Sampling Based Stability Analysis of Switched Linear Systems

1 Introduction

As the computational power of the computers has increased, so is the complexity of dynamical systems we can model with them. Today, the industrial dynamical systems does not only consist of simple differential or difference equations; they are multimodal, hybrid, and often contain variety of subcomponents such as lookup tables, delay differential equations, actuation delays and thermodynamic models. The current modeling paradigm is also highly distributed, in the sense that, some of the model subcomponents are developed by different parties and therefore their internal structure is partially, if not completely unknown to the end user. Hence, it is often virtually impossible to obtain simple analytical formulas for today's industrial scale dynamical system models. However, thanks to the advances in our computational power again, in spite of this increased complexity, for a notable subset of these models we can still efficiently perform simulations. Therefore, it is a natural question to ask whether we can provide formal analyses about certain properties of these systems based solely on the collected simulation data. In this paper, we focus on the analysis of one of the most important properties of dynamical systems in the context of control theory: stability.

More formally, given a dynamical system as in:

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

where, $x_k \in \mathbb{R}^n$, k is index of time. Let $y_k := x_{k+1}$ We ask the following question: *given N input-output pairs,* (x_1, y_1), (x_2, y_2), *…,* (x_N, y_N) *such that* $y_k = f(k, x_k)$, *what can we say about the stability of the system* (**??**)*?* The answer is immediate when (**??**) is a linear time-invariant system, since we can simply identify the system by n linearly independent output traces. In this paper, we seek the answer to this question for switched linear systems for which the problem immediately becomes nontrivial. A switched linear systems is in the form:

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

where, $\sigma : \mathbb{N} \to \{1, 2, \ldots, m\}$ is the switching sequence and $A_{\sigma(k)} \in \mathcal{M}$, for all σ and k. Aside from their theoretical value, switched systems model the behavior of dynamical systems in the presence of 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.

Assessing the stability of nonlinear systems by leveraging simulations has been an active area of research in the recent years. Simulation data has been used in both construction and verification of Lyapunov functions. Topcu et.al. [?] and Kapinski et.al. [?] construct Lyapunov function candidates using the simulation traces, however to be able to formally verify the constructed Lyapunov function, they require the knowledge of the full dynamics. In [?] and [?] Bobiti and Lazar address this and provide sampling based probabilistic and deterministic guarantees of a given Lyapunov function candidate. The presented method requires the knowledge of how fast the output of the system can change as the initial condition changes, and moreover the number of required samples increases exponentially in the dimension of the state, n.

The stability of switched systems closely relates closely to the *joint spectral radius* (JSR) of the matrices appearing in (2). Under certain conditions deciding stability amounts to deciding whether JSR less than one or not. There has been a lot of work on developing algorithms to approximate this quantity, when the matrices appearing in (2) are known. Therefore, our work is also connected to the identification of switched systems, since once the system (2) is identified one can then apply these well-established results. However, there are two main reasons behind our quest to directly work on input-output pairs and bypassing the identification phase: (1) Even when $\mathcal{M}$ is known, approximating the JSR is NP-hard [?]. (2) Identifying the set $\mathcal{M}$ is also NP-hard. Therefore, the existing identification techniques can identify $\mathcal{M}$ up to an approximation error. As a result, how to relate this identification error to an error on the stability of (2) is still nontrivial.

In this paper, we present an algorithm to approximate the JSR of a switched linear system from N input-output pairs. This algorithm provides an upper bound on the JSR with a user-defined confidence level. As the number of samples increases, this bound gets tight. Moreover, we characterize with a closed form expression what the exact trade-off between the tightness of this bound and the number of samples is. In order to understand the quality of our technique, the algorithm also provides a deterministic lower-bound.

