Data Driven Stability Analysis of Black-box Switched Linear Systems

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.

1 Introduction

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

More formally, we consider a dynamical system as in:

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

where, $x_k \in X$ is the state and $k \in \mathbb{N}$ is the time index. 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), ..., (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 [?]; 2) Identification can potentially introduce approximation errors, and can be algorithmically hard as well. Again, this is the case for switched systems [?]. 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 [?, ?, ?, ?]. In particular, the initial idea behind this paper was influenced by the recent efforts in [?, ?], and [?] 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 [?] and [?], or via sampling-based techniques as in [?]. This also relates to almost-Lyapunov functions introduced in [?], which presents a relaxed notion of stability proved via Lyapunov functions decreasing everywhere except on a small set. Note that, since we do not

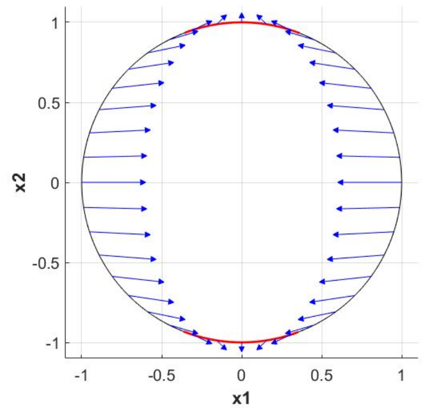

Fig. 1: A simple dynamics and the level set of an “almost Lyapunov function”. Even though this function decreases at almost all points in its level set, almost all trajectories diverge to infinity.

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 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 infinity 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 [?].

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 [?]. 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 [?, ?, ?]), and provides an upper bound on the JSR with a user-defined confidence level. As N increases, this bound gets tighter. Moreover, with a closed form expression, we characterize what is the exact trade-off between the tightness of this bound and the number of samples. In order to understand the quality of our upper bound, the algorithm also provides a deterministic lower bound. Finally, we provide an asymptotic guarantee on the gap between the upper and the lower bound, for large N.

The organization of the paper is as follows: In Section 2, we introduce the notations and provide the necessary background in stability of switched systems. In Section 3, we present a deterministic lower bound for the JSR. Section 4 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 5 and conclude in Section 6, while hinting at our related future work.

2 Preliminaries

2.1 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 $\mu(Z) = 1$.

2.2 Stability of Switched Linear Systems

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

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

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:*

$$f(convhull\ (X)) \subset convhull\ (f(X)).$$

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 (from [?])**.** *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, [?])**.** *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, [?])**.** *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: $\forall\gamma > 0$, $\rho\left(\mathcal{M}/\gamma\right) = \mathcal{M}/\gamma$).

3 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. *[?, 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 [?]), and they constitute a stability guarantee for switched systems as we formalize next.

Theorem 3.2. *[?, 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 straightforward lower bound.

The two theorems above provide us with a *converse Lyapunov result:* if there exists a CQLF, then our system is stable. If, on the contrary, there is no such stability guarantee, one may conclude a lower bound on the JSR. By combining these two results, one obtains an approximation algorithm for the JSR: the upper bound γ^* obtained above is within an error factor of $1/\sqrt(n)$ of the true value. It turns out that one can still refine this technique, in order to improve the error factor, and asymptotically get rid of it. This is a well known technique for the 'white-box' computation of the JSR, which we summarize in the following corollary:

Corollary 3.3. *For any finite set of matrices such that $\rho(\mathcal{M}) < \frac{1}{\sqrt[2l]{n}}$, there exists a Common Quadratic Lyapunov Function (CQLF) for*

$$\mathcal{M}^l := \{A_{i_1}, \ldots, A_{i_l} : A_i \in \mathcal{M}\},$$

that is, a $P \succ 0$ such that:

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

Proof. It is easy to see from the definition of the JSR that

$$\rho(\mathcal{M}^l) = \rho(\mathcal{M})^l.$$

Thus, applying Theorem 3.1 to $\mathcal{M}^l$, one directly obtains the corollary. □

We are now able to proceed with the main result of this section, which leverages the fact that necessary conditions for the existence of a CQLF for (3) can be obtained by considering a finite number of pairs $(x_i, x_{i+l}) \in \mathbb{R}^n$. Recall that in our setting, we assume that we observe pairs of the form (x_k, x_{k+1}) (or (x_k, x_{k+l}) for some $l \in \mathbb{N}$) but we do not observe the mode applied to the system during this time step.

In the next theorem, we provide a lower bound based on these observations, whose accuracy depends on the 'horizon' l.

