# Filling in the Gaps in Raphael's Idea

## 1 Preliminaries

We consider the usual Hilbert finite normed vector space $(\mathbb{R}^n, \ell_2)$, $n \in \mathbb{N}_{>0}$, $\ell_2$ the classical euclidean norm. We denote a unit ball in $\mathbb{R}^n$ with $\mathrm{B}$ and unit sphere in $\mathbb{R}^n$ of radius $r$ as $\mathrm{S}$. We only denote the radius $r$ explicitly as in $\mathrm{B}_r$ and $\mathrm{S}_r$, 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 $\lambda$ by $\mathcal{H}_\lambda$.

For the rest of the write-up, we denote the set of indices of the modes as $M = \{1, 2, \dots, m\}$, where $m \in \mathbb{N}_{>0}$ is the number of the modes. We denote the joint spectral radius of the set of matrices $\{A_1, A_2, \dots A_m\}$ by $\rho$. Let us consider $X = \mathrm{S} \times M$ the Cartesian product of the unit sphere $\mathrm{S}$ 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 $\mathrm{S}$ 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 $\sigma$-algebra $\Sigma$, this parametrization also induces a $\sigma$-algebra on $\mathrm{S}$, $\Sigma_{\mathrm{S}}$. Hence, a measure $\mu$ on the measurable space $(\mathbb{R}^n, \Sigma)$ defines a measure $\mu_{\mathrm{S}}$ on the measurable space $(\mathrm{S}, \Sigma_{\mathrm{S}})$. This measure can be seen as push-forward $\phi_{i*}(\mu)$ of $\mu$ 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 $\sigma$-algebra and Lebesgue measure in $\mathbb{R}^n$, we obtain a $\sigma$-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 $\sigma^{n-1}$ 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 $\sigma$-algebra $\Sigma_{\mathrm{S}}$ and $M$ with the classical $\sigma$-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 $\mu_{\mathrm{S}}$ on $(\mathrm{S}, \Sigma_{\mathrm{S}})$ and an unsigned finite measure[1] $\mu_M$ on $(M, \Sigma_M)$ with $\mathrm{supp}(\mu_M) = M$. In other words, $\forall\, k \in M$, $\mu_M(\{k\}) > 0$.

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

## 2 Optimization Problem

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

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

Note that if $P$ is a solution to (1), then so is $\alpha P$ for any $\alpha \in \mathbb{R}_{>0}$.Therefore, we can rewrite (1) as the following optimization problem:

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

We define the linear isomorphism $\Phi$ 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 (2) as:

$$\begin{aligned} &\text{find} && p \\ &\text{subject to} && f(p, x) \le 0, \forall x \in X. \end{aligned} \tag{3}$$

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

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

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

**Proposition 2.1.** *The optimization problem* (3) *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 $p$. □

Note that the optimization problem (3) has infinitely many constraints. We next consider the following optimization problem where we sample $N$ constraints of (3) 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 $\mathrm{Opt}(\omega)$:

$$\begin{aligned} &\text{find} && p \\ &\text{subject to} && f(p,x) \leq 0, \forall x \in \omega. \end{aligned} \tag{4}$$

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 (3). 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 2.2** (from [?]). *Consider the optimization problem Opt($\omega$) given in (4). Let Assumption 1 and Assumption 2 hold. Then, for all $\epsilon \in (0,1)$ the following holds:*

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

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 (2) to the measure of all points in $X$.

We now state our main theorem, which is based on Theorem 2.2 and devote the next section to proving it step by step. We denote by $\gamma^*$, the optimum value of the following optimization problem:

$$\begin{aligned} \min_{P,\gamma} \quad & \gamma \\ \text{subject to} \quad & (A_i s)^T P (A_i s) \leq \gamma^2 s^T P s, \quad \forall i = \{1, 2, \ldots, m\}, \forall\, s \in \mathrm{S}, \\ & P \succ 0. \end{aligned} \tag{5}$$

**Theorem 2.3** (Main Theorem). *For any $\eta > 0$, given $N \geq n+1$ and $\beta \in [0,1)$, we can compute $\delta < \infty$ such that with probability at least $\beta$, $\rho \leq \delta(1+\eta)\gamma^*$. Moreover, as $N \to \infty$, $\delta \to 1$.*

## 3 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 $V_{\mathrm{S}}, V_M$ be its projections on S and $M$, respectively.

