Data Driven Stability Analysis of Black-box Switched Linear Systems with Probabilistic Guarantees

***Abstract*— We explore the general problem of deciding stability of a "black-box" dynamical system, that is a system whose model equations are not known. The only information in our possession is a set of observations, that is, couples of vectors of the type $(x(k), x(k+1))$. We adopt a probabilistic approach, and focus on switched systems, which are a widely used model for many complex systems, and are well-known to be hard to analyze, even if the equations of the model are known.**

We show that, for a given (randomly generated) set of observations, one can give a stability guarantee on the system, 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 a geometrical analysis, combining chance-constrained optimization theory with stability analysis tools 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 a more general class of dynamical systems as well. 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., $y_i = f(k, x_i)$) 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 dynamics, yet we can observe f by observing couples of points (x_k, x_{k+1}) as defined in (1).

A potential approach to this problem is first to 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 input-output pairs and bypass the identification phase: 1) Even when the function f is known, in general, the stability analysis is a very difficult problem [3], Any reference recommendations here?. 2) Identification can potentially introduce approximation errors, and can be algorithmically hard as well. Again, this is the case for switched systems [12]. A fortiori, the combination of these two steps in an efficient and robust way seems far from obvious.

The initial idea behind this paper was influenced by the recent efforts in [18], [11], and [4] on using simulation traces to find Lyapunov functions for systems with known dynamics.Which paper of Sayan Mitra are we talking about? 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 [18] and [11], 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 many 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,

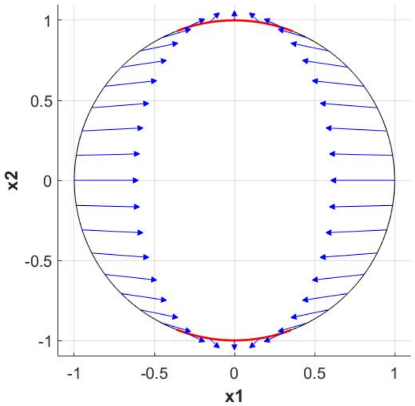

Fig. 1. A simple dynamics and the level set of an 'almost Lyapunov function'. Even though the function decreases at almost all points in the state space, all trajectories are unstable.

one can easily construct an example with one stable and one unstable eigenvalues for which even though almost all trajectories diverge to the infinity, it is possible to construct a Lyapunov function candidate whose level sets are contracting everywhere except on a small set. 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 Figure I. Moreover, the size of this "violating set" can be arbitrarily small based on the magnitude of the unstable eigenvalue.

In this paper, we take a step to close this gap by focusing on switched linear systems. Identifying and deciding the stability of arbitrary switched linear systems is NP-hard [10]. Aside from their theoretical value, switched systems are a popular model for many complex systems, as for instance dynamics with (known or unknown) varying parameters. These parameters can model internal properties of the dynamical system such as uncertainties, look-up tables, values in a discrete register as well as exogenous inputs provided by a controller in a closed-loop control system [14], [8].

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 [10]. 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.

We note that, aside from the stability analysis, black-box setting has been adopted by several researchers in the analysis and control of dynamical systems in the recent years. For example, in [2], the authors consider the tuning a black-box plant by choosing suitable inputs. In [1], by exciting a black-box system with different inputs, authors generate test cases to challenge the system with respect to a given design specification. Anything more here?

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, which are matrices $P \in \mathcal{S}^n$ such that $\forall\ x \in \mathbb{R}^n \setminus \{0\}$, $x^T P x > 0$, 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. 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 B (respectively 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 S by Π_{S}.

