"The Capacity of the Relay Channel": Solution to Cover's Problem in the Gaussian Case

Xiugang Wu, *Member, IEEE*, Leighton Pate Barnes, *Student Member, IEEE*, and Ayfer Özgür, *Member, IEEE*

***Abstract*— Consider a memoryless relay channel, where the relay is connected to the destination with an isolated bit pipe of capacity C_0. Let $C(C_0)$ denote the capacity of this channel as a function of C_0. What is the critical value of C_0, such that $C(C_0)$ first equals $C(\infty)$? This is a long-standing open problem posed by Cover and named "The Capacity of the Relay Channel," in *Open Problems in Communication and Computation*, Springer-Verlag, 1987. In this paper, we answer this question in the Gaussian case and show that $C(C_0)$ cannot equal to $C(\infty)$ unless $C_0 = \infty$, regardless of the SNR of the Gaussian channels. This result follows as a corollary to a new upper bound we develop on the capacity of this channel. Instead of "single-letterizing" expressions involving information measures in a high-dimensional space as is typically done in converse results in information theory, our proof directly quantifies the tension between the pertinent n-letter forms. This is done by translating the information tension problem to a problem in high-dimensional geometry. As an intermediate result, we develop an extension of the classical isoperimetric inequality on a high-dimensional sphere, which can be of interest in its own right.**

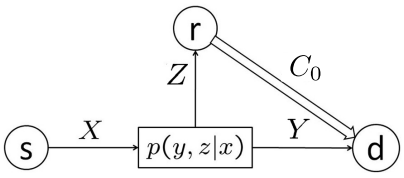
***Index Terms*— Relay channel, capacity, information inequality, geometry, isoperimetric inequality, concentration of measure.**

I. Problem Setup and Main Result

IN 1987, Thomas M. Cover formulated a seemingly simple question in *Open Problems in Communication and Computation*, Springer-Verlag [2], which he called "The Capacity of the Relay Channel". This problem, not much longer than a single page in [2], remains open to date. His problem statement, taken verbatim from [2] with only a few minor notation changes, is as follows:

The Capacity of the Relay Channel

Consider the following seemingly simple discrete memoryless relay channel: Here Z and Y are conditionally independent and conditionally identically distributed given X, that is, $p(z, y|x) = p(z|x)p(y|x)$. Also, the channel from Z to Y does not interfere with Y. A $(2^{nR}, n)$ code for this channel is a map $X^n : [1 : 2^{nR}] \to \mathcal{X}^n$, a relay function $f_n : \mathcal{Z}^n \to [1 : 2^{nC_0}]$



and a decoding function $g_n : \mathcal{Y}^n \times [1 : 2^{nC_0}] \to [1 : 2^{nR}]$. The probability of error is given by

$$P_e^{(n)} = Pr(g_n(Y^n, f_n(Z^n)) \neq M),$$

where the message M is uniformly distributed over $[1 : 2^{nR}]$ and

$$p(m, y^n, z^n) = 2^{-nR} \prod_{i=1}^{n} p(y_i|x_i(m)) \prod_{i=1}^{n} p(z_i|x_i(m)).$$

Let $C(C_0)$ be the supremum of achievable rates R for a given C_0, that is, the supremum of the rates R for which $P_e^{(n)}$ can be made to tend to zero. We note the following facts:

1. $C(0) = \sup_{p(x)} I(X; Y)$.
2. $C(\infty) = \sup_{p(x)} I(X; Y, Z)$.
3. *$C(C_0)$ is a nondecreasing function of C_0.*

What is the critical value of C_0 such that $C(C_0)$ first equals $C(\infty)$?

A. Main Result

As is customary in network information theory, Cover formulates the problem for discrete memoryless channels. However, the same question clearly applies to channels with continuous input and output alphabets, and in particular when the channels from the source to the relay and the destination are Gaussian, which is the canonical model for wireless relay channels. More formally, assume

$$\begin{cases} Z = X + W_1 \\ Y = X + W_2 \end{cases}$$

with the transmitted signal being constrained to average power P, i.e.,

$$\|x^n(m)\|^2 \leq nP, \quad \forall m \in [1 : 2^{nR}], \tag{1}$$

and $W_1, W_2 \sim \mathcal{N}(0, N)$ representing Gaussian noises that are independent of each other and X. See Fig. 1.

For this Gaussian relay channel, it is easy to observe that[1]

$$C(\infty) = \frac{1}{2} \log \left(1 + \frac{2P}{N}\right).$$

[1] All logarithms throughout the paper are to base two.

Manuscript received January 8, 2017; revised October 27, 2017 and September 10, 2018; accepted September 15, 2018. Date of publication October 22, 2018; date of current version December 19, 2018. The work was supported in part by NSF under Award CCF-1704624 and in part by the Center for Science of Information, an NSF Science and Technology Center, under Grant Agreement CCF-0939370. The work of X. Wu was done when he was with Stanford University. This paper was presented in part at the 2016 Allerton Conference on Communication, Control, and Computing [1].

X. Wu is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: xwu@udel.edu).

L. P. Barnes and A. Özgür are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: lpb@stanford.edu; aozgur@stanford.edu).

Communicated by M. Wigger, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2876892

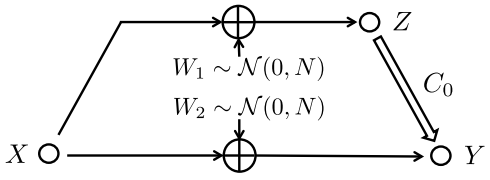


Fig. 1. Symmetric Gaussian relay channel.

Let C_0^* denote the threshold in Cover's problem, i.e.

$$C_0^* := \inf\{C_0 : C(C_0) = C(\infty)\}. \tag{2}$$

For the Gaussian model, there is no known scheme that allows to achieve $C(\infty)$ at a finite C_0 regardless of the parameters of the channels, i.e. the signal to noise power ratio (SNR) P/N. Therefore, from an achievability perspective we only have the trivial bound

$$C_0^* \leq \infty.$$

On the converse side, any upper bound on the capacity of this channel can be used to establish a lower bound on C_0^*. The only upper bound on the capacity of this channel (prior to our work in [5] and [6] preceding the current paper) was the celebrated cut-set bound developed by Cover and El Gamal in 1979 [10]. It yields the following lower bound on C_0^*:

$$C_0^* \geq \frac{1}{2}\log\left(1+\frac{2P}{N}\right) - \frac{1}{2}\log\left(1+\frac{P}{N}\right).$$

Note that the cut-set bound does not preclude achieving $C(\infty)$ at finite C_0. Moreover, it is interesting to note that as P/N decreases to zero, this lower bound decreases to zero. This implies a sharp dichotomy between the current achievability and converse results for this problem, which becomes even more apparent in the limit when SNR goes to zero: the cut-set bound does not preclude achieving $C(\infty)$ at diminishing C_0 if $C(\infty)$ itself is diminishing, while from an achievability perspective we need $C_0 = \infty$ regardless of the SNRs of the channels (apart from the trivial case when P/N is exactly equal to 0). The main result of our paper is to show that $C_0^* = \infty$ regardless of the parameters of the problem, answering Cover's long-standing question for the canonical Gaussian model.

Theorem 1: For the symmetric Gaussian relay channel depicted in Fig. 1, $C_0^* = \infty$.

This theorem follows immediately from the following theorem which establishes a new upper bound on the capacity of this channel for any C_0.

Theorem 2: For the symmetric Gaussian relay channel depicted in Fig. 1, the capacity $C(C_0)$ satisfies

$$C(C_0) \leq \frac{1}{2}\log\left(1+\frac{P}{N}\right) + \sup_{\theta\in\left[\arcsin(2^{-C_0}),\frac{\pi}{2}\right]} \min\left\{\begin{array}{c} C_0 + \log\sin\theta, \\ \min\limits_{\omega\in\left(\frac{\pi}{2}-\theta,\frac{\pi}{2}\right]} h_\theta(\omega)\end{array}\right\}$$

where

$$h_\theta(\omega) := \tfrac{1}{2}\log\left(\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})\sin^2\theta}{(P+N)(\sin^2\theta-\cos^2\omega)}\right).$$

In Fig. 2 we plot this upper bound (label: New bound) under three different SNR values of the Gaussian channels, together with the cut-set bound [10] and an upper bound on the capacity of this channel we have previously derived in [6] (label: Old bound). For reference, we also provide the rate achieved by a compress-and-forward relay strategy (label: C-F), which employs Gaussian input distribution at the source combined with Gaussian quantization and Wyner-Ziv binning at the relay.[2] The flat levels at which the cut-set bound and our old bound saturate in these plots precisely correspond to $C(\infty)$. Note that while these earlier bounds reach $C(\infty)$ at finite C_0 values, hence leading to finite lower bounds on C_0^*, our new bound remains bounded away from $C(\infty)$ in all the three plots. Indeed, it can be formally shown that the new bound remains bounded away from $C(\infty)$ (the flat level in the plots) at any finite C_0 value. We prove this formally in the proof of Theorem 1.

While in this paper we restrict our attention to the symmetric case, an assumption imposed by Cover in his original formulation of the problem given above, our methods and results also extend to the asymmetric case. In [8], we show that when the relay's and the destination's observations are corrupted by independent Gaussian noises of different variances, it is still true that $C_0^* = \infty$ regardless of the channel parameters. The extension to this asymmetric case heavily builds on the methods and results we develop in this paper for the symmetric case. Interestingly, the symmetric case, which Cover seems to somewhat arbitrarily assume in his problem formulation, turns out to be the canonical case for our proof technique. We also provide a solution to Cover's problem for binary symmetric channels in [9] using a similar approach.

B. Technical Approach

There are two basic aspects in an information-theoretic characterization of an operational problem: the so-called achievability result and converse result. An achievability result establishes what is possible in a given setting, while the converse result distinguishes what is impossible. The ideal situation is when these two results match, in which case an information limit is born. The most famous example goes back to Shannon and the inception of the field: Reliable communication is possible over a noisy channel if, and only if, the rate of transmission does not exceed the capacity of the channel [18].

Over the last two decades, there has been significant leap forward in developing achievable schemes for multi-user problems, ranging from schemes based on interference alignment and distributed MIMO, to lattice-based techniques, to strategies inspired by network coding and linear deterministic models. This stands in fairly stark contrast to the set of converse arguments in the information theorist's toolkit. Almost all converse arguments rely on a few fundamental tools

[2] In the low SNR regime, we can achieve higher rates using bursty compress-and-forward [21], as demonstrated in the left-most plot of Fig. 2. Note that since we still impose the Gaussian restriction on the input and quantization distributions for bursty compress-forward, the resultant rates are not concave in C_0 and can be further improved by time sharing.

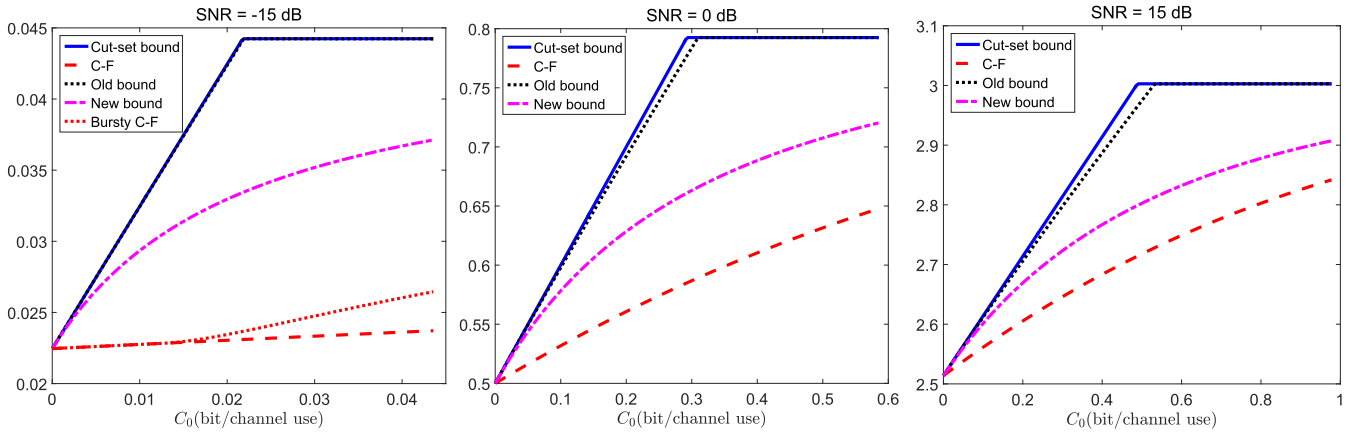


Fig. 2. Upper bounds and achievable rates for the Gaussian relay channel.

that go back to the early years of the field: information measure calculus (e.g., chain rules, non-negativity of divergence), Fano's inequality, and the entropy power inequality. The typical converse program follows from a clever application of these tools to "single-letterize" an expression involving information measures in a high-dimensional space (so called n-letter forms), with the possible introduction of auxiliary random variables as needed.

In this paper, we take a different approach. Instead of focusing on single-letterizing pertinent n-letter forms, we aim to directly quantify the tension between them. To do this, we lift the problem to an even higher dimensional space and study the geometry of the typical sequences generated independently and identically (i.i.d.) from these n-dimensional distributions. We establish non-trivial geometric properties satisfied by these typical sequences, which are then translated to inequalities satisfied by the original n-dimensional information measures. This notion of "typicality", connecting information measures associated with a distribution to probabilities of long i.i.d. sequences generated from this distribution, is a standard tool in establishing achievability results in information theory but to the best of our knowledge has been rarely used in proving converse results in network information theory, with only a few examples such as the work of Zhang [11] from 1988 and our recent works [3]–[7].

To study the geometry of the typical sequences, we use classical tools from high-dimensional geometry, such as the isoperimetric inequality [14], measure concentration [12], and rearrangement and symmetrization theory [13], [25]. We also prove a new geometric result which can be regarded as an extension of the classical isoperimetric inequality on a high-dimensional sphere and can be of interest in its own right. Note that the classical isoperimetric inequality on the sphere states that among all sets on the sphere with a given measure (area), the spherical cap has the smallest boundary or more generally the smallest neighborhood [16]. As an intermediate result in this paper, we show that the spherical cap not only minimizes the measure of its neighborhood, but roughly speaking, also minimizes the measure of its intersection with the neighborhood of a randomly chosen point on the sphere.

The incorporation of geometric insight in information theory is not new. Formulating the problem of determining the communication capacity of channels as a problem in high-dimensional geometry is indeed one of Shannon's most important insights that has led to the conception of the field. In his classical paper "Communication in the presence of noise", 1949 [17], Shannon develops a geometric representation of any point-to-point communication system, and then uses this geometric representation to derive the capacity formula for the AWGN channel. His converse proof is based on a sphere-packing argument, which relies on the notion of sphere hardening (i.e. measure concentration) in high-dimensional space. Our approach resembles Shannon's approach in [17] in that the main argument in our proof is also a packing argument; however, instead of packing smaller spheres in a larger sphere, we pack (quantization) regions of some minimal measure (and unknown shape) inside a spherical cap. The key ingredient in our packing argument is the extended isoperimetric inequality we develop, which guarantees that each of these quantization regions has some minimal intersection with the spherical cap. Also, note that we do not directly study the geometry of the codewords as in [17], but rather use geometry in an indirect way to solve an n-letter information tension problem.

C. Organization of the Paper

The remainder of the paper is organized as follows. In Section II, we review some basic definitions and results for high-dimensional spheres, and state our main geometric result in Theorem 7, which can be regarded as an extension of the classical isoperimetric inequality on the sphere. In Section III, we introduce some typicality lemmas and combine them with Theorem 7 to prove a key information inequality stated in Theorem 8. The proofs of our main theorems, Theorem 1 and 2, are almost immediate given Theorem 8 and are provided in Section IV.

Appendices A and B are then devoted to the proof of Theorem 7 and the proofs of the typicality lemmas introduced in Section III, respectively. The proofs of these typicality lemmas require us to derive formulas and exponential characterizations for the area/volume of various high dimensional sets including

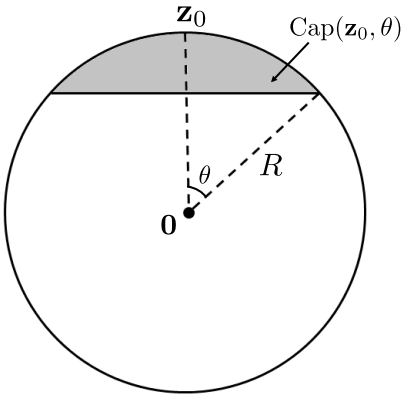


Fig. 3. A spherical cap with angle θ.

balls, spherical caps, shell caps, and intersections of such sets. We derive these characterizations in Appendix C.

II. Geometry of High-Dimensional Spheres

In this section, we summarize some basic definitions and results for high-dimensional spheres and present our main geometric result which can be regarded as an extension of the classical isoperimetric inequality on high-dimensional spheres. This result is the key to proving the information inequality we present in the next section, which in turn is the key to proving Theorems 1 and 2.

A. Basic Results on High-Dimensional Spheres

We now summarize some basic results on high-dimensional spheres that will be referred to later in the paper.