The organization of the paper is follows: Section introduces our notation and definitions from the switched linear system literature to present our results, Section formalizes the problem definition, Section provides an algorithm to compute a deterministic lower bound and

2 Preliminaries

We consider the usual Hilbert finite normed vector space $(\mathbb{R}^n, \ell_2)$, $n \in \mathbb{N}_{>0}$, ℓ_2 the classical euclidean norm. We denote a unit ball in $\mathbb{R}^n$ with B and unit sphere in $\mathbb{R}^n$ of radius r as S. We only denote the radius r explicitly as in B_r and S_r, when r is different than 1. We denote the set of real symmetric matrices of size n by $\mathbb{S}^n$, and the set of linear functions in $\mathbb{R}^n$ by $\mathcal{L}(\mathbb{R}^n)$. We denote the ellipsoid described by the matrix $P \in \mathbb{S}^n$ as E_P. We denote the homothety of ratio λ by $\mathcal{H}_\lambda$.

For the rest of the write-up, we denote the set of indices of the modes as $M = \{1, 2, \ldots, m\}$, where $m \in \mathbb{N}_{>0}$ is the number of the modes. We denote the joint spectral radius of the set of matrices $\{A_1, A_2, \ldots A_m\}$ by ρ. Let us consider $X = \mathrm{S} \times M$ the Cartesian product of the unit sphere S with M. Every element of X can be written as $x = (s_x, k_x)$ with $s_x \in \mathrm{S}$ and $k_x \in M$. For notational simplicity, we drop the subscript x whenever it is clear from the context.

We define the projections:

$$\pi_{\mathrm{S}} : \mathrm{S} \times M \to \mathrm{S}, (s, k) \mapsto s$$

$$\pi_M : \mathrm{S} \times M \to M, (s, k) \mapsto k.$$

It is well-known that S is a $n-1$ embedded submanifold of $\mathbb{R}^n$, and can thus be seen as an image of an atlas (collection) of smooth maps $\phi_i : U \to \mathrm{S}$, $U \in \mathbb{R}^n$ called charts. It has the topology inherited from its ambient space $\mathbb{R}^n$. If $\mathbb{R}^n$ is provided with a σ-algebra Σ, this parametrization also induces a σ-algebra on S, Σ_{S}. Hence, a measure μ on the measurable space $(\mathbb{R}^n, \Sigma)$ defines a measure μ_{S} on the measurable space $(\mathrm{S}, \Sigma_{\mathrm{S}})$. This measure can be seen as push-forward $\phi_{i*}(\mu)$ of μ by the charts, i.e., $\phi_{i*}(\mu)(A) = \mu(\phi_i^{-1}(A))$ for any $A \in \Sigma_{\mathrm{S}}$. In particular, with the classical Borel σ-algebra and Lebesgue measure in $\mathbb{R}^n$, we obtain a σ-algebra $\mathcal{B}_{\mathrm{S}}$ with $A \in \mathcal{B}_{\mathrm{S}}$ if and only if the sector tA, $t \in [0, 1]$ is in $\mathcal{B}_{\mathbb{R}^n}$; and the classical spherical measure commonly denoted by σ^{n-1} and defined by

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

We can notice that $\sigma^{n-1}(\mathrm{S}) = 1$.

We assume now that S is provided with a σ-algebra Σ_{S} and M with the classical σ-algebra associated to finite sets: $\Sigma_M = \wp(M)$, where $\wp(M)$ is the power set of M.

We consider an unsigned finite spherical measure μ_{S} on $(\mathrm{S}, \Sigma_{\mathrm{S}})$ and an unsigned finite measure[1] μ_M on (M, Σ_M) with $\mathrm{supp}(\mu_M) = M$. In other words, $\forall\, k \in M$, $\mu_M(\{k\}) > 0$.

We denote the product σ-algebra $\Sigma_{\mathrm{S}} \bigotimes \Sigma_M$ engendered by Σ_{S} and Σ_M: $\Sigma = \sigma(\pi_{\mathrm{S}}^{-1}(\Sigma_{\mathrm{S}}), \pi_M^{-1}(\Sigma_M))$. On this set, we define the product measure $\mu = \mu_{\mathrm{S}} \otimes \mu_M$ which is an unsigned finite measure on X.

