Data Driven Stability Analysis of Black-box Switched Linear Systems

Ayca Balkan[1] and Raphael Jungers[2] and Joris Kenanian[1] and Paulo Tabuada[1]

***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.

I. 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. In this paper, we focus on switched systems, but 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: Given N pairs, (x_1, y_1), (x_2, y_2), $\ldots$, (x_N, y_N) belonging to the behavior of the system (1), (i.e., $y_k = f(k, x_k)$) for some k), what can we say about the stability of the system (1)? For the rest of the paper, we use the term *black-box* to refer to models where we do not have access to the model, i.e., to f, yet we can indirectly learn information about f by observing 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: 1) Even when the function f is known, in general, stability analysis is a very difficult problem [3]; 2) Identification can potentially introduce approximation errors, and can be algorithmically hard as well. Again, this is the case for switched systems [13]. 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 [2], [1], [9], [8]. In particular, the initial idea behind this paper was influenced by the recent efforts in [19], [12], 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 [19] 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 2D linear system, the connection between these two beliefs is nontrivial. In fact, one

This work is supported by ??

[1]A. Balkan, J.Kenanian, and Paulo Tabuada are with University of California, Los Angeles.

[2]R. Jungers is with ICTEAM Institute, at Université catholique de Louvain.

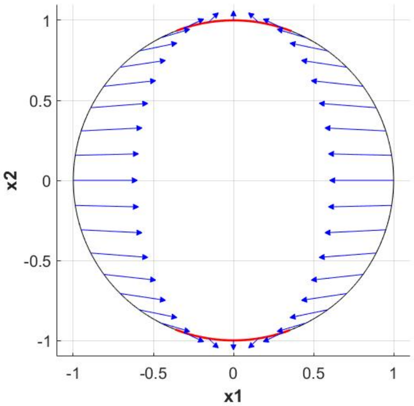

Fig. 1. A simple dynamics and the level set of an "almost Lyapunov function". Even though this function decreases at almost all points in its level set, almost all trajectories diverge to infinity.

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. For example, the system

$$x^+ = \begin{bmatrix} 0.14 & 0 \\ 0 & 1.35 \end{bmatrix} x,$$

admits a Lyapunov function candidate on the unit circle except on the two red areas shown in Fig. 1. Moreover, the size of this "violating set" can be made arbitrarily small by changing the magnitude of the unstable eigenvalue. Nevertheless, the only trajectories that do not diverge to infinite are those starting on the stable eigenspace that has zero measure.

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: identifying and deciding stability from of arbitrary switched linear systems is NP-hard [11].

The stability of switched systems is closely related to the *joint spectral radius* (JSR) of the matrices modeling the dynamics in each mode. Deciding stability amounts to deciding whether the JSR is less than one [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], [16], [7]), 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 lower bound, for large N.

The organization of the paper is as follows: In Section II, we introduce the notations and provide the necessary background in stability of switched systems. In Section III, we present a deterministic lower bound for the JSR. Section IV presents the main contribution of this paper where we provide a probabilistic stability guarantee for a given switched system, based on finite observations. We experimentally demonstrate the performance of the presented techniques in Section V and conclude in Section VI, while hinting at our related future work.

II. Preliminaries

A. Notation

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 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 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 $\Pi_{\mathbb{S}} := x/\|x\|$.

For an ellipsoid centered at the origin, and for any of its subsets $\mathcal{A}$, the *sector* defined by $\mathcal{A}$ is the subset

$$\{t\mathcal{A},\ t \in [0,1]\} \subset\ \mathbb{R}^n.$$

We denote by $E_P^{\mathcal{A}}$ the sector induced by $\mathcal{A} \subset E_P$. In the particular case of the unit sphere, we instead write $\mathbb{S}^{\mathcal{A}}$. We can notice that $E_P^{E_P}$ is the volume in $\mathbb{R}^n$ defined by E_P: $E_P^{E_P} = \{x \in \mathbb{R}^n : x^T P x \leq 1\}$.

We consider in this work the classical unsigned and finite uniform spherical measure on $\mathbb{S}$, denoted by σ^{n-1}. It is associated to $\mathcal{B}_{\mathbb{S}}$, the spherical Borelian σ-algebra, and is derived from the Lebesgue measure λ. We have $\mathcal{B}_{\mathbb{S}}$ defined by $\mathcal{A} \in \mathcal{B}_{\mathbb{S}}$ if and only if $\mathbb{S}^{\mathcal{A}} \in \mathcal{B}_{\mathbb{R}^n}$. The spherical measure σ^{n-1} is defined by

$$\forall\ \mathcal{A} \in \mathcal{B}_{\mathbb{S}},\ \sigma(\mathcal{A}) = \frac{\lambda(\mathbb{S}^{\mathcal{A}})}{\lambda(\mathbb{B})}.$$

In other words, the spherical measure of a subset of the sphere is related to the Lebesgue measure of the sector of the unit ball it induces. Notice that $\sigma^{n-1}(\mathbb{S}) = 1$. Since $P \in \mathcal{S}^n_{++}$, it can be written in its Choleski form

$$P = L^T L, \tag{2}$$

where L is an upper triangular matrix. Note that, L^{-1} maps the elements of $\mathbb{S}$ to E_P. Then, we define the measure on the ellipsoid σ_P on the σ-algebra $\mathcal{B}_{E_P} := L^{-1}\mathcal{B}_{\mathbb{S}}$, where $\forall\, \mathcal{A} \in \mathcal{B}_{E_P}$, $\sigma_P(\mathcal{A}) = \sigma^{n-1}(L\mathcal{A})$.