(i) **Isoperimetric Inequality:** Let $\mathbb{S}^{m-1} \subseteq \mathbb{R}^m$ denote the $(m-1)$-sphere of radius R, i.e.,

$$\mathbb{S}^{m-1} = \left\{\mathbf{z} \in \mathbb{R}^m : \|\mathbf{z}\| = R\right\},$$

equipped with the rotation invariant (Haar) measure $\mu = \mu_{m-1}$ that is normalized such that

$$\mu(\mathbb{S}^{m-1}) = \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} R^{m-1},$$

i.e. the usual surface area. Let $\mathbb{P}(A)$ denote the probability of a set or event A with respect to the corresponding Haar probability measure, i.e. the normalized Haar measure such that $\mathbb{P}(\mathbb{S}^{m-1}) = 1$. A spherical cap is defined as a ball on $\mathbb{S}^{m-1}$ in the geodesic metric (or simply the angle) $\angle(\mathbf{z}, \mathbf{y}) = \arccos(\langle \mathbf{z}/R, \mathbf{y}/R\rangle)$, i.e.,

$$\mathrm{Cap}(\mathbf{z}_0, \theta) = \left\{\mathbf{z} \in \mathbb{S}^{m-1} : \angle(\mathbf{z}_0, \mathbf{z}) \le \theta\right\}.$$

See Fig. 3. We will often say that an arbitrary set $A \subseteq \mathbb{S}^{m-1}$ has an effective angle θ if $\mu(A) = \mu(C)$, where $C = \mathrm{Cap}(\mathbf{z}_0, \theta)$ for some arbitrary $\mathbf{z}_0 \in \mathbb{S}^{m-1}$.

The following proposition is the so-called isoperimetric inequality, which was first proved by Levy in 1951 [14]. (See also [16].) It states the intuitive fact that among all sets on the sphere with a given measure, the spherical cap has the smallest boundary, or more generally the smallest neighborhood. This is formalized as follows:

Proposition 3: For any arbitrary set $A \subseteq \mathbb{S}^{m-1}$ such that $\mu(A) = \mu(C)$, where $C = \mathrm{Cap}(\mathbf{z}_0, \theta) \subseteq \mathbb{S}^{m-1}$ is a spherical cap, it holds that

$$\mu(A_t) \ge \mu(C_t), \quad \forall t \ge 0,$$

where A_t is the t-neighborhood of A, defined as

$$A_t = \left\{\mathbf{z} \in \mathbb{S}^{m-1} : \min_{\mathbf{z}' \in A} \angle(\mathbf{z}, \mathbf{z}') \le t\right\},$$

and similarly

$$C_t = \left\{\mathbf{z} \in \mathbb{S}^{m-1} : \min_{\mathbf{z}' \in C} \angle(\mathbf{z}, \mathbf{z}') \le t\right\} = \mathrm{Cap}(\mathbf{z}_0, \theta + t).$$

(ii) **Measure Concentration:** Measure concentration on the sphere refers to the fact that most of the measure of a high-dimensional sphere is concentrated around any equator. The following elementary result capturing this phenomenon will be used later in the paper when we prove the extended isoperimetric inequality.

Proposition 4: Given any $\epsilon, \delta > 0$, there exists some $M(\epsilon, \delta)$ such that for any $m \ge M(\epsilon, \delta)$ and any $\mathbf{z} \in \mathbb{S}^{m-1}$,

$$\mathbb{P}\left(\angle(\mathbf{z}, \mathbf{Y}) \in [\pi/2 - \epsilon, \pi/2 + \epsilon]\right) \ge 1 - \delta, \tag{3}$$

where $\mathbf{Y} \in \mathbb{S}^{m-1}$ is distributed according to the Haar probability measure.

Proof: Let $\mathbf{e}_1 = (R, 0, \ldots, 0)$. Note for any $\mathbf{z} \in \mathbb{S}^{m-1}$, the distribution of $\angle(\mathbf{z}, \mathbf{Y})$ is the same as the distribution of $\angle(\mathbf{e}_1, \mathbf{Y})$, since $\mathbf{z}$ can be written in the form $\mathbf{z} = U\mathbf{e}_1$, where U is an orthogonal matrix, and the distribution of $\mathbf{Y}$ is rotation-invariant. Therefore, without loss of generality, we can assume $\mathbf{z} = \mathbf{e}_1$. Since $\langle \mathbf{e}_1/R, \mathbf{Y}/R\rangle = Y_1/R$, we have $E[\langle \mathbf{e}_1/R, \mathbf{Y}/R\rangle] = E[Y_1]/R = 0$; we also have $E[\langle \mathbf{e}_1/R, \mathbf{Y}/R\rangle^2] = E[Y_1^2]/R^2 = 1/m$ because $E[Y_1^2] = \cdots = E[Y_m^2]$ and $E[Y_1^2] + \cdots + E[Y_m^2] = R^2$. Therefore by Chebyshev's inequality, for any $\mu > 0$,

$$\mathbb{P}(|\langle \mathbf{e}_1/R, \mathbf{Y}/R\rangle| \ge \mu) \le \frac{1}{m\mu^2}.$$

Recalling that $\angle(\mathbf{e}_1, \mathbf{Y}) = \arccos(\langle \mathbf{e}_1/R, \mathbf{Y}/R\rangle)$ and noting that the R.H.S. of the above inequality can be made arbitrarily small by choosing m to be sufficiently large, we have proved the proposition. ∎

(iii) **Blowing-Up Lemma:** The above measure concentration result combined with the isoperimetric inequality immediately yields the following result:

Proposition 5: Let $A \subseteq \mathbb{S}^{m-1}$ be an arbitrary set and $C = \mathrm{Cap}(\mathbf{z}_0, \theta) \subseteq \mathbb{S}^{m-1}$ be a spherical cap such that $\mu(A) = \mu(C)$, i.e. A has an effective angle of θ. Then for any $\epsilon > 0$ and m sufficiently large,

$$\mathbb{P}(A_{\frac{\pi}{2} - \theta + \epsilon}) \ge 1 - \epsilon. \tag{4}$$

Proof: If $A = \mathrm{Cap}(\mathbf{z}_0, \theta)$, $\mathbb{P}(A_{\frac{\pi}{2} - \theta + \epsilon}) \ge 1 - \epsilon$ due to Proposition 4. If A is not a spherical cap, then $\mathbb{P}(A_{\frac{\pi}{2} - \theta + \epsilon}) \ge P(C_{\frac{\pi}{2} - \theta + \epsilon})$ where $C = \mathrm{Cap}(\mathbf{z}_0, \theta)$, due to the isoperimetric inequality in Proposition 3. ∎

If we take A to be a half sphere, this result says that most of the measure of the sphere is concentrated around the boundary of this half-sphere, i.e. an equator, which is the result in Proposition 4. However, due to the isoperimetric inequality, Proposition 5 allows us to make the stronger statement that the measure is concentrated around the boundary of any set with probability $1/2$. While the elementary results we establish above suggest that this concentration takes place at a polynomial speed in the dimension m, it can be shown that the measure concentrates around the boundary of any set with probability $1/2$ exponentially fast in the dimension m; see [15].

B. Extended Isoperimetry on the Sphere and the Shell

An almost equivalent way to state the blowing-up lemma in Proposition 5 is the following: Let $A \subseteq \mathbb{S}^{m-1}$ be an arbitrary set with effective angle $\theta > 0$. Then for any $\epsilon > 0$ and sufficiently large m,

$$\mathbb{P}\left(\mu\left(A \cap \mathrm{Cap}\left(\mathbf{Y}, \frac{\pi}{2} - \theta + \epsilon\right)\right) > 0\right) > 1 - \epsilon, \qquad (5)$$

where $\mathbf{Y}$ is distributed according to the normalized Haar measure on $\mathbb{S}^{m-1}$. In words, if we take a $\mathbf{y}$ uniformly at random on the sphere and draw a spherical cap of angle slightly larger than $\frac{\pi}{2} - \theta$ around it, this cap will intersect the set A with high probability. This statement is almost equivalent to (4) since the $\mathbf{y}$'s for which the intersection has non-zero measure lie in the $\frac{\pi}{2} - \theta + \epsilon$-neighborhood of A. Note that similarly to Proposition 5, this statement would trivially follow from measure concentration on the sphere (Proposition 4) if A were known to be a spherical cap, and it holds for any A due to the isoperimetric inequality in Proposition 3. By building on the Riesz rearrangement inequality [25], we prove the following extended result:

Theorem 6: Let $A \subseteq \mathbb{S}^{m-1}$ be any arbitrary subset of $\mathbb{S}^{m-1}$ with effective angle $\theta > 0$, and let $V = \mu(\mathrm{Cap}(\mathbf{z}_0, \theta) \cap \mathrm{Cap}(\mathbf{y}_0, \omega))$ where $\mathbf{z}_0, \mathbf{y}_0 \in \mathbb{S}^{m-1}$ with $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and $\theta + \omega > \pi/2$. (See Fig. 4.) Then for any $\epsilon > 0$, there exists an $M(\epsilon)$ such that for $m > M(\epsilon)$,

$$\mathbb{P}\left(\mu(A \cap \mathrm{Cap}(\mathbf{Y}, \omega + \epsilon)) > (1 - \epsilon)V\right) \geq 1 - \epsilon,$$

where $\mathbf{Y}$ is a random vector on $\mathbb{S}^{m-1}$ distributed according to the normalized Haar measure.

If A itself is a cap, then the statement in Theorem 6 is straightforward and follows from the fact that $\mathbf{Y}$ with high probability will be concentrated around the equator at angle $\pi/2$ from the pole of A (Proposition 4). Therefore, as m gets large for almost all $\mathbf{Y}$, the intersection of the two spherical caps will be given by V. See Fig. 4. The statement, however, is stronger than this and holds for any arbitrary set A, analogous to the isoperimetric inequality in (5). It states that no matter what the set A is, if we take a random point on the sphere and draw a cap of angle slightly larger than ω centered at this point, for any $\omega > \pi/2 - \theta$, then with high probability the intersection of the cap with the set A would be at least as large as the intersection we would get if A were a spherical cap. In this sense, Theorem 6 can be regarded as an extension of the isoperimetric inequality in Proposition 3, even though the latter can be stated purely geometrically and implies the weaker probabilistic statement in (5), while our result is inherently probabilistic.

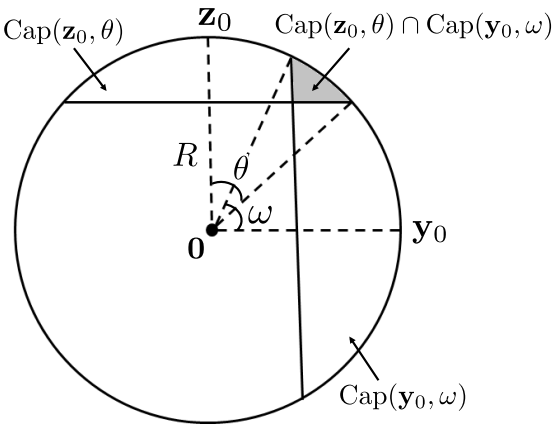


Fig. 4. Intersection of two spherical caps.

Theorem 6 is in fact a special case of a more general theorem that is true for subsets on a spherical shell. Let

$$\mathbb{L}^m = \{\mathbf{y} \in \mathbb{R}^m : \ R_L \leq \|\mathbf{y}\| \leq R_U\}$$

be this shell, where $0 \leq R_L \leq R_U$. A cap on this shell with pole $\mathbf{z}_0$ and angle θ can be defined as a ball in terms of the angle:

$$\angle(\mathbf{y}, \mathbf{z}) = \arccos\left(\frac{\mathbf{y} \cdot \mathbf{z}}{\|\mathbf{y}\|\|\mathbf{z}\|}\right)$$

on the shell, i.e.,

$$\mathrm{ShellCap}(\mathbf{z}_0, \theta) = \left\{\mathbf{z} \in \mathbb{L}^m : \angle(\mathbf{z}_0, \mathbf{z}) \leq \theta\right\}.$$

Let $|A|$ denote the standard m-dimensional Euclidean measure of a subset $A \subseteq \mathbb{L}^m$. We will say that an arbitrary set $A \subseteq \mathbb{L}^m$ has effective angle $\theta > 0$ if its measure is equal to that of a shell cap of angle θ, i.e. $|A| = |\mathrm{ShellCap}(\mathbf{z}_0, \theta)|$ for some $\mathbf{z}_0 \in \mathbb{L}^m$. We will also say that a probability measure $\mathbb{P}$ for subsets of $\mathbb{L}^m$ is rotationally invariant if $\mathbb{P}(A) = \mathbb{P}(UA)$ for any orthogonal matrix U, where UA denotes the image of the set A under the linear transformation U. The following more general theorem holds in the shell setting.

Theorem 7: Let $A \subseteq \mathbb{L}^m$ be any arbitrary subset of $\mathbb{L}^m$ with effective angle $\theta > 0$, and let $V = |\mathrm{ShellCap}(\mathbf{z}_0, \theta) \cap \mathrm{ShellCap}(\mathbf{y}_0, \omega)|$ where $\mathbf{z}_0, \mathbf{y}_0 \in \mathbb{L}^m$ with $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and $\theta + \omega > \pi/2$. Then for any $\epsilon > 0$, there exists an $M(\epsilon)$ such that for $m > M(\epsilon)$,

$$\mathbb{P}\left(|A \cap \mathrm{ShellCap}(\mathbf{Y}, \omega + \epsilon)| > (1 - \epsilon)V\right) \geq 1 - \epsilon,$$

where $\mathbf{Y}$ is a random vector drawn from any rotationally invariant probability measure on $\mathbb{L}^m$.

We prove Theorems 6 and 7 in Appendix A. Note that $M(\epsilon)$ in these two results depends only on ϵ—in particular it does not depend on the radius parameters for $\mathbb{L}^m$ and $\mathbb{S}^{m-1}$, respectively, which means that these two results also apply if the radius parameters depend on the dimension m. In the following section, we will be mainly interested in the case

when the radius parameters scale in the square-root of the dimension.

III. Information Tension in a Symmetric Markov Chain

In this section, we prove an inequality between information measures in a certain type of Markov chain, which can be of interest in its own right. The proof of this inequality builds on Theorem 7 from the previous section. As we will see in Section IV, the main theorems in this paper, i.e. Theorems 1 and 2, are almost immediate given this result. We now state this result in the following theorem.

Theorem 8: Consider a Markov chain $I_n - Z^n - X^n - Y^n$ where X^n, Y^n and Z^n are n-length random vectors and $I_n = f_n(Z^n)$ is a deterministic mapping of Z^n to a set of integers. Assume moreover that Z^n and Y^n are i.i.d. white Gaussian vectors given X^n, i.e. $Z^n, Y^n \sim \mathcal{N}(X^n, N\, I_{n\times n})$ where $I_{n\times n}$ denotes the identity matrix, $E[\|X^n\|^2] = nP$, and $H(I_n|X^n) = -n \log \sin\theta_n$ for some $\theta_n \in [0, \pi/2]$. Then the following inequality holds for any n,

$$H(I_n|Y^n) \leq n \cdot \min_{\omega\in(\frac{\pi}{2}-\theta_n, \frac{\pi}{2}]} \frac{1}{2}\log\left(\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega)}\right). \quad (6)$$

Note that $H(I_n|Y^n)$ is trivially lower bounded by $H(I_n|X^n)$ for any Markov chain $I_n - Z^n - X^n - Y^n$. The above theorem says that if $I_n - Z^n - X^n - Y^n$ satisfies the conditions of the theorem, then $H(I_n|Y^n)$ can also be upper bounded in terms of $H(I_n|X^n)$. In particular, it provides an upper bound on $H(I_n|Y^n)$ in terms of $\theta_n = \arcsin\, 2^{-\frac{1}{n}H(I_n|X^n)}$. It can be easily verified that this upper bound on $H(I_n|Y^n)$ is decreasing with increasing θ_n, or equivalently decreasing with decreasing $H(I_n|X^n)$, and implies that $H(I_n|Y^n) \to 0$ as $H(I_n|X^n) \to 0$.

We next turn to proving Theorem 8. The reader who is interested in seeing how this theorem leads to Theorems 1 and 2, without seeing its own proof, can jump to Section IV. In order to prove Theorem 8, we will first establish some properties that are satisfied with high probability by long i.i.d. sequences generated from the source distribution (I_n, Z^n, X^n, Y^n) satisfying the assumptions of the theorem. We now state and discuss these properties in Section III-A and then use them to prove Theorem 8 in Section III-B.

A. Typicality Lemmas

Assume (I_n, Z^n, X^n, Y^n) satisfy the assumptions of Theorem 8. Consider the B-length i.i.d. sequence

$$\{(I_n(b), Z^n(b), X^n(b), Y^n(b))\}_{b=1}^B, \quad (7)$$

where for any $b \in [1:B]$, $(I_n(b), Z^n(b), X^n(b), Y^n(b))$ has the same distribution as (I_n, Z^n, X^n, Y^n). For notational convenience, in the sequel we write the B-length sequence $[X^n(1), X^n(2), \ldots, X^n(B)]$ as $\mathbf{X}$ and similarly define $\mathbf{Y}, \mathbf{Z}$ and $\mathbf{I}$; note that we have $\mathbf{I} = [f_n(Z^n(1)), f_n(Z^n(2)), \ldots, f_n(Z^n(B))] =: f(\mathbf{Z})$. Also let $\text{Shell}(\mathbf{c}, r_1, r_2)$ denote the spherical shell

$$\text{Shell}(\mathbf{c}, r_1, r_2) := \left\{\mathbf{a} \in \mathbb{R}^{nB} : r_1 \leq \|\mathbf{a}-\mathbf{c}\| \leq r_2\right\},$$

and let $\text{Ball}(\mathbf{c}, r)$ denote the Euclidean ball

$$\text{Ball}(\mathbf{c}, r) := \left\{\mathbf{a} \in \mathbb{R}^{nB} : \|\mathbf{a}-\mathbf{c}\| \leq r\right\}.$$

We next state several properties that $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{I}$ satisfy with high probability when B is large. The proofs of these properties are given in Appendix B.

