \binom{N}{j} \varepsilon^j (1-\varepsilon)^{N-j}$, and $V(\omega_N)$ is the set $\{(x, \mathbf{j}) \in Z_l : (\mathbb{A}_l x)^T P(\omega_N) \mathbb{A}_l x > (\gamma^*_{\omega_N})^{2l} x^T P(\omega_N) x\}$ i.e., it is the set of constraints that are violated by the optimal solution of $Opt(\omega_N)$.

Corollary 11 *Consider a set of matrices $\mathcal{M}$, γ^* optimal solution of (5) and matrix $P \succ 0$ optimal solution of $Opt(\omega_N)$. Then P satisfies:*

$$(\mathbf{A}_l x)^T P \mathbf{A}_l x \leq (\gamma^*)^{2l} x^T P x, \forall x \in \mathbb{S} \setminus \tilde{\mathbb{S}}, \forall \mathbf{j} \in M^l \tag{8}$$

with $\tilde{\mathbb{S}} = \pi_{\mathbb{S}}(V(\omega_N)) \subset \mathbb{S}$ such that $\sigma^{n-1}(\tilde{\mathbb{S}}) \leq \varepsilon m^l$.

The proof of Corollary 11 is based on straightforward arguments on measures, and is given in [technical report]. This result allows us to only consider set of violating points on the sphere from now. Note that, this result is conservative: the case where we have the equality $\sigma(\tilde{\mathbb{S}}) = \varepsilon m^l$ corresponds to the case where we have only observed one mode for the entirety of the trace and have then minimal knowledge on the system for a given ε.

The above results allow us to conclude, from a finite number of observations, that with probability β (where β goes to 1 as N goes to infinity), the required property is actually satisfied for the complete sphere $\mathbb{S}$, except on

[1] Theorem 3.3 in [6] requires $\mathrm{Opt}(\omega_N)$ to satisfy the following technical assumptions:

(1) When the problem $\mathrm{Opt}(\omega_N)$ admits an optimal solution, this solution is unique.

(2) Problem $\mathrm{Opt}(\omega_N)$ is nondegenerate with probability 1.

Here, the first assumption can be enforced if required by adding a tie-breaking rule to $\mathrm{Opt}(\omega_N)$ as explained in Appendix A in [7], while the second assumption can be lifted, as explained in PART 2b in [8], thanks to the introduction of a "constraint heating".

a small set of measure at most $\tilde{\varepsilon} = \varepsilon m^l$. This means that, the ellipsoid E_P computed by $\mathrm{Opt}(\omega_N)$ is "almost invariant" except on a set of measure bounded by $\tilde{\varepsilon}$. This can be represented in the case $n = 2$ by the following plot, where the red points of E_P are points that might violate the contractivity constraint. Here, the set of red points has measure at most $\tilde{\varepsilon}$.

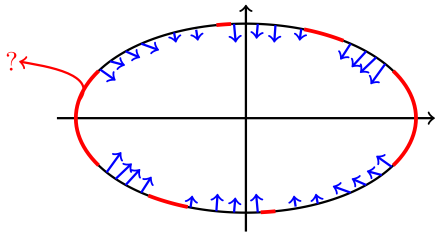

Fig. 1. Representation of the "partial invariance property" obtained by application of the results in Theorem 10. A priori, we know nothing about the images of the red points. Our goal is to convert this partial invariance property into a global stability property.

Thus, we are left with the following question:

Problem 12 *What can we conclude on the JSR if the Lyapunov property is satisfied by all points, except a set of measure $\tilde{\varepsilon}$?*

We will answer to this question by considering the largest ellipsoid included in the convex hull of the points of E_P that satisfy the invariance property. This ellipsoid will indeed satisfy our invariance property thanks to the following key property of switched linear systems.

Property 4.1 *The dynamics given in (2) is convexity-preserving, meaning that for any set of points $X \subset \mathbb{R}^n$, we have $f(\mathit{convhull}\,(X)) \subset \mathit{convhull}\,(f(X))$.*

Of course, for a fixed measure $\tilde{\varepsilon}$, this largest ellipsoid will depend on the distribution of points of E_P that violate the constraint. In order to obtain a guarantee on our upper bound, we will look for the smallest such ellipsoid obtained over all possible sets $V(\omega_N)$ of measure $\tilde{\varepsilon}$. In the particular case where $E_P = \mathbb{S}$, we benefit of the following tool.