3 Optimization Problem

We are interested in solving the following optimization problem for a given $\gamma \in (0, 1)$:

$$\begin{aligned} &\text{find} && P \\ &\text{subject to} && (A_i s)^T P (A_i s) \leq \gamma^2 s^T P s, \quad \forall i = \{1, 2, \ldots, m\}, \forall\, s \in \mathrm{S}, \\ &&& P \succ 0. \end{aligned} \tag{3}$$

Note that if P is a solution to (3), then so is αP for any $\alpha \in \mathbb{R}_{>0}$.Therefore, we can rewrite (3) as the following optimization problem:

$$\begin{aligned} &\text{find} && P \\ &\text{subject to} && (A_k s)^T P (A_k s) \leq \gamma^2 s^T P s, \quad \forall i = \{1, 2, \ldots, m\}, \forall\, s \in \mathrm{S}, \\ &&& P \succeq I. \end{aligned} \tag{4}$$

We define the linear isomorphism Φ as the natural mapping $\Phi : \mathbb{R}^{\frac{n(n+1)}{2}} \to \mathbb{S}^n$. Using this mapping, for a fixed $\gamma \in (0, 1]$ we can rewrite (4) as:

$$\begin{aligned} &\text{find} && p \\ &\text{subject to} && f(p, x) \leq 0, \forall x \in X. \end{aligned} \tag{5}$$

where $f(p, x) = \max(f_1(p, x), f_2(p))$, and

$$\begin{aligned} f_1(p, x) &:= (A_k s)^T \Phi(p) (A_k s) - \gamma^2 s^T \Phi(p) s \\ f_2(p) &:= \lambda_{\max}(\Phi(-p)) + 1. \end{aligned}$$

[1] Recall that the support of a measure μ defined on a measurable space (X, Σ) is $\mathrm{supp}(\mu) = \overline{\{A \in \Sigma | \mu(A) > 0\}}$

Proposition 3.1. *The optimization problem* (5) *is convex.*

Proof. The function $f_1(p, x)$ is clearly convex in p for a fixed $x \in X$. The function $\lambda_{max} : \mathbb{S}^n \to \mathbb{R}$ maps a symmetric positive matrix to its maximum eigenvalue. It is well-known that the function $\lambda_{\max}$ is a convex function of P. [?]. This means that, $p \mapsto \Phi(\lambda_{\max}(p))$ is convex in p. Moreover, maximum of convex functions is also convex, which shows that $f(p, x)$ is convex in p. $\square$

Note that the optimization problem (5) has infinitely many constraints. We next consider the following optimization problem where we sample N constraints of (5) independently and identically with the probability measure $\mathbb{P}(A) = \frac{\mu(A)}{\mu(X)}, \forall\, A \in \Sigma$, where $N \geq d + 1$, and $d := \frac{n(n+1)}{2}$. We denote this sampling by $\omega := \{x_1, x_2, \ldots, x_N\} \subset X$, and obtain the following convex optimization problem $\mathrm{Opt}(\omega)$:

$$\begin{aligned} &\text{find} && p \\ &\text{subject to} && f(p, x) \leq 0, \forall x \in \omega. \end{aligned} \tag{6}$$

Let $p^*(\omega)$ be the solution of $\mathrm{Opt}(\omega)$. We are interested in the probability of $p^*(\omega)$ violating at least one constraint in the original problem (5). Therefore, we define constraint violation property next.