Lemma 9: For any $\delta > 0$ and B sufficiently large, we have

$$\Pr(E_1) \geq 1-\delta$$

and

$$\Pr(E_2) \geq 1-\delta,$$

where E_1 and E_2 are defined to be the following two events respectively:

$$\left\{\mathbf{Z} \in \text{Shell}\left(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)}\right)\right\}, \quad (8)$$

and

$$\left\{\mathbf{Y} \in \text{Shell}\left(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)}\right)\right\}. \quad (9)$$

The proof of this lemma is a simple application of the law of large numbers and is included in Appendix B-A. The lemma simply states that when B is large, $\mathbf{Y}$ and $\mathbf{Z}$ will concentrate in a thin nB-dimensional shell of radius $\sqrt{nB(P+N)}$.

Lemma 10: Given any $\epsilon > 0$ and a pair of $(\mathbf{x}, \mathbf{i})$, let $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ be a set of $\mathbf{z}$'s defined as[3]

$$\begin{aligned} S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) := \Big\{&\mathbf{z} \in f^{-1}(\mathbf{i}) : \\ &\|\mathbf{x}-\mathbf{z}\| \in [\sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}] \quad (10)\\ &\mathbf{z} \in \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\epsilon)}\right) \quad (11)\\ &2^{nB(\log\sin\theta_n - \epsilon)} \leq p(f(\mathbf{z})|\mathbf{x}) \leq 2^{nB(\log\sin\theta_n+\epsilon)}\Big\} \quad (12)\end{aligned}$$

where $\theta_n = \arcsin\, 2^{-\frac{1}{n}H(I_n|X^n)}$ as in Theorem 8. Then for B sufficiently large, there exists a set $S_\epsilon(X^n, I_n)$ of $(\mathbf{x}, \mathbf{i})$ pairs, such that

$$\Pr((\mathbf{X}, \mathbf{I}) \in S_\epsilon(X^n, I_n)) \geq 1-\sqrt{\epsilon}, \quad (13)$$

and for any $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$,

$$\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) \geq 2^{nB(\log\sin\theta_n - 2\epsilon)}. \quad (14)$$

This lemma establishes the existence of a high probability set $S_\epsilon(X^n, I_n)$ of $(\mathbf{x}, \mathbf{i})$ sequences, and a conditional typical set $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ for each $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$ such that $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ satisfies some natural properties. Note that all properties in the definition of $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ as well as (14) are analogous to properties of strongly typical sets as stated in [21, Ch. 2]. However, the notion of strong typicality does not apply to the current case since Z^n and Y^n are continuous random vectors and X^n may or may not be continuous. Nevertheless, analogous properties can still be proved in this case; see the proof of this lemma in Appendix B-B.

[3]Note that under this definition of $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$, if a pair $(\mathbf{x}, \mathbf{i})$ doesn't satisfy $2^{nB(\log\sin\theta_n-\epsilon)} \leq p(\mathbf{i}|\mathbf{x}) \leq 2^{nB(\log\sin\theta_n+\epsilon)}$, then the set $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ is empty because no $\mathbf{z}$ can satisfy the condition in (12).

The following result has a slightly different flavor from the previous two lemmas in that it is simply a corollary of Theorem 7 from Section II.

Corollary 11: For any N, ϵ such that $N > \epsilon > 0$, consider the spherical shell in $\mathbb{R}^m$

$$\text{Shell}\left(\mathbf{0}, \sqrt{m(N-\epsilon)}, \sqrt{m(N+\epsilon)}\right) = \left\{\mathbf{y} \in \mathbb{R}^m : \sqrt{m(N-\epsilon)} \le \|\mathbf{y}\| \le \sqrt{m(N+\epsilon)}\right\}.$$

Let $A \subseteq \text{Shell}\left(\mathbf{0}, \sqrt{m(N-\epsilon)}, \sqrt{m(N+\epsilon)}\right)$ be an arbitrary subset on this shell with volume

$$|A| \ge 2^{\frac{m}{2}\log 2\pi e(N+\epsilon)\sin^2\theta}, \tag{15}$$

where $\theta \in (0, \pi/2)$. For any $\omega \in (\pi/2 - \theta, \pi/2]$ and m sufficiently large, we have

$$\Pr\left(\left|A \cap \text{Ball}\left(\mathbf{Y}, 2\sqrt{m(N+\epsilon)}\sin\frac{\omega+\epsilon}{2} + 2\sqrt{m\epsilon}\right)\right| \ge 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon]}\right) \ge 1 - \epsilon, \tag{16}$$

where $\mathbf{Y}$ is drawn from any rotationally invariant distribution on the Shell $\left(\mathbf{0}, \sqrt{m(N-\epsilon)}, \sqrt{m(N+\epsilon)}\right)$.

This is a simple corollary of Theorem 7 when applied to a specific shell and a subset A of this shell with measure prescribed by (15). The prescribed measure means that A has an effective angle (asymptotically) greater than or equal to θ. The corollary follows by observing that due to the triangle inequality (see also Fig. 5), for any $\mathbf{y}$ in the shell, ShellCap$(\mathbf{y}, \omega + \epsilon)$ considered in Theorem 7 is contained in the Euclidean ball

$$\text{Ball}\left(\mathbf{y}, 2\sqrt{m(N+\epsilon)}\sin\frac{\omega+\epsilon}{2} + 2\sqrt{m\epsilon}\right).$$

The lower bound on the intersection volume in (16) follows from an explicit characterization of

$$V = |\text{ShellCap}(\mathbf{z}_0, \theta) \cap \text{ShellCap}(\mathbf{y}_0, \omega)|$$

in Theorem 7, where $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and $\theta + \omega > \pi/2$; see Appendix C-B, and in particular Lemma 16, for this characterization. A formal proof of Corollary 11 is given in Appendix B-C.

The above corollary together with Lemma 10 leads to the following lemma.

Lemma 12: For any $\delta > 0$ and B sufficiently large, we have

$$\Pr(E_3) \ge 1 - \delta,$$

where E_3 is defined to be the following event:

$$\left\{\left|f^{-1}(\mathbf{I}) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right) \cap \text{Ball}\left(\mathbf{Y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right)\right| \ge 2^{nB[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n - \cos^2\omega)) - \delta]}\right\} \tag{17}$$

in which $f^{-1}(\mathbf{I}) := \{\mathbf{a} \in \mathbb{R}^{nB} : f(\mathbf{a}) = \mathbf{I}\}$ and $\omega \in (\pi/2 - \theta_n + \delta, \pi/2]$.

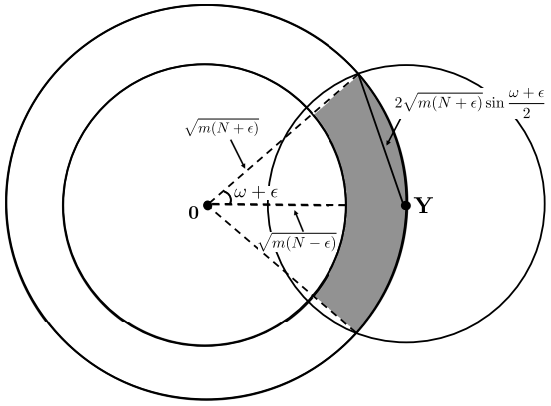


Fig. 5. Euclidean ball contains the shell cap.

This lemma can also be regarded as a typicality lemma as it states a property satisfied by $(\mathbf{I}, \mathbf{Y})$ pair with high probability when B is large. However, this is a non-trivial property. The lemma follows by first fixing a pair $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$ and showing that the volume of the set $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ defined in Lemma 10 can be lower bounded by

$$2^{\frac{nB}{2}\log(2\pi eN\sin^2\theta_n)},$$

up to the first order term in the exponent. Since by definition $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ is a subset of the shell

$$\text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right),$$

and given $\mathbf{X} = \mathbf{x}$, $\mathbf{Y}$ is isotropic Gaussian (therefore rotationally invariant around $\mathbf{x}$ when constrained to this shell), we can apply Corollary 11 to the above shell by choosing the set A to be $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$. This allows us to conclude that

$$\Pr\left(\left|S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) \cap \text{Ball}\left(\mathbf{Y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \epsilon\right)}\right)\right| \ge 2^{nB\left[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n - \cos^2\omega)) - \epsilon\right]} \middle| \mathbf{X} = \mathbf{x}\right) \ge 1 - \epsilon. \tag{18}$$

The conclusion of Lemma 12 then follows by observing that by definition

$$S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) \subseteq f^{-1}(\mathbf{i}) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\epsilon)}\right),$$

and removing the conditioning with respect to $\mathbf{X}$ in (18). The formal proof of Lemma 12 is given in Appendix B-D.

B. Proof of Theorem 8

We are now ready to prove Theorem 8, which mainly builds on Lemma 12. Consider a $\mathbf{Y}$ that with high probability lies in the ball with center $\mathbf{0}$ and approximate radius $\sqrt{nB(P+N)}$, and draw another ball around $\mathbf{Y}$ of approximate radius $\sqrt{nBN4\sin^2\frac{\omega}{2}}$ and intersect this ball with the original ball; equivalently, this corresponds to considering a cap around $\mathbf{Y}$ of angle ϕ on the original ball (see Fig. 6). Lemma 12 asserts that this cap around $\mathbf{Y}$ will have a certain minimal intersection volume with $f^{-1}(\mathbf{I})$. In other words, there is a subset of this cap with certain minimal volume that is mapped

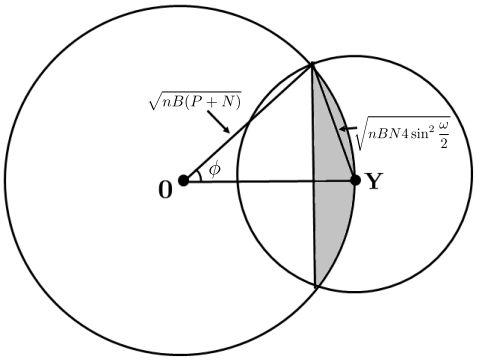


Fig. 6. A spherical cap with angle $\phi = 2\arcsin\sqrt{\frac{N\sin^2\frac{\omega}{2}}{P+N}}$.

to $\mathbf{I}$. This naturally lends itself to a packing argument: the number of distinct $\mathbf{I}$ values plausible under a given $\mathbf{Y}$ can be upper bounded by the ratio between the volume of the cap around $\mathbf{Y}$ and the minimal intersection volume occupied for each distinct $\mathbf{I}$. This in turn leads to a bound on $H(\mathbf{I}|\mathbf{Y})$.

We now proceed with the formal proof. Consider the indicator function

$$F = \mathbb{I}(E_1, E_2, E_3)$$

where $\mathbb{I}(\cdot)$ is defined as

$$\mathbb{I}(A) = \begin{cases} 1 & \text{if } A \text{ holds} \\ 0 & \text{otherwise,} \end{cases}$$

and the events E_1, E_2 and E_3 are as given by (8), (9) and (17) respectively. Obviously, by the union bound, we have

$$\Pr(F = 1) \geq 1 - 3\delta$$

for any $\delta > 0$ and B sufficiently large, and therefore

$$\begin{aligned} H(\mathbf{I}|\mathbf{Y}) &\leq H(\mathbf{I}, F|\mathbf{Y}) \\ &= H(F|\mathbf{Y}) + H(\mathbf{I}|\mathbf{Y}, F) \\ &\leq H(\mathbf{I}|\mathbf{Y}, F) + 1 \\ &= \Pr(F = 1) H(\mathbf{I}|\mathbf{Y}, F = 1) \\ &\quad + \Pr(F = 0) H(\mathbf{I}|\mathbf{Y}, F = 0) + 1 \\ &\leq H(\mathbf{I}|\mathbf{Y}, F = 1) + 3\delta nBC_0 + 1. \end{aligned} \tag{19}$$

To bound $H(\mathbf{I}|\mathbf{Y}, F = 1)$, it suffices to bound $H(\mathbf{I}|\mathbf{Y} = \mathbf{y}, F = 1)$ for any

$$\mathbf{y} \in \text{Shell}\left(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)}\right). \tag{20}$$

For this, we apply a packing argument as follows. Consider a ball centered at any $\mathbf{y}$ satisfying (20) and of radius $\sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}$, i.e.,

$$\text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right),$$

where ω satisfies

$$\pi/2 - \theta_n + \delta < \omega \leq \pi/2.$$

We now use the following lemma (whose proof is included in Appendix C-C) to upper bound the volume of the intersection between this ball and Ball $\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right)$, i.e.,

$$\left|\text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right)\right|.$$

Lemma 13: Let $\text{Ball}(\mathbf{c}_1, \sqrt{mR_1})$ and $\text{Ball}(\mathbf{c}_2, \sqrt{mR_1})$ be two balls in $\mathbb{R}^m$ with $\|\mathbf{c}_1 - \mathbf{c}_2\| = \sqrt{mD}$, where D satisfies $(\sqrt{R_1} - \sqrt{R_2})^2 < D < (\sqrt{R_1} + \sqrt{R_2})^2$. Then for any $\epsilon > 0$ and m sufficiently large, we have

$$\left|\text{Ball}(\mathbf{c}_1, \sqrt{mR_1}) \cap \text{Ball}(\mathbf{c}_2, \sqrt{mR_1})\right| \leq 2^{m\left(\frac{1}{2}\log \pi e \lambda(R_1, R_2, D) + \epsilon\right)}$$

where

$$\lambda(R_1, R_2, D) := \frac{2R_1D + 2R_1R_2 + 2DR_2 - R_1^2 - R_2^2 - D^2}{2D}.$$

Using the above lemma, we have for B sufficiently large,

$$\begin{aligned} &\left|\text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right)\right| \\ &\leq 2^{nB\left[\frac{1}{2}\log \pi e \lambda\left(P+N+\delta, N\left(4\sin^2\frac{\omega}{2}+\delta\right), \|\mathbf{y}\|\right)+\delta\right]} \\ &= 2^{nB\left[\frac{1}{2}\log \pi e \lambda\left(P+N, 4N\sin^2\frac{\omega}{2}, P+N\right)+\delta_1\right]} \\ &= 2^{nB\left[\frac{1}{2}\log \frac{8\pi eN\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{P+N}+\delta_1\right]}, \end{aligned}$$

for some $\delta_1 \to 0$ as $\delta \to 0$, where the first inequality is an immediate application of Lemma 13, the first equality follows from the fact that

$$\mathbf{y} \in \text{Shell}\left(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)}\right)$$

and the continuity of the function $\lambda(R_1, R_2, D)$ in its arguments, and the second equality follows from a simple evaluation of $\lambda\left(P+N, 4N\sin^2\frac{\omega}{2}, P+N\right)$.

On the other hand, the condition $F = 1$ (c.f. the definition of E_3 in Lemma 12) also ensures that

$$\left|f^{-1}(\mathbf{I}) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta}\right) \cap \text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right)\right| \geq 2^{nB[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n - \cos^2\omega)) - \delta]}.$$

Since $f^{-1}(\mathbf{i})$ are disjoint sets for different $\mathbf{i}$, given $F = 1$ and $\mathbf{Y} = \mathbf{y}$, the number of different possible values for $\mathbf{I}$ can be upper bounded by the ratio between

$$\left|\text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right)\right|$$

and

$$2^{nB[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n-\cos^2\omega))-\delta]},$$

which can be further upper bounded by

$$2^{nB\left[\frac{1}{2}\log\frac{8\pi eN\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{P+N}-\frac{1}{2}\log(2\pi eN(\sin^2\theta_n-\cos^2\omega))+\delta+\delta_1\right]} = 2^{nB\left[\frac{1}{2}\log\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n-\cos^2\omega)}+\delta_2\right]},$$

where $\delta_2 \to 0$ as $\delta \to 0$. This immediately implies the following upper bound on $H(\mathbf{I}|\mathbf{Y}=\mathbf{y}, F=1)$ and therefore $H(\mathbf{I}|\mathbf{Y}, F=1)$,

$$H(\mathbf{I}|\mathbf{Y}, F=1) \le nB\left[\frac{1}{2}\log\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n-\cos^2\omega)}+\delta_2\right],$$

which combined with (19) yields that

$$H(\mathbf{I}|\mathbf{Y}) \le nB\left[\frac{1}{2}\log\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n-\cos^2\omega)}+\delta_2\right] + 3\delta nBC_0 + 1.$$

Dividing both sides of the above inequality by B and noting that

$$H(\mathbf{I}|\mathbf{Y}) = \sum_{b=1}^{B} H(I_n(b)|Y^n(b)) = BH(I_n|Y^n),$$

we have

$$H(I_n|Y^n) \le n\left(\frac{1}{2}\log\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n-\cos^2\omega)}+\delta_2+3\delta C_0+\frac{1}{nB}\right), \tag{21}$$

which holds for any

$$\omega \in (\pi/2-\theta_n+\delta, \pi/2]. \tag{22}$$

Since δ, δ_2 and $\frac{1}{nB}$ in (21)–(22) can all be made arbitrarily small by choosing B sufficiently large, we obtain

$$H(I_n|Y^n) \le n\left(\frac{1}{2}\log\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n-\cos^2\omega)}\right), \tag{23}$$

for any $\omega \in \left(\frac{\pi}{2}-\theta_n, \frac{\pi}{2}\right]$. This completes the proof of Theorem 8.

IV. Proofs of Theorems 1 and 2

We now prove Theorem 2 by using Theorem 8, and use Theorem 2 to prove Theorem 1.