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

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

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 $\sigma^{n-1}(\mathrm{S}) = 1$. Since $P \in \mathcal{S}^n_{++}$, it can be written in its Choleski form $P = UDU^{-1}$, where D is the diagonal matrix of its eigenvalues and U is an orthogonal matrix (see e.g. [5]). Let

$$L := UD^{1/2}U^{-1}. \tag{2}$$

Note that, L^{-1} maps the elements of S to E_P. Then, we define the measure on the ellipsoid σ_P on the σ-algebra $\mathcal{B}_{E_P} := L^{-1}\mathcal{B}_{\mathrm{S}}$, 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 $Z = \mathrm{S} \times M$ as the Cartesian product of the unit sphere and M. We denote the product σ-algebra $\mathcal{B}_{\mathrm{S}} \bigotimes \Sigma_M$ generated by $\mathcal{B}_{\mathrm{S}}$ and Σ_M: $\Sigma = \sigma(\pi_{\mathrm{S}}^{-1}(\mathcal{B}_{\mathrm{S}}), \pi_M^{-1}(\Sigma_M))$, where $\pi_{\mathrm{S}} : Z \to \mathrm{S}$ 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 Linear Switched Systems

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

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

with switching sequence $\tau : \mathbb{N} \to M$. There are two important properties of linear switched 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 [9] Given a finite set of matrices $\mathcal{M} \subset \mathbb{R}^{n\times n}$, its *joint spectral radius* (JSR) is given by

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

Property 2.3 (Corollary 1.1, [10]): 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, [10]): 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: $\forall \gamma > 0$, $\rho\left(\frac{\mathcal{M}}{\gamma}\right) = \frac{\rho(\mathcal{M})}{\gamma}$).

III. A Deterministic Lower Bound for 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: [9, 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 the switched system as the following theorem formalizes:

Theorem 3.2: [9, 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 smaller the γ is, a tighter 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 $i \in \{1, \ldots N\}$. 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:

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

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

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

By Remark 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 it is, is not possible since it involves infinitely many constraints. Nevertheless, we are going to show that the solution of the Optimization problem (5) allows us to not only compute a lower bound, but also a (probabilistic) upper bound. We now analyze the relationship between the solutions of the optimization problem (4) and the following optimization problem with finitely many constraints:

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

where $Z := \mathrm{S} \times M$, $\eta > 0$, 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 S. This is thanks 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. 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 translate in our terms a classical result from random convex optimization.

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. If Opt(ω_N) satisfies the following assumptions:

1) When the problem Opt(ω_N) admits an optimal solution, this solution is unique.
2) Problem Opt(ω_N) is nondegenerate[1] with probability one.

Then, for all $\epsilon \in (0, 1]$ the following holds:

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

where

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

i.e., the set of constraints that are violated by the optimal solution of Opt(ω_N).

Before proceeding to the proof, we note that this theorem is an immediate application of Theorem 3.3 in [6]. Note that, the assumptions given in the statement of the theorem are of technical nature. That is, if any of the two does not hold for the optimization problem Opt(ω_N), it is possible to construct a slightly modified optimization problem for which they indeed hold and work with this modified optimization problem instead. We refer the interested reader to [6] for a more detailed discussion of such modification techniques.

Proof: Note that the optimization problem Opt(ω_N) can be written as:

$$\begin{aligned} \min_{P,t} \quad & t \\ \text{s.t.} \quad & g_{\gamma^*}(P, z) \leq 0, \forall\ z \in Z \end{aligned} \tag{8}$$

where $g_{\gamma^*}(P, z) = \max(g_1(P, z), g_2(P), g_3(P))$, and

$$\begin{aligned} g_1(P, z) &= (A_j z)^T P(A_j z) - {\gamma^*}^2 z^T Pz \\ g_2(P) &= \lambda_{\max}(-P) + 1. \\ g_3(P, t) &= \lambda_{\max}(P) - t. \end{aligned}$$

■

The objective function of (8) is linear while each constraint is convex in P for all $z \in Z$. Also note that, the set of decision variables is in $\mathbb{R}^{\frac{n(n+1)}{2}+1}$. Then, we can invoke Theorem 3.3 in [6] with the optimization problem (8) to conclude the statement of the theorem, with $d = \frac{n(n+1)}{2} + 1$.