Definition 13 *We define the* spherical cap *on $\mathbb{S}$ for a given hyperplane $c^T x = k$ as $\mathcal{C}_{c,k} := \{x \in \mathbb{S} : c^T x > k\}$.*

We now define the function

$$\Delta : \begin{cases} \wp(\mathbb{S}) \to [0, 1] \\ X \mapsto \sup\{r : r\mathbb{B} \subset \mathrm{convhull}\ (\mathbb{S} \setminus X)\}. \end{cases} \tag{9}$$

The following proposition tells us that Δ is minized when X is a spherical cap, i.e., the minimal radius δ of the largest sphere $\mathbb{S}_\delta$ included in $\mathbb{S} \setminus X$ will be reached when X is a spherical cap.

Proposition 14 *Let $\mathcal{X}_{\tilde{\varepsilon}} = \{X \subset \mathbb{S} : \sigma^{n-1}(X) \leq \tilde{\varepsilon}\}$. Then, for any $\tilde{\varepsilon} \in [0, 1]$, the function $\Delta(X)$ attains its minimum over $\mathcal{X}_{\tilde{\varepsilon}}$ for some X which is a spherical cap.*

A proof of Proposition 14 is given in technical report. By homogeneity of the system, we have $x \in \tilde{\mathbb{S}} \iff -x \in \tilde{\mathbb{S}}$, which implies that the minimal δ will in fact occur when the set of violating points is the union of two symmetric spherical caps, each of measure $\frac{\tilde{\varepsilon}}{2}$.

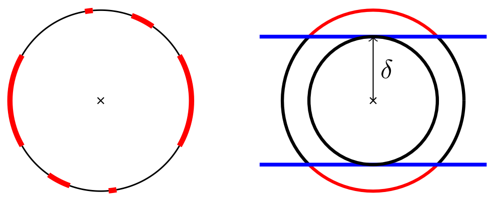

Fig. 2. On the left, case when the ellipse in Fig. 4.1 is a sphere. On the right, case giving minimal δ. The set of points violating the invariance constraint (in red) is the union of two spherical caps, each of measure $\tilde{\varepsilon}$.

Remark 15 *When $\varepsilon \geq \frac{1}{m^l}$, we have $\tilde{\varepsilon} \geq 1$ and $\delta(\tilde{\varepsilon}) = 0$: the upper we can give for the JSR is then only $+\infty$.*

4.2 A global upper bound

In this section, we give an answer to Problem 12 in Theorem 16. In order to use the properties for spheres given in the previous section, we will have to relate $E_{P(\omega_N)}$ to $\mathbb{S}$. In order to do thi, we apply a change of coordinates bringing E_P to $\mathbb{S}$. Since $P \in \mathcal{S}^n_{++}$, it can be written in its Cholesky form

$$P = L^T L, \tag{10}$$

where L is an upper triangular matrix. Note that, L maps the elements of E_P to $\mathbb{S}$.

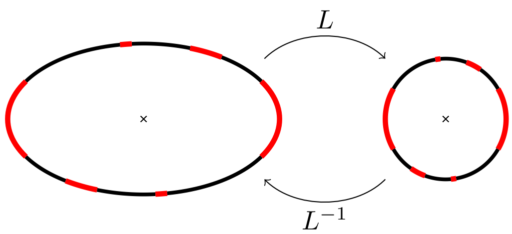

Fig. 3. Change of coordinates to bring our problem back to the case of the unit sphere.

Theorem 16 *Let $\gamma^* \in \mathbb{R}_{>0}$. Consider a set of matrices $\mathcal{M}$, and a matrix $P \succ 0$ optimal solution of (8) for some $\tilde{\mathbb{S}} \subset \mathbb{S}$ where $\sigma^{n-1}(\tilde{\mathbb{S}}) \leq \tilde{\varepsilon}$. Then, we have*

$$\rho(\mathcal{M}) \leq \frac{\gamma^*}{\sqrt[l]{\delta\left(\frac{\tilde{\varepsilon}\kappa(P)}{2}\right)}}$$

with $\kappa(P)=\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}$ *and* $\delta(x)=\sqrt{1-I^{-1}\left(2x;\frac{n-1}{2},\frac{1}{2}\right)}$ *(I is the regularized incomplete beta function).*

Proof. *i)* Since we have seen in the previous section a technique to solve the spherical case, we first bring our problem to the spherical case. To do so, we perform the change of coordinates defined as in (10) by $L\in\mathcal{L}(\mathbb{R}^n)$ which maps the ellipsoid E_P to the sphere $\mathbb{S}$. By defining $\bar{A}_{j_i}=LA_{j_i}L^{-1}$, and $\bar{\mathbf{A}}_l=\bar{A}_{j_l}\bar{A}_{j_{l-1}}\dots\bar{A}_{j_1}$, problem (8) becomes:

$$(\bar{\mathbf{A}}_l x)^T\bar{\mathbf{A}}_l x\le(\gamma^*)^{2l}x^Tx,\ \forall x\in L^{-1}(\mathbb{S}\setminus\tilde{\mathbb{S}}),\forall\mathbf{j}\in M^l. \tag{11}$$