Theorem 3.4. *For an arbitrary $l \in \mathbb{N}$, and a given uniform sampling:*

$$\omega_N := \{(x_1, j_{1,1}, \dots, j_{1,l}), (x_2, j_{2,1}, \dots, j_{2,l}), \dots, (x_N, j_{N,1}, \dots, j_{N,l})\} \subset \mathbb{R}^n \times M^l,$$

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

$$y_i = A_{j_{i,1}} \dots A_{j_{i,l}} x_i \ \forall (x_i, y_i) \in W_{\omega_N}.$$

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

$$\begin{aligned} & min_P \quad \gamma \\ & s.t. \quad (y_i)^T P (y_i) \leq \gamma^{2l} x_i^T P x_i, \ \forall (x_i, y_i) \in W_{\omega_N} \\ & \qquad\quad P \succ 0, \ \gamma \geq 0. \end{aligned} \tag{5}$$

Then, we have:

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

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

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

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

By Property 2.4, this means that there exists no CQLF for the scaled set of matrices $\frac{\mathcal{M}^l}{(\gamma^*(\omega_N)-\epsilon)^l}$. Since the above inequality is true for every $\epsilon \geq 0$, using Theorem 3.1, and the fact that $\rho(\mathcal{M}^l) = \rho(\mathcal{M})^l$, we conclude:

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

□

Remark 3.1. *In the above discussion, we introduce the concept of 'l-step CQLF', and showed that it allows to refine the initial $1/\sqrt(n)$ approximation provided by the CQLF method. In the switching systems literature, there are other techniques for refining this approximation, as for instance replacing the LMIs in Theorem 3.1 by Sum-Of-Squares (SOS) constraints; see [?] or [?, Theorem 2.16]. It seems that the concept of l-step CQLF is better suited for our purpose, as we briefly discuss below. We leave for further work a more systematic analysis of the behaviours of the different refining techniques.*

4 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 an l-step CQLF which is valid with probability at least β. Note that, the existence of an l-step CQLF implies $\rho \leq 1$ due to Theorem 3.2. Even though the solution of the optimization problem in (4) provides a stability guarantee, 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 theoretical optimization problem (4) and the practical version, with finitely many constraints. Even though in practice, one would solve the optimization problem (5) as suggested above, for the sake of rigor and clarity of our proofs, we introduce a slighly different optimization problem. In this optimization problem, an objective function is considered, and a 'regularization parameter' $\eta > 0$ is added. As the reader will see, we will derive results valid for arbitrarily small values of η, and so this will not hamper the practical accuracy of our technique, but it will allow us to derive a theoretical asymptotic guarantee (i.e. for large number of observations).

$$\begin{aligned} \min_P \quad & \lambda_{\max}(P) \\ \text{s.t.} \quad & (A_{j_1} \dots A_{j_l} x)^T P (A_{j_1} \dots A_{j_l} x) \leq ((1+\eta)\gamma^*(\omega_N))^{2l} x^T P x, \\ & \qquad\qquad\qquad\qquad \forall (x, j_1 \dots j_l) \in \omega_N \subset Z, \\ & P \succeq I, \end{aligned} \tag{6}$$

where $Z := \mathbb{S} \times M^l$, $\eta > 0$, 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 an arbitrary set enclosing the origin, e.g., $\mathbb{S}$.

For the rest of the discussion, we refer to the optimization problem (6) by $\mathrm{Opt}(\omega_N)$. 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 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 $\mathrm{Opt}(\omega_N)$ is a feasible solution to (4) has been widely studied in the literature [?]. 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, [?])**.** *Let d be the dimension of $Opt(\omega_N)$ and $N \geq d+1$. Consider the optimization problem $Opt(\omega_N)$ given in* (6)*, where ω_N is a uniform random 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

$$V(\omega_N) = \{z \in Z : (A_{j_1} \dots A_{j_l} z)^T P(\omega_N)(A_{j_1} \dots A_{j_l} z) > \gamma^* z^T P(\omega_N) z\},$$

i.e., it is the set of constraints that are violated by the optimal solution of $Opt(\omega_N)$.

Theorem 4.1 states that the optimal solution of the sampled problem $\mathrm{Opt}(\omega_N)$ 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. This means that, the ellipsoid computed by $\mathrm{Opt}(\omega_N)$ is “almost invariant" except a set of measure ϵ. The rest of this section builds on this, and it has two important intermediate results leading us to our main theorem. By exploiting Property 2.2, we first show in Lemma 4.2 how one can compute an upper bound on the JSR when this "almost invariant" set is the unit sphere, i.e., $\mathbb{S}$.

Lemma 4.2. *Let $\epsilon \in (0,1)$ and $\gamma \in \mathbb{R}_{>0}$. Consider the set of matrices $\mathcal{M}$ and $A \in \mathcal{M}$ satisfying:*

$$(A_{j_1} \dots A_{j_l} x)^T (A_{j_1} \dots A_{j_l} x) \leq \gamma^{2l} x^T x, \quad \forall\, x \in \mathbb{S} \setminus \mathbb{S}', \forall\, j \in, M, \tag{8}$$

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

$$\rho(\mathcal{M}) \leq \frac{\gamma}{\sqrt[l]{\alpha(\epsilon)}}$$

where $\alpha(\epsilon)$ is given in (25).