For $m \in \mathbb{N}_{>0}$, we denote by M the set $M = \{1, 2, \ldots, m\}$. Set M is provided with the classical σ-algebra associated to the finite sets: $\Sigma_M = \wp(M)$, where $\wp(M)$ is the set of subsets of M. We consider the uniform measure μ_M on (M, Σ_M).

We define $Z = \mathbb{S} \times M$ as the Cartesian product of the unit sphere and M. We denote the product σ-algebra $\mathcal{B}_{\mathbb{S}} \bigotimes \Sigma_M$ generated by $\mathcal{B}_{\mathbb{S}}$ and Σ_M: $\Sigma = \sigma(\pi_{\mathbb{S}}^{-1}(\mathcal{B}_{\mathbb{S}}), \pi_M^{-1}(\Sigma_M))$, where $\pi_{\mathbb{S}} : Z \to \mathbb{S}$ and $\pi_M : Z \to M$ are the standard projections. On this set, we define the product measure $\mu = \sigma^{n-1} \otimes \mu_M$. Note that, μ is a uniform measure on Z and $\mu(Z) = 1$.

B. 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:

$$x_{k+1} = f(k, x_k), \tag{3}$$

with $f(k, x_k) = A_{\tau(k)} x_k$ and switching sequence $\tau : \mathbb{N} \to M$. There are two important properties of switched linear systems that we exploit in this paper.

Property 2.1: Let $\xi(x, k, \tau)$ denote the state of the system (3) at time k starting from the initial condition x and with switching sequence τ. The dynamical system (3) is homogeneous: $\xi(\gamma x, k, \tau) = \gamma \xi(x, k, \tau)$.

Property 2.2: The dynamics given in (3) is convexity-preserving, meaning that for any set of points $X \subset \mathbb{R}^n$ we have:

$$f(\text{convhull } (X)) \subset \text{convhull } (f(X)).$$

The joint spectral radius of the set of matrices $\mathcal{M}$ closely relates to the stability of the system (3) and is defined as follows:

Definition 2.1: [10] Given a finite set of matrices $\mathcal{M} \subset \mathbb{R}^{n \times n}$, its *joint spectral radius* (JSR) is given by

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

Property 2.3 (Corollary 1.1, [11]): Given a finite set of matrices $\mathcal{M}$, the corresponding switched dynamical system is stable if and only if $\rho(\mathcal{M}) < 1$.

Property 2.4 (Proposition 1.3, [11]): 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: $\forall \gamma > 0$, $\rho\left(\mathcal{M}/\gamma\right) = \mathcal{M}/\gamma$).

III. A Deterministic Lower Bound for the JSR

We start by computing a lower bound for ρ which is based on the following theorem from the switched linear systems literature.

Theorem 3.1: [10, Theorem 2.11] For any finite set of matrices such that $\rho(\mathcal{M}) < \frac{1}{\sqrt{n}}$, there exists a Common Quadratic Lyapunov Function (CQLF) for $\mathcal{M}$, that is, a $P \succ 0$ such that:

$$\forall\ A \in \mathcal{M},\ A^T P A \preceq P.$$

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 3.2: [10, Prop. 2.8] Consider a finite set of matrices $\mathcal{M}$. If there exist a $\gamma \geq 0$ and $P \succ 0$ such that

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

then $\rho(\mathcal{M}) \leq \gamma$.

Note that the smaller γ is, the tighter is the upper bound we get on $\rho(\mathcal{M})$. Therefore, we can consider, in particular, the optimal solution γ^* of the following optimization problem:

$$\begin{array}{ll} \min_{\gamma, P} & \gamma \\ \text{s.t.} & (Ax)^T P (Ax) \leq \gamma^2 x^T P x,\ \forall\ A \in \mathcal{M}, \forall\, x \in \mathbb{R}^n, \\ & P \succ 0. \end{array} \tag{4}$$

Even though this upper bound is more difficult to obtain in a black-box setting where only a finite number of observations are available, in this section we leverage Theorem 3.1 in order to derive a straight-forward lower bound.

The following theorem shows that the existence of a CQLF for (3) can be checked by considering N pairs $(x_i, j_i) \in \mathbb{R}^n \times M$, where $i \in \{1, \dots N\}$. Recall that in our setting, we assume that we observe pairs of the form (x_k, x_{k+1}), but we do not observe the mode applied to the system during this time step.

Theorem 3.3: For a given uniform sampling:

$$\omega_N := \{(x_1, j_1), (x_2, j_2), \dots, (x_N, j_N)\} \subset \mathbb{R}^n \times M,$$

let $W_{\omega_N} = \{(x_1, y_1), ..., (x_N, y_N)\}$ be the corresponding available observations, which satisfy

$$y_i = A_{j_i} x_i\ \forall (x_i, y_i) \in W_{\omega_N}.$$

Also let $\gamma^*(\omega_N)$ be the optimal solution of the following optimization problem:

$$\begin{array}{ll} \min_{P} & \gamma \\ \text{s.t.} & (y_i)^T P (y_i) \leq \gamma^2 x_i^T P x_i,\ \forall\, i : 1 \leq i \leq N. \\ & P \succ 0 \end{array} \tag{5}$$

Then, we have:

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

Note that, (5) can be efficiently solved by semidefinite programming and bisection on the variable γ (see [5]).