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 that lead us to our main theorem. In Theorem 4.2, we first show how to map the measure of the violated constraints on Z to the measure of violating points on the unit sphere,

[1] Informally, the problem Optω_N is nondegenerate, when there are no redundant support constraints.

S. We next show via Lemma 4.4 how one can compute an upper bound on the JSR by working on S. We then tie this Lemma with Theorem 4.1 to prove the main result of this section.

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

$$(A_j x)^T P(A_j x) \leq \gamma^2 x^T P x,\ \forall (x,j) \in Z \setminus V, \tag{9}$$

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:

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

for some $\mathrm{S}' \subset \mathrm{S}$ such that:

$$\sigma(\mathrm{S}') \leq m\epsilon\kappa(P),$$

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

Proof: Note that $V \subset \Sigma$. Let $V_{\mathrm{S}} = \pi_{\mathrm{S}}(V)$ 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 $\mathcal{S}_i \in \Sigma_{\mathrm{S}}$. We notice that $V_{\mathrm{S}} = \sqcup_{1\leq i\leq 2^m}\mathcal{S}_i$, and

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

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:

$$\sigma^{n-1}(V_{\mathrm{S}}) \leq \frac{\mu(V)}{\frac{1}{m}} \leq m\epsilon. \tag{10}$$

This means that

$$(A_j x)^T P(A_j x) \leq \gamma^2 x^T P x,\ \forall x \in \mathrm{S} \setminus V_S, \forall m \in M, \tag{11}$$

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

We then perform the change of coordinates defined by $L^{-1} \in \mathcal{L}(\mathbb{R}^n)$ which maps S to E_P defined as in (2). We can then rewrite (11) in this new coordinates system as in:

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

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

$$(\bar{A}_j x)^T(\bar{A}_j x) \leq \gamma^2 x^T x,\ \forall x \in \mathrm{S}\setminus\Pi_S(L^{-1}(V_S)), \forall m \in M. \tag{13}$$

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

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

which means the following holds:

$$\begin{aligned}
\sigma^{n-1}(V') &\leq \frac{1}{\lambda_{\min}(L^{-1})^n}\lambda(C) \\
&\leq \frac{1}{\lambda_{\min}(L^{-1})^n}\lambda(L^{-1}(S^{V_{\mathrm{S}}})) \\
&= \frac{|\det(L^{-1})|}{\lambda_{\min}(L^{-1})^n}\lambda(S^{V_S}), \qquad (14) \\
&= \sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}\sigma^{n-1}(V_S) \qquad (15)
\end{aligned}$$

where (14) 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. [17]). Putting together (11), (15), and (13), we get the statement of the theorem where $S' = \Pi_S(L^{-1}(V_S))$. ■

Lemma 4.3: Let $\epsilon \in (0, 1]$. Then, we can compute $\alpha(\epsilon)$ satisfying:

$$\alpha(\epsilon) = \sup_{\substack{\mathrm{S}'\subset\mathrm{S}:\\ \sigma^{n-1}(\mathrm{S}')\leq\epsilon}} \{r : r\mathrm{B} \subset \text{convhull } (\mathrm{S} \setminus \mathrm{S}'). \tag{16}$$

where we recall that B (respectively S) denote the unit ball (respectively sphere) centered at the origin.

Proof: See Appendix C. ■

Lemma 4.4: Let $\epsilon \in (0, 1]$ and $\gamma \in \mathbb{R}_{>0}$. Consider the set of matrices and $A \in \mathcal{M}$ satisfying:

$$(A_j x)^T(A_j x) \leq \gamma x^T x, \quad \forall x \in \mathrm{S} \setminus \mathrm{S}', \forall j \in, M, \tag{17}$$

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

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

where $\alpha(\epsilon)$ is defined as in (16).