Constraint violation probability [?] The constraint violation probability is defined as:

$$\mathcal{V}^*(\omega) = \begin{cases} \mathbb{P}\{x \in X : f(p^*(\omega), x) > 0\} & \text{if } \omega \in X^{N*}, \\ 1, & \text{otherwise} \end{cases}$$

where $X^{N*} := \{\omega \in X^N : \text{the solution of } \mathrm{Opt}(\omega) \text{ exists}\}$. Note that, since we have $\mathbb{P}(A) = \frac{\mu(A)}{\mu(X)}$, we can rewrite this as:

$$\mathcal{V}^*(\omega) = \begin{cases} \frac{\mu\{x \in X : f(p^*(\omega), x) > 0\}}{\mu(X)} & \text{if } \omega \in X^{N*}, \\ 1, & \text{otherwise} \end{cases}$$

We make the following assumptions on the problem $\mathrm{Opt}(\omega)$:

1. Uniqueness of solution: Note that this can be enforced by adding a tie-break rule of at most $\frac{n(n-1)}{2}$ convex conditions discriminating our solutions.

2. Nondegenaracy: with probability 1, there is no redundancy in the constraint obtained from the sampling.

The following theorem from [?] explicitly gives a relationship between $V^*(\omega)$ and N, n.

Theorem 3.2 (from [?])**.** *Consider the optimization problem Opt(ω) given in* (6)*. Let Assumption 1 and Assumption 2 hold. Then, for all $\epsilon \in (0,1)$ the following holds:*

$$\mathbb{P}^N\{\{\mathcal{V}^*(\omega) \le \epsilon\} \cap X^{N*}\} \ge 1 - \sum_{j=0}^{d} \binom{N}{j} \epsilon^j (1-\epsilon)^{N-j}.$$

Note that $\epsilon = 1 - I^{-1}(\beta, N-d, d+1)$ and can be interpreted as the ratio of the measure of points in X that might violate at least one of the constraints in (4) to the measure of all points in X.

We now state our main theorem, which is based on Theorem 3.2 and devote the next section to proving it step by step. We denote by γ^*, the optimum value of the following optimization problem:

$$\begin{aligned} \min_{P,\gamma} \quad & \gamma \\ \text{subject to} \quad & (A_i s)^T P (A_i s) \le \gamma^2 s^T P s, \quad \forall i = \{1,2,\ldots,m\}, \forall\, s \in \mathrm{S}, \\ & P \succ 0. \end{aligned} \tag{7}$$

Theorem 3.3 (Main Theorem)**.** *For any $\eta > 0$, given $N \ge n+1$ and $\beta \in [0,1)$, we can compute $\delta < \infty$ such that with probability at least β, $\rho \le \delta(1+\eta)\gamma^*$. Moreover, as $N \to \infty$, $\delta \to 1$.*

4 Relating the measure of bad sets

For a given sampling $\omega \in X^{N*}$, let $V := \{x \in X : f(p^*(\omega), x) > 0\}$, i.e., the set of points for which at least one constraint is violated, and V_{S}, V_M be its projections on S and M, respectively.

Lemma 4.1. $\mu_S(V_S) \le \frac{\mu(V)}{m_1}$*, where $m_1 = \min\{\mu_M(\{k\}), k \in M\}$.*

Proof. Let $A \subset X$, $A_{\mathrm{S}} = \pi_{\mathrm{S}}(A)$ and $A_M = \pi_M(A)$. We notice that Σ_M is the disjoint union of its 2^m elements $\{B_i, i \in \{1,2,\ldots 2^m\}\}$. Then A is the disjoint union $A = \sqcup_{1 \le i \le 2^m}(A_i, B_i)$ where $A_i = \pi_M^{-1}(B_i) \in \mathrm{S}$. We notice that $A_{\mathrm{S}} = \sqcup_{1 \le i \le 2^m} A_i$, and

$$\mu_{\mathrm{S}}(A_{\mathrm{S}}) = \sum_{1 \le i \le 2^m} \mu_{\mathrm{S}}(A_i).$$