Proof: Let $\epsilon > 0$. By definition of γ^*, there exists no matrix $P \in \mathcal{S}^n_{++}$ such that:

$$(Ax)^T P (Ax) \leq (\gamma^*(\omega_N) - \epsilon)^2 x^T P x, \quad \forall x \in \mathbb{R}^n, \forall\, A \in \mathcal{M}.$$

By Property 2.4, this means that there exists no CQLF for the scaled set of matrices $\frac{\mathcal{M}}{(\gamma^*(\omega_N) - \epsilon)}$. Then, using Theorem 3.1, we conclude:

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

■

IV. A Probabilistic Stability Guarantee

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 CQLF which is valid with probability at least β. Note that, the existence of a CQLF implies $\rho \leq 1$ due to Theorem 3.2. Even though the solution of the optimization problem in (4) provides a CQLF, solving this problem as stated is not possible since it involves infinitely many constraints. Nevertheless, we show that the solution of the optimization problem (5) allows us to not only compute a lower bound, but also a (probabilistic) upper bound on the JSR. We now analyze the relationship between the solutions of the optimization problem (4) and the following optimization problem with finitely many constraints:

$$\begin{array}{ll} \min_P & \lambda_{\max}(P) \\ \text{s.t.} & (A_j x)^T P (A_j x) \leq ((1+\eta)\gamma^*(\omega_N))^2 x^T P x, \\ & \qquad\qquad \forall (x,j) \in \omega_N \subset Z, \\ & P \succeq I, \end{array} \tag{6}$$

where $Z := \mathbb{S} \times M$, $\eta > 0$, and $\gamma^*(\omega_N)$ is the optimal solution to the optimization problem (5). Recall that ω_N is an N-uniform random sampling of the set Z. Note that, instead of the set $\mathbb{R}^n$ we sample on the unit sphere $\mathbb{S}$. This is due to Property 2.1, since it implies that it is sufficient to show the decrease of a CQLF on a set enclosing the origin, e.g., $\mathbb{S}$.

For the rest of the discussion, we refer to the optimization problem (6) by Opt(ω_N). We denote its optimal solution by $P(\omega_N)$. We drop the explicit dependence of P on ω_N when it is clear from the context. There are a few points that are worth noting about (6). Firstly, due to Property 2.1, we are can replace the constraint $P \succ 0$ with the constraint $P \succeq I$. Moreover, for reasons that will become clear later in the discussion, we chose the objective function as $\lambda_{\max}(P)$, instead of solving a feasibility problem in P. Lastly, the additional η factor is introduced to ensure strict feasibility of (6), which will be helpful in the following discussion.

The curious question whether the optimal solution of the sampled problem Opt(ω_N) 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 (6), 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 4.1 (adapted from Theorem 3.3, [6]): Let d be the dimension of Opt(ω_N) and $N \geq d+1$. Consider the optimization problem Opt(ω_N) given in (6), where ω_N is a uniform random sampling of the set Z. Then, for all $\epsilon \in (0,1)$ the following holds:

$$\mu^N\{\omega_N \in Z^N : \mu(V(\omega_N)) \leq \epsilon\} \geq 1 - \sum_{j=0}^{d} \binom{N}{j} \epsilon^j (1-\epsilon)^{N-j}, \tag{7}$$

where μ^N denotes the product probability measure on Z^N, and $V(\omega_N)$ is defined by

$$V(\omega_N) = \{z \in Z : (A_j z)^T P(\omega_N)(A_j z) > \gamma^* z^T P(\omega_N) z\},$$

i.e., it is the set[1] of constraints that are violated by the optimal solution of Opt(ω_N).

[1]Consider the function $h : z \mapsto (A_{\pi_M(z)}\pi_{\mathbb{S}}(z))^T P (A_{\pi_M(z)}\pi_{\mathbb{S}}(z)) - \gamma^2 \pi_S(z)^T P \pi_S(z)$. The constraint violation set is defined by $V = h^{-1}((0, +\infty))$. Due to continuity of h, this set is an element of Σ.

Theorem 4.1 states that the optimal solution of the sampled problem Opt(ω_N) violates no more than an ϵ fraction of the constraints in the original optimization problem (4) with probability β, where β goes to 1 as N goes to infinity.

The rest of this section has two important intermediate results leading us to our main theorem. In Proposition 4.2, we first show how to map the measure of the violated constraints on Z to the measure of violating constraints on the unit sphere, $\mathbb{S}$. This is thanks to the homogeneity of the dynamics stated in Property 2.1. By exploiting Property 2.2, we next show in Lemma 4.3 how one can compute an upper bound on the JSR by working on $\mathbb{S}$. We then tie these lemmas with Theorem 4.1 to prove the main result of this section.