By using the homogeneity of the dynamics, we have:

$$\begin{aligned}&(\bar{\mathbf{A}}_l x)^T\bar{\mathbf{A}}_l x\le(\gamma^*)^{2l}x^Tx,\ \forall x\in L^{-1}(\mathbb{S}\setminus\tilde{\mathbb{S}})\\ &\quad\implies(\bar{\mathbf{A}}_l x)^T\bar{\mathbf{A}}_l x\le(\gamma^*)^{2l}x^Tx,\ \forall x\in\Pi_{\mathbb{S}}\left(L^{-1}(\mathbb{S}\setminus\tilde{\mathbb{S}})\right),\end{aligned}$$

and therefore we can rewrite (11) as:

$$(\bar{\mathbf{A}}_l x)^T\bar{\mathbf{A}}_l x\le(\gamma^*)^{2l}x^Tx,\forall x\in\mathbb{S}\setminus\Pi_{\mathbb{S}}(L^{-1}\tilde{\mathbb{S}}),\forall\mathbf{j}\in M^l. \tag{12}$$

ii) We now show how to relate $\sigma^{n-1}(\Pi_{\mathbb{S}}(L^{-1}(\tilde{\mathbb{S}})))$ to $\sigma^{n-1}(\tilde{\mathbb{S}})$, measure of the violating set in the initial coordinates. Consider $\mathbb{S}^{\tilde{\mathbb{S}}}$, the sector of $\mathbb{B}$ defined by $\tilde{\mathbb{S}}$. We denote $C:=L^{-1}(\tilde{\mathbb{S}})$ and $C':=\Pi_{\mathbb{S}}(L^{-1}(\tilde{\mathbb{S}}))$. We have $\Pi_{\mathbb{S}}(C)=C'$ and $\mathbb{S}^{C'}\subset\frac{1}{\lambda_{\min(L^{-1})}}C$. This leads to:

$$\sigma^{n-1}(C')=\lambda\left(\mathbb{S}^{C'}\right)\le\lambda\left(\frac{1}{\lambda_{\min(L^{-1})}}E^C_{P'}\right).$$

Then, the following holds:

$$\begin{aligned}\sigma^{n-1}(C')\le\frac{\lambda(E^C_{P'})}{\lambda_{\min}(L^{-1})^n}&\le\frac{\lambda\left(L^{-1}(\mathbb{S}^{\tilde{\mathbb{S}}})\right)}{\lambda_{\min}(L^{-1})^n}\\ &=\frac{|\det(L^{-1})|}{\lambda_{\min}(L^{-1})^n}\lambda\left(\mathbb{S}^{\tilde{\mathbb{S}}}\right) &(13)\\ &=\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}\sigma^{n-1}(\tilde{\mathbb{S}}) &(14)\end{aligned}$$

where (13) follows from the fact that $\lambda(Q(X))=|\det(Q)|\lambda(X)$, for any set $X\subset\mathbb{R}^n$ and $Q\in\mathcal{L}(\mathbb{R}^n)$ (see e.g. [19]). Hence, we have

$$(\bar{\mathbf{A}}_l x)^T\bar{\mathbf{A}}_l x\le(\gamma^*)^{2l}x^Tx,\forall x\in\mathbb{S}\setminus\mathbb{S}',\forall\mathbf{j}\in M^l, \tag{15}$$

with $\mathbb{S}'=\Pi_{\mathbb{S}}(L^{-1}\tilde{\mathbb{S}})$ and $\sigma^{n-1}(\mathbb{S}')=\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}\sigma^{n-1}(\tilde{\mathbb{S}})=\kappa(P)\tilde{\varepsilon}$.

iii) For a such given set $\mathbb{S}'$, we look for the largest sphere included in convhull $(\mathbb{S}\setminus\mathbb{S}')$. By homogeneity of the system, this sphere is centered at the origin, and we denote by α its radius. By (15), l-traces initialized on $\mathbb{S}\setminus\mathbb{S}'$ will be in $(\gamma^*)^l\mathbb{B}$:

$$\bar{\mathbf{A}}_l\left(\mathbb{S}\setminus\mathbb{S}'\right)\subset(\gamma^*)^l\mathbb{B},\ \forall\mathbf{j}\in M^l.$$