We have

$$\begin{aligned}\mu(A) = \mu(\sqcup_{1\leq i\leq 2^m}(A_i, B_i)) &= \sum_{1\leq i\leq 2^m} \mu((A_i, B_i)) \\ &= \sum_{1\leq i\leq 2^m} \mu_{\mathrm{S}} \otimes \mu_M((A_i, B_i)) \\ &= \sum_{1\leq i\leq 2^m} \mu_{\mathrm{S}}(A_i)\mu_M(B_i).\end{aligned}$$

Let m_1 be the minimum value of μ_M on its atoms: $m_1 = \min\{\mu_M(\{k\}), k \in M\}$ (recall that $m_1 > 0$). Then since $\forall\ i$, $\mu_M(B_i) \geq m_1$, we have

$$\mu_{\mathrm{S}}(A_{\mathrm{S}}) \leq \frac{\mu(A)}{m_1}. \tag{8}$$

This proves our statement by taking $A = V_{\mathrm{S}}$. □

Corollary 4.2. *When the modes are sampled from the set M uniformly random,*

$$\mu_S(V_S) \leq m\mu(V).$$

We consider the linear transformation mapping S to E_P that denoted by $L \in \mathcal{L}(\mathbb{R}^n)$. Note that since $P \in \mathbb{S}^n$, it can be written in its Choleski form $P = UDU^{-1}$, where D diagonal matrix of its eigenvalues, and $U \in O_n(\mathbb{R})$. We define $D^{1/2}$ the positive square root of D as the matrix $\mathrm{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n})$. Then, the positive square root of P is $VD^{1/2}V$. This means that, $L = P^{1/2}$. For the rest of the write-up, we denote

$$V' := \Pi_{\mathrm{S}}(L^{-1}(V_{\mathrm{S}})),$$

and show how to upper bound $\sigma^{n-1}(V')$ in terms of $\mu(V)$.

Lemma 4.3. *Let ψ a smooth change of coordinates in $\mathbb{R}^n$ and $\mathcal{D} \subset S$, whose image under ψ is $\mathcal{D}' \subset \psi(S)$. Let μ_S be a positive spherical measure induced by a measure μ on $\mathbb{R}^n$. Let Σ_E and μ_E be the σ-algebra and the measure induced from Σ_S and μ_S on the ellipsoid $E = \psi(S)$. Then*

$$\mu_E(\psi(V_S)) = |\det(\psi)|\mu_S(V_S), \tag{9}$$

where $\psi \in \mathcal{L}(\mathbb{R}^n)$.

Proof. We have $\mu_{\mathrm{S}}(\mathcal{D}) = \int_{x\in\mathcal{D}} \mathbb{1}_D(x)\ d\mu_{\mathrm{S}}(x)$, $\mu_{\mathrm{S}} = \{\phi_{i*}(\mu)\}_i$ and

$$\mu(\mathcal{D}') = \int_{y\in\mathcal{D}'} \mathbb{1}_{D'}(y)\ d\mu(y) = \int_{x\in\mathcal{D}} \mathbb{1}_{x\in\mathcal{D}} |\det J(\phi(x))|\ d\mu(x).$$

This gives

$$\mu_E(\mathcal{D}') = \int_{y\in\mathcal{D}'} \mathbb{1}_{D'}(y)\ d\mu_E(y) = \int_{x\in\mathcal{D}} \mathbb{1}_{x\in\mathcal{D}} |\det J(\psi(x))|\ d\mu_{\mathrm{S}}(x).$$

In particular, if $\psi \in \mathcal{L}(\mathbb{R}^n)$, then $\forall x \in \mathbb{R}^n$, $\det(J(\psi(x)) = \det(\psi)$ and

$$\mu_E(\mathcal{D}') = \int_{y\in\mathcal{D}'} \mathbb{1}_{D'}(y)\ d\mu_E(y) = |\det(\psi)| \int_{x\in\mathcal{D}} \mathbb{1}_{x\in\mathcal{D}}\ d\mu_{\mathrm{S}}(x).$$

This proves the statement of the lemma when $\mathcal{D} = V_{\mathrm{S}}$. □

Definition Let X be a Hilbert space, A a nonempty subset of X and $\psi : A \to X$. Then ψ is called firmly nonexpansive if

$$\forall x, y \in A, \|\psi(x) - \psi(y)\|^2 + \|(\mathrm{Id} - \psi)(x) - (\mathrm{Id} - \psi)(y)\|^2 \leq \|x - y\|^2,$$

where Id denotes the identity function from X to X.