Proposition 4.2: Let $\gamma \in \mathbb{R}_{>0}$. Consider a set of matrices $A \in \mathcal{M}$, and a matrix $P \succ 0$ satisfying:

$$(A_j x)^T P (A_j x) \leq \gamma^2 x^T P x, \ \forall\, (x,j) \in Z \setminus V, \tag{8}$$

for some $V \subset Z$ where $\mu(V) \leq \epsilon$. Then, by defining L as in (2) and $\bar{A}_j = L^{-1} A_j L$, one also has:

$$(\bar{A}_j x)^T (\bar{A}_j x) \leq \gamma^2 x^T x, \ \forall\, x \in \mathbb{S} \setminus \mathbb{S}', \forall\, j \in M,$$

for some $\mathbb{S}' \subset \mathbb{S}$ such that: $\sigma(\mathbb{S}') \leq m\epsilon\kappa(P)$, where

$$\kappa(P) = \sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}.$$

Proof: Recall that $V \in \Sigma$. Let $V_{\mathbb{S}} = \pi_{\mathbb{S}}(V)$ and $V_M = \pi_M(V)$. We know that Σ_M is the disjoint union of its 2^m elements $\{\mathcal{M}_i, i \in \{1, 2, \ldots 2^m\}\}$. Then V can be written as the disjoint union $V = \sqcup_{1 \leq i \leq 2^m}(\mathcal{S}_i, \mathcal{M}_i)$ where $\mathcal{S}_i \in \Sigma_{\mathbb{S}}$. We notice that $V_{\mathbb{S}} = \sqcup_{1 \leq i \leq 2^m} \mathcal{S}_i$, and

$$\sigma^{n-1}(V_{\mathbb{S}}) = \sum_{1 \leq i \leq 2^m} \sigma^{n-1}(\mathcal{S}_i).$$

We have

$$\begin{aligned} \mu(V) &= \mu(\sqcup_{1 \leq i \leq 2^m}(\mathcal{S}_i, \mathcal{M}_i)) = \sum_{1 \leq i \leq 2^m} \mu(\mathcal{S}_i, \mathcal{M}_i) \\ &= \sum_{1 \leq i \leq 2^m} \sigma^{n-1} \otimes \mu_M(\mathcal{S}_i, \mathcal{M}_i) \\ &= \sum_{1 \leq i \leq 2^m} \sigma^{n-1}(\mathcal{S}_i)\mu_M(\mathcal{M}_i). \end{aligned}$$

Note that we have $\min_{j \in M} \mu_M(\{j\}) = \frac{1}{m}$. Then since $\forall\, i$, $\mu_M(\mathcal{M}_i) \geq \frac{1}{m}$, we get:

$$\sigma^{n-1}(V_{\mathbb{S}}) \leq \frac{\mu(V)}{\frac{1}{m}} \leq m\epsilon. \tag{9}$$

This means that

$$(A_j x)^T P (A_j x) \leq \gamma^2 x^T P x, \ \forall\, x \in \mathbb{S} \setminus V_{\mathbb{S}}, \forall\, m \in M, \tag{10}$$

where $\sigma^{n-1}(V_{\mathbb{S}}) \leq m\epsilon$.

We then perform the change of coordinates defined by $L^{-1} \in \mathcal{L}(\mathbb{R}^n)$ which maps $\mathbb{S}$ to E_P, defined as in (2). We can then rewrite (10) in this new coordinates system as in:

$$(\bar{A}_j x)^T (\bar{A}_j x) \leq \gamma^2 x^T x, \ \forall\, x \in E_P \setminus L^{-1}(V_{\mathbb{S}}), \forall\, m \in M. \tag{11}$$

Due to the the homogeneity of the dynamics described in Property 2.1, this implies:

$$(\bar{A}_j x)^T(\bar{A}_j x) \leq \gamma^2 x^T x,\ \forall\, x \in \mathbb{S}\setminus\Pi_{\mathbb{S}}(L^{-1}(V_{\mathbb{S}})),\ \forall\, m \in M. \tag{12}$$

We now show how to relate $\sigma^{n-1}(V_{\mathbb{S}})$ to $\sigma^{n-1}(\Pi_{\mathbb{S}}(L^{-1}(V_{\mathbb{S}})))$. Consider $\mathbb{S}^{V_{\mathbb{S}}}$, the sector of $\mathbb{B}$ defined by $V_{\mathbb{S}}$. We denote $C := L^{-1}(\mathbb{S}^{V_{\mathbb{S}}})$ and $V' := \Pi_{\mathbb{S}}(L^{-1}(V_{\mathbb{S}}))$. We have $\Pi_{\mathbb{S}}(C) = V'$ and $\mathbb{S}^{V'} \subset \mathcal{H}_{1/\lambda_{\min}(L^{-1})}(C)$, where $\mathcal{H}$ is the homothety of ratio $1/\lambda_{\min}(L^{-1})$. This leads to:

$$\sigma^{n-1}(V') = \lambda(\mathbb{S}^{V'}) \leq \lambda(\mathcal{H}_{1/\lambda_{\min}(L^{-1})}(C)).$$

Then, the following holds:

$$\begin{aligned}
\sigma^{n-1}(V') &\leq \frac{1}{\lambda_{\min}(L^{-1})^n}\lambda(C)\\
&\leq \frac{1}{\lambda_{\min}(L^{-1})^n}\lambda(L^{-1}(\mathbb{S}^{V_{\mathbb{S}}}))\\
&= \frac{|\det(L^{-1})|}{\lambda_{\min}(L^{-1})^n}\lambda(\mathbb{S}^{V_{\mathbb{S}}}), &(13)\\
&= \sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}\sigma^{n-1}(V_{\mathbb{S}}) &(14)
\end{aligned}$$

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. [18]). Putting together (10), (12), and (14) we get the statement of the theorem where $\mathbb{S}' = \Pi_{\mathbb{S}}(L^{-1}(V_{\mathbb{S}}))$. ■

Lemma 4.3: Let $\epsilon \in (0,1)$ and $\gamma \in \mathbb{R}_{>0}$. Consider the set of matrices $\mathcal{M}$ and $A \in \mathcal{M}$ satisfying:

$$(A_j x)^T(A_j x) \leq \gamma x^T x, \quad \forall\, x \in \mathbb{S}\setminus\mathbb{S}', \forall\, j \in, M, \tag{15}$$

where $\mathbb{S}' \subset \mathbb{S}$ and $\sigma^{n-1}(\mathbb{S}') \leq \epsilon$, then we have:

$$\rho(\mathcal{M}) \leq \frac{\gamma}{\alpha(\epsilon)}$$

where $\alpha(\epsilon)$ is defined as in (20).