Property 4.1 also implies: $\bar{\mathbf{A}}_l\left(\text{convhull}\left(\mathbb{S}\setminus\mathbb{S}'\right)\right)\subset\text{convhull}\left(\bar{\mathbf{A}}_l(\mathbb{S}\setminus\mathbb{S}')\right)\subset(\gamma^*)^l\mathbb{B},\forall\mathbf{j}\in M^l$. Since $\alpha\mathbb{S}\subset\text{convhull}\left(\mathbb{S}\setminus\mathbb{S}'\right)$, then $\forall\mathbf{j}\in M^l$, $\bar{\mathbf{A}}_l\left(\alpha\mathbb{S}\right)=\alpha\bar{\mathbf{A}}_l\left(\mathbb{S}\right)\subset\text{convhull}\left(\bar{\mathbf{A}}_l(\mathbb{S}\setminus\mathbb{S}')\right)\subset(\gamma^*)^l\mathbb{B}$, which implies that $\bar{\mathbf{A}}_l(\mathbb{S})\subset\frac{(\gamma^*)^l}{\alpha}\mathbb{B}$.

We consider $\delta(\varepsilon):=\inf_{X\in\mathcal{X}_\varepsilon}\sup\{r:r\mathbb{B}\subset\text{convhull}\left(\mathbb{S}\setminus X\right)\}$ where $\mathcal{X}_\varepsilon=\{X\subset\mathbb{S}:\sigma^{n-1}(X)\le\varepsilon\}$. For a given ε, $\delta(\varepsilon)$ is the smallest value of α over all possible sets $\mathbb{S}'$ of measure ε. Proposition 14 tells us that in our case this δ is reached when $\mathbb{S}'$ is the union of two symmetric spherical caps of measure $\frac{\tilde{\varepsilon}\kappa(P)}{2}$. Moreover, for a given spherical cap of area measure ε, we can compute a closed form expression of the radius of the corresponding largest ball. We have $\delta(\varepsilon)=\sqrt{1-I^{-1}\left(2\varepsilon;\frac{n-1}{2},\frac{1}{2}\right)}$ (see [technical report for a detailed proof]).

iv) Finally, we even have by homogeneity that (12) implies

$$(\mathbf{A}_l x)^TP\mathbf{A}_l x\le(\gamma^*)^{2l}x^TPx,\forall x\in\mathbb{S}\setminus\Pi_{\mathbb{S}}(L^{-1}\tilde{\mathbb{S}}),\forall\,\mathbf{j}\in M^l$$

which, combined with the latter point, gives us $\rho(\mathcal{M}^l)\le\frac{\gamma^l}{\delta(\frac{\tilde{\varepsilon}\kappa(P)}{2})}$, hence $\rho(\mathcal{M})\le\frac{\gamma^*}{\sqrt[l]{\delta\left(\frac{\tilde{\varepsilon}\kappa(P)}{2}\right)}}$.

4.3 Main Theorem

We are now ready to prove our main theorem by putting together all the above pieces. For a given level of confidence β, we prove that the upper bound $\gamma^*(\omega_N)$, which is valid solely on finitely many observations, is in fact a true upper bound, at the price of increasing it by the factor $\frac{1}{\sqrt[l]{\delta(\beta,\omega_N)}}$. Moreover, as expected, this factor gets smaller as we increase N and decrease β.

Theorem 17 *Consider an n-dimensional switched linear system as in (2) and a uniform random sampling $\omega_N\subset Z_l$, where $N\ge\frac{n(n+1)}{2}+1$. Let $\gamma^*(\omega_N)$ be the optimal solution to (6). Then, for any given $\beta\in[0,1)$ and $\eta>0$, we can compute $\delta(\beta,\omega_N)$, such that with probability at least β we have:*

$$\rho\le\frac{\gamma^*(\omega_N)(1+\eta)}{\sqrt[l]{\delta(\beta,\omega_N)}},$$

where $\lim_{N\to\infty} \delta(\beta, \omega_N) = 1$ *with probability* 1.

Proof. By definition of $\gamma^*(\omega_N)$ we have

$$(\mathbf{A}_l x)^T P \mathbf{A}_l x \leq (\gamma^*(1+\eta))^{2l} x^T P x, \ \forall (x, \mathbf{j}) \in \omega_N.$$

for some $P \succ 0$. Then, by taking $\beta = 1 - I(1-\varepsilon; N-d, d+1)$ in Theorem 10, with probability at least β the following holds:

$$(\mathbf{A}_l x)^T P \mathbf{A}_l x \leq ((\gamma^*(1+\eta))^{2l}\, x^T P x, \\ \forall\ (x, \mathbf{j}) \in Z_l \setminus V(\omega_N).$$