A. Proof of Theorem 2

Suppose a rate R is achievable. Then there exists a sequence of $(2^{nR}, n)$ codes such that the average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$. Let the relay's transmission be denoted by $I_n = f_n(Z^n)$. By standard information theoretic arguments, for this sequence of codes we have

$$\begin{aligned} nR &= H(M) \\ &= I(M; Y^n, I_n) + H(M|Y^n, I_n) \\ &\le I(X^n; Y^n, I_n) + n\mu \qquad (24) \\ &= I(X^n; Y^n) + I(X^n; I_n|Y^n) + n\mu \\ &= I(X^n; Y^n) + H(I_n|Y^n) - H(I_n|X^n) + n\mu \qquad (25) \\ &\le nI(X_Q; Y_Q) + H(I_n|Y^n) - H(I_n|X^n) + n\mu \qquad (26) \\ &\le \frac{n}{2}\log\left(1+\frac{P}{N}\right) + H(I_n|Y^n) - H(I_n|X^n) + n\mu, \qquad (27) \end{aligned}$$

for any $\mu > 0$ and n sufficiently large. In the above, (24) follows from applying the data processing inequality to the Markov chain $M - X^n - (Y^n, I_n)$ and Fano's inequality, (25) uses the fact that $I_n - X^n - Y^n$ form a Markov chain and thus $H(I_n|X^n, Y^n) = H(I_n|X^n)$, (26) follows by defining the time sharing random variable Q to be uniformly distributed over $[1:n]$, and (27) follows because

$$\begin{aligned} E[X_Q^2] &= \frac{1}{2^{nR}}\sum_{m=1}^{2^{nR}}\frac{1}{n}\sum_{i=1}^{n}x_i^2(m) \\ &= \frac{1}{n}\frac{1}{2^{nR}}\sum_{m=1}^{2^{nR}}\|x^n(m)\|^2 \\ &\le P. \end{aligned}$$

Given (27), the standard way to proceed would be to upper bound the first entropy term by $H(I_n|Y^n) \le H(I_n) \le nC_0$ and lower bound the second entropy term $H(I_n|X^n)$ simply by 0. This would lead to the so-called multiple-access bound in the well-known cut-set bound on the capacity of this channel [10]. However, as we already point out in our previous works[3]–[7], this leads to a loose bound since it does not capture the inherent tension between how large the first entropy term can be and how small the second one can be. Instead, we can use Theorem 8 to more tightly upper bound the difference $H(I_n|Y^n) - H(I_n|X^n)$ in (27).

We start by verifying that the random variables I_n, X^n, Z^n and Y^n associated with a code of blocklength n satisfy the conditions in Theorem 8. It is trivial to observe that they satisfy the required Markov chain condition and Z^n and Y^n are i.i.d. Gaussian given X^n due to the channel structure. Also assume that

$$E[\|X^n\|^2] = \frac{1}{2^{nR}}\sum_{m=1}^{2^{nR}}\|x^n(m)\|^2 = nP'$$

with $P' \le P$, and assume that $H(I_n|X^n) = -n\log\sin\theta_n$. Then, applying Theorem 8 to the random variables associated

with a code for the relay channel, we have

$$\begin{aligned} &H(I_n|Y^n) \\ &\quad\le n \cdot \min_{\omega\in(\frac{\pi}{2}-\theta_n,\frac{\pi}{2}]} \frac{1}{2}\log\left(\frac{4\sin^2\frac{\omega}{2}(P'+N-N\sin^2\frac{\omega}{2})}{(P'+N)(\sin^2\theta_n-\cos^2\omega)}\right) \\ &\quad\le n \cdot \min_{\omega\in(\frac{\pi}{2}-\theta_n,\frac{\pi}{2}]} \frac{1}{2}\log\left(\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n-\cos^2\omega)}\right), \end{aligned}$$

and therefore,

$$H(I_n|Y^n) - H(I_n|X^n) \le n \cdot \min_{\omega\in(\frac{\pi}{2}-\theta_n,\frac{\pi}{2}]} h_{\theta_n}(\omega) \tag{28}$$

where $h_{\theta_n}(\omega)$ is defined as

$$h_{\theta_n}(\omega) = \frac{1}{2}\log\left(\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})\sin^2\theta_n}{(P+N)(\sin^2\theta_n-\cos^2\omega)}\right), \tag{29}$$

in which $\theta_n = \arcsin\ 2^{-\frac{1}{n}H(I_n|X^n)}$ satisfies

$$\theta_0 := \arcsin(2^{-C_0}) \le \arcsin 2^{-\frac{1}{n}H(I_n|X^n)} = \theta_n \le \frac{\pi}{2}. \tag{30}$$

Plugging (28) into (27), we conclude that for any achievable rate R,

$$R \le \frac{1}{2}\log\left(1+\frac{P}{N}\right) + \min_{\omega\in(\frac{\pi}{2}-\theta_n,\frac{\pi}{2}]} h_{\theta_n}(\omega) + \mu. \tag{31}$$

At the same time, for any achievable rate R, we also have

$$R \le \frac{1}{2}\log\left(1+\frac{P}{N}\right) + C_0 + \log\sin\theta_n + \mu, \tag{32}$$

which simply follows from (27) by upper bounding $H(I_n|Y^n)$ with nC_0 and plugging in the definition of θ_n. Therefore, if a rate R is achievable, then for any $\mu > 0$ and n sufficiently large it should simultaneously satisfy both (31) and (32) for some θ_n that satisfies the condition in (30). This concludes the proof of the theorem.

B. Proof of Theorem 1

In order to show that Theorem 1 follows from Theorem 2, consider the following bound on $C(C_0)$ implied by Theorem 2:

$$C(C_0) \le \frac{1}{2}\log\left(1+\frac{P}{N}\right) + \sup_{\theta\in[\arcsin(2^{-C_0}),\frac{\pi}{2}]} \min_{\omega\in(\frac{\pi}{2}-\theta,\frac{\pi}{2}]} h_\theta(\omega). \tag{33}$$

With θ_0 defined as $\arcsin(2^{-C_0})$, we can upper bound the right-hand side of (33) to obtain

$$C(C_0) \le \frac{1}{2}\log\left(1+\frac{P}{N}\right) + \sup_{\theta\in[\theta_0,\frac{\pi}{2}]} \min_{\omega\in(\frac{\pi}{2}-\theta_0,\frac{\pi}{2}]} h_\theta(\omega).$$

Also because given any fixed $\omega \in \left(\frac{\pi}{2}-\theta_0,\frac{\pi}{2}\right]$, $h_\theta(\omega) \le h_{\theta_0}(\omega)$ for any $\theta \in [\theta_0, \pi/2]$, we further have

$$C(C_0) \le \frac{1}{2}\log\left(1+\frac{P}{N}\right) + \min_{\omega\in(\frac{\pi}{2}-\theta_0,\frac{\pi}{2}]} h_{\theta_0}(\omega). \tag{34}$$

The significance of the function $h_{\theta_0}(\omega)$ is that for any $\theta_0 > 0$,

$$h_{\theta_0}\left(\frac{\pi}{2}\right) = \frac{1}{2}\log\left(\frac{2P+N}{P+N}\right), \tag{35}$$

and $h_{\theta_0}(\omega)$ is increasing at $\omega = \frac{\pi}{2}$, or more precisely,

$$h'_{\theta_0}\left(\frac{\pi}{2}\right) = \frac{P}{(2P+N)\ln 2} > 0.$$

Therefore, as long as $\theta_0 > 0$, which is the case when C_0 is finite, the minimization of $h_{\theta_0}(\omega)$ with respect to ω in (34) yields a value strictly smaller than $h_{\theta_0}\left(\frac{\pi}{2}\right)$ in (35). This would allow us to conclude that the capacity $C(C_0)$ for any finite C_0 is strictly smaller than $\frac{1}{2}\log\left(1+\frac{2P}{N}\right)$.

We now formalize the above argument. Using the definition of the derivative, one obtains

$$h'_{\theta_0}\left(\frac{\pi}{2}\right) = \lim_{\Delta\to 0} \frac{h_{\theta_0}\left(\frac{\pi}{2}\right) - h_{\theta_0}\left(\frac{\pi}{2}-\Delta\right)}{\Delta}.$$

Therefore, there exists a sufficiently small $\Delta_1 > 0$ such that $0 < \Delta_1 < \theta_0$ and

$$\left|\frac{h_{\theta_0}\left(\frac{\pi}{2}\right) - h_{\theta_0}\left(\frac{\pi}{2}-\Delta_1\right)}{\Delta_1} - h'_{\theta_0}\left(\frac{\pi}{2}\right)\right| \le \frac{h'_{\theta_0}\left(\frac{\pi}{2}\right)}{2}.$$

For such Δ_1 we have

$$\begin{aligned} h_{\theta_0}\left(\frac{\pi}{2}-\Delta_1\right) &\le h_{\theta_0}\left(\frac{\pi}{2}\right) - \frac{\Delta_1\ h'_{\theta_0}\left(\frac{\pi}{2}\right)}{2} \\ &= \frac{1}{2}\log\left(\frac{2P+N}{P+N}\right) - \frac{P\Delta_1}{2(2P+N)\ln 2}, \end{aligned}$$

which further implies that

$$\min_{\omega\in(\frac{\pi}{2}-\theta_0,\frac{\pi}{2}]} h_{\theta_0}(\omega) \le \frac{1}{2}\log\left(\frac{2P+N}{P+N}\right) - \frac{P\Delta_1}{2(2P+N)\ln 2}. \tag{36}$$

Combining (34) and (36) we obtain that for any finite C_0, there exists some $\Delta_1 > 0$ such that

$$C(C_0) \le \frac{1}{2}\log\left(1+\frac{2P}{N}\right) - \frac{P\Delta_1}{2(2P+N)\ln 2}. \tag{37}$$

This proves Theorem 1.

V. Conclusion

We have proved a new upper bound on the capacity of the Gaussian relay channel and solved a problem posed by Cover in [2], which has remained open since 1987. The derivation of our upper bound focuses on directly characterizing the tension between information measures of pertinent n-letter random variables. In particular, this is done via the following steps:

- we first use "typicality" to translate the information tension problem to a problem regarding the geometry of the typical sets of these n-letter random variables;
- we then use results and tools in the (broadly defined) field of concentration of measure, in particular rearrangement theory, to establish non-trivial geometric properties for these typical sets;
- we finally use these geometric properties to construct a packing argument, which leads to an inequality between the original n-letter information measures.

In contrast, the typical program for proving converses in network information theory focuses on "single-letterizing" n-letter information measures. This makes it difficult to invoke

tools from geometry and concentration of measure, which in retrospect appear well-suited for quantifying information tensions that lie at the hearth of network problems. Indeed, to the best of our knowledge, the use of concentration of measure in information theory has been mostly limited to establishing strong converses for problems whose capacity is already known (see [12], [26]), and it has been rarely used to derive first-order results, i.e. bounds on the capacity of multi-user networks. Our proof suggests that measure concentration, in particular geometric inequalities and their functional counterparts, can have a bigger role to play in network information theory. It would be interesting to better understand this role and see if the program developed in this paper can be used to prove converses for other open problems in network information theory.

APPENDIX A
PROOFS OF EXTENDED ISOPERIMETRIC INEQUALITIES

In this appendix, we prove the extended isoperimetric inequalities on the sphere and on the shell, as stated in Theorems 6 and 7 respectively. In particular, we will first prove the shell case and then show that the sphere case follows as a corollary.

A. Preliminaries

We begin with some preliminaries that will be used in the proofs. Our main tool for proving Theorems 6 and 7 is the symmetric decreasing rearrangement of functions on the sphere, along with a version of the Riesz rearrangement inequality on the sphere due to Baernstein II and Taylor [25].

For any measurable function $f : \mathbb{S}^{m-1} \to \mathbb{R}$ and pole $\mathbf{z}_0$, the symmetric decreasing rearrangement of f about $\mathbf{z}_0$ is defined to be the function $f^* : \mathbb{S}^{m-1} \to \mathbb{R}$ such that $f^*(\mathbf{y})$ depends only on the angle $\angle(\mathbf{y}, \mathbf{z}_0)$, is nonincreasing in $\angle(\mathbf{y}, \mathbf{z}_0)$, and has super-level sets of the same Haar measure as f, i.e.

$$\mu\big(\{\mathbf{y} : f^*(\mathbf{y}) > d\}\big) = \mu\big(\{\mathbf{y} : f(\mathbf{y}) > d\}\big)$$

for all d. The function f^* is unique up to its value on sets of measure zero.

One important special case is when the function $f = 1_A$ is the characteristic function for a subset A. The function 1_A is just the function such that

$$1_A(\mathbf{y}) = \begin{cases} 1 & \mathbf{y} \in A \\ 0 & \text{otherwise.} \end{cases}$$

In this case, 1_A^* is equal to the characteristic function associated with a spherical cap of the same size as A. In other words, if A^* is a spherical cap about the pole $\mathbf{z}_0$ such that $\mu(A^*) = \mu(A)$, then $1_A^* = 1_{A^*}$.

Lemma 14 (Baernstein II and Taylor [25]): Let K be a nondecreasing bounded measurable function on the interval $[-1, 1]$. Then for all functions $f, g \in L^1(\mathbb{S}^{m-1})$,

$$\int_{\mathbb{S}^{m-1}} \left(\int_{\mathbb{S}^{m-1}} f(\mathbf{z}) K\left(\langle \mathbf{z}/R, \mathbf{y}/R \rangle\right) d\mathbf{z} \right) g(\mathbf{y}) d\mathbf{y} \leq \int_{\mathbb{S}^{m-1}} \left(\int_{\mathbb{S}^{m-1}} f^*(\mathbf{z}) K\left(\langle \mathbf{z}/R, \mathbf{y}/R \rangle\right) d\mathbf{z} \right) g^*(\mathbf{y}) d\mathbf{y}.$$

For any $f \in L^1(\mathbb{S}^{m-1})$, define

$$\psi(\mathbf{y}) = \int_{\mathbb{S}^{m-1}} f(\mathbf{z}) K\left(\langle \mathbf{z}/R, \mathbf{y}/R \rangle\right) d\mathbf{z}$$

to be the inner integral in Lemma 14. When applying Lemma 14 we will use test functions g that are characteristic functions. Let $g = 1_C$ where $C = \{\mathbf{y} : \psi(\mathbf{y}) > d\}$ for some d (i.e. C is a super-level set of ψ). For a fixed measure $\mu(C)$, the left-hand side of the inequality from Lemma 14 will be maximized by this choice of C. With this choice we have the following equality:

$$\int_{\mathbb{S}^{m-1}} \psi(\mathbf{y}) 1_C(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{S}^{m-1}} \psi^*(\mathbf{y}) 1_C^*(\mathbf{y}) d\mathbf{y} = \int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y}.$$

This follows from the layer-cake decomposition for any nonnegative and measurable function ψ in that

$$\begin{aligned} \int_{\mathbb{S}^{m-1}} \psi(\mathbf{y}) 1_C(\mathbf{y}) d\mathbf{y} &= \int_C \psi(\mathbf{y}) d\mathbf{y} \\ &= \int_C \int_0^\infty 1_{\{\psi(\mathbf{y}) > t\}} dt d\mathbf{y} \\ &= \int_0^\infty \int_C 1_{\{\psi(\mathbf{y}) > t\}} d\mathbf{y} dt \\ &= \int_0^\infty \int_{\mathbb{S}^{m-1}} 1_{\{\psi(\mathbf{y}) > \max(t,d)\}} d\mathbf{y} dt \\ &= \int_0^\infty \int_{\mathbb{S}^{m-1}} 1_{\{\psi^*(\mathbf{y}) > \max(t,d)\}} d\mathbf{y} dt \\ &= \int_0^\infty \int_{C^*} 1_{\{\psi^*(\mathbf{y}) > t\}} d\mathbf{y} dt \\ &= \int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y}. \end{aligned} \tag{38}$$

Using this equality and our choice for g we will rewrite the inequality from Lemma 14 as

$$\int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y} \leq \int_{C^*} \bar{\psi}(\mathbf{y}) d\mathbf{y} \tag{39}$$

where

$$\bar{\psi}(\mathbf{y}) = \int_{\mathbb{S}^{m-1}} f^*(\mathbf{z}) K\left(\langle \mathbf{z}/R, \mathbf{y}/R \rangle\right) d\mathbf{z}.$$

Note that both $\psi^*(\mathbf{y})$ and $\bar{\psi}(\mathbf{y})$ are spherically symmetric. More concretely, they both depend only on the angle $\angle(\mathbf{y}, \mathbf{z}_0)$, so in an abuse of notation we will write $\bar{\psi}(\alpha)$ and $\psi^*(\alpha)$ where $\alpha = \angle(\mathbf{y}, \mathbf{z}_0)$.

For convenience we will define a measure ν by

$$d\nu(\phi) = A_{m-2}(R \sin\phi) R d\phi$$

where $A_m(R)$ denotes the Haar measure of the m-sphere with radius R. We do this so that an integral like

$$\int_{\mathbb{S}^{m-1}} \psi^* d\mathbf{y} = \int_0^\pi \psi^*(\phi) A_{m-2}(R \sin\phi) R d\phi$$

can be expressed as

$$\int_0^{\pi} \psi^* d\nu.$$

B. Proof of Theorem 7 (The Shell Case)

Let $A \subseteq \mathbb{L}^m$ be a given subset with effective angle θ. In order to apply Lemma 14, note that

$$\begin{aligned}
&|A \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)| \\
&= \int_{\mathbb{R}^m} 1_{A\cap \text{ShellCap}(\mathbf{y},\omega+\epsilon)}(\mathbf{z})\, d\mathbf{z} \\
&= \int_{\mathbb{S}^{m-1}} \left(\int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} 1_{A\cap \text{ShellCap}(\mathbf{y},\omega+\epsilon)} \left(\frac{r}{R}\mathbf{z}\right)\, dr \right) d\mathbf{z}
\end{aligned}$$

by using spherical coordinates, so that if we define

$$f_A(\mathbf{z}) = \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} 1_A \left(\frac{r}{R}\mathbf{z}\right)\, dr \tag{40}$$

for $A \subseteq \mathbb{L}^m$ and

$$K(\cos\alpha) = \begin{cases} 0 & \omega+\epsilon < \alpha \le \pi \\ 1 & 0 \le \alpha \le \omega+\epsilon, \end{cases}$$

then

$$\begin{aligned}
\psi(\mathbf{y}) &= |A \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)| \\
&= \int_{\mathbb{S}^{m-1}} f_A(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle) d\mathbf{z}.
\end{aligned}$$

Both ψ and f_A can be thought of as functions on the sphere $\mathbb{S}^{m-1}$. Let ψ^*, f_A^* be their respective symmetric decreasing rearrangements about a pole $\mathbf{z}_0$. Define

$$\bar{\psi}(\mathbf{y}) = \int_{\mathbb{S}^{m-1}} f_A^*(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle) d\mathbf{z}$$

so that by definition we have (39).