Theorem 4.4 (from [?]). *Let C be a nonempty closed convex subset of X, then the convex projector on C, Π_C, is firmly nonexpansive.*

Corollary 4.5.

$$\|\Pi_C(x) - \Pi_C(y)\| \leq \|x - y\| \qquad \forall\, x, y \in C. \tag{10}$$

Lemma 4.6.

$$\mu_S(\Pi_S(L^{-1}(V_S))) \leq \det(L^{-1}) \left(\frac{1}{\lambda_{\min}(L^{-1})}\right)^n \mu_S(V_S). \tag{11}$$

Proof. Note that the mapping Π_{S} can be seen as the composition of the Π_{S_r} for some $r > 0$, and $\mathcal{H}_{\frac{1}{r}}$. Let $E' := L^{-1}(\mathrm{S})$, then when $r < \min_{x\in E'} \|x\| = \lambda_{\min}(L^{-1})$ we have

$$\Pi_{\mathrm{S}_{\lambda_{\min}}}(x) = \Pi_{\mathrm{B}_{\lambda_{\min}}}(x) \qquad \forall\, x \in E'.$$

This shows that the restriction of $\Pi_{\mathrm{S}_{\lambda_{\min}}}$ to E' is a convex projector.

Then by Corollary 4.5

$$\|\Pi_{\mathrm{S}_{\lambda_{\min}}}(x) - \Pi_{\mathrm{S}_{\lambda_{\min}}}(y)\| \leq \|x-y\|, \qquad \forall\, x, y \in E'. \tag{12}$$

This shows that 1 is a Lipschitz constant of the function $\Pi_{\mathrm{S}_{\lambda_{\min}}}$ on E'.

By composing $\Pi_{\mathrm{S}_{\lambda_{\min}}}$ with $\mathcal{H}_{\frac{1}{\lambda_{\min}}}$, we obtain Π_{S}. Since the Lipschitz constant of composition of two functions can be bounded by the multiplication of Lipschitz constants of each function, the Lipschitz constant of Π_S on E' is $\frac{1}{\lambda_{\min}}$, which means that:

$$\|\Pi_{\mathrm{S}}(x) - \Pi_{\mathrm{S}}(y)\| \leq \frac{1}{\lambda_{\min}}\|x-y\|, \qquad \forall\, x, y \in E'. \tag{13}$$

Note that, the inequality in (13) is an equality when x is in the eigenspace of $\lambda_{\min}$ and $y = -x$.

Recall that for any smooth Lipschitz function ϕ with Lipschitz constant, $\mathrm{Lip}(\phi)$, we have for all x, $|\det(J(\phi(x))| \leq \mathrm{Lip}(\phi)^n$. Combining this with (13) and Lemma 4.6, we get the statement of the lemma. □

Theorem 4.7. $\sigma^{n-1}(V') \leq m\epsilon\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}$, *where* $\mu(V) = \epsilon$.

Proof. By taking μ_{S} as the uniform spherical measure σ^{n-1}, and combining Corollary 4.2 with Lemma 4.6 we get the statement of the theorem. □

5 Relating ϵ to δ

We denote $\epsilon' := \frac{\epsilon}{2}\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}$, where the additional factor $\frac{1}{2}$ follows from the homogeneity of the dynamics. In this section, we show how to relate ϵ' to δ in the statement of the Theorem 3.3. We start by a few definitions that will help us along the way. Let d be a distance on $\mathbb{R}^n$. We define the distance between a set $X \subset \mathbb{R}^n$ and a point $p \in \mathbb{R}^n$ is $d(X,p) := \inf_{x\in X} d(x,p)$.