Proof: Note that, (15) implies that: $A_j(\mathbb{S}\setminus\mathbb{S}') \subset \gamma\mathbb{B}$. Using Property 2.2 this also implies:

$$A_j\text{convhull }(\mathbb{S}\setminus\mathbb{S}') \subset \text{convhull }(A_j(\mathbb{S}\setminus\mathbb{S}')) \subset \gamma\mathbb{B}.$$

Then, by Lemma 1.5 in Appendix C, we have:

$$A_j(\alpha(\epsilon)\mathbb{B}) \subset A_j(\text{convhull }(\mathbb{S}\setminus\mathbb{S}')) \subset \gamma\mathbb{B}, \quad \forall\, j \in M,$$

where $\alpha(\epsilon)$ is given in (24). Therefore, we get:

$$\alpha(\epsilon)A_j(\mathbb{B}) \subset \gamma\mathbb{B},$$

which implies that $\rho(\mathcal{M}) \leq \frac{\gamma}{\alpha(\epsilon)}$. ■

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 $\frac{1}{\delta(\beta,\omega_N)}$. Moreover, as expected, this factor gets smaller as we increase N and decrease β.

Theorem 4.4: Consider an n-dimensional switched linear system as in (3) and a uniform random sampling $\omega_N \subset Z$, where $N \geq \frac{n(n+1)}{2} + 1$. 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:

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

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

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

$$(A_j x)^T P(A_j x) \leq (\gamma^*(1+\eta))^2 x^T P x, \quad \forall\,(x,j) \in \omega_N$$

for some $P \succ 0$. Then, by rewriting Theorem 4.1 we also have:

$$\mu^N\left\{\omega_N \in Z^N : \mu(V(\omega_N)) \leq \epsilon\right\} \geq 1 - I(1-\epsilon; N-d, d+1), \tag{16}$$

where $I(\ell; a, b)$ is the regularized incomplete beta function. Let $\epsilon(\beta, N) = 1 - I^{-1}(1-\beta; N-d, d+1)$. Then, by Theorem 4.1, with probability at least β the following holds:

$$(A_j x)^T P(A_j x) \leq (\gamma^*(1+\eta))^2 x^T P x, \quad \forall (x,j) \in Z\setminus V.$$

By Theorem 4.2, this implies that with probability at least β the following also holds:

$$(\bar{A}_j x)^T(\bar{A}_j x) \leq \gamma^2 x^T x,\ \forall\, x \in \mathbb{S}\setminus\mathbb{S}', \forall\, j \in M,$$

for some $\mathbb{S}'$ where $\sigma^{n-1}(\mathbb{S}') \leq m\epsilon\kappa(P)$. Then, applying Lemma 4.3, we can compute

$$\delta(\beta,\omega_N) = \alpha(\epsilon'(\beta,N)),$$

where

$$\epsilon'(\beta,N) = \frac{1}{2}m\kappa(P)\epsilon(\beta,N) \tag{17}$$

such that with probability at least β we have:

$$\bar{A}_j\mathbb{B} \subset \frac{\gamma^*(\omega_N)(1+\eta)}{\delta(\beta,\omega_N)}\mathbb{B},\ \forall\, j \in M,$$

By Property 2.4, this means that with probability at least β:

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

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 due to the homogeneity of the system described in Property 2.1, which implies that $x \in V_{\mathbb{S}} \iff -x \in V_{\mathbb{S}}$. We refer the interested reader to Appendix D for the second part of this proof, namely showing that $\lim_{N\to\infty}\delta(\beta,\omega_N) = 1$. ■

V. Experimental Results

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 $N := 15 + 50k$, $k \in \{0, \ldots, 10\}$, according to Theorem 3.3 and Theorem 4.4, respectively. We illustrate the average performance of our algorithm over 10 different runs in Fig. 2 and Fig. 3. Fig. 2 shows the evolution of $\delta(\beta,N)$ as N

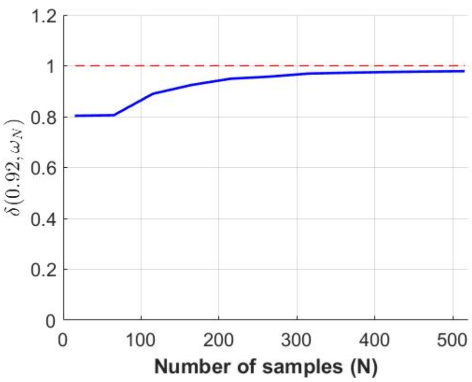

Fig. 2. Evolution of δ with increasing N.