with $\mu_l(V(\omega_N)) \leq \varepsilon$ and $\varepsilon(\beta, N) = 1 - I^{-1}(1-\beta; N-d, d+1)$. Thanks to Corollary 11, we even have

$$(\mathbf{A}_l x)^T P \mathbf{A}_l x \leq ((\gamma^*(1+\eta))^{2l}\, x^T P x, \forall x \in \mathbb{S} \backslash \tilde{\mathbb{S}}, \forall \mathbf{j} \in M^l$$

with $\tilde{\mathbb{S}} = \pi_{\mathbb{S}}(V)$ and $\sigma^{n-1}(\tilde{\mathbb{S}}) \leq \varepsilon m^l$. Then by Theorem 16, we can compute $\delta(\beta, \omega_N) = \delta(\varepsilon'(\beta, N))$, where

$$\varepsilon'(\beta, N) = \frac{1}{2}\varepsilon(\beta, N) m^l \kappa(P) \tag{16}$$

such that with probability at least β we have:

$$\rho \leq \frac{\gamma^*(\omega_N)(1+\eta)}{\sqrt[l]{\delta(\beta, \omega_N)}},$$

which completes the proof of the first part of the theorem. Note that, the ratio $\frac{1}{2}$ introduced in the expression of ε' is, as we have already mentioned, due to the homogeneity of the system.

Let us prove now that $\lim_{N\to\infty} \delta(\beta, \omega_N) = 1$ with probability 1. We recall that, $\delta(\beta, \omega_N) = \alpha\left(\varepsilon(\beta, \omega_N) m^l \kappa(P(\omega_N))\right)$. We start by showing that $\kappa(P(\omega_N))$ is uniformly bounded in N. The optimization problem $\mathrm{Opt}(\omega_N)$ given in (6), with $\gamma^*(\omega_N)$ replaced by $\gamma^*(Z_l)(1+\frac{\eta}{2})$ is strictly feasible, and thus admits a finite optimal value K for some solution $P_{\eta/2}$. Note that, $\lim_{N\to\infty} \gamma^*(\omega_N) = \gamma^*(Z_l)$ with probability 1. Thus, for large enough N, $\gamma^*(\omega_N)(1+\eta) > \gamma^*(Z_l)(1+\frac{\eta}{2})$. This also means that, for large enough N, $\mathrm{Opt}(\omega_N)$ admits $P_{\eta/2}$ as a feasible solution and thus the optimal value of $\mathrm{Opt}(\omega_N)$ is bounded by K. In other words, $\lambda_{\max}(P(\omega_N)) \leq K$. Moreover, since $\lambda_{\max}(P(\omega_N)) \geq 1$, we also have $\det(P(\omega_N)) \geq 1$, which means that

$$\kappa(P(\omega_N)) = \sqrt{\frac{\lambda_{\max}(P(\omega_N))^n}{\det(P(\omega_N))}} \leq \sqrt{K^n} \tag{17}$$

We next show that for a fixed $\beta \in [0,1)$, $\lim_{N\to\infty} \varepsilon(\beta, N) = 0$. Note that, $\varepsilon(\beta, N)$ is intrinsically defined by $1-\beta = \sum_{j=0}^{d} \binom{N}{j} \varepsilon^j (1-\varepsilon)^{N-j}$. We can then upper bound the term $1-\beta$ as in:

$$1 - \beta \leq (d+1) N^d (1-\varepsilon)^{N-d}. \tag{18}$$

We prove $\lim_{N\to\infty} \varepsilon(\beta, N) = 0$ by contradiction. Assume that $\lim_{N\to\infty} \varepsilon(\beta, N) \neq 0$. This means that, there exists some $c > 0$ such that $\varepsilon(\beta, N) > c$ infinitely often. Then, consider the subsequence N_k such that $\varepsilon(\beta, N_k) > c,\ \forall k$. Then, by (18) we have for any $k \in \mathbb{N}$:

$$1-\beta \leq (d+1) N_k^d (1-\varepsilon)^{N_k - d} \leq (d+1) N_k^d (1-c)^{N_k - d}.$$

Note that $\lim_{k\to+\infty} (d+1) N_k^d (1-c)^{N_k-d} = 0$, which implies that there exists a k' such that:

$$(d+1) N_{k'}^d (1-c)^{N'_k - d} < 1 - \beta,$$

which is a contradiction. Therefore, we must have $\lim_{N\to\infty} \varepsilon(\beta, N) = 0$. Putting this together with (17), we get: $\lim_{N\to\infty} m^l \kappa(P(\omega_N)) \varepsilon(\beta, \omega_N) = 0$. By the continuity of the function I^{-1} this also implies: $\lim_{N\to\infty} \alpha\left(\varepsilon(\beta, \omega_N) m^l \kappa(P(\omega_N))\right) = 1$.