Proof: Note that, (17) implies that: $A_j(\mathrm{S} \setminus \mathrm{S}') \subset \gamma\mathrm{B}$. Using Property 2.2 this also implies:

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

By Lemma 4.3 we have:

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

where $\alpha(\epsilon)$ is defined as in (25). Therefore, we get:

$$\alpha(\epsilon)A_j(\mathrm{B}) \subset \gamma\mathrm{B}.$$

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

Theorem 4.5: Consider an n-dimensional switching 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:

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

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

$$\mathbb{P}^N \{\mu(V(\omega_N)) \le \epsilon\} \ge 1 - I(1-\epsilon; N-d, d+1), \quad (18)$$

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:

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

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

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

for some S' where $\sigma^{n-1}(S') \le m\epsilon\kappa(P)$. Then, applying Lemma 4.4, we can compute

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

where

$$\epsilon'(\beta, N) = m\kappa(P)\epsilon(\beta, N) \quad (19)$$

such that with probability at least β we have:

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

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

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

which completes the proof of the first part of the theorem. 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 condifence level, $\beta = 0.92$ and compute both upper and lower bounds on the JSR for $N := 15 + 50k$, $k \in \{0, \ldots, 10\}$. We demonstrate the average performance of our algorithm over 10 different runs in Figure 2 and Figure 3. Figure 2 demonstrates the evolution of $\delta(\beta, N)$ as N increases. We observe that δ converges to 1 as expected. In Figure 3, we plot the upper bound and lower bound for the JSR of the system computed by Theorem 4.5 and Theorem 3.3, respectively. To demonstrate the performance of our technique, we also provide the JSR approximated by the JSR toolbox [19], 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.

We next test our algorithm with a 4-dimensional system, with 5 modes. We observe a slower convergence of δ as well as upper and lower bounds on ρ. [to add plots here, but they are not super nice: we do not see the convergence to 1 of δ: even after 10000 points, δ is below 0.85].

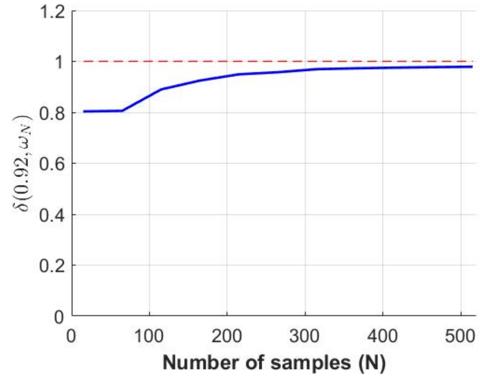

Fig. 2. Evolution of δ along N.

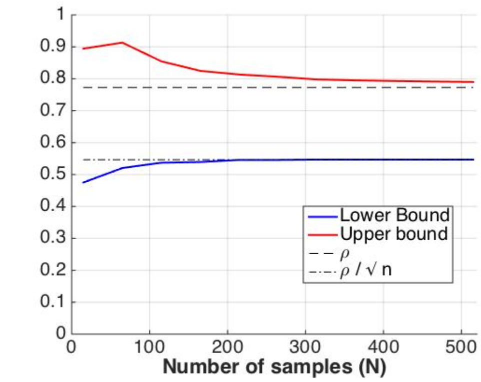

Fig. 3. Evolution of the bounds along N.

Finally, we randomly generate $10,000$ 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 xxx positive tests, out of $10,000$, which gives us a probability of ?? 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 conservative.

VI. Conclusions

In this work, we have investigated the question of how one can conclude stability of a system, by only observing trajectories, without having any mathematical description of the dynamics ruling the system. Our goal, motivated by both practical and theoretical considerations, was to avoid combining identification of the system with classical stability analysis techniques. Indeed, in real-world applications, often the true equations describing the dynamics might be extremely complex or nonstandard, or might even not be well defined. Moreover, identification of such complex systems is typically hard, and it is not clear how observation or computation errors would propagate in such a two-step strategy.