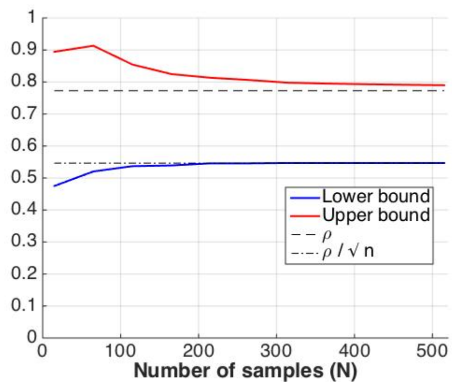

Fig. 3. Evolution of the upper and lower bounds on the JSR with increasing N, for $\beta = 0.92$.

increases. We illustrate that δ converges to 1 as expected. In Fig. 3, we plot the upper bound and lower bound for the JSR of the system computed by Theorem 4.4 and Theorem 3.3, respectively. To demonstrate the performance of our technique, we also provide the JSR approximated by the JSR toolbox [20], which turns out to be 0.7727. As can be seen, the upper bound approaches to a close vicinity of the real JSR with approximately 250 samples. In addition, the lower bound converges to $\frac{\rho}{\sqrt{n}}$ 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 $N = 5,000 - 10,000$ 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 of factor $\frac{\rho}{\sqrt{n}}$ 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].

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 techniques is greater than the actual JSR of the system. We get 9858 positive tests, out of $10,000$, which gives us a probability of 0.9858 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.

VI. 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 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.

Appendix

A. Notation and Background

Before proceeding to the main lemmas we use to prove Lemma 1.5, we first introduce some necessary definitions and related background.

Let d be a distance on $\mathbb{R}^n$. The distance between a set $X \subset \mathbb{R}^n$ and a point $p \in \mathbb{R}^n$ is $d(X, p) := \inf_{x \in X} d(x, p)$. Note that the map $p \mapsto d(X, p)$ is continuous on $\mathbb{R}^n$. Given a set $X \subset \mathbb{R}^n$, ∂X denotes the boundary of set X.

Definition 1.1: We define the *spherical cap* on $\mathbb{S}$ for a given hyperplane $c^T x = k$ as:

$$\mathcal{C}_{c,k} := \{x \in \mathbb{S} : c^T x > k\}.$$

Remark 1.1: Consider the spherical caps $\mathcal{C}_{c,k_1}$ and $\mathcal{C}_{c,k_2}$ such that $k_1 > k_2$, then we have:

$$\sigma^{n-1}(\mathcal{C}_{c,k_1}) < \sigma^{n-1}(\mathcal{C}_{c,k_2}).$$

Definition 1.2: A *supporting hyperplane* of a set $X \subset \mathbb{R}^n$ is a hyperplane $\{x : c^T x = k\}$ that has the following two properties:

- $X \subset \{x : c^T x \leq k\}$ or $X \subset \{x : c^T x \geq k\}$.
- $X \cap \{x : c^T x = k\} \neq \emptyset$.

Remark 1.2: [5] Consider a convex set $X \subset \mathbb{R}^n$. For every $x \in \partial X$, there exists a supporting hyperplane containing x. Moreover, if X is a smooth manifold, then this supporting hyperplane is unique.

Remark 1.3: The distance between the point $x = 0$ and the hyperplane $c^T x = k$ is $\frac{|k|}{\|c\|}$.

We now define the function $\Delta : \wp(\mathbb{S}) \to [0,1]$ as:

$$\Delta(X) := \sup\{r : r\mathbb{B} \subset \text{convhull } (\mathbb{S} \setminus X)\}. \tag{18}$$

Note that, $\Delta(X)$ can be rewritten as:

$$\Delta(X) = d(\partial\text{convhull } (\mathbb{S} \setminus X), 0). \tag{19}$$

Lemma 1.1: Consider the spherical cap $\mathcal{C}_{c,k}$. We have:

$$\Delta(\mathcal{C}_{c,k}) = \min\left(1, \frac{|k|}{\|c\|}\right).$$

Proof: Note that:

$$\text{convhull } (\mathbb{S} \setminus X) = \left\{x \in \mathbb{B} : c^T x \le k\right\}.$$

Then the following equalities hold:

$$\begin{aligned}
\Delta(X) &= d(\partial\text{convhull } (\mathbb{S} \setminus X), 0) \\
&= \min(d(\partial\mathbb{B}, 0), d(\partial\{x : c^T x \le k\}, 0)) \\
&= \min(d(\mathbb{S}, 0), d(\{x : c^T x = k\}, 0)) \\
&= \min\left(1, \frac{|k|}{\|c\|}\right).
\end{aligned}$$

■

Corollary 1.2: Consider the spherical caps $\mathcal{C}_{c,k_1}$ and $\mathcal{C}_{c,k_2}$ such that $k_1 \le k_2$. Then we have:

$$\Delta(\mathcal{C}_{c,k_1}) \le \Delta(\mathcal{C}_{c,k_2}).$$

B. Preliminary Results

Lemma 1.3: For any set $X \subset \mathbb{S}$, there exist c and k such that $\mathcal{C}_{c,k}$ satisfies: $\mathcal{C}_{c,k} \subset X$, and $\Delta(\mathcal{C}_{c,k}) = \Delta(X)$.