5 Experimental Results

5.1 Algorithm

Raphael's other upper bound: There is no conservatism in multiplying ε by m^l, as in the worst case this really happens: if $\varepsilon = 1/m$, it is well possible that one mode is totally forgotten, and that our δ must be equal to zero (because then all points are bad points). However, when multiplying by $\kappa(P)$, we are conservative, because this bound is exactly reached only at a single point on the ellipsoid. So, when we derive an upper bound on the size of the bad points on the sphere after changing the coordinates, we could also derive a "lower bound on the size of the good points". This gives a second upper bound on the size of the bad points by taking the complement. This one can never be larger than 1. Algorithm: maybe explain we take for the bounds for the bisection on γ, 0 and U, with U the max of the norms $y^T y$, and that for some precision η, we run the algorithm at most $\lceil \log_2(U/\eta) \rceil$.

5.2 Experimental results

To do: add Ayca's new plots

We illustrate our technique on a two-dimensional switched system with 4 modes. We fix the confidence level, $\beta = 0.92$, and compute the lower and upper bounds on the JSR for $N := 15 + 15k$, $k \in \{0, \ldots, 23\}$, according to Theorem 9 and Theorem 17, respectively. We illustrate the average performance of our algorithm over 10 different runs in Fig. 4 and Fig. 5. Fig. 4 shows

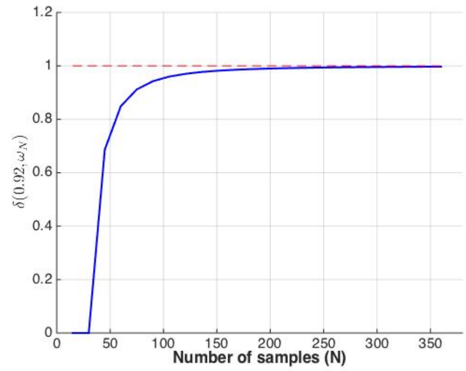

Fig. 4. Evolution of δ with increasing N.

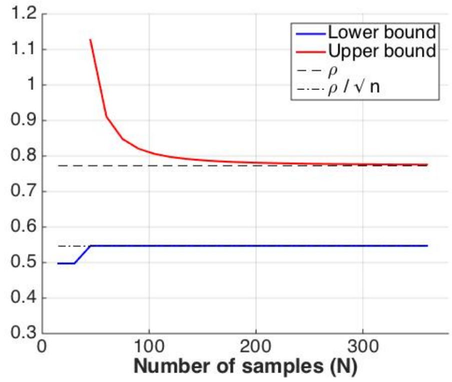

Fig. 5. Evolution of the upper and lower bounds on the JSR with increasing N, for $\beta = 0.92$.

the evolution of $\delta(\beta, N)$ as N increases. We illustrate that δ converges to 1 as expected. In Fig. 5, we plot the upper bound and lower bound for the JSR of the system computed by Theorem 17 and Theorem 9, respectively. To demonstrate the performance of our technique, we also provide the JSR approximated by the JSR toolbox [21], which turns out to be 0.7727. Note that, the plot for the upper bound starts from $N = 45$. This is due to the fact for $N = 15$, and $N = 30$, $\delta(\beta, \omega_N) = 0$, hence it is not possible to compute a nontrivial upper bound for these small values of N. As can be seen, the upper bound approaches to a close vicinity of the real JSR with approximately 200 samples. In addition, the gap between the upper and lower bound converges to a multiplicative factor of $\frac{\rho}{\sqrt{n}}$ as expected.

Note that, if we increase the dimension of the switched system, the convergence of δ to 1 will become much slower. We confirmed this via experiments up to dimension $n = 6$. For example, for dimension $n = 4$, it took $N = 5,000$ to $N = 10,000$ points to reach $\delta = 0.9$. We nevertheless observe convergence of the upper bound to $\rho(\mathcal{M})$, and convergence of the lower bound to $\frac{\rho(\mathcal{M})}{\sqrt{n}}$. The gap between these two limits is $\frac{\rho}{\sqrt{n}}$ and could be improved by considering a more general class of common Lyapunov functions, such as those that can be described by sum-of-squares polynomials [17]. We leave this for future work.

Finally, we randomly generate $10,000$ test cases with systems of dimension between 2 and 7, number of modes between 2 and 5, and size of samples N between 30 and 800. We take $\beta = 0.92$ and we check if the upper bound computed by our technique is greater than the actual JSR of the system. We get 9873 positive tests, out of $10,000$, which gives us a probability of 0.9873 of the correctness of the upper bound computed. Note that, this probability is significantly above the provided β. This is expected, since our techniques are based on worst-case analysis and thus fairly conservative.

