The Geometry of the Relay Channel

Xiugang Wu, Leighton Pate Barnes and Ayfer Özgür
Stanford University, Stanford, CA 94305
Email: {x23wu, lpb, aozgur}@stanford.edu

***Abstract*—Consider a memoryless relay channel, where the channel from the relay to the destination is 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 our recent work, we answered this question in the case when the channels from the source to the relay and destination are symmetric, which is the original assumption imposed by Cover, and when these channels are Gaussian. We showed that $C(C_0)$ can not equal to $C(\infty)$ unless $C_0 = \infty$, regardless of the SNR of the Gaussian channels, while the cut-set bound would suggest that $C(\infty)$ can be achieved at finite C_0.**

In this paper, we show that our techniques for solving Cover's problem can be naturally extended to the general Gaussian case, where the channels from the source to the relay and destination may be asymmetric, and prove an upper bound on the capacity $C(C_0)$ of a general Gaussian relay channel for any C_0. This upper bound immediately implies that our previous conclusion, i.e. $C(C_0)$ can not equal to $C(\infty)$ unless $C_0 = \infty$, also holds in the asymmetric case. Our approach is geometric and relies on a strengthening of the isoperimetric inequality on the sphere by using the Riesz rearrangement inequality.

I. Introduction, Problem Setup and Results

Consider a simple relay channel as depicted in Fig. 1, where the source's input X is received by the relay Z and the destination Y through a channel $p(y,z|x)$, and the relay Z can communicate to the destination Y via an isolated bit pipe of capacity C_0. A $(2^{nR}, n)$ code for this channel consists of an encoding function $X^n : [1 : 2^{nR}] \to \mathcal{X}^n$, 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 average probability of error of the code is defined as

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

where the message M is assumed to be uniformly drawn from the message set $[1 : 2^{nR}]$. A rate R is said to be achievable if there exists a sequence of $(2^{nR}, n)$ codes such that the average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$. The capacity of the relay channel is the supremum of all achievable rates, denoted by $C(C_0)$. Obviously, $C(C_0)$ is a non-decreasing function in C_0, and we have $C(0) = \max_{p(x)} I(X;Y)$ and $C(\infty) = \max_{p(x)} I(X;Y,Z)$.

Characterizing $C(C_0)$ for this channel has proven to be very difficult after decades of significant research efforts [1]–[3].

The work was supported in part by NSF award CCF-1514538 and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370.

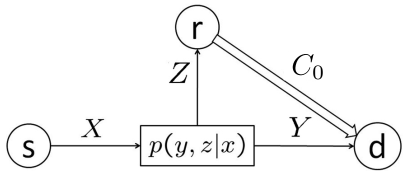


Fig. 1. A relay channel.

A seemingly less demanding question that we believe still captures a lot of the difficulties in the problem is the one posed by Cover, and indeed named "The Capacity of the Relay Channel" by him, in *Open Problems in Communication and Computation*, Springer-Verlag, 1987 [4]. Specifically, assuming Z and Y are conditionally independent and identically distributed given X, Cover asked: "*What is the critical value of C_0 such that $C(C_0)$ first equals $C(\infty)$?*" In other words, we are interested in characterizing

$$C_0^* := \inf\{C_0 : C(C_0) = C(\infty)\}, \tag{1}$$

i.e., the minimum rate needed for the Z-Y link so that the maximum possible rate $C(\infty)$, corresponding to full cooperation between the relay and the destination, can be achieved.

A. Main Result

In our recent work [5], we answered this long-standing open question in the Gaussian case. In particular, consider a Gaussian relay channel as depicted in Fig. 2, in which

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

with X being constrained to average power P, i.e.[1],

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

and $W_1 \sim \mathcal{N}(0, N_1), W_2 \sim \mathcal{N}(0, N_2)$ representing Gaussian noises that are independent of each other and X. When Z and Y are conditionally i.i.d. given X, i.e. when $N_1 = N_2 =: N$, we observe that

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

Moreover, the cut-set bound [2] yields the following lower bound on C_0^*:

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

[1]Note that this constraint is less stringent than requiring $\|x^n(m)\|^2 \leq nP$, $\forall m \in [1 : 2^{nR}]$.

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

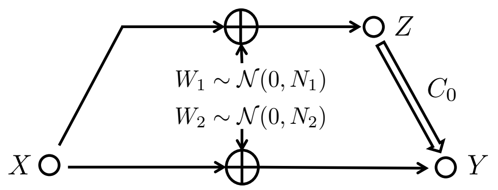


Fig. 2. A Gaussian relay channel.

which may lead one to suspect that $C(\infty)$ could be achieved at finite C_0. Surprisingly, however, it turns out that $C(C_0)$ cannot equal $C(\infty)$ unless $C_0 = \infty$, regardless of the SNR $= \frac{P}{N}$ of the Gaussian channels. Formally, we have the following theorem.