Proof: Let $\tilde{X} :=$ convhull $(S \backslash X)$. Since d is continuous and the set $\partial\tilde{X}$ is compact, there exists a point $x^* \in \partial\tilde{X}$, such that:

$$\Delta(X) = d(\partial X_S, 0) = \min_{x \in \partial\tilde{X}} d(x, 0) = d(x^*, 0).$$

Next, consider the supporting hyperplane of $\tilde{X}$ at x^*, which we denote by $\left\{x : c^T x = k\right\}$. Note that this supporting hyperplane is a supporting hyperplane of the ball $(\Delta(X)\mathbb{B})$ at x^* since we have:

$$\partial(\Delta(X)\mathbb{B}) \subset \partial\tilde{X} \subset \left\{x : c^T x = k\right\}.$$

By Remark 1.2, this implies that $\left\{x : c^T x = k\right\}$ is in fact the unique supporting hyperplane at x^*. Then we have:

$$\Delta(X) = d(x^*, 0) = d(\{x : c^T x = k\}, 0) = \min\left(1, \frac{|k|}{\|c\|}\right).$$

Now, consider the spherical cap $\mathcal{C}_{c,k}$. Then, by Lemma 1.1 we have $\Delta(\mathcal{C}_{c,k}) = \min\left(1, \frac{|k|}{\|c\|}\right)$. Therefore, $\Delta(X) = \Delta(\mathcal{C}_{c,k})$.

We next show $\mathcal{C}_{c,k} \subset X$. We prove this by contradiction. Assume $x \in \mathcal{C}_{c,k}$ and $x \notin X$. Note that, if $x \notin X$, then $x \in \mathbb{S} \setminus X \subset$ convhull $(\mathbb{S} \setminus X)$. Since $x \in \mathcal{C}_{c,k}$, we have $c^T x > k$. But due to the fact that $x \in$ convhull $(\mathbb{S} \setminus X)$, we also have $c^T x \le k$, which leads to a contradiction. Therefore, $\mathcal{C}_{c,k} \subset X$. ■

Proposition 1.4: Let $\mathcal{X}_\epsilon = \{X \subset \mathbb{S} : \sigma^{n-1}(X) = \epsilon\}$. Then, for any $\epsilon \in (0,1)$, the function $\Delta(X)$ attains its minimum over X_ϵ for some X which is a spherical cap.

Proof: We prove this via contradiction. Assume that there exists no spherical cap in $\mathcal{X}_\epsilon$ such that $\Delta(X)$ attains its minimum. This means there exists an $X^* \in \mathcal{X}_\epsilon$, where X^* is not a spherical cap and $\arg\min_{X \in \mathcal{X}_\epsilon}(\Delta(X)) = X^*$. By Lemma 1.3, we can construct a spherical cap $\mathcal{C}_{c,k}$ such that $\mathcal{C}_{c,k} \subset X^*$ and $\mathcal{C}_{c,k} = \Delta(X^*)$. Note that, we further have $\mathcal{C}_{c,k} \subsetneq X^*$, since X^* is assumed not to be a spherical cap. This means that, there exists a spherical cap $\sigma^{n-1}(\mathcal{C}_{c,k})$ such that $\sigma^{n-1}(\mathcal{C}_{c,k}) < \epsilon$.

Then, the spherical cap $\mathcal{C}_{c,\tilde{k}}$ with $\sigma^{n-1}(\mathcal{C}_{c,\tilde{k}}) = \epsilon$, satisfies $\tilde{k} < k$ by Remark 1.1. This implies

$$\Delta(\mathcal{C}_{c,\tilde{k}}) < \Delta(\mathcal{C}_{c,k}) = \Delta(X^*)$$

by Corollary 1.2. Therefore, $\Delta(\mathcal{C}_{c,\tilde{k}}) < \Delta(X^*)$. This is a contradiction since we initially assumed that $\Delta(X)$ attains its minimum over $\mathcal{X}_\epsilon$ at X^*. ■

C. Main Lemma

Lemma 1.5: Let $\epsilon \in (0,1]$. Then, we can compute $\alpha(\epsilon)$ satisfying:

$$\alpha(\epsilon) = \sup_{\substack{\mathbb{S}' \subset \mathbb{S}: \\ \sigma^{n-1}(\mathbb{S}') \le \epsilon}} \{r : r\mathbb{B} \subset \text{convhull } (\mathbb{S} \setminus \mathbb{S}')\}. \tag{20}$$

where we recall that $\mathbb{B}$ (respectively $\mathbb{S}$) denote the unit ball (respectively unit sphere).

Proof: Let the function $\Delta(X)$ be defined as in (18). Then by Lemma 1.4 we know that:

$$\Delta(X_\epsilon) \ge \Delta(\mathcal{C}_{c,k}), \tag{21}$$

for some spherical cap $\mathcal{C}_{c,k} \subset \mathbb{S}$, where $\sigma^{n-1}(\mathcal{C}_{c,k}) = \epsilon$. It is known (see e.g. [14]) that the area of such $\mathcal{C}_{c,k}$, is given by the equation:

$$\sigma^{n-1}(\mathcal{C}_{c,k}) = \frac{\pi^{d/2}}{\Gamma[\frac{d}{2}]} I\left(1 - \Delta(X_\epsilon)^2; \frac{d-1}{2}, \frac{1}{2}\right), \tag{22}$$

where I is the regularized incomplete beta function. Since, $\sigma^{n-1}(X_\epsilon) \le \epsilon$, we get the following set of equations:

$$\begin{aligned}
\frac{\epsilon\Gamma[\frac{d}{2}]}{\pi^{d/2}} &\le I\left(1 - \Delta(X_\epsilon)^2; \frac{d-1}{2}, \frac{1}{2}\right) \\
1 - \Delta(\mathcal{C}_{c,k})^2 &\le I^{-1}\left(\frac{\epsilon\Gamma(\frac{d}{2})}{\pi^{d/2}}; \frac{d-1}{2}, \frac{1}{2}\right) \\
\Delta(\mathcal{C}_{c,k})^2 &\ge 1 - I^{-1}\left(\frac{\epsilon\Gamma(\frac{d}{2})}{\pi^{d/2}}; \frac{d-1}{2}, \frac{1}{2}\right)
\end{aligned} \tag{23}$$