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

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

Proposition 5.1 (see e.g. [?]). *The distance between the point* $x = 0$ *and the hyperplane* $c^T x = k$ *is* $\frac{|k|}{\|c\|}$.

We define the function $\Delta : 2^{\mathrm{S}} \to [0,1]$ as:

$$\Delta(X) := \sup\{r : \mathrm{B}_r \subseteq \text{convhull } (\mathrm{S} \setminus X)\}. \tag{14}$$

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

$$\Delta(X) = d(\partial \text{convhull } (\mathrm{S} \setminus X), 0). \tag{15}$$

Lemma 5.2. $\Delta(\mathcal{C}_{c,k}) = \min\left(1, \frac{|k|}{\|c\|}\right)$.

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

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

□

Corollary 5.3. $\Delta(\mathcal{C}_{c,k_1}) < \Delta(\mathcal{C}_{c,k_2})$ *when* $k_1 < k_2$.

Lemma 5.4. $\sigma^{n-1}(\mathcal{C}_{c,k_1}) < \sigma^{n-1}(\mathcal{C}_{c,k_2})$, *for* $k_1 > k_2$.

Proof. convhull $(\mathrm{S} \setminus \{x \in \mathrm{S} : c^T x > k_1\}) \subseteq$ convhull $(\mathrm{S} \setminus \{x \in \mathrm{S} : c^T x > k_2\})$, for $k_1 > k_2$. □

Now we are ready to present the following lemma which is the key to proving our main result.

Lemma 5.5. *For any set $X \subseteq S$, there exists c and k such that $\mathcal{C}_{c,k}$ satisfies:*

$$\mathcal{C}_{c,k} \subseteq X,$$

and

$$\Delta(\mathcal{C}_{c,k}) = \Delta(X). \tag{16}$$

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

$$\Delta(X) = d(\partial X_S, 0) = \inf_{x \in \partial X_S} d(x, 0) = \min_{x \in \partial X_S} d(x, 0) = d(x^*, 0). \tag{17}$$

Next, consider the supporting hyperplane of X_S at x^*, which we denote by $\left\{x : c^T x = k\right\}$. Note that, this supporting hyperplane is unique because it is also a supporting hyperplane of the ball $\mathrm{B}_{\Delta(X)}$ at x^* as well, which is unique. This can be seen from the fact that:

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

Then we have:

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

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

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

We now prove our main result.

Theorem 5.6. *Let $X_{\epsilon'} = \{X \subset S : \sigma^{n-1}(X) = \epsilon'\}$. Then, for any $\epsilon' \in (0,1)$, the function $\Delta(X)$ attains its minimum over $X_{\epsilon'}$ for some X which is a spherical cap.*

Proof. We prove this via contradiction. Assume that there exists no spherical cap in $X_{\epsilon'}$ such that $\Delta(X)$ attains its minimum. This means there exists an $X^* \in X_{\epsilon'}$, where X^* is not a spherical cap and $\arg\min_{X \in X_{\epsilon'}}(\Delta(X)) = X^*$. By Lemma 5.5 we can construct a spherical cap $\mathcal{C}_{c,k}$ such that $\mathcal{C}_{c,k} \subseteq X^*$ and $\mathcal{C}_{c,k} = \Delta(X^*)$. Note that, we further have $\mathcal{C}_{c,k} \subset X^*$, 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 $\tilde{k} < k$, due to Lemma 5.4. This implies $\Delta(\mathcal{C}_{c,\tilde{k}}) < \Delta(\mathcal{C}_{c,k}) = \Delta(X^*)$ due to Lemma 5.3. Therefore, $\Delta(\mathcal{C}_{c,\tilde{k}}) < \Delta(X^*)$. This is a contradiction since we initially assumed that $\Delta(X)$ attains its minimum over $X_{\epsilon'}$ at X^*. □