Theorem 1.1 ([5]): For the Gaussian relay channel depicted in Fig. 2, $C_0^* = \infty$ when $N_1 = N_2$.

In this paper, we further consider the general Gaussian case when Z and Y are independently Gaussian given X but not necessarily have the same variance, i.e. N_1 and N_2 may be different. We show that our approach developed in [5] for solving Cover's problem can be naturally extended to this general case and leads to an upper bound on the capacity $C(C_0)$ of the channel for any C_0 under any configuration of (P, N_1, N_2). This upper bound is formally stated in the following theorem.

Theorem 1.2: For a general Gaussian relay channel as depicted in Fig. 2, the capacity $C(C_0)$ satisfies

$$\begin{cases} C(C_0) \leq \frac{1}{2}\log\left(1+\frac{P}{N_2}\right) + C_0 + \log\sin\theta & (3) \\ C(C_0) \leq \frac{1}{2}\log\left(1+\frac{P}{N_2}\right) + \min\limits_{\omega\in\left(\frac{\pi}{2}-\theta,\frac{\pi}{2}\right]} h_\theta(\omega), & (4) \end{cases}$$

for some $\theta \in \left[\arcsin 2^{-C_0}, \frac{\pi}{2}\right]$, where $h_\theta(\omega)$ is defined to be

$$\frac{1}{2}\log\left(\frac{\left[P(N_1+N_2)+N_1N_2\sin^2\omega - 2P\sqrt{N_1N_2}\cos\omega\right]\sin^2\theta}{N_1(P+N_2)(\sin^2\theta-\cos^2\omega)}\right).$$

In Fig. 3 we plot the upper bound in Theorem 1.2 (label: New bound) for the symmetric case when $N_1 = N_2 = N$ under three different values of the SNR $= \frac{P}{N}$ of the Gaussian channels, together with the celebrated cut-set bound [2] and an upper bound on the capacity of this channel we have previously derived in [10] (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. Note that from these figures one can visually observe that the new upper bound reaches the value $C(\infty)$ only as $C_0 \to \infty$, which leads to the conclusion in Theorem 1.1. In fact, as we will show in the next section, the new upper bound in Theorem 1.2 directly bounds the capacity $C(C_0)$ of a general Gaussian relay channel away from $C(\infty)$ for any finite C_0, and thus resolves Cover's problem regarding C_0^* for the more general asymmetric case. In particular, we have the following corollary.

Corollary 1.1: For a general Gaussian relay channel as depicted in Fig. 2, $C_0^* = \infty$.

Due to space constraints, in this paper we only provide a proof sketch for Theorem 1.2. The proof details can be found in the long version posted on arxiv [6], however note that this long version only focuses on the symmetric case of $N_1 = N_2$.

B. Approach

Our approach builds on the method we developed in our earlier work [7]–[11] for characterizing information tensions in a Markov chain by using high-dimensional geometry. The main idea is to study the geometry of the high-dimensional typical sets associated with the random variables in the Markov chain and then translate this high-dimensional geometry to information inequalities for the random variables. The main geometric tool employed in our previous work [7]–[11] was the so-called blowing-up lemma. In the current paper, our main geometric ingredient is a strengthening of the isoperimetric inequality on a high-dimensional sphere, which we developed in [5] by building on the Riesz rearrangement inequality. The classical isoperimetric inequality on the sphere states that among all sets on the sphere with a given volume the spherical cap has the smallest boundary or more generally the smallest volume of neighborhood. In [5] (see also [6]), we showed that the spherical cap is the extremal set not only in terms of minimizing the volume of its neighborhood, but roughly speaking also in terms of minimizing its total intersection volume with a ball drawn around a randomly chosen point on the sphere.

II. Proofs

The proof of Theorem 1.2 follows from the below lemma, which is the main technical focus of this paper and whose proof is outlined in Section II-C. We now state this lemma and show how it leads to the bound in Theorem 1.2, which is then used to establish Corollary 1.1.

Lemma 2.1: Let I_n be an integer random variable and X^n, Y^n and Z^n be n-length random vectors which form the Markov chain $I_n - Z^n - X^n - Y^n$. Assume moreover that Z^n and Y^n are independent white Gaussian vectors given X^n such that $Z^n \sim \mathcal{N}(X^n, N_1 I_{n\times n})$ and $Y^n \sim \mathcal{N}(X^n, N_2 I_{n\times n})$ where $I_{n\times n}$ denotes the identity matrix, and $E[\|X^n\|^2] = nP$, and $I_n = f_n(Z^n)$ is a deterministic mapping of Z^n to a set of integers. Let $H(I_n|X^n)$ be denoted by $-n\log\sin\theta_n$, i.e., define $\theta_n := \arcsin 2^{-\frac{1}{n}H(I_n|X^n)}$. Then the inequality (5) holds for any n.