For these reasons, we aim at understanding how the observation of well-behaved trajectories *intrinsically* implies stability of a system. As argued in the introduction, it is easily seen that some standing assumption should be made on the system, in order to allow for any sort of nontrivial certificate, solely from a finite number of observations. The novelty of our contribution is twofold: First, we use as such a standing assumption that the system is a switching linear system. This assumption covers a wide range of systems of interest nowadays, and to our knowledge no 'black-box stability' result was available so far on switching systems. Second, we apply powerful techniques from chance constrained optimization. The application is not obvious, and relies on geometric properties of linear switching systems. We believe that this guarantee is quite powerful, in view of the hardness of the general problem, and in the future, we plan to investigate how to generalize it 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 4.3, we first introduce the necessary preliminary 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 $X \subset \mathbb{R}^n$, we denote ∂X denotes the boundary of set X.

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

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

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 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 $\Delta : \wp(\mathrm{S}) \to [0,1]$ as:

$$\Delta(X) := \sup\{r : r\mathrm{B} \subset \text{convhull } (\mathrm{S} \setminus X)\}. \tag{20}$$

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

$$\Delta(X) = d(\partial\text{convhull } (\mathrm{S} \setminus X), 0). \tag{21}$$

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:

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

Then the following equalities hold:

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

■

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

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

B. Preliminary Results

Lemma 1.3: For any set $X \subset \mathrm{S}$, 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 $\tilde{X} := \text{convhull } (S \setminus X)$. 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 $(\Delta(X)\mathrm{B})$ at x^* since we have:

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

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 $x \in \mathrm{S} \setminus X \subset \text{convhull } (\mathrm{S} \setminus X)$. Since $x \in \mathcal{C}_{c,k}$, we have $c^T x > k$. But due to the fact that $x \in \text{convhull } (\mathrm{S} \setminus X)$, we also have $c^T x \leq k$, which leads to a contradiction. Therefore, $\mathcal{C}_{c,k} \subset X$. ■

Proposition 1.4: Let $\mathcal{X}_\epsilon = \{X \subset \mathrm{S} : \sigma^{n-1}(X) = \epsilon\}$. 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. Proof of Lemma 4.3

Proof: Let the function $\Delta(X)$ be defined as in (20). Then by Lemma 1.4 we know that:

$$\Delta(X_\epsilon) \geq \Delta(\mathcal{C}_{c,k}), \tag{22}$$

for some spherical cap $\mathcal{C}_{c,k} \subset \mathrm{S}$, where $\sigma^{n-1}(\mathcal{C}_{c,k}) = \epsilon$. It is known (see e.g. [13]) that the area of such $\mathcal{C}_{c,k}$, is given by the equation:

$$\sigma^{n-1}(\mathcal{C}_{c,k}) = \frac{\pi^{d/2}}{\Gamma[\frac{d}{2}]} I\left(1 - \Delta(X_\epsilon)^2; \frac{d-1}{2}, \frac{1}{2}\right), \tag{23}$$

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

$$\begin{aligned} \frac{\epsilon\Gamma[\frac{d}{2}]}{\pi^{d/2}} &\leq I\left(1 - \Delta(X_\epsilon)^2; \frac{d-1}{2}, \frac{1}{2}\right) \\ 1 - \Delta(\mathcal{C}_{c,k})^2 &\leq I^{-1}\left(\frac{\epsilon\Gamma(\frac{d}{2})}{\pi^{d/2}}; \frac{d-1}{2}, \frac{1}{2}\right) \\ \Delta(\mathcal{C}_{c,k})^2 &\geq 1 - I^{-1}\left(\frac{\epsilon\Gamma(\frac{d}{2})}{\pi^{d/2}}; \frac{d-1}{2}, \frac{1}{2}\right) \end{aligned} \tag{24}$$

The inequalities (24) and (22) imply the inclusion given in (16), where

$$\alpha(\epsilon) = \sqrt{1 - I^{-1}\left(\frac{\epsilon\Gamma(\frac{d}{2})}{\pi^{d/2}}; \frac{d-1}{2}, \frac{1}{2}\right)}. \tag{25}$$