5.3 Networked Control System

We now consider a linear time-invariant control system given as $x_{k+1} = Ax_k + Bu_k$, with control law of the form $u_k = Kx_k$. Matrices A, B and K are unknown. The open-loop system is unstable with eigenvalues at $\{0.45, 1.1\}$. The controller stabilizes the system by bringing its eigenvalues to $\{0.8, -0.7\}$.

The control input is transmitted over a wireless communication channel that is utilized by ℓ users, including the controller. The communication between the users and the recipients is performed based on the IEEE 802.15.4 MAC layer protocol [?], which is used in some of the proposed standards for control over wireless, e.g., WirelessHART [?]. This MAC layer integrates both guaranteed slots and contention based slots. In this example, we consider a beacon-enabled mode of the MAC protocol. In this setup, a centralized control user periodically synchronizes and configures all the users. This period is named Beacon Interval. This interval is divided into two subintervals: active and inactive period. The active period is divided into 6 slots. The first 2 slots correspond to the contention access period (CAP), and the next 4 slots correspond to the collusion free period (CFP). In the CAP, the users can only send their message if the channel is "idle" with carrier-sense multiple access with collision avoidance (CSMA/CA). In the CFP however, each user has guaranteed time slots, during which there are no packet losses. In our example, the third and fourth slots are designated for the controller, while the fifth and sixth slots are allocated to the other users. Finally, during the inactive period, all users enter a low-power mode to save energy. We illustrate the overall structure of this communication protocol in Fig. 6. We now want to decide whether the resulting closed-loop network control system is stable by simulating it starting from different initial conditions.

Note that, the closed-loop dynamics of the underlying system when the controller is active is $A_c = A + BK$. Then, we can model the overall networked control system by the switched linear system $x_{k+1} = \bar{A}x_k$, where $\bar{A} \in \mathcal{M}$ and $\mathcal{M} = \{A^2A_c^2A^4, A_cAA_c^2A^4, AA_c^3A^4, A_c^4A^4\}$.

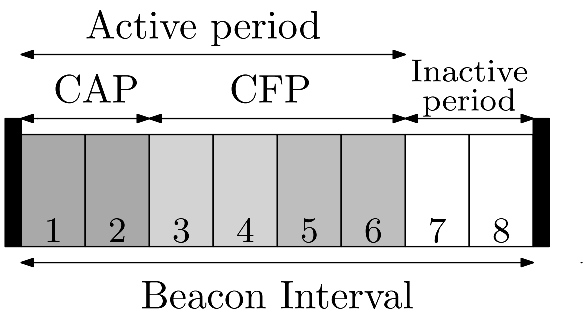

Fig. 6. The time allocation structure of the modified IEEE 802.15.4 MAC layer.

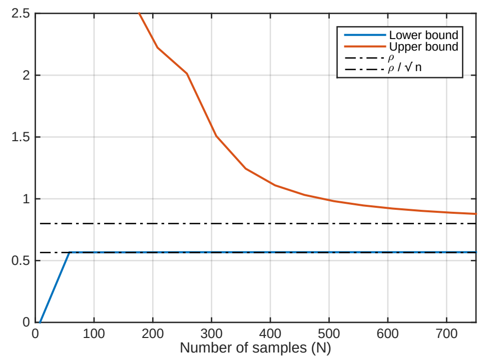

Fig. 7. The evolution of the computed upper and lower bounds on the JSR with respect to the number of simulations collected from the networked control system.

Note that, each element of $\mathcal{M}$ corresponds to a different utilization of the CFP by the users. For example, the mode defined by $A_cAA_c^2A^4$ is active when the first slot in the CFP is assigned to the controller and the second slot is assigned to the other users. We assume that all of the users using the channel have an equal probability of being assigned to a time slot during the CFP. Therefore, the probability of each mode in $\mathcal{M}$ being active is $\left\{\frac{1}{(\ell-1)^2}, \frac{1}{\ell(\ell-1)}, \frac{1}{(\ell-1)\ell}, \frac{1}{\ell^2}\right\}$, i.e., these modes will not be active with the same probability. Hence, we make use of Remark 6 and update our bounds accordingly. Fig. 5.3 shows the computed bounds. As can be seen, approximately after 500 samples, the upper bound on the JSR drops below 1, which lets us decide that the given closed-loop networked control system is stable, with probability 0.95.