A. Proof of Theorem 1.2

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

$$\begin{aligned} nR &\leq I(X^n; Y^n, I_n) + n\mu & (6) \\ &= I(X^n;Y^n) + I(X^n; I_n|Y^n) + n\mu & \\ &= I(X^n;Y^n) + H(I_n|Y^n) - H(I_n|X^n) + n\mu & (7) \\ &\leq nI(X_Q;Y_Q) + H(I_n|Y^n) - H(I_n|X^n) + n\mu & (8) \end{aligned}$$

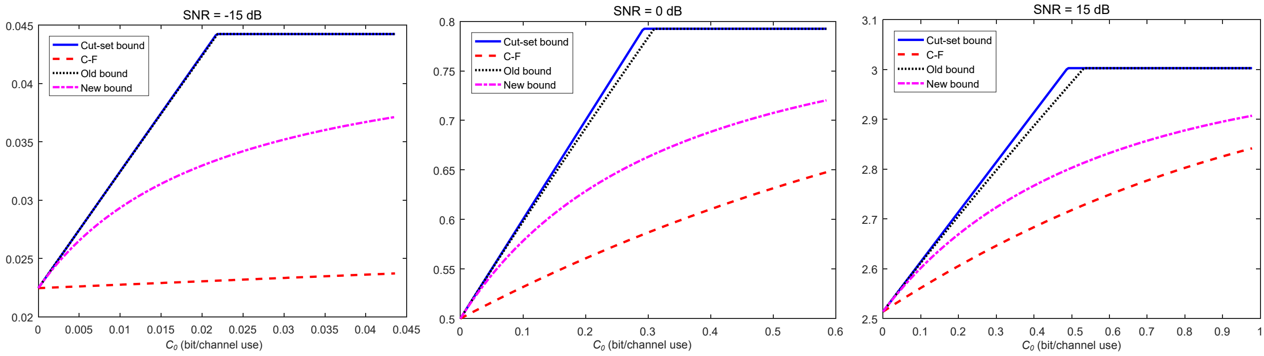


Fig. 3. Upper bounds and achievable rates for the Gaussian relay channel.

$$H(I_n|Y^n) \le n \cdot \min_{\omega \in \left(\frac{\pi}{2}-\theta_n, \frac{\pi}{2}\right]} \frac{1}{2} \log \left(\frac{P(N_1+N_2) + N_1 N_2 \sin^2\omega - 2P\sqrt{N_1 N_2}\cos\omega}{N_1(P+N_2)(\sin^2\theta_n - \cos^2\omega)} \right). \tag{5}$$

$$\le \frac{n}{2} \log\left(1 + \frac{P}{N_2}\right) + H(I_n|Y^n) - H(I_n|X^n) + n\mu, \tag{9}$$

for any $\mu > 0$ and n sufficiently large. In the above, (6) follows from Fano's inequality, (7) uses the fact that $I_n - X^n - Y^n$ form a Markov chain and thus $H(I_n|X^n, Y^n) = H(I_n|X^n)$, (8) follows by defining the time sharing random variable Q to be uniformly distributed over $[1:n]$, and (9) follows because $E[X_Q^2] \le P$.

Now we use Lemma 2.1 to upper bound the difference $H(I_n|Y^n) - H(I_n|X^n)$ in (9). 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 the lemma. It is trivial to observe that they satisfy the required Markov chain condition and Z^n and Y^n are independently Gaussian given X^n due to the channel structure. Note also that without loss of generality we can assume that the code satisfies the average power constraint in (2) with equality, because given a $(2^{nR}, n)$ code with average probability of error $P_e^{(n)}$ and $E[\|X^n\|^2] = nP' < nP$, we can always scale up the codewords by a factor of $\sqrt{nP/nP'}$ and achieve an average probability of error smaller than or equal to $P_e^{(n)}$.

Therefore, applying Lemma 2.1 to the random variables associated with a code for the relay channel, we can bound the difference of the two entropy terms in (9) and conclude that for any achievable rate R,

$$R \le \frac{1}{2}\log\left(1+\frac{P}{N_2}\right) + \min_{\omega \in \left(\frac{\pi}{2}-\theta_n, \frac{\pi}{2}\right]} h_{\theta_n}(\omega) + \mu, \tag{10}$$

where $h_{\theta_n}(\omega)$ is defined similarly as in Theorem 1.2, with $\theta_n := \arcsin 2^{-\frac{1}{n}H(I_n|X^n)}$ satisfying

$$\theta_0 := \arcsin 2^{-C_0} \le \arcsin 2^{-\frac{1}{n}H(I_n|X^n)} = \theta_n \le \frac{\pi}{2}. \tag{11}$$

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

$$R \le \frac{1}{2}\log\left(1+\frac{P}{N_2}\right) + C_0 + \log\sin\theta_n + \mu, \tag{12}$$

which simply follows from (9) by upper bounding $H(I_n|Y^n)$ with nC_0 and plugging in the definition of θ_n. Combining (10) and (12) concludes the proof of the theorem.