The inequality (39) allows to compare ψ and $\bar{\psi}$, but we require a way to compare ψ with the function arising from a shell cap of angle θ. Let

$$A' = \text{ShellCap}(\mathbf{z}_0, \theta)$$

and

$$\bar{\bar{\psi}}(\mathbf{y}) = |A' \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)|.$$

We will show that

$$\int_{C^*} \bar{\psi}(\mathbf{y}) d\mathbf{y} \le \int_{C^*} \bar{\bar{\psi}}(\mathbf{y}) d\mathbf{y} \tag{41}$$

so that along with (39),

$$\int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y} \le \int_{C^*} \bar{\bar{\psi}}(\mathbf{y}) d\mathbf{y}. \tag{42}$$

To show the inequality (41) note

$$\begin{aligned}
&\int_{C^*} \bar{\psi}(\mathbf{y}) d\mathbf{y} \\
&\quad = \int_{\mathbb{S}^{m-1}} \int_{\mathbb{S}^{m-1}} 1_{C^*}(\mathbf{y}) f_A^*(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle) d\mathbf{y} d\mathbf{z} \\
&\quad = \int_{\mathbb{S}^{m-1}} f_A^*(\mathbf{z}) \left(\int_{\mathbb{S}^{m-1}} 1_{C^*}(\mathbf{y}) K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle) d\mathbf{y} \right) d\mathbf{z}.
\end{aligned} \tag{43}$$

The term inside the parentheses is the measure of the intersection between the cap C^* centered at $\mathbf{z}_0$ and a cap of angle $\omega+\epsilon$ centered at $\mathbf{z}$. This intersection measure is a function only of the angle $\angle(\mathbf{z}_0, \mathbf{z})$ and is nonincreasing in that angle. Consider functions $f : \mathbb{S}^{m-1} \to \mathbb{R}$ with $0 \le f(\mathbf{z}) \le \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr$ and $\int f(\mathbf{z}) d\mathbf{z} = |A|$. Both f_A^* and $f_{A'}$ satisfy these properties and moreover $f_{A'}$ is extremal in the sense that $f_{A'}(\mathbf{z}) = \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr$ when $\angle(\mathbf{z}_0, \mathbf{z}) \le \theta$ and 0 when $\angle(\mathbf{z}_0, \mathbf{z}) > \theta$. Therefore (43) is maximized by replacing f_A^* with $f_{A'}$, and

$$\begin{aligned}
&\int_{C^*} \bar{\psi}(\mathbf{y}) d\mathbf{y} \\
&\quad = \int_{\mathbb{S}^{m-1}} f_A^*(\mathbf{z}) \left(\int_{\mathbb{S}^{m-1}} 1_{C^*}(\mathbf{y}) K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle) d\mathbf{y} \right) d\mathbf{z} \\
&\quad \le \int_{\mathbb{S}^{m-1}} f_{A'}(\mathbf{z}) \left(\int_{\mathbb{S}^{m-1}} 1_{C^*}(\mathbf{y}) K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle) d\mathbf{y} \right) d\mathbf{z} \\
&\quad = \int_{C^*} \bar{\bar{\psi}}(\mathbf{y}) d\mathbf{y}.
\end{aligned}$$

Equipped with (42), we are now ready to finish the proof of Theorem 7. Proposition 4 implies that for any $0 < \epsilon < 1$, there exists an $M(\epsilon)$ such that for $m > M(\epsilon)$ we have

$$\mathbb{P}\left(\angle(\mathbf{z}_0, \mathbf{Y}) \in [\pi/2 - \epsilon, \pi/2 + \epsilon]\right) \ge 1 - \frac{\epsilon^2}{2} \tag{44}$$

where $\mathbf{Y}$ is drawn from any rotationally invariant distribution on $\mathbb{L}^m$. Because the random quantity $|A \cap \text{ShellCap}(\mathbf{Y}, \omega+\epsilon)|$ depends only on the direction of $\mathbf{Y}$, and not on its magnitude, we can instead consider $\mathbf{Y}$ to be distributed according to the Haar measure on $\mathbb{S}^{m-1}$. The constant $M(\epsilon)$ is determined only by the concentration of measure phenomenon cited above, and it does not depend on any parameters in the problem other than ϵ. From now on, let us restrict our attention to dimensions $m > M(\epsilon)$. Due to the triangle inequality for the geodesic metric, for $\mathbf{y}$ such that $\angle(\mathbf{z}_0, \mathbf{y}) \in [\pi/2 - \epsilon, \pi/2 + \epsilon]$ we have

$$A' \cap \text{ShellCap}(\mathbf{y}_0, \omega) \subseteq A' \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)$$

where $\mathbf{y}_0$ is such that $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$. Therefore,

$$\bar{\bar{\psi}}(\angle(\mathbf{z}_0, \mathbf{y})) = |A' \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)| \ge V \tag{45}$$

for all for $\mathbf{y}$ such that $\angle(\mathbf{z}_0, \mathbf{y}) \in [\pi/2 - \epsilon, \pi/2 + \epsilon]$ and

$$\begin{aligned}
\mathbb{P}\left(\bar{\bar{\psi}}(\mathbf{Y}) \ge V\right) &= \mathbb{P}\left(|A' \cap \text{ShellCap}(\mathbf{Y}, \omega+\epsilon)| \ge V\right) \\
&\ge 1 - \frac{\epsilon^2}{2} \\
&\ge 1 - \frac{\epsilon}{2}.
\end{aligned} \tag{46}$$

To prove the theorem, we need to show that

$$\begin{aligned}
&\mathbb{P}\left(\psi(\mathbf{Y}) > (1-\epsilon)V\right) \\
&\quad = \mathbb{P}\left(|A \cap \text{ShellCap}(\mathbf{Y}, \omega+\epsilon)| > (1-\epsilon)V\right) \\
&\quad \ge 1 - \epsilon
\end{aligned} \tag{47}$$

for any arbitrary set $A \subseteq \mathbb{L}^m$. Recall that by the definition of a decreasing symmetric rearrangement, we have

$$\mathbb{P}\left(\psi^*(\mathbf{Y}) > d\right) = \mathbb{P}\left(\psi(\mathbf{Y}) > d\right)$$

for any threshold d and this implies

$$\mathbb{P}\left(\psi^*(\mathbf{Y}) \le (1-\epsilon)V\right) = \mathbb{P}\left(\psi(\mathbf{Y}) \le (1-\epsilon)V\right). \quad (48)$$

Therefore, the desired statement in (47) can be equivalently written as

$$\mathbb{P}\left(\psi^*(\mathbf{Y}) \le (1-\epsilon)V\right) \le \epsilon. \quad (49)$$

Turning to proving (49), recall that by the definition of a decreasing symmetric rearrangement, $\psi^*(\alpha)$ is nonincreasing in the angle $\alpha = \angle(\mathbf{y}, \mathbf{z}_0)$ over the interval $0 \le \alpha \le \pi$. Let β be the smallest value such that $\psi^*(\beta) = (1-\epsilon)V$, or more explicitly,

$$\beta = \inf\{\alpha : \psi^*(\alpha) \le (1-\epsilon)V\}.$$

If $\beta \ge \pi/2+\epsilon$, then (49) would follow trivially from (44) and the fact that $\psi^*(\alpha)$ would be greater than $(1-\epsilon)V$ for all $0 < \alpha < \pi/2+\epsilon$. We will therefore assume that $0 < \beta < \pi/2+\epsilon$. It remains to show that even if this is the case, we have (49).

By the definition of β and the fact that ψ^* is nonincreasing,

$$\begin{aligned}\mathbb{P}\left(\psi^*(\mathbf{Y}) \le (1-\epsilon)V\right) &= \frac{1}{A_{m-1}(R)}\int_\beta^\pi d\nu \\ &= \frac{1}{A_{m-1}(R)}\int_\beta^{\max\{\beta,\frac{\pi}{2}-\epsilon\}} d\nu \\ &\quad + \frac{1}{A_{m-1}(R)}\int_{\max\{\beta,\frac{\pi}{2}-\epsilon\}}^{\frac{\pi}{2}+\epsilon} d\nu \\ &\quad + \frac{1}{A_{m-1}(R)}\int_{\frac{\pi}{2}+\epsilon}^{\pi} d\nu. \quad (50)\end{aligned}$$

To bound the first and third terms of (50) note that

$$\frac{1}{A_{m-1}(R)}\int_\beta^{\max\{\beta,\frac{\pi}{2}-\epsilon\}} d\nu + \frac{1}{A_{m-1}(R)}\int_{\frac{\pi}{2}+\epsilon}^{\pi} d\nu \le \frac{\epsilon^2}{2} \quad (51)$$

$$\le \frac{\epsilon}{2} \quad (52)$$

as a consequence of (44). In order to bound the second term, we establish the following chain of (in)equalities which will be justified below.

$$\begin{aligned}&\frac{1}{A_{m-1}(R)}\int_{\frac{\pi}{2}+\epsilon}^{\pi} d\nu \\ &\quad \ge \frac{1}{(1-\epsilon)VA_{m-1}(R)}\int_{\frac{\pi}{2}+\epsilon}^{\pi}(\psi^* - \bar{\bar{\psi}})d\nu \quad (53)\\ &\quad = \frac{1}{(1-\epsilon)VA_{m-1}(R)}\int_0^{\frac{\pi}{2}+\epsilon}(\bar{\bar{\psi}} - \psi^*)d\nu \quad (54)\\ &\quad \ge \frac{1}{(1-\epsilon)VA_{m-1}(R)}\int_\beta^{\frac{\pi}{2}+\epsilon}(\bar{\bar{\psi}} - \psi^*)d\nu \quad (55)\\ &\quad \ge \frac{\epsilon}{(1-\epsilon)A_{m-1}(R)}\int_{\max\{\beta,\frac{\pi}{2}-\epsilon\}}^{\frac{\pi}{2}+\epsilon} d\nu \quad (56)\\ &\quad \ge \frac{\epsilon}{A_{m-1}(R)}\int_{\max\{\beta,\frac{\pi}{2}-\epsilon\}}^{\frac{\pi}{2}+\epsilon} d\nu \quad (57)\end{aligned}$$

Combining (57) with (51) reveals that the second term in (50) is also bounded by $\epsilon/2$, therefore

$$\mathbb{P}\left(\psi^*(\mathbf{Y}) \le (1-\epsilon)V\right)$$

must be bounded by ϵ, which proves Theorem 7.

The first inequality (53) is a consequence of the fact that over the range of the integral, ψ^* is less than or equal to $(1-\epsilon)V$ and $\bar{\bar{\psi}}$ is non-negative. The equality in (54) follows from

$$\int_0^\pi \psi^* d\nu = \int_0^\pi \bar{\bar{\psi}} d\nu,$$

which is itself a consequence of (38) with $C = \mathbb{S}^{m-1}$ and

$$\begin{aligned}\int_{\mathbb{S}^{m-1}} \psi(\mathbf{y})d\mathbf{y} &= \int_{\mathbb{S}^{m-1}}\int_{\mathbb{S}^{m-1}} f_A(\mathbf{z})K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle)d\mathbf{z}d\mathbf{y} \\ &= \int\int K(\langle \mathbf{y}/R, \mathbf{z}/R\rangle)d\mathbf{y}\, f_A(\mathbf{z})d\mathbf{z} \\ &= \int \mu(\mathrm{Cap}(\mathbf{y},\omega))\, f_A(\mathbf{z})d\mathbf{z} \\ &= \mu(\mathrm{Cap}(\mathbf{y},\omega))|A| \\ &= \int \mu(\mathrm{Cap}(\mathbf{y},\omega))\, f_{A'}(\mathbf{z})d\mathbf{z} \\ &= \int\int f_{A'}(\mathbf{z})K(\langle \mathbf{z}/R, \mathbf{y}/R\rangle)d\mathbf{z}d\mathbf{y} \\ &= \int_{\mathbb{S}^{m-1}} \bar{\bar{\psi}}(\mathbf{y})d\mathbf{y}. \quad (58)\end{aligned}$$

Next we have (55) which is due to the rearrangement inequality (42) when C is the super-level set $\{\mathbf{y} : \psi(\mathbf{y}) > (1-\epsilon)V\}$. By the definition of a symmetric decreasing rearrangement, $\mu(\{\mathbf{y} : \psi(\mathbf{y}) > (1-\epsilon)V\}) = \mu(\{\mathbf{y} : \psi^*(\mathbf{y}) > (1-\epsilon)V\})$, and the set on the right-hand side is an open or closed spherical cap of angle β. Thus C^* is a spherical cap with angle β and the rearrangement inequality (42) gives

$$\int_0^\beta \psi^* d\nu \le \int_0^\beta \bar{\bar{\psi}} d\nu.$$

Finally, for the inequality (56), we first replace the lower integral limit with $\max\{\beta, \pi/2-\epsilon\} \ge \beta$. Then $\bar{\bar{\psi}} \ge V$ over the range of the integral due to (45). Additionally, $\psi^* \le (1-\epsilon)V$ over the range of the integral, and the inequality follows.

C. Proof of Theorem 6 (The Sphere Case)

Given any $A \subseteq \mathbb{S}^{m-1}$ with effective angle $\theta > 0$, construct a corresponding

$$A_{\text{shell}} = \left\{\mathbf{y} \in \mathbb{L}^m : R\frac{\mathbf{y}}{\|\mathbf{y}\|} \in A\right\}.$$

The set A_{shell} also has effective angle θ as a subset of $\mathbb{L}^m$ since

$$\begin{aligned}|A_{\text{shell}}| &= \int_{\mathbb{R}^m} 1_{A_{\text{shell}}}(\mathbf{z})\, d\mathbf{z} \\ &= \int_{\mathbb{S}^{m-1}} \left(\int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} 1_{A_{\text{shell}}}\left(\frac{r\mathbf{z}}{R}\right)\, dr \right) d\mathbf{z} \\ &= \int_{\mathbb{S}^{m-1}} 1_A(\mathbf{z}) d\mathbf{z} \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \\ &= \mu(A) \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \\ &= \mu(\text{Cap}(\mathbf{y}, \theta)) \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \\ &= \int_{\mathbb{R}^m} 1_{\text{ShellCap}(\mathbf{y},\theta)}(\mathbf{z})\, d\mathbf{z} \\ &= |\text{ShellCap}(\mathbf{y}, \theta)|.\end{aligned}$$

For any $\epsilon > 0$, we can apply Theorem 7 to find an $M(\epsilon)$ such that for $m > M(\epsilon)$,

$$\mathbb{P}\left(|A_{\text{shell}} \cap \text{ShellCap}(\mathbf{Y}, \omega + \epsilon)| > (1-\epsilon) V_{\text{shell}}\right) \geq 1 - \epsilon, \tag{59}$$

where $V_{\text{shell}} = |\text{ShellCap}(\mathbf{z}_0, \theta) \cap \text{ShellCap}(\mathbf{y}_0, \omega)|$ with $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$. Because the set $\text{ShellCap}(\mathbf{y}, \omega)$ depends only on the direction of $\mathbf{y}$, and not on its magnitude, the probability in (59) is the same whether we consider $\mathbf{Y}$ to be uniformly distributed on $\mathbb{S}^{m-1}$ or from some rotationally invariant probability distribution on $\mathbb{L}^m$. Using spherical coordinates, we have

$$\begin{aligned}&|A_{\text{shell}} \cap \text{ShellCap}(\mathbf{y}, \omega + \epsilon)| \\ &= \int_{\mathbb{R}^m} 1_{A_{\text{shell}} \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)}(\mathbf{z})\, d\mathbf{z} \\ &= \int_{\mathbb{S}^{m-1}} \left(\int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} 1_{A_{\text{shell}} \cap \text{ShellCap}(\mathbf{y}, \omega+\epsilon)}\left(\frac{r\mathbf{z}}{R}\right)\, dr \right) d\mathbf{z} \\ &= \int_{\mathbb{S}^{m-1}} 1_{A \cap \text{Cap}(\mathbf{y}, \omega+\epsilon)}(\mathbf{z}) d\mathbf{z} \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \\ &= \mu(A \cap \text{Cap}(\mathbf{y}, \omega + \epsilon)) \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr\end{aligned}$$

and similarly,

$$\begin{aligned}&|\text{ShellCap}(\mathbf{z}_0, \theta) \cap \text{ShellCap}(\mathbf{y}_0, \omega)| \\ &\quad = \int_{\mathbb{S}^{m-1}} 1_{\text{Cap}(\mathbf{z}_0,\theta) \cap \text{Cap}(\mathbf{y}_0,\omega)}(\mathbf{z}) d\mathbf{z} \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \\ &\quad = \mu(\text{Cap}(\mathbf{z}_0, \theta) \cap \text{Cap}(\mathbf{y}_0, \omega)) \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr.\end{aligned}$$

By dividing out the $\int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr$ term, (59) implies

$$\mathbb{P}\left(\mu(A \cap \text{Cap}(\mathbf{Y}, \omega + \epsilon)) > (1-\epsilon) V\right) \geq 1 - \epsilon \tag{60}$$

where $V = \mu(\text{Cap}(\mathbf{z}_0, \theta) \cap \text{Cap}(\mathbf{y}_0, \omega))$ as desired.

Appendix B
Proofs of Typicality Lemmas

Here we prove the typicality lemmas presented in Section III-A.