Proof. Note that, (8) implies that: $A_j(\mathbb{S} \setminus \mathbb{S}') \subset \gamma^l \mathbb{B}$. Using Property 2.2 this also implies:

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

Then, by Lemma .5 in Appendix .3, we have:

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

by definition of $\alpha(\epsilon)$ given in (24). Therefore, we get:

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

which implies that $\rho(\mathcal{M}^l) \leq \frac{\gamma^l}{\alpha(\epsilon)}$ and hence $\rho(\mathcal{M}) \leq \frac{\gamma}{\sqrt[l]{\alpha(\epsilon)}}$. ☐

We now know how to compute an upper bound on the JSR when the "almost invariant" ellipsoid is $\mathbb{S}$. Thanks to Property 2.4, if this is not the case, we can simply perform a change of coordinates mapping this ellipsoid to $\mathbb{S}$ and compute the JSR in the new coordinates system instead. To do this, in the next theorem, we bound the measure of violating constraints on $\mathbb{S}$ after the change of coordinates, in terms of the measure of the violated constraints on $\mathbb{S} \times M$ in the original coordinates.

Proposition 4.3. *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* $\mu(V) \le \epsilon$. *Then, by defining* L *as in* (2) *and* $\bar{A}_j = L^{-1} A_j L$, *one also has:*

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

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

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

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

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

We have

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

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

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

This means that

$$(A_j x)^T P (A_j x) \le \gamma^2 x^T P x, \; \forall\, x \in \mathbb{S} \setminus V_{\mathbb{S}}, \forall\, m \in M, \tag{11}$$

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

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 (19) in this new coordinates system as in:

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

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

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

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:

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

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. [?]). Putting together (19), (13), and (15) we get the statement of the theorem where $\mathbb{S}' = \Pi_{\mathbb{S}}(L^{-1}(V_{\mathbb{S}}))$. $\square$

Proposition 4.4. *Let $\gamma \in \mathbb{R}_{>0}$. Consider a set of matrices $A \in \mathcal{M}$, and a matrix $P \succ 0$ satisfying:*

$$(A_j x)^T P(A_j x) \leq \gamma^2 x^T P x, \ \forall\, (x, j) \in Z \setminus V, \tag{16}$$

for some $V \subset Z$ where $\mu(Z \setminus V) \geq 1 - \epsilon$. 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: $\sigma(\mathbb{S}') \leq 1 - \epsilon\kappa(P)$, where

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

Proof. Let $\bar{V}_{\mathbb{S}} = \pi_{\mathbb{S}}(Z \setminus V)$. 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 $Z \setminus V$ can be written as the disjoint union $Z \setminus V = \sqcup_{1 \leq i \leq 2^m}(\mathcal{S}_i, \mathcal{M}_i)$ where $\mathcal{S}_i \in \Sigma_{\mathbb{S}}$. We notice that $\bar{V}_{\mathbb{S}} = \sqcup_{1 \leq i \leq 2^m} \mathcal{S}_i$, and

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

We have

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

Note that we have $\mu_M(\mathcal{M}_i) \leq 1$, and we get:

$$\mu(Z \setminus V) \leq \sigma^{n-1}(\bar{V}_{\mathbb{S}}), \tag{17}$$

and therefore

$$\sigma^{n-1}(\bar{V}_{\mathbb{S}}) \geq \mu(Z \setminus V) \geq 1 - \epsilon. \tag{18}$$

This means that

$$(A_j x)^T P (A_j x) \leq \gamma^2 x^T P x, \ \forall\, x \in \mathbb{S} \setminus V_{\mathbb{S}}, \forall\, m \in M, \tag{19}$$

where $\sigma^{n-1}(S \setminus V_{\mathbb{S}}) \geq 1 - \epsilon$. By similar arguments to the previous theorem we get:

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

for some $\mathbb{S}' \subset \mathbb{S}$ such that: $\sigma(\mathbb{S}' \setminus \mathbb{S}) \geq (1 - \epsilon)\kappa_{\min}(P)$, where

$$\kappa_{\min}(P) = \sqrt{\frac{\lambda_{\min}(P)^n}{\det(P)}},$$

which means

$$\sigma(\mathbb{S}') \leq 1 - (1 - \epsilon)\kappa_{\min}(P).$$

$\square$

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

$$\mu^N \left\{ \omega_N \in Z^N : \mu(V(\omega_N)) \leq \epsilon \right\} \geq 1 - I(1-\epsilon; N-d, d+1), \tag{20}$$

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 4.4, 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 4.2, we can compute

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

where

$$\epsilon'(\beta,N) = \frac{1}{2} m\kappa(P)\epsilon(\beta,N) \tag{21}$$

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 .4 for the second part of this proof, namely showing that $\lim_{N\to\infty} \delta(\beta,\omega_N) = 1$ with probability 1.