6 Conclusions

In this paper, we investigated the question of how one can conclude stability of a dynamical system when a model is not available and, instead, we have randomly generated state measurements. Our goal is to understand how the observation of well-behaved trajectories *intrinsically* implies stability of a system. It is not surprising that we need some standing assumptions on the system, in order to allow for any sort of nontrivial stability certificate solely from a finite number of observations.

The novelty of our contribution is twofold: First, we use as standing assumption that the unknown system can be described by a switching linear system. This assumption covers a wide range of systems of interest, and to our knowledge no such "black-box" result has been available so far on switched systems. Second, we apply powerful techniques from chance constrained optimization. The application is not obvious, and relies on geometric properties of linear switched systems.

We believe that this guarantee is quite powerful, in view of the hardness of the general problem. In the future, we plan to investigate how to generalize our results to more complex or realistic systems. We are also improving the numerical properties of our technique by incorporating sum-of-squares optimization, and relaxing the sampling assumptions on the observations.

Remark 18 *In the above discussion, we introduce the concept of 'l-step CQLF', and showed that it allows to refine the initial* $1/\sqrt(n)$ *approximation provided by the CQLF method. In the switching systems literature, there are other techniques for refining this approximation, as for instance replacing the LMIs in Theorem* **??** *by Sum-Of-Squares (SOS) constraints; see [18] or [11, Theorem 2.16]. It seems that the concept of* l*-step CQLF is better suited for our purpose, as we briefly discuss below. We leave for further work a more systematic analysis of the behaviours of the different refining techniques.*

Acknowledgements

Acknowledgements, NSF etc.

References

[1] A. Balkan, P. Tabuada, J. V. Deshmukh, X. Jin, and J. Kapinski. Underminer: A framework for automatically identifying non-converging behaviors in black box system models. In *Proceedings of the Thirteenth ACM International Conference on Embedded Software (EMSOFT)*, EMSOFT '16, pages 1–10, 2016.

[2] F. Blanchini, G. Fenu, G. Giordano, and F. A. Pellegrino. Model-free plant tuning. *IEEE Transactions on Automatic Control*, 62(6):2623–2634, 2017.

[3] V. D. Blondel and J. N. Tsitsiklis. Complexity of stability and controllability of elementary hybrid systems. *Automatica*, 35(3):479–489, 1999.

[4] R. Bobiti and M. Lazar. A delta-sampling verification theorem for discrete-time, possibly discontinuous systems. In *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, pages 140–148. ACM, 2015.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[6] Calafiore G. C. Random convex programs. *SIAM Journal on Optimization*, 20(6):3427–3464, 2010.

[7] G. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[8] M. C. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.

[9] P. S. Duggirala, S. Mitra, and M. Viswanathan. Verification of annotated models from executions. In *Proceedings of the Eleventh ACM International Conference on Embedded Software*, EMSOFT '13, pages 26:1–26:10, 2013.

[10] Z. Huang and S. Mitra. Proofs from simulations and modular annotations. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, HSCC '14, pages 183–192. ACM, 2014.

[11] R. Jungers. The joint spectral radius: theory and applications. In *Lecture Notes in Control and Information Sciences*, volume 385. Springer-Verlag, 2009.

[12] J. Kapinski, J. V. Deshmukh, S. Sankaranarayanan, and N. Arechiga. Simulation-guided lyapunov analysis for hybrid dynamical systems. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, pages 133–142. ACM, 2014.

[13] A. Kozarev, J. Quindlen, J. How, and U. Topcu. Case studies in data-driven verification of dynamical systems. In *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, HSCC '16, pages 81–86. ACM, 2016.

[14] F. Lauer. On the complexity of switching linear regression. *Automatica*, 74:80–83, 2016.

[15] S. Liu, D. Liberzon, and V. Zharnitsky. On almost lyapunov functions for non-vanishing vector fields. In *Proceedings of the 55th IEEE Conference on Decision and Control (CDC)*, pages 5557–5562, 2016.

[16] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2006.

[17] A. Papachristodoulou and S. Prajna. On the construction of lyapunov functions using the sum of squares decomposition. In *Proceedings of 41st IEEE Conference on Decision and Control (CDC)*, pages 3482–3487, 2002.

[18] P. A. Parrilo and A. Jadbabaie. Approximation of the joint spectral radius using sum of squares. *Linear Algebra and its Applications*, 428(10):2385–2402, 2008.

[19] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., 3rd edition, 1987.

[20] U. Topcu, A. Packard, and P. Seiler. Local stability analysis using simulations and sum-of-squares programming. *Automatica*, 44(10):2669–2675, 2008.

[21] G. Vankeerberghen, J. Hendrickx, and R. Jungers. Jsr: A toolbox to compute the joint spectral radius. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, HSCC '14, pages 151–156. ACM, 2014.