A. Proof of Lemma 9

Recalling that $\mathbf{Z} = [Z^n(1), Z^n(2), \ldots, Z^n(B)]$, we have

$$\|\mathbf{Z}\|^2 = \sum_{b=1}^{B} \|Z^n(b)\|^2.$$

Therefore by the weak law of large numbers, for any $\delta > 0$ and B sufficiently large we have

$$\Pr\left(\left| \frac{1}{B}\|\mathbf{Z}\|^2 - E[\|Z^n\|^2] \right| \leq \delta \right) \geq 1 - \delta,$$

i.e.,

$$\Pr(\|\mathbf{Z}\|^2 \in [nB(P + N - \delta), nB(P + N + \delta)]) \geq 1 - \delta,$$

since by assumption $E[\|X^n\|^2] = nP$ and thus $E[\|Z^n\|^2] = n(P + N)$. Because $\mathbf{Z}$ and $\mathbf{Y}$ are identically distributed, the above relation also holds with $\|\mathbf{Z}\|^2$ replaced by $\|\mathbf{Y}\|^2$. This completes the proof of the lemma.

B. Proof of Lemma 10

We now present the proof of Lemma 10. By the law of large numbers and Lemma 9, we have for any $\epsilon > 0$ and sufficiently large B,

$$\Pr((\mathbf{X}, \mathbf{Z}) \in S_\epsilon(X^n, Z^n)) \geq 1 - \epsilon$$

where

$$\begin{aligned}&S_\epsilon(X^n, Z^n) \\ &:= \Big\{ (\mathbf{x}, \mathbf{z}) : \|\mathbf{x} - \mathbf{z}\| \in [\sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}], \\ &\qquad \mathbf{z} \in \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\epsilon)}\right), \\ &\qquad 2^{nB(\log \sin\theta_n - \epsilon)} \leq p(f(\mathbf{z})|\mathbf{x}) \leq 2^{nB(\log \sin\theta_n + \epsilon)} \Big\}.\end{aligned}$$

Note that in terms of $S_\epsilon(X^n, Z^n)$, the set $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ in Lemma 10 can be simply written as

$$S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) = \{\mathbf{z} : f(\mathbf{z}) = \mathbf{i}, (\mathbf{x}, \mathbf{z}) \in S_\epsilon(X^n, Z^n)\}.$$

Therefore, for B sufficiently large, we have

$$\begin{aligned}\Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{X}, \mathbf{I})) &= \Pr(f(\mathbf{Z}) = \mathbf{I}, (\mathbf{X}, \mathbf{Z}) \notin S(X^n, Z^n)) \\ &\leq \epsilon.\end{aligned}$$

On the other hand, defining $S_\epsilon(X^n, I_n) := \{(\mathbf{x}, \mathbf{i}) : \Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) \geq 1 - \sqrt{\epsilon}\}$, we have

$$\begin{aligned}&\Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{X}, \mathbf{I})) \\ &\quad = \sum_{(\mathbf{x},\mathbf{i}) \in S_\epsilon(X^n, I_n)} \Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) p(\mathbf{x}, \mathbf{i}) \\ &\qquad + \sum_{(\mathbf{x},\mathbf{i}) \notin S_\epsilon(X^n, I_n)} \Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) p(\mathbf{x}, \mathbf{i}) \\ &\quad \geq \sqrt{\epsilon} \cdot \Pr(S_\epsilon^c(X^n, I_n)).\end{aligned}$$

Therefore, we have for B sufficiently large,

$$\Pr(S_\epsilon^c(X^n, I_n)) \leq \frac{\epsilon}{\sqrt{\epsilon}} = \sqrt{\epsilon},$$

and thus

$$\Pr(S_\epsilon(X^n, I_n)) \geq 1 - \sqrt{\epsilon},$$

which proves (13).

To prove (14), consider any $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$. From the definition of $S_\epsilon(X^n, I_n)$, $\Pr(S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) \geq 1 - \sqrt{\epsilon}$. Therefore, $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ must be nonempty, i.e., there exists at least one $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$. Consider any $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$. By the definition of $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$, we have $f(\mathbf{z}) = \mathbf{i}$ and $(\mathbf{x}, \mathbf{z}) \in S_\epsilon(X^n, Z^n)$. Then, it follows from the definition of $S_\epsilon(X^n, Z^n)$ that

$$2^{nB(\log \sin\theta_n - \epsilon)} \leq p(f(\mathbf{z})|\mathbf{x}) = p(\mathbf{i}|\mathbf{x}) \leq 2^{nB(\log \sin\theta_n + \epsilon)}.$$

This further implies that

$$\begin{aligned}
&\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) \\
&\quad = \frac{\Pr(f(\mathbf{Z}) = \mathbf{i}|\mathbf{x})\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, f(\mathbf{Z}) = \mathbf{i})}{\Pr(f(\mathbf{Z}) = \mathbf{i}|\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}), \mathbf{x})} \\
&\quad = p(\mathbf{i}|\mathbf{x})\Pr(S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) \\
&\quad \geq 2^{nB(\log \sin\theta_n - \epsilon)}(1 - \sqrt{\epsilon}) \\
&\quad \geq 2^{nB(\log \sin\theta_n - 2\epsilon)}
\end{aligned}$$

for sufficiently large B, which concludes the proof of (14) and Lemma 10.

C. Proof of Corollary 11

Let the effective angle of A be denoted by θ', i.e.,

$$|A| = |\text{ShellCap}(\mathbf{z}_0, \theta')|$$

for some

$$\mathbf{z}_0 \in \text{Shell}\left(\mathbf{0}, \sqrt{m(N-\epsilon)}, \sqrt{m(N+\epsilon)}\right),$$

where

$$\begin{aligned}
&\text{ShellCap}(\mathbf{z}_0, \theta') \\
&:= \left\{\mathbf{z} \in \text{Shell}(\mathbf{0}, \sqrt{m(N-\epsilon)}, \sqrt{m(N+\epsilon)}) : \angle(\mathbf{z}_0, \mathbf{z}) \leq \theta'\right\}.
\end{aligned}$$

Then using the formula for the volume of a shell cap (c.f. Appendix C-A and in particular (66)), we have

$$|A| \leq 2^{\frac{m}{2}[\log(2\pi e(N+\epsilon)\sin^2\theta') + \epsilon_1]}$$

for some $\epsilon_1 \to 0$ as $m \to \infty$. Recall that by assumption

$$|A| \geq 2^{\frac{m}{2}[\log(2\pi e(N+\epsilon)\sin^2\theta)]},$$

and we hence have

$$\theta' \geq \theta - \epsilon_2$$

for some $\epsilon_2 \to 0$ as $m \to \infty$.

We now apply Theorem 7 to this specific shell and subset A. First, using the formula of the intersection volume of two shell caps (c.f. Appendices C-B and in particular Lemma 16), we have

$$\begin{aligned}
&|\text{ShellCap}(\mathbf{z}_0, \theta') \cap \text{ShellCap}(\mathbf{y}_0, \omega)| \\
&\quad \geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta' - \cos^2\omega)) - \epsilon_3]} \\
&\quad \geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon_4]}
\end{aligned}$$

for some $\epsilon_3, \epsilon_4 \to 0$ as $m \to \infty$, where $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and $\theta' + \omega > \pi/2$. Then Theorem 7 asserts that for any $\omega \in (\pi/2 - \theta', \pi/2]$ and m sufficiently large,

$$\begin{aligned}
&\Pr\Big(|A \cap \text{ShellCap}(\mathbf{Y}, \omega + \epsilon)| \\
&\qquad \geq (1-\epsilon)2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon_4]}\Big) \geq 1 - \epsilon,
\end{aligned}$$

where $\mathbf{Y}$ is a random vector drawn from any rotationally invariant distribution on the shell. Since $\pi/2 - \theta' \leq \pi/2 - \theta + \epsilon_2$, the condition $\omega \in (\pi/2 - \theta', \pi/2]$ in the above can be replaced with the weaker condition $\omega \in (\pi/2 - \theta + \epsilon_2, \pi/2]$. Now by choosing m sufficiently large we can make ϵ_2, ϵ_4 and $\frac{2}{m}\log(1-\epsilon)$ as small as desired, so we have

$$\begin{aligned}
&\Pr\left(|A \cap \text{ShellCap}(\mathbf{Y}, \omega + \epsilon)| \geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon]}\right) \\
&\qquad \geq 1 - \epsilon,
\end{aligned}$$

for any $\omega \in (\pi/2 - \theta, \pi/2]$ and m sufficiently large. Finally, observe that for any $\mathbf{y}$ in the considered shell,

$$\begin{aligned}
&\text{ShellCap}(\mathbf{y}, \omega + \epsilon) \\
&\qquad \subseteq \text{Ball}\left(\mathbf{y}, 2\sqrt{m(N+\epsilon)}\sin\frac{\omega + \epsilon}{2} + 2\sqrt{m\epsilon}\right).
\end{aligned}$$

This simply follows from the geometry illustrated in Fig. 5 combined with the triangle inequality and the fact that the thickness of the shell can be trivially bounded by $2\sqrt{m\epsilon}$. Therefore, we can conclude that

$$\begin{aligned}
&\Pr\left(\left|A \cap \text{Ball}\left(\mathbf{Y}, 2\sqrt{m(N+\epsilon)}\sin\frac{\omega+\epsilon}{2} + 2\sqrt{m\epsilon}\right)\right|\right. \\
&\qquad \left. \geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon]}\right) \geq 1 - \epsilon
\end{aligned}$$

for any $\omega \in (\pi/2 - \theta, \pi/2]$ and m sufficiently large. This completes the proof of Corollary 11.

D. Proof of Lemma 12

Fix $\epsilon > 0$ and consider a pair $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$. From Lemma 10, we have

$$\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) \geq 2^{nB(\log \sin\theta_n - 2\epsilon)},$$

for B sufficiently large. We also have

$$\begin{aligned}
\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) &\leq |S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})| \sup_{\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})} p(\mathbf{z}|\mathbf{x}) \\
&\leq |S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})| 2^{-nB\left(\frac{1}{2}\log 2\pi eN - \epsilon_1\right)},
\end{aligned}$$

for some $\epsilon_1 \to 0$ as $\epsilon \to 0$, where $p(\mathbf{z}|\mathbf{x})$ refers to the conditional density of $\mathbf{z}$ given $\mathbf{x}$. The second inequality in the above follows because for any $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$, we have

$$\|\mathbf{x} - \mathbf{z}\| \in [\sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}],$$

and therefore using the fact that $\mathbf{Z}$ is Gaussian distributed given $\mathbf{x}$, we have for any $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$,

$$\begin{aligned}
p(\mathbf{z}|\mathbf{x}) &= \frac{1}{(2\pi N)^{\frac{nB}{2}}} e^{-\frac{\|\mathbf{z}-\mathbf{x}\|^2}{2N}} \\
&\leq 2^{-\frac{nB(N-\epsilon)}{2N}\log e - \frac{nB}{2}\log 2\pi N} \\
&= 2^{-nB\left(\frac{1}{2}\log 2\pi eN - \epsilon_1\right)}
\end{aligned}$$

where $\epsilon_1 \to 0$ as $\epsilon \to 0$. Therefore, for B sufficiently large, the volume of $S_\epsilon(Z^n|\mathbf{x},\mathbf{i})$ can be lower bounded by

$$|S_\epsilon(Z^n|\mathbf{x},\mathbf{i})| \geq 2^{nB\left(\frac{1}{2}\log(2\pi eN\sin^2\theta_n)-2\epsilon-\epsilon_1\right)}.$$

Let θ_n' be defined such that

$$\log 2\pi e(N+\epsilon)\sin^2\theta_n' = \frac{1}{2}\log(2\pi eN\sin^2\theta_n) - 2\epsilon - \epsilon_1.$$

Obviously, we have $\theta_n' \leq \theta_n$ and $\theta_n' \to \theta_n$ as $\epsilon \to 0$. Noting that $S_\epsilon(Z^n|\mathbf{x},\mathbf{i})$ is a subset of

$$\text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right),$$

by Corollary 11, for any $\omega \in (\pi/2 - \theta_n', \pi/2]$ we have

$$\Pr\left(\left|S_\epsilon(Z^n|\mathbf{x},\mathbf{i}) \cap \text{Ball}\left(\mathbf{U}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2}+\epsilon_2\right)}\right)\right| \geq 2^{nB\left[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n-\cos^2\omega))-\epsilon_3\right]}\right) \geq 1-\epsilon \quad (61)$$

for any $\mathbf{U}$ drawn from a rotationally invariant distribution around $\mathbf{x}$ on $\text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right)$, where ϵ_2 is defined such that

$$\sqrt{nBN\left(4\sin^2\frac{\omega}{2}+\epsilon_2\right)} = 2\sqrt{nB(N+\epsilon)}\sin\frac{\omega+\epsilon}{2} + 2\sqrt{m\epsilon},$$

and ϵ_3 is defined such that

$$\frac{1}{2}\log(2\pi eN(\sin^2\theta_n - \cos^2\omega)) - \epsilon_3 = \frac{1}{2}\log(2\pi eN(\sin^2\theta_n' - \cos^2\omega)) - \epsilon,$$

and both ϵ_2 and ϵ_3 tend to zero as ϵ goes to zero.

We now translate the bound (61) on the probability involving a rotationally invariantly distributed $\mathbf{U}$ on the shell to a bound on the probability involving $\mathbf{Y}$. Define $\mathcal{Y}_{(\mathbf{x},\mathbf{i})}$ to be the following set of $\mathbf{y}$:

$$\left\{\mathbf{y} : \left|S_\epsilon(Z^n|\mathbf{x},\mathbf{i}) \cap \text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2}+\epsilon_2\right)}\right)\right| \geq 2^{nB\left[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n-\cos^2\omega))-\epsilon_3\right]}\right\}.$$

Then we have for $(\mathbf{x},\mathbf{i}) \in S_\epsilon(X^n, I_n)$ and B sufficiently large,

$$\begin{aligned}
&\Pr(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}|\mathbf{x}) \\
&\geq \Pr\Big(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}, \\
&\qquad \mathbf{Y} \in \text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right)\Big|\mathbf{x}\Big) \\
&= \Pr\left(\mathbf{Y} \in \text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right)\Big|\mathbf{x}\right) \\
&\quad \times \Pr\Big(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}\Big|\mathbf{x}, \\
&\qquad \mathbf{Y} \in \text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right)\Big) \\
&\geq (1-\epsilon)\Pr\Big(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}\Big|\mathbf{x}, \\
&\qquad \mathbf{Y} \in \text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right)\Big) \\
&\geq (1-\epsilon)^2,
\end{aligned}$$

where the second inequality simply follows by applying the law of large numbers in a manner similar to the proof of Lemma 9, and the last inequality follows from combining (61) and the fact that if $\mathbf{x}$ is known and $\mathbf{Y}$ is restricted to $\text{Shell}\left(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}\right)$ then $\mathbf{Y}$ is rotationally invariant around $\mathbf{x}$ on this shell.

Since by definition $S_\epsilon(Z^n|\mathbf{x},\mathbf{i})$ is a subset of $f^{-1}(\mathbf{i}) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\epsilon)}\right)$, we have

$$\begin{aligned}
&\left|f^{-1}(\mathbf{i}) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\epsilon)}\right) \cap \text{Ball}\left(\mathbf{y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2}+\epsilon_2\right)}\right)\right| \\
&\geq 2^{nB\left[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n-\cos^2\omega))-\epsilon_3\right]}
\end{aligned}$$

for any $\mathbf{y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}$, and therefore for B sufficiently large,

$$\begin{aligned}
&\Pr\Bigg(\Bigg|f^{-1}(\mathbf{I}) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\epsilon)}\right) \cap \text{Ball}\left(\mathbf{Y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2}+\epsilon_2\right)}\right)\Bigg| \\
&\qquad \geq 2^{nB\left[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n-\cos^2\omega))-\epsilon_3\right]}\Bigg) \\
&\geq \sum_{(\mathbf{x},\mathbf{i})} \Pr(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}|\mathbf{x})\, p(\mathbf{x},\mathbf{i}) \\
&\geq \sum_{(\mathbf{x},\mathbf{i}) \in S_\epsilon(X^n, I_n)} \Pr(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x},\mathbf{i})}|\mathbf{x})\, p(\mathbf{x},\mathbf{i}) \\
&\geq (1-\epsilon)^2(1-\sqrt{\epsilon}) \\
&\geq 1 - 4\sqrt{\epsilon},
\end{aligned}$$

for any $\omega \in (\pi/2 - \theta_n', \pi/2]$. Finally, choosing $\delta = \max\{4\sqrt{\epsilon}, \epsilon_2, \epsilon_3, \theta_n - \theta_n'\}$ concludes the proof of Lemma 12. Note that by choosing B sufficiently large, ϵ and therefore δ can be made arbitrarily small.

Appendix C
Miscellaneous Results in High-Dimensional Geometry

This appendix derives some miscellaneous results in high-dimensional geometry, including the surface area (volume) of a spherical (shell) cap, the surface area (volume) of the intersection of two spherical (shell) caps, and the volume of the intersection of two balls.

A. Surface Area (Volume) of a Spherical (Shell) Cap

We first derive the surface area (volume) formula for a spherical (shell) cap. See also [23].