**Lemma 3.1.** $\mu_S(V_S) \leq \frac{\mu(V)}{m_1}$, *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 $\Sigma_M$ is the disjoint union of its $2^m$ elements $\{B_i, i \in \{1, 2, \dots 2^m\}\}$. Then $A$ is the disjoint union $A = \sqcup_{1\leq i\leq 2^m}(A_i, B_i)$ where $A_i = \pi_M^{-1}(B_i) \in \mathrm{S}$. We notice that $A_{\mathrm{S}} = \sqcup_{1\leq i\leq 2^m} A_i$, and

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

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 $\mu_M$ 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

$$\mu_{\mathrm{S}}(A_{\mathrm{S}}) \leq \frac{\mu(A)}{m_1}. \tag{6}$$

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

**Corollary 3.2.** *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 $\mathrm{diag}(\sqrt{d_1}, \dots, \sqrt{d_n})$. 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 3.3.** *Let $\psi$ a smooth change of coordinates in $\mathbb{R}^n$ and $\mathcal{D} \subset S$, whose image under $\psi$ is $\mathcal{D}' \subset \psi(S)$. Let $\mu_S$ be a positive spherical measure induced* *by a measure* $\mu$ *on* $\mathbb{R}^n$*. Let* $\Sigma_E$ *and* $\mu_E$ *be the* $\sigma$*-algebra and the measure induced from* $\Sigma_S$ *and* $\mu_S$ *on the ellipsoid* $E = \psi(S)$*. Then*

$$\mu_E(\psi(V_S)) = |\det(\psi)|\mu_S(V_S), \tag{7}$$

*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 $\psi$ 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 3.4** (from [?])**.** *Let* $C$ *be a nonempty closed convex subset of* $X$*, then the convex projector on* $C$*,* $\Pi_C$*, is firmly nonexpansive.*

**Corollary 3.5.**

$$\|\Pi_C(x) - \Pi_C(y)\| \leq \|x - y\| \qquad \forall\, x, y \in C. \tag{8}$$

**Lemma 3.6.**

$$\mu_S(\Pi_S(L^{-1}(V_S))) \leq \det(L^{-1}) \left(\frac{1}{\lambda_{\min}(L^{-1})}\right)^n \mu_S(V_S). \tag{9}$$

*Proof.* Note that the mapping $\Pi_{\mathrm{S}}$ can be seen as the composition of the $\Pi_{\mathrm{S}_r}$ 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 3.5

$$\|\Pi_{\mathrm{S}_{\lambda_{\min}}}(x) - \Pi_{\mathrm{S}_{\lambda_{\min}}}(y)\| \leq \|x - y\|, \qquad \forall\; x, y \in E'. \tag{10}$$

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 $\Pi_{\mathrm{S}}$. 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 $\Pi_S$ on $E'$ is $\frac{1}{\lambda_{\min}}$, which means that:

$$\|\Pi_{\mathrm{S}}(x) - \Pi_{\mathrm{S}}(y)\| \leq \frac{1}{\lambda_{\min}} \|x - y\|, \qquad \forall\, x, y \in E'. \tag{11}$$

Note that, the inequality in (11) is an equality when $x$ is in the eigenspace of $\lambda_{\min}$ and $y = -x$.

Recall that for any smooth Lipschitz function $\phi$ with Lipschitz constant, $\mathrm{Lip}(\phi)$, we have for all $x$, $|\det(J(\phi(x))| \leq \mathrm{Lip}(\phi)^n$. Combining this with (11) and Lemma 3.6, we get the statement of the lemma. □

**Theorem 3.7.** *$\sigma^{n-1}(V') \leq m\epsilon\sqrt{\frac{\lambda_{\max}(P)^n}{\det(P)}}$, where $\mu(V) = \epsilon$.*

*Proof.* By taking $\mu_{\mathrm{S}}$ as the uniform spherical measure $\sigma^{n-1}$, and combining Corollary 3.2 with Lemma 3.6 we get the statement of the theorem. □

## 4 Relating $\epsilon$ to $\delta$

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 $\epsilon'$ to $\delta$ in the statement of the Theorem 2.3. We start by a few definitions that will help us along the way.

**Supporting Hyperplane** The hyperplane $H = \{x | c^T x = k\}$ is a supporting hyperplane of the set $X \subseteq \mathbb{R}^n$ if $X$ is contained either in the set $\{x | c^T x \leq k\}$ or $\{x | c^T x \geq k\}$ and $X$ contains one point $z \in X$. Then $H$ is the supporting hyperplane of $X$ at $z$. We denote the set of all supporting hyperplanes of a set $X$ at point $z$ by $H_z(X)$.

**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 4.1.** *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 $d(X,p) := \inf_{x \in X} d(x,p)$. Then if $X$ is compact, we have:*

$$d(X,p) = \min_{x \in X} d(x,p).$$

*Proof.* This is due to the fact that the function $d$ is continuous and therefore attains its minimum on the compact set $X$. $\square$

**Proposition 4.2** (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:

$$\Delta(X) = \sup\{r : \mathrm{B}_r \subseteq \text{convhull } (\mathrm{S} \setminus X)\}. \tag{12}$$

**Lemma 4.3.** $\Delta(\mathcal{C}_{c,k}) = \min\left(1, \frac{|k|}{\|c\|}\right)$.

*Proof.* Note that convhull $(\mathrm{S} \setminus \mathcal{C}_{c,k}) = \left\{x \in \mathrm{B} : c^T x \le k\right\}$. Let $X_{\mathcal{C}_{c,k}} :=$ convhull $(\mathrm{S} \setminus \mathcal{C}_{c,k})$.

$$\begin{aligned} \Delta(X) &= d(\partial X_{\mathcal{C}_{c,k}}, 0) \\ &= \min(d(\mathrm{S}, 0), d(\{x : c^T x = k\}, 0)) \\ &= \min\left(1, \frac{|k|}{\|c\|}\right). \end{aligned}$$