Theorem 5.7. *Given a spherical cap $\mathcal{C}_{c,k} \subseteq S$ such that $\sigma^{n-1}(\mathcal{C}_{c,k}) = \epsilon'$,*

$$\Delta(\mathcal{C}_{c,k}) = \sqrt{(1-\alpha)},$$

where $\alpha := I^{-1}\left(\frac{\epsilon'\Gamma(\frac{d}{2})}{\pi^{d/2}}, \frac{d-1}{2}, \frac{1}{2}\right)$ and $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$. Here I^{-1} is the inverse incomplete beta function, i.e., $I^{-1}(y,a,b) = x$ where $I_x(a,b) = y$.

Proof. Let $h := 1 - \Delta(\mathcal{C}_{c,k})$. It is well known [?] that the area of the spherical cap $\mathcal{C}_{c,k} \subseteq \mathrm{S}$ is given by the equation:

$$\epsilon' = \sigma^{n-1}(\mathcal{C}_{c,k}) = \frac{\pi^{d/2}}{\Gamma[\frac{d}{2}]} I_{2h-h^2}\left(\frac{d-1}{2}, \frac{1}{2}\right), \tag{18}$$

where I is the incomplete beta function. From this, we get the following set of equations:

$$\begin{aligned} \frac{\epsilon'\Gamma[\frac{d}{2}]}{\pi^{d/2}} &= I_{2h-h^2}\left(\frac{d-1}{2}, \frac{1}{2}\right) \\ 2h - h^2 &= I^{-1}\left(\frac{\epsilon'\Gamma(\frac{d}{2})}{\pi^{d/2}}, \frac{d-1}{2}, \frac{1}{2}\right) \\ 2h - h^2 &= \alpha \\ h^2 - 2h + \alpha &= 0. \end{aligned} \tag{19}$$

From (19), we get $h = 1 \pm \sqrt{(1-\alpha)}$. Since $h \leq 1$, we conclude that $\Delta(\mathcal{C}_{c,k}) = \sqrt{(1-\alpha)}$. Note that, $\Delta(\mathcal{C}_{c,k})$ only depends on ϵ for fixed n. □

Corollary 5.8. *For a fixed $\beta \in (0,1)$, $\lim_{N\to\infty} \delta_\beta(N) = 1$.*

Proof. We first prove that $\lim_{N\to\infty} \epsilon_\beta(N) = 0$. Note that, we can upper bound $1-\beta$ as follows:

$$\begin{pmatrix} 1 - \beta = \sum_{j=0}^{d} N \\ j\epsilon^j(1-\epsilon)^{N-j} \leq (d+1)N^d(1-\epsilon)^{N-d}. \end{pmatrix} \tag{20}$$

We prove $\lim_{N\to\infty} \epsilon_\beta(N) = 0$ by contradiction and assume that $\lim_{N\to\infty} \epsilon_\beta(N) \neq 0$. This means that, there exists some $\delta > 0$ such that $\epsilon_\beta(N) > \delta$ infinitely often. Then, consider the subsequence N_k such that $\epsilon_\beta(N_k) > \delta$, $\forall\, k$. By (20) we have:

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

Note that $\lim_{k\to\infty}(d+1)N_k^d(1-\delta)^{N_k-d} = 0$. Therefore, there exists a k' such that, we have $(d+1)N_{k'}^d(1-\delta)^{N_k'-d} < 1-\beta$, which is a contradiction. Therefore, we must have $\lim_{N\to\infty} \epsilon_\beta(N) = 0$.

Showing I^{-1} in its first parameter, $\delta = \sqrt{1-\alpha}$ tends to 1 as $\epsilon \to 0$.

□