Let $C \subseteq \mathbb{S}^{m-1}$ be a spherical cap with angle θ on the $(m-1)$-sphere of radius $R = \sqrt{mN}$. The area $\mu(C)$ of C can be written as

$$\mu(C) = \int_0^\theta A_{m-2}(R\sin\rho)Rd\rho$$

where $A_{m-2}(R\sin\rho)$ is the total surface area of the $(m-2)$-sphere of radius $R\sin\rho$. Plugging in the expression for the

surface area of an $(m-2)$-sphere leads to

$$\mu(C) = \frac{2\pi^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)}(mN)^{\frac{m-2}{2}} \int_0^\theta \sin^{m-2}\rho \, d\rho.$$

We now characterize the exponent of $\mu(C)$. First, by Stirling's approximation, $\frac{2\pi^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)}(mN)^{\frac{m-2}{2}}$ in the above can be bounded as

$$2^{\frac{m}{2}[\log(2\pi eN)-\epsilon_1]} \le \frac{2\pi^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)}(mN)^{\frac{m-2}{2}} \le 2^{\frac{m}{2}[\log(2\pi eN)+\epsilon_1]} \tag{62}$$

for some $\epsilon_1 \to 0$ as $m \to \infty$. Also for m sufficiently large, we have

$$\begin{aligned}\int_0^\theta \sin^{m-2}\rho d\rho &= \int_0^\theta 2^{\frac{m-2}{2}\log\sin^2\rho} d\rho \\ &\ge \int_{\theta-\frac{1}{m}}^\theta 2^{\frac{m-2}{2}\log\sin^2\rho} d\rho \\ &\ge \frac{1}{m} 2^{\frac{m-2}{2}\log\sin^2(\theta-\frac{1}{m})} \\ &\ge 2^{\frac{m}{2}\left(\log\sin^2\theta-\epsilon_2\right)}\end{aligned}$$

and

$$\begin{aligned}\int_0^\theta \sin^{m-2}\rho d\rho &= \int_0^\theta 2^{\frac{m-2}{2}\log\sin^2\rho} d\rho \\ &\le \theta \cdot 2^{\frac{m-2}{2}\log\sin^2\theta} \\ &\le 2^{\frac{m}{2}\left(\log\sin^2\theta+\epsilon_2\right)}\end{aligned}$$

for some $\epsilon_2 \to 0$ as $m \to \infty$. Therefore, the area $\mu(C)$ can be bounded as

$$2^{\frac{m}{2}[\log(2\pi eN\sin^2\theta)-\epsilon]} \le \mu(C) \le 2^{\frac{m}{2}[\log(2\pi eN\sin^2\theta)+\epsilon]} \tag{63}$$

for some $\epsilon \to 0$ as $m \to \infty$.

Now suppose that $C = \text{ShellCap}(\mathbf{z}_0, \theta)$ is a shell cap on

$$\text{Shell}\left(\mathbf{0}, \sqrt{m(N-\delta)}, \sqrt{m(N+\delta)}\right)$$

where $\|\mathbf{z}_0\| = \sqrt{m(N-\delta)}$. Let $R_L = \sqrt{m(N-\delta)}$, $R_U = \sqrt{m(N+\delta)}$ and define $\mathbb{S}_{R_L}^{m-1}$ to be the $m-1$ sphere of radius R_L with Haar measure μ_{R_L}. We use spherical coordinates to integrate over the surface areas of the individual caps that make up the shell cap,

$$\begin{aligned}|C| &= \int_{\mathbb{R}^m} 1_{\text{ShellCap}(\mathbf{z}_0,\theta)} \, d\mathbf{z} \\ &= \int_{\mathbb{S}_{R_L}^{m-1}} \left(\int_{R_L}^{R_U} \left(\frac{r}{R_L}\right)^{m-1} 1_{\text{ShellCap}(\mathbf{z}_0,\theta)}\left(\frac{r}{R_L}\mathbf{z}\right) dr \right) d\mathbf{z} \\ &= \int_{\mathbb{S}_{R_L}^{m-1}} 1_{\text{Cap}(\mathbf{z}_0,\theta)}(\mathbf{z}) d\mathbf{z} \int_{R_L}^{R_U} \left(\frac{r}{R_L}\right)^{m-1} dr \\ &= \mu_{R_L}(\text{Cap}(\mathbf{z}_0,\theta)) \int_{R_L}^{R_U} \left(\frac{r}{R_L}\right)^{m-1} dr \end{aligned} \tag{64}$$

where the integral term on the right is bounded as

$$\int_{R_L}^{R_U} \left(\frac{r}{R_L}\right)^{m-1} dr \ge \left(\sqrt{m(N+\delta)} - \sqrt{m(N-\delta)}\right). \tag{65}$$

Together with (64), (63) and (65) imply

$$|C| \ge 2^{\frac{m}{2}[\log(2\pi e(N-\delta)\sin^2\theta)-\epsilon]}$$

for sufficiently large m. In a similar way,

$$|C| \le 2^{\frac{m}{2}[\log(2\pi e(N+\delta)\sin^2\theta)+\epsilon]},$$

and therefore

$$2^{\frac{m}{2}[\log(2\pi e(N-\delta)\sin^2\theta)-\epsilon]} \le |C| \le 2^{\frac{m}{2}[\log(2\pi e(N+\delta)\sin^2\theta)+\epsilon]} \tag{66}$$

where $\epsilon \to 0$ as $m \to \infty$.

B. Surface Area (Volume) of the Intersection of Two Spherical (Shell) Caps

Recall $\mathbb{S}^{m-1} \subset \mathbb{R}^m$ is the $(m-1)$-sphere of radius $R = \sqrt{mN}$. Let

$$C_i = \text{Cap}(\mathbf{v}_i, \theta_i) = \{\mathbf{v} \in \mathbb{S}^{m-1} : \angle(\mathbf{v}, \mathbf{v}_i) \le \theta_i\}, \quad i = 1, 2$$

be two spherical caps on $\mathbb{S}^{m-1}$ such that $\angle(\mathbf{v}_1, \mathbf{v}_2) = \frac{\pi}{2}$, $\theta_i \le \frac{\pi}{2}$, and $\theta_1 + \theta_2 > \frac{\pi}{2}$. We have the following lemma that characterizes the intersection measure $\mu(C_1 \cap C_2)$ of these two caps.

Lemma 15: For any $\epsilon > 0$ there exists an $M(\epsilon)$ such that for $m > M(\epsilon)$,

$$\mu(C_1 \cap C_2) \le 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta_1 - \cos^2\theta_2))+\epsilon]}$$

and

$$\mu(C_1 \cap C_2) \ge 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta_1 - \cos^2\theta_2))-\epsilon]}.$$

Proof: To prove this lemma, we will first derive the surface area formula for the intersection of the above two caps (see also [24]), and then characterize the exponent of this area.

1) Deriving the Surface Area Formula: Consider the points $\mathbf{v} \in \mathbb{S}^{m-1}$ such that

$$\angle(\mathbf{v}_1, \mathbf{v}) = \theta_1$$

and

$$\angle(\mathbf{v}_2, \mathbf{v}) = \theta_2.$$

These points satisfy the linear relations

$$\langle \mathbf{v}_1, \mathbf{v} \rangle = R^2 \cos\theta_1$$

and

$$\langle \mathbf{v}_2, \mathbf{v} \rangle = R^2 \cos\theta_2,$$

and therefore all such $\mathbf{v}$ lie in the unique $m-1$ dimensional subspace H defined by

$$\left\langle \frac{\mathbf{v}_1}{\cos\theta_1} - \frac{\mathbf{v}_2}{\cos\theta_2}, \mathbf{v} \right\rangle = 0.$$

The angle between the hyperplane H and the vector $\mathbf{v}_2$ is

$$\phi = \frac{\pi}{2} - \arccos\left(\frac{1}{R\sqrt{\frac{1}{\cos^2\theta_1} + \frac{1}{\cos^2\theta_2}}} \left\langle \frac{\mathbf{v}_1}{\cos\theta_1} - \frac{\mathbf{v}_2}{\cos\theta_2}, \mathbf{v}_2 \right\rangle \right)$$

and because $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal and $\|\mathbf{v}_2\| = R$,

$$\phi = \frac{\pi}{2} - \arccos\left(\frac{1}{\cos\theta_2\sqrt{\frac{1}{\cos^2\theta_1}+\frac{1}{\cos^2\theta_2}}}\right) = \arctan\left(\frac{\cos\theta_1}{\cos\theta_2}\right).$$

The approach will be as follows. Divide the intersection $C_1 \cap C_2$ into two parts C^+ and C^- that are on either side of the hyperplane H. More concretely,

$$C^+ = \left\{\mathbf{v} \in C_1 \cap C_2 : \left\langle \mathbf{v}, \frac{\mathbf{v}_1}{\cos\theta_1} - \frac{\mathbf{v}_2}{\cos\theta_2}\right\rangle \geq 0\right\}$$

and

$$C^- = \left\{\mathbf{v} \in C_1 \cap C_2 : \left\langle \mathbf{v}, \frac{\mathbf{v}_1}{\cos\theta_1} - \frac{\mathbf{v}_2}{\cos\theta_2}\right\rangle < 0\right\}.$$

Each part C^+ and C^- can be written as a union of lower dimensional spherical caps. We will find the measure of each part by integrating the measures of these lower dimensional caps.

The measure of the cap C_2 can be expressed as the integral

$$\mu(C_2) = \int_0^{\theta_2} A_{m-2}(R\sin\rho)Rd\rho$$

where $A_{m-2}(R\sin\rho)$ is the surface area of the $(m-2)$-sphere with radius $R\sin\rho$. If we consider a single $(m-2)$-sphere at some angle ρ, then the hyperplane H divides that $(m-2)$-sphere into two spherical caps. The claim is that each of these $m-2$ dimensional caps that is on the side of H with $\mathbf{v}_1$ is contained in C^+ (and those on the side with $\mathbf{v}_2$ are contained in C^-). Furthermore, all points in C^+ are in one of these $m-2$ dimensional caps. The claim follows because

$$\left\langle \mathbf{v}, \frac{\mathbf{v}_1}{\cos\theta_1} - \frac{\mathbf{v}_2}{\cos\theta_2}\right\rangle \geq 0$$

implies

$$\cos\theta_2\cos\left(\angle(\mathbf{v},\mathbf{v}_1)\right) \geq \cos\theta_1\cos\left(\angle(\mathbf{v},\mathbf{v}_2)\right)$$

and since $\angle(\mathbf{v},\mathbf{v}_2) \leq \theta_2$ and $\cos\left(\angle(\mathbf{v},\mathbf{v}_2)\right) \geq \cos\theta_2$, this implies

$$\cos\theta_2\cos\left(\angle(\mathbf{v},\mathbf{v}_1)\right) \geq \cos\theta_1\cos\theta_2.$$

Finally, this implies $\angle(\mathbf{v},\mathbf{v}_1) \leq \theta_1$, $\mathbf{v} \in C_1$, and $\mathbf{v} \in C^+$.

Note that for $\rho < \phi$, the $(m-2)$-sphere at angle ρ is entirely on the $\mathbf{v}_2$ side of H, and does not need to be included when computing the measure of C^+. This establishes the fact that

$$\mu(C^+) = \int_\phi^{\theta_2} C_{m-2}^{\theta_\rho}(R\sin\rho)Rd\rho$$

where $C_{m-2}^{\theta_\rho}(R\sin\rho)$ is the surface area of an $m-2$ dimensional spherical cap defined by angle θ_ρ on the $(m-2)$-sphere of radius $R\sin\rho$. Writing

$$\cos\theta_\rho = \frac{h}{R\sin\rho}$$

note that h is the distance from the center of the $(m-2)$-sphere at angle ρ to the $m-2$ dimensional hyperplane that divides the sphere into two caps. Furthermore, since the $(m-2)$-sphere has center $(R\cos\rho)\mathbf{v}_2$, we have

$$\tan\phi = \frac{h}{R\cos\rho}.$$

Therefore,

$$\theta_\rho = \arccos\left(\frac{\tan\phi}{\tan\rho}\right).$$

Combining this with the corresponding result for $\mu(C^-)$ yields

$$\begin{aligned}\mu(C_1\cap C_2) &= \mu(C^+) + \mu(C^-)\\ &= \int_\phi^{\theta_2} C_{m-2}^{\arccos\left(\frac{\tan\phi}{\tan\rho}\right)}(R\sin\rho)Rd\rho\\ &\quad + \int_{\frac{\pi}{2}-\phi}^{\theta_1} C_{m-2}^{\arccos\left(\frac{\tan(\pi/2-\phi)}{\tan\rho}\right)}(R\sin\rho)Rd\rho.\end{aligned}$$

This expression can be rewritten using known expressions for the area of a spherical cap in terms of the regularized incomplete beta function as

$$\mu(C_1\cap C_2) = J(\phi,\theta_2) + J(\pi/2-\phi,\theta_1),$$

where $J(\phi,\theta_2)$ is defined as

$$J(\phi,\theta_2) = \frac{(\pi mN)^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)}\int_\phi^{\theta_2}(\sin^{m-2}\rho)I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)d\rho \tag{67}$$

and $J(\pi/2-\phi,\theta_1)$ is defined similarly. Here in (67), $I_x(a,b)$ is the regularized incomplete beta function, given by

$$I_x(a,b) = \frac{B(x;a,b)}{B(a,b)}, \tag{68}$$

where $B(x;a,b)$ and $B(a,b)$ are the incomplete beta function and the complete beta function respectively:

$$B(x;a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$$
$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

2) Characterizing the Exponent: We now lower and upper bound $J(\phi,\theta_2)$ with exponential functions. First, using Stirling's approximation, $\frac{(\pi mN)^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)}$ on the R.H.S. of (67) can be bounded as

$$2^{\frac{m}{2}[\log(2\pi eN)-\epsilon_1]} \leq \frac{(\pi mN)^{\frac{m-1}{2}}}{\Gamma\left(\frac{m-1}{2}\right)} \leq 2^{\frac{m}{2}[\log(2\pi eN)+\epsilon_1]} \tag{69}$$

for some $\epsilon_1 \to 0$ as $m \to \infty$.

Now consider

$$I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)$$

inside the integral on the R.H.S. of (67). In light of (68), it can be written as

$$I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right) = \frac{B\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2;\frac{m-2}{2},\frac{1}{2}\right)}{B\left(\frac{m-2}{2},\frac{1}{2}\right)}. \tag{70}$$

For the denominator in (70), by Stirling's approximation, we have

$$B\left(\frac{m-2}{2},\frac{1}{2}\right)\sim\Gamma\left(\frac{1}{2}\right)\left(\frac{m-2}{2}\right)^{-\frac{1}{2}}.$$

For the numerator in (70), we have

$$\begin{aligned}
&B\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2;\frac{m-2}{2},\frac{1}{2}\right)\\
&\quad=\int_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-4}{2}}(1-t)^{-\frac{1}{2}}dt\\
&\quad\geq\int_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-4}{2}}dt\\
&\quad=\frac{2}{m-2}t^{\frac{m-2}{2}}\Big|_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\\
&\quad=\frac{2}{m-2}\left[1-\left(\frac{\tan\phi}{\tan\rho}\right)^2\right]^{\frac{m-2}{2}}\\
&\quad\geq 2^{\frac{m}{2}\left[\log\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2\right)-\epsilon_2\right]},
\end{aligned}$$

for some $\epsilon_2\to 0$ as $m\to\infty$, and

$$\begin{aligned}
&B\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2;\frac{m-2}{2},\frac{1}{2}\right)\\
&\quad=\int_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-4}{2}}(1-t)^{-\frac{1}{2}}dt\\
&\quad\leq\int_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-4}{2}}\left(1-\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2\right)\right)^{-\frac{1}{2}}dt\\
&\quad=\frac{\tan\rho}{\tan\phi}\int_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-4}{2}}dt\\
&\quad\leq\frac{\tan\theta_2}{\tan\phi}\int_0^{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-4}{2}}dt\\
&\quad=\frac{2\tan\theta_2}{(m-2)\tan\phi}\left[1-\left(\frac{\tan\phi}{\tan\rho}\right)^2\right]^{\frac{m-2}{2}}\\
&\quad\leq 2^{\frac{m}{2}\left[\log\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2\right)+\epsilon_3\right]},
\end{aligned}$$

for some $\epsilon_3\to 0$ as $m\to\infty$. Also noting that

$$\sin^{m-2}\rho=2^{\frac{m-2}{2}\log\sin^2\rho}$$

with $\rho\in[\phi,\theta_2]$, we can bound the integrand in (67) as

$$\begin{aligned}
&\left(\sin^{m-2}\rho\right)I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)\\
&\quad\geq 2^{\frac{m}{2}\left[\log\left((\sin^2\rho)\left(1-\left(\frac{\tan\phi}{\tan\rho}\right)^2\right)\right)-\epsilon_4\right]}\\
&\quad=2^{\frac{m}{2}\left[\log\left(\sin^2\rho-\tan^2\phi\cos^2\rho\right)-\epsilon_4\right]}
\end{aligned}$$

and

$$\begin{aligned}
&\left(\sin^{m-2}\rho\right)I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)\\
&\qquad\leq 2^{\frac{m}{2}\left[\log\left(\sin^2\rho-\tan^2\phi\cos^2\rho\right)+\epsilon_4\right]}
\end{aligned}$$

for some $\epsilon_4\to 0$ as $m\to\infty$. For sufficiently large m,

$$\begin{aligned}
&\int_\phi^{\theta_2}(\sin^{m-2}\rho)I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)d\rho\\
&\quad\geq\int_{\theta_2-\frac{1}{m}}^{\theta_2}(\sin^{m-2}\rho)I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)d\rho\\
&\quad\geq\int_{\theta_2-\frac{1}{m}}^{\theta_2}2^{\frac{m}{2}\left[\log\left(\sin^2\rho-\tan^2\phi\cos^2\rho\right)-\epsilon_4\right]}d\rho\\
&\quad\geq\frac{1}{m}2^{\frac{n}{2}\left[\log\left(\sin^2(\theta_2-\frac{1}{m})-\tan^2\phi\cos^2(\theta_2-\frac{1}{m})\right)-\epsilon_4\right]}\\
&\quad\geq 2^{\frac{m}{2}\left[\log\left(\sin^2\theta_2-\tan^2\phi\cos^2\theta_2\right)-\epsilon_5\right]}\\
&\quad=2^{\frac{m}{2}\left[\log\left(\sin^2\theta_2-\cos^2\theta_1\right)-\epsilon_5\right]},
\end{aligned}$$

and

$$\begin{aligned}
&\int_\phi^{\theta_2}(\sin^{m-2}\rho)I_{1-\left(\frac{\tan\phi}{\tan\rho}\right)^2}\left(\frac{m-2}{2},\frac{1}{2}\right)d\rho\\
&\qquad\leq 2^{\frac{m}{2}\left[\log\left(\sin^2\theta_2-\cos^2\theta_1\right)+\epsilon_5\right]}
\end{aligned}$$

for some $\epsilon_5\to 0$ as $m\to\infty$.