■

D. Proof of $\lim_{N\to\infty} \delta(\beta, \omega_N) = 1$ in Theorem 4.5

Recall that $\delta(\beta, \omega_N) = \alpha(m\kappa(P(\omega_N)\epsilon(\beta, N))$. We prove the limit of $\delta(\beta, \omega_N)$ in two parts. We first prove that $\kappa(P(\omega_N))$ is uniformly bounded in N. We then show that $\lim_{N\to\infty} \epsilon(\beta, N) = 0$. Then, the result of the theorem follows since $\lim_{\epsilon\to 0} \alpha(\epsilon) = 1$, which is immediate from the definition of $\alpha(\epsilon)$.

Lemma 1.5: We consider a sampling $\omega(N)$, where $N \in \mathbb{N}_{>0}$ and $N \geq d$. We assume that ω_N is sequential, i.e.,

$$\forall\ d \leq N_1 < N_2,\ \omega_{N_1} \subset \omega_{N_2}.$$

Let $P(\omega_N)$ be the optimal solution to the optimization problem Opt(ω_N). Then, $\kappa(P(\omega_N))$ is uniformly bounded in N.

Proof: We first define the following optimization problem:

$$\begin{aligned} \min_P \quad & \lambda_{\max}(P) \\ \text{s.t.} \quad & (A_j x)^T P (A_j x) \leq (1+\eta)\gamma^2 x^T P x,\ \forall (x, j) \in \omega_N, \\ & P \succeq I, \end{aligned} \tag{26}$$

where we denote its optimal solution by $\lambda_{\max}(\gamma, \omega_N)$. Note that, for all $t \in \mathbb{Z}$ such that $K < t \leq N$ we have $\gamma^*(\omega_t) \leq \gamma^*(\omega_N)$. Also note that,

$$\lambda_{\max}(\gamma^*(\omega_N), \omega_N) \leq \lambda_{\max}(\gamma^*(\omega_t), \omega_N).$$

But note that, there exists a $c > 0$ such that $\lambda_{\max}(\gamma^*(\omega_t), \omega_N) < c$ since the problem (26) is strictly feasible for any γ such that $\gamma \leq \gamma^*$. This implies: $\lambda_{\max}(\gamma^*(\omega_t), \omega_N) \leq c$. Since $\det(P(\omega_N)) \geq 1$, this means

$$\kappa(P(\omega_N)) = \sqrt{\frac{\lambda_{\max}(P(\omega_N))^n}{\det(P(\omega_N))}} \leq \sqrt{c^n},$$

which completes the proof of this lemma. ■

Note that, even though we proved that $\kappa(P(\omega_N))$ is uniformly bounded for the special case of sequential sampling, a similar probabilistic result can be proved even when new sample points are chosen for each value of N.

We now show that for a fixed $\beta \in (0, 1]$ $\lim_{N\to\infty} \Phi(\beta, N) = 0$. Note that, this is equivalent to showing that for a fixed $\beta \in (0, 1]$, $\lim_{N\to\infty} \epsilon(\beta, N) = 0$, where $\epsilon(\beta, N)$ is also 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 $1 - \beta$ as in:

$$1 - \beta \leq (d+1)N^d(1-\epsilon)^{N-d}. \tag{27}$$

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 (27) 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, we have

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

which is a contradiction. Therefore, we must have $\lim_{N\to\infty} \epsilon(\beta, N) = 0$. Putting together this with Lemma 1.5, we get: $\lim_{N\to\infty} m\kappa(P(\omega_N))\epsilon(\beta, \omega_N) = 0$, which implies 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] S.S. Ge and Z. Sun. *Switched linear systems: Control and design*. Communications and control engineering series. Springer, 2005.

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

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

[11] 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.

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

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

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

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

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

[19] 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.