The inequalities (23) and (21) imply the inclusion given in (20), where

$$\alpha(\epsilon) = \sqrt{1 - I^{-1}\left(\frac{\epsilon\Gamma(\frac{d}{2})}{\pi^{d/2}}; \frac{d-1}{2}, \frac{1}{2}\right)}. \tag{24}$$

■

D. *Proof of* $\lim_{N\to\infty} \delta(\beta, \omega_N) = 1$

Note that, the optimization problem (6), with $\gamma^*(\omega_N)$ replaced by $\gamma^*(Z) + \frac{\eta}{2}$ is strictly feasible, and thus admits a finite optimal value K for some solution $P_{\eta/2}$ Note that, $\lim_{N\to\infty} \gamma^*(\omega_N) = \gamma^*(Z)$. Thus, for N large enough, (6) admits $P_{\eta/2}$ as a feasible solution, and thus its optimal value 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

$$\kappa(P(\omega_N)) = \sqrt{\frac{\lambda_{\max}(P(\omega_N))^n}{\det(P(\omega_N))}} \leq \sqrt{K^n}. \qquad (25)$$

We now show that for a fixed $\beta \in (0,1)$ $\lim_{N\to\infty} \epsilon(\beta, N) = 0$. Note that, $\epsilon(\beta, N)$ is intrinsically defined by the following equation:

$$1 - \beta = \sum_{j=0}^{d} \binom{N}{j} \epsilon^j (1-\epsilon)^{N-j}.$$

We can then upper bound the term $1 - \beta$ as in:

$$1 - \beta \leq (d+1)N^d(1-\epsilon)^{N-d}. \qquad (26)$$

We prove $\lim_{N\to\infty} \epsilon(\beta, N) = 0$ by contradiction. Assume that $\lim_{N\to\infty} \epsilon(\beta, N) \neq 0$. This means that, there exists some $c > 0$ such that $\epsilon(\beta, N) > c$ infinitely often. Then, consider the subsequence N_k such that $\epsilon(\beta, N_k) > c$, $\forall\, k$. Then, by (26) we have:

$$1-\beta \leq (d+1)N_k^d(1-\epsilon)^{N_k-d} \leq (d+1)N_k^d(1-c)^{N_k-d}\,\forall\, k \in \mathbb{N}.$$

Note that $\lim_{k\to\infty}(d+1)N_k^d(1-c)^{N_k-d} = 0$. Therefore, 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 $\lim_{N\to\infty} \epsilon(\beta, N) = 0$.

Putting together this with (25), we get:

$$\lim_{N\to\infty} m\kappa(P(\omega_N))\epsilon(\beta, \omega_N) = 0.$$

By the continuity of the function I^{-1} this implies: $\lim_{N\to\infty} \alpha(m\kappa(P(\omega_N))\epsilon(\beta, \omega_N)) = 1$.

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 *2016 International Conference on Embedded Software (EMSOFT)*, pages 1–10, Oct 2016.

[2] F. Blanchini, G. Fenu, G. Giordano, and F. A. Pellegrino. Model-free plant tuning. *IEEE Transactions on Automatic Control*, PP(99):1–1, 2016.

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

[4] Ruxandra Bobiti and Mircea 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*, HSCC '15, pages 140–148, New York, NY, USA, 2015. ACM.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

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

[7] 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.

[8] Parasara Sridhar Duggirala, Sayan Mitra, and Mahesh 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, Piscataway, NJ, USA, 2013. IEEE Press.

[9] Zhenqi Huang and Sayan 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, New York, NY, USA, 2014. ACM.

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

[11] Raphaël Jungers. *The joint spectral radius*, volume 385 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 2009. Theory and applications.

[12] James Kapinski, Jyotirmoy V. Deshmukh, Sriram Sankaranarayanan, and Nikos Arechiga. Simulation-guided lyapunov analysis for hybrid dynamical systems. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, HSCC '14, pages 133–142, New York, NY, USA, 2014. ACM.

[13] Fabien Lauer. On the complexity of switching linear regression. *arXiv preprint arXiv:1510.06920*, 2015.

[14] S. Li. Concise formulas for the area and volume of a hyperspherical cap. *Asian Journal of Mathematics & Statistics*, 4:66–70, 2011.

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

[16] Arkadi Nemirovski and Alexander 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*, pages 3482–3487, Las Vegas, Nevada USA, December 2002.

[18] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[19] Ufuk Topcu, Andrew Packard, and Peter Seiler. Local stability analysis using simulations and sum-of-squares programming. *Automatica*, 44(10):2669 – 2675, 2008.

[20] Guillaume Vankeerberghen, Julien Hendrickx, and Raphaël M. 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, New York, NY, USA, 2014. ACM.