Combining this with (69), we can bound $J(\phi,\theta_2)$ as

$$\begin{aligned}
2^{\frac{m}{2}\left[\log 2\pi eN\left(\sin^2\theta_2-\cos^2\theta_1\right)-\epsilon_6\right]}&\leq J(\phi,\theta_2)\\
&\leq 2^{\frac{m}{2}\left[\log 2\pi eN\left(\sin^2\theta_2-\cos^2\theta_1\right)+\epsilon_6\right]}
\end{aligned}$$

for some $\epsilon_6\to 0$ as $m\to\infty$.

Due to symmetry, we can also bound $J(\pi/2-\phi,\theta_1)$ as

$$\begin{aligned}
2^{\frac{m}{2}\left[\log 2\pi eN\left(\sin^2\theta_1-\cos^2\theta_2\right)-\epsilon_6\right]}&\leq J(\pi/2-\phi,\theta_1)\\
&\leq 2^{\frac{m}{2}\left[\log 2\pi eN\left(\sin^2\theta_1-\cos^2\theta_2\right)+\epsilon_6\right]}.
\end{aligned}$$

Noting that $\sin^2\theta_2-\cos^2\theta_1=\sin^2\theta_1-\cos^2\theta_2$, we have

$$\begin{aligned}
\mu(C_1\cap C_2)&\geq J(\phi,\theta_2)+J(\pi/2-\phi,\theta_1)\\
&\geq 2^{\frac{m}{2}\left[\log 2\pi eN\left(\sin^2\theta_1-\cos^2\theta_2\right)-\epsilon\right]}
\end{aligned}$$

and

$$\mu(C_1\cap C_2)\leq 2^{\frac{m}{2}\left[\log 2\pi eN\left(\sin^2\theta_1-\cos^2\theta_2\right)+\epsilon\right]}$$

for some $\epsilon\to 0$ as $m\to\infty$. This completes the proof of the lemma. ■

We now utilize Lemma 15 to characterize the volume of the intersection of two shell caps. Consider a spherical shell

$$\text{Shell}\,(\mathbf{0},R_L,R_U)$$

with $R_L=\sqrt{m(N-\delta)}$, $R_U=\sqrt{m(N+\delta)}$ and two caps on this shell, i.e. $S_1=\text{ShellCap}(\mathbf{z}_0,\theta)$ and $S_2=\text{ShellCap}(\mathbf{y}_0,\omega)$, where $\angle(\mathbf{z}_0,\mathbf{y}_0)=\pi/2$ and $\theta+\omega>\pi/2$. The following lemma bounds the intersection volume $|S_1\cap S_2|$ of these two shell caps.

Lemma 16: For any $\epsilon>0$ there exists an $M(\epsilon)$ such that for $m>M(\epsilon)$,

$$|S_1\cap S_2|\geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta-\cos^2\omega))-\epsilon]}$$

and

$$|S_1\cap S_2|\leq 2^{\frac{m}{2}[\log(2\pi e(N+\delta)(\sin^2\theta-\cos^2\omega))+\epsilon]}.$$

Proof: Using spherical coordinates, we have

$$\begin{aligned}|S_1 \cap S_2| &= \int_{\mathbb{R}^m} 1_{S_1\cap S_2}(\mathbf{z})\, d\mathbf{z} \\ &= \int_{\mathbb{S}^{m-1}} \left(\int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} 1_{S_1\cap S_2}\left(\frac{r}{R}\mathbf{z}\right)\, dr \right) d\mathbf{z} \\ &= \int_{\mathbb{S}^{m-1}} 1_{\mathrm{Cap}(\mathbf{z}_0,\theta)\cap \mathrm{Cap}(\mathbf{y}_0,\omega)}(\mathbf{z}) d\mathbf{z} \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \\ &= \mu(\mathrm{Cap}(\mathbf{z}_0,\theta)\cap \mathrm{Cap}(\mathbf{y}_0,\omega)) \int_{R_L}^{R_U} \left(\frac{r}{R}\right)^{m-1} dr \end{aligned} \tag{71}$$

where the integral term on the right is bounded as

$$\begin{aligned}\int_{\sqrt{m(N-\delta)}}^{\sqrt{m(N+\delta)}} \left(\frac{r}{R}\right)^{m-1} dr &\geq \int_{\sqrt{mN}}^{\sqrt{m(N+\delta)}} \left(\frac{r}{R}\right)^{m-1} dr \\ &\geq \sqrt{m(N+\delta)} - \sqrt{mN}. \end{aligned} \tag{72}$$

Given $\epsilon > 0$, set $M = \max\{M_1, M_2\}$ where M_1 is given by Lemma 15 to ensure

$$\mu(\mathrm{Cap}(\mathbf{z}_0,\theta)\cap \mathrm{Cap}(\mathbf{y}_0,\omega)) \geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon/2]}$$

and M_2 is chosen to be sufficiently large so that the right-hand side of (72) satisfies

$$\sqrt{m(N+\delta)} - \sqrt{mN} \geq 2^{-m\epsilon}.$$

Together with (71), this implies

$$|S_1 \cap S_2| \geq 2^{\frac{m}{2}[\log(2\pi eN(\sin^2\theta - \cos^2\omega)) - \epsilon]}$$

for $m > M$.

For the inequality in the other direction, define $\mathbb{S}_{R_U}^{m-1}$ to be the $m-1$ sphere of radius R_U with Haar measure μ_{R_U}. Then

$$\begin{aligned}&|S_1 \cap S_2| \\ &= \int_{\mathbb{R}^m} 1_{S_1\cap S_2}(\mathbf{z})\, d\mathbf{z} \\ &= \int_{\mathbb{S}_{R_U}^{m-1}} \left(\int_{R_L}^{R_U} \left(\frac{r}{R_U}\right)^{m-1} 1_{S_1\cap S_2}\left(\frac{r}{R_U}\mathbf{z}\right)\, dr \right) d\mathbf{z} \\ &= \int_{\mathbb{S}_{R_U}^{m-1}} 1_{\mathrm{Cap}(\mathbf{z}_0,\theta)\cap \mathrm{Cap}(\mathbf{y}_0,\omega)}(\mathbf{z}) d\mathbf{z} \int_{R_L}^{R_U} \left(\frac{r}{R_U}\right)^{m-1} dr \\ &= \mu_{R_U}(\mathrm{Cap}(\mathbf{z}_0,\theta)\cap \mathrm{Cap}(\mathbf{y}_0,\omega)) \int_{R_L}^{R_U} \left(\frac{r}{R_U}\right)^{m-1} dr \end{aligned} \tag{73}$$

where the integral term on the right is bounded as

$$\int_{\sqrt{m(N-\delta)}}^{\sqrt{m(N+\delta)}} \left(\frac{r}{R_U}\right)^{m-1} dr \leq \sqrt{m(N+\delta)} - \sqrt{m(N-\delta)}. \tag{74}$$

Given $\epsilon > 0$, set $M = \max\{M_1, M_2\}$ where M_1 is given by Lemma 15 to ensure

$$\mu_{R_U}(\mathrm{Cap}(\mathbf{z}_0,\theta)\cap \mathrm{Cap}(\mathbf{y}_0,\omega)) \leq 2^{\frac{m}{2}[\log(2\pi e(N+\delta)(\sin^2\theta_1 - \cos^2\theta_2)) + \epsilon/2]}$$

and M_2 is chosen to be sufficiently large so that the right-hand side of (74) satisfies

$$\sqrt{m(N+\delta)} - \sqrt{m(N-\delta)} \leq 2^{m\epsilon}.$$

Together with (73), this implies

$$|S_1 \cap S_2| \leq 2^{\frac{m}{2}[\log(2\pi e(N+\delta)(\sin^2\theta - \cos^2\omega)) + \epsilon]}$$

for $m > M$. ∎

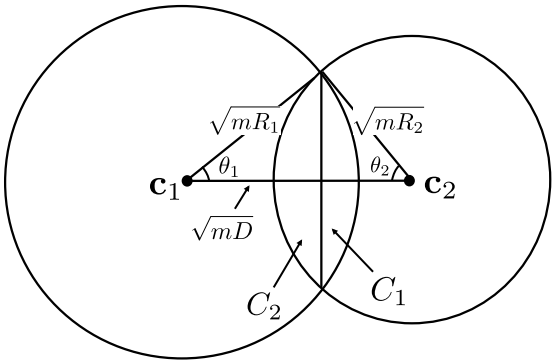


Fig. 7. Intersection of two balls.

C. Volume of the Intersection of Two Balls

Proof of Lemma 13: The intersection of Ball $(\mathbf{c}_1, \sqrt{mR_1})$ and Ball$(\mathbf{c}_2, \sqrt{mR_1})$ consists of two caps: C_1 and C_2, as depicted in Fig. 7. To bound the volume of Ball$(\mathbf{c}_1, \sqrt{mR_1}) \cap$ Ball$(\mathbf{c}_2, \sqrt{mR_1})$, we will bound $|C_1|$ and $|C_2|$ respectively.

We first bound $|C_1|$. By the cosine formula, we have

$$\begin{aligned}\cos\theta_1 &= \frac{mR_1 + mD - mR_2}{2\sqrt{mR_1}\sqrt{mD}} \\ &= \frac{R_1 + D - R_2}{2\sqrt{R_1 D}}\end{aligned}$$

and therefore

$$\begin{aligned}\sin^2\theta_1 &= 1 - \cos^2\theta_1 \\ &= 1 - \frac{(R_1 + D - R_2)^2}{4\ R_1 D} \\ &= \frac{2R_1D + 2R_1R_2 + 2DR_2 - R_1^2 - R_2^2 - D^2}{4R_1D}.\end{aligned}$$

From Appendix C-A, we have for any $\epsilon > 0$ and m sufficiently large,

$$\begin{aligned}|C_1| &\leq 2^{m\left(\frac{1}{2}\log 2\pi e R_1 \sin^2\theta_1 + \frac{\epsilon}{2}\right)} \\ &= 2^{m\left(\frac{1}{2}\log \pi e \lambda(R_1,R_2,D) + \frac{\epsilon}{2}\right)}\end{aligned}$$

where

$$\lambda(R_1, R_2, D) := \frac{2R_1D + 2R_1R_2 + 2DR_2 - R_1^2 - R_2^2 - D^2}{2D}.$$

Similarly, we have

$$\begin{aligned}\sin^2\theta_2 &= 1-\cos^2\theta_2\\ &= 1-\frac{(R_2+D-R_1)^2}{4\ R_2D}\\ &= \frac{2R_1D+2R_1R_2+2DR_2-R_1^2-R_2^2-D^2}{4R_2D}\end{aligned}$$

and therefore

$$\begin{aligned}|C_2| &\le 2^{m\left(\frac{1}{2}\log 2\pi e R_2\sin^2\theta_2+\frac{\epsilon}{2}\right)}\\ &= 2^{m\left(\frac{1}{2}\log \pi e\lambda(R_1,R_2,D)+\frac{\epsilon}{2}\right)}.\end{aligned}$$

Combining the above, we obtain

$$\begin{aligned}&\left|\mathrm{Ball}(\mathbf{c}_1,\sqrt{mR_1})\cap\mathrm{Ball}(\mathbf{c}_2,\sqrt{mR_1})\right|\\ &\quad=|C_1|+|C_2|\\ &\quad\le 2^{m\left(\frac{1}{2}\log \pi e\lambda(R_1,R_2,D)+\epsilon\right)}\end{aligned}$$

for any $\epsilon>0$ and m sufficiently large. ∎

ACKNOWLEDGEMENT

The authors would like to acknowledge inspiring discussions with Liang-Liang Xie within a preceding collaboration [4]. They would also like to thank the anonymous reviewers and the Associate Editor for many valuable comments that helped improve the presentation of this paper.

REFERENCES

[1] X. Wu, L. P. Barnes, and A. Özgür, "Cover's open problem: 'The capacity of the relay channel,'" in *Proc. 54th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Monticello, IL, USA, 2016, pp. 148–155.

[2] T. M. Cover, "The capacity of the relay channel," in *Open Problems in Communication and Computation*, T. M. Cover and B. Gopinath, Eds. New York, NY, USA: Springer-Verlag, 1987, pp. 72–73.

[3] X. Wu and A. Özgür, "Improving on the cut-set bound via geometric analysis of typical sets," in *Proc. Int. Zurich Seminar Commun.*, 2016, pp. 1675–1679.

[4] X. Wu, A. Özgür, and L.-L. Xie, "Improving on the cut-set bound via geometric analysis of typical sets," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2254–2277, Apr. 2017.

[5] X. Wu and A. Özgür, "Cut-set bound is loose for Gaussian relay networks," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Monticello, IL, USA, Sep./Oct. 2015, pp. 1135–1142.

[6] X. Wu and A. Özgür, "Cut-set bound is loose for Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1023–1037, Feb. 2018.

[7] X. Wu and A. Özgür, "Improving on the cut-set bound for general primitive relay channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, Jul. 2016, pp. 1675–1679.

[8] X. Wu, L. P. Barnes, and A. Özgür, "The geometry of the relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 2017, pp. 2233–2237.

[9] L. P. Barnes, X. Wu, and A. Özgür, "A solution to cover's problem for the binary symmetric relay channel: Geometry of sets on the Hamming sphere," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2017, pp. 844–851.

[10] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[11] Z. Zhang, "Partial converse for a relay channel," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 1106–1110, Sep. 1988.

[12] M. Raginsky and I. Sason, "Concentration of measure inequalities in information theory, communications, and coding," *Found. Trends Commun. Inf. Theory*, vol. 10, nos. 1–2, pp. 1–246, Oct. 2013.

[13] A. Burchard. (Jun. 2009). *A Short Course Rearrangement Inequalities*. [Online]. Available: http://www.math.utoronto.ca/almut/rearrange.pdf

[14] P. Lévy, *Problmès Concrets d'Analyse Fonctionnelle*, 2nd ed. Paris, France: Gautheir-Villars, 1951.

[15] J. Matoušek, *Lectures on Discrete Geometry*, vol. 212. Berlin, Germany: Springer, 2002.

[16] G. Schechtman, "Concentration, results and applications," in *Handbook of the Geometry of Banach Spaces*, vol. 2, W. B. Johnson and J. Lindenstrauss, Eds. Amsterdam, The Netherlands: North Holland, 2003, pp. 1603–1634.

[17] C. E. Shannon, "Communication in the presence of noise," *Proc. Inst. Radio Eng.*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[18] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[20] K. Marton, "Bounding $\bar{d}$-distance by informational divergence: A method to prove measure concentration," *Ann. Probab.*, vol. 24, no. 2, pp. 857–866, 1996.

[21] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[22] M. Talagrand, "Transportation cost for Gaussian and other product measures," *Geometric Funct. Anal.*, pp. 587–600, 1996.

[23] S. Li, "Concise formulas for the area and volume of a hyperspherical cap," *Asian J. Math. Statist.*, vol. 4, no. 1, pp. 66–70, 2011.

[24] Y. Lee and W. C. Kim, "Concise formulas for the surface area of the intersection of two hyperspherical caps," KAIST, Daejeon, South Korea, Tech. Rep., 2014.

[25] A. Baernstein II and B. A. Taylor, "Spherical rearrangements, subharmonic functions, and $*$-functions in n-space," *Duke Math. J.*, vol. 43, no. 2, pp. 245–268, 1976.

[26] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.

Xiugang Wu (M'14) received the B.Eng. degree with honors in electronics and information engineering from Tongji University, Shanghai, China, in 2007, and the M.A.Sc and Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2009 and 2014, respectively. He was a postdoctoral fellow in the Department of Electrical Engineering, Stanford University, Stanford, CA, during 2015–2018. He is currently an assistant professor at the University of Delaware, Newark, DE, where he is jointly appointed in the Department of Electrical and Computer Engineering and the Department of Computer and Information Sciences. His research interests are in information theory, networks, data science, and the interplay between them. He is a recipient of the 2017 NSF Center for Science of Information (CSoI) Postdoctoral Fellowship.

Leighton Pate Barnes (S'17) received a B.S. in Mathematics '13, B.S. in Electrical Science and Engineering '13, and M.Eng. in Electrical Engineering and Computer Science '15, all from the Massachusetts Institute of Technology. While there, he received the Harold L. Hazen Award for excellence in teaching. He is currently a Ph.D. candidate in the Department of Electrical Engineering at Stanford University, where he studies geometric extremal problems applied to information theory, communication, and estimation.

Ayfer Özgür (M'06) received her B.Sc. degrees in electrical engineering and physics from Middle East Technical University, Turkey, in 2001 and the M.Sc. degree in communications from the same university in 2004. From 2001 to 2004, she worked as hardware engineer for the Defense Industries Development Institute in Turkey. She received her Ph.D. degree in 2009 from the Information Processing Group at EPFL, Switzerland. In 2010 and 2011, she was a post-doctoral scholar at the same institution. She is currently an Assistant Professor in the Electrical Engineering Department at Stanford University where she is a Hoover and Gabilan Fellow. Her current research interests include distributed communication and learning, wireless systems, and information theory. Dr. Özgür received the EPFL Best Ph.D. Thesis Award in 2010, the NSF CAREER award in 2013 and the Okawa Foundation Research Grant in 2018.