□

5 Experimental Results

6 Conclusions

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

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

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

.1 Notation and Background

Before proceeding to the main lemmas we use to prove Lemma .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. *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. *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 .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 .2 (see e.g. [**?**]). *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 .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:

$$\Delta(X) := \sup\{r : r\mathbb{B} \subset \text{convhull } (\mathbb{S} \setminus X)\}. \tag{22}$$

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

$$\Delta(X) = d(\partial \text{convhull } (\mathbb{S} \setminus X), 0). \tag{23}$$

Lemma .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:

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

Then the following equalities hold:

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

□

Corollary .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:*

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

.2 Preliminary Results

Lemma .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 $(S \setminus X)$. 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 $\left\{x : c^Tx = k\right\}$. Note that this supporting hyperplane is a supporting hyperplane of the ball $\Delta(X)\mathbb{B}$ at x^* since we have:

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

By Remark .2, this implies that $\left\{x : c^Tx = k\right\}$ is in fact the unique supporting hyperplane at x^*. Then we have:

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

Now, consider the spherical cap defined by this supporting hyperplane, i.e., $\mathcal{C}_{c,k}$. Then, by Lemma .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 $c^Tx > k$. But due to the fact that $x \in$ convhull $(\mathbb{S}\setminus X)$, we also have $c^Tx \le k$, which leads to a contradiction. Therefore, $\mathcal{C}_{c,k} \subset X$. □

Proposition .4. *Let $\mathcal{X}_\epsilon = \{X \subset \mathbb{S} : \sigma^{n-1}(X) \le \epsilon\}$. Then, for any $\epsilon \in (0,1)$, the function $\Delta(X)$ attains its minimum over $\mathcal{X}_\epsilon$ 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 .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. This implies

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

by Corollary .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^*. □

.3 Main Lemma

Lemma .5. *Let $\epsilon \in (0,1)$ and $\alpha : (0,1) \to \mathbb{R}_{\ge 0}$ be defined by:*

$$\alpha(\epsilon) := \sup_{X\in\mathcal{X}_\epsilon} \{r : r\mathbb{B} \subset \mathit{convhull}\ (\mathbb{S} \setminus X)\}, \tag{24}$$

where $\mathcal{X}_\epsilon = \{X \subset \mathbb{S} : \sigma^{n-1}(X) \le \epsilon\}$. Then, $\alpha(\epsilon)$ is given by the formula:

$$\alpha(\epsilon) = \sqrt{1 - I^{-1}\left(\epsilon; \frac{n-1}{2}, \frac{1}{2}\right)}, \tag{25}$$

where I is the regularized incomplete beta function.

Proof. By Proposition .4 we know that:

$$\alpha(\epsilon) = \Delta(\mathcal{C}_{c,k}), \tag{26}$$

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. [?]) that the area of such $\mathcal{C}_{c,k}$, is given by the equation:

$$\sigma^{n-1}(\mathcal{C}_{c,k}) = I\left(1 - \Delta(\mathcal{C}_{c,k})^2; \frac{n-1}{2}, \frac{1}{2}\right). \tag{27}$$

Since, $\sigma^{n-1}(\mathcal{C}_{c,k}) = \epsilon$, we get the following set of equations:

$$\begin{aligned}\epsilon &= I\left(1-\Delta(\mathcal{C}_{c,k})^2; \frac{n-1}{2}, \frac{1}{2}\right)\\ 1-\Delta(\mathcal{C}_{c,k})^2 &= I^{-1}\left(\epsilon; \frac{n-1}{2}, \frac{1}{2}\right)\\ \Delta(\mathcal{C}_{c,k})^2 &= 1-I^{-1}\left(\epsilon; \frac{n-1}{2}, \frac{1}{2}\right). \qquad (28)\end{aligned}$$

Then, the inequalities (26) and (28) imply the inclusion given in (24). $\square$

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

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

$$\kappa(P(\omega_N)) = \sqrt{\frac{\lambda_{\max}(P(\omega_N))^n}{\det(P(\omega_N))}} \leq \sqrt{K^n}. \qquad (29)$$

We next 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:

$$1-\beta \leq (d+1)N^d(1-\epsilon)^{N-d}. \qquad (30)$$

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 (30) 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 this together with (29), 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 also implies: $\lim_{N\to\infty}\alpha(m\kappa(P(\omega_N))\epsilon(\beta,\omega_N)) = 1$.