$\square$

**Corollary 4.4.** $\Delta(\mathcal{C}_{c,k_1}) < \Delta(\mathcal{C}_{c,k_2})$ *when* $k_1 < k_2$.

**Lemma 4.5.** $\sigma^{n-1}(\mathcal{C}_{c,k_1}) < \sigma^{n-1}(\mathcal{C}_{c,k_2})$*, 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 $k_1 > k_2$. $\square$

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

**Lemma 4.6.** *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*

$$\Delta(\mathcal{C}_{c,k}) = \Delta(X). \tag{13}$$

*Proof.* Let $X_S :=$ convhull $(S \setminus X)$. Note that when $X_S = \emptyset$, the statement of the theorem holds trivially. So for the rest of the proof we assume that $X_S \neq \emptyset$. Since the distance function $d$ continuous and the set $\partial X_S$ is compact there exists a point $x^* \in \partial X_S$, such that:

$$\Delta(X) = d(\partial X_S, 0) = \inf_{x \in \partial X_S} d(x, 0) = \min_{x \in \partial X_S} d(x, 0) = d(x^*, 0). \tag{14}$$

Then, by the supporting hyperplane theorem [**?**], there exists a supporting hyperplane $\left\{x : c^T x = k\right\}$ of $\partial X_S$ such that $x^* \in \left\{x : c^T x = k\right\}$. Note that:

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

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

We now prove our main result.

**Theorem 4.7.** *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 4.6 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 4.5. This implies $\Delta(\mathcal{C}_{c,\tilde{k}}) < \Delta(\mathcal{C}_{c,k}) = \Delta(X^*)$ due to Lemma 4.4. 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 4.8.** *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:

$$\epsilon' = \sigma^{n-1}(\mathcal{C}_{c,k}) = \frac{\pi^{d/2}}{\Gamma[\frac{d}{2}]} I_{2h-h^2}\left(\frac{d-1}{2}, \frac{1}{2}\right), \tag{15}$$

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

$$\begin{aligned} \frac{\epsilon'\Gamma[\frac{d}{2}]}{\pi^{d/2}} &= I_{2h-h^2}\left(\frac{d-1}{2}, \frac{1}{2}\right) \\ 2h - h^2 &= I^{-1}\left(\frac{\epsilon'\Gamma(\frac{d}{2})}{\pi^{d/2}}, \frac{d-1}{2}, \frac{1}{2}\right) \\ 2h - h^2 &= \alpha \\ h^2 - 2h + \alpha &= 0. \end{aligned} \tag{16}$$

From (16), 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 $\epsilon$ for fixed $n$. $\square$

**Corollary 4.9.** *For a given* $\beta = 1 - \sum_{j=0}^{d}\binom{N}{j}\epsilon^j(1-\epsilon)^{N-j}$*, we have*

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

*Proof.* We first show that $\lim_{N\to\infty} \epsilon'_\beta(N) = 0$.

$$\begin{aligned} 1 - \beta = \sum_{j=0}^{d}\binom{N}{j}\epsilon^j(1-\epsilon)^{N-j} &\leq (d+1)N^d(1-\epsilon)^N \\ \log\left(\frac{1-\beta}{d+1}\right) &\leq d\log N + N\log(1-\epsilon(N)) \\ \frac{\log\left(\frac{1-\beta}{d+1}\right)}{d\log N} &\leq 1 + \frac{N\log(1-\epsilon(N))}{d\log N} \\ \lim_{N\to+\infty}\frac{\log\left(\frac{1-\beta}{d+1}\right)}{d\log N} &< \lim_{N\to+\infty}\frac{N\log(1-\epsilon(N))}{d\log N} \\ 0 &< \lim_{N\to+\infty}\frac{N\log(1-\epsilon(N))}{d\log N} \end{aligned} \tag{17}$$

We now prove by contradiction that (17) implies that $\epsilon(N) \to 0$. Assume that $\lim_{N\to+\infty} \epsilon(N) \neq 0$. Let $g_1(N) := \frac{N}{d \log N}$ and $g_2(N) := \log(1-\epsilon(N))$. Note that $\lim_{N\to+\infty} g_1(N) = \infty$ and since $\epsilon(N) \in (0,1)$, we have $g_2(N) < 0$ for all $N$. Since $\lim_{N\to+\infty} \epsilon(N) \neq 0$, for there exists a subsequence $n_k$ of $\mathbb{N}$ such that $\lim_{k\to\infty} g_2(n_k) = -\epsilon$ for some $\epsilon > 0$. This means that there exists a subsequence $n_k$ of $\mathbb{N}$ such that $\lim_{k\to\infty} g_1(n_k)g_2(n_k) = -\infty$, which implies that either the limit $\lim_{N\to+\infty} \frac{N\log(1-\epsilon(N))}{d \log N}$ does not exist or it is $-\infty$, which is a contradiction since we have (17).

By continuity and monotonicity of $I^{-1}$ in its first parameter, $\delta = \sqrt{1-\alpha}$ tends to 1 as $\epsilon \to 0$.

$\square$