B. *Proof of Corollary 1.1*

To see Theorem 1.2 implies Corollary 1.1, consider the second bound (4) on $C(C_0)$ in Theorem 1.2. Note that for any $\theta > 0$, the function $h_\theta(\omega)$ in (4) satisfies

$$\begin{aligned} h_\theta\left(\frac{\pi}{2}\right) &= \frac{1}{2}\log\frac{P(N_1+N_2)+N_1N_2}{N_1(P+N_2)} \\ &= C(\infty) - \frac{1}{2}\log\left(1+\frac{P}{N_2}\right) \end{aligned} \tag{13}$$

and $h_\theta(\omega)$ is increasing at $\frac{\pi}{2}$, or more precisely, $h'_\theta\left(\frac{\pi}{2}\right) > 0$. Therefore, as long as $\theta > 0$, which is the case when C_0 is finite, the minimization of $h_\theta(\omega)$ with respect to ω in (4) yields a value strictly smaller than $h_\theta\left(\frac{\pi}{2}\right)$ in (13), and thus $C(C_0)$ for any finite C_0 is strictly smaller than $C(\infty)$.

C. *Proof Outline for Lemma 2.1*

Recall that Lemma 2.1 upper bounds $H(I_n|Y^n)$ in terms of $H(I_n|X^n)$ in a Markov chain $I_n - Z^n - X^n - Y^n$, where Z^n and Y^n are independent Gaussian vectors given X^n, $E[\|X^n\|^2] = nP$ and $I_n = f_n(Z^n)$ is a deterministic mapping of Z^n to a set of integers. Interestingly, although this result itself is stated for random variables in n dimensions, to prove it we will take a lifting step to a higher dimensional space and look at the i.i.d. extensions of the random variables X^n, Y^n, Z^n and I_n. Specifically, we consider the following B-length i.i.d. sequence

$$\{(X^n(b), Y^n(b), Z^n(b), I_n(b))\}_{b=1}^B, \tag{14}$$

where for any $b \in [1:B]$, $(X^n(b), Y^n(b), Z^n(b), I_n(b))$ has the same distribution as (X^n, Y^n, Z^n, I_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 here we have

$$\mathbf{I} = [f_n(Z^n(1)), f_n(Z^n(2)), \ldots, f_n(Z^n(B))] =: f(\mathbf{Z}).$$

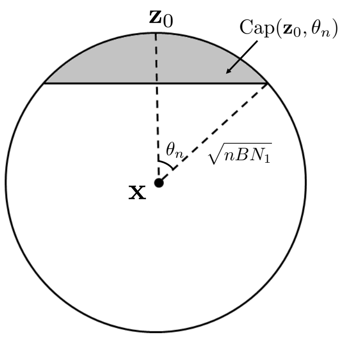


Fig. 4. A spherical cap with angle θ_n.

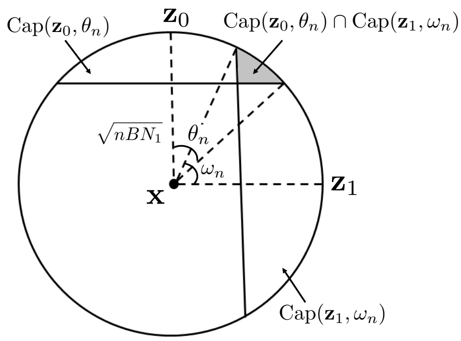


Fig. 5. Intersection of two spherical caps.

Now with a standard typicality argument, it can be shown that if $H(I_n|X^n)$ is fixed to be $-n\log\sin\theta_n$, then as B gets large for any typical $(\mathbf{x}, \mathbf{i})$ pair, we have

$$p(\mathbf{i}|\mathbf{x}) = \Pr(f(\mathbf{Z}) = \mathbf{i}|\mathbf{x}) \doteq 2^{nB\log\sin\theta_n}. \tag{15}$$

Since given $\mathbf{x}$, typical $\mathbf{z}$ sequences will be approximately uniformly distributed on an ϵ-thin spherical shell centered at $\mathbf{x}$ and of radius $\sqrt{nBN_1}$, denoted as

$$\begin{aligned}&\text{Shell}\left(\mathbf{x}, \sqrt{nB(N_1 \pm \epsilon)}\right)\\ &:= \left\{\mathbf{a} \in \mathbb{R}^{nB} : \|\mathbf{a}-\mathbf{x}\| \in \left[\sqrt{nB(N_1-\epsilon)}, \sqrt{nB(N_1+\epsilon)}\right]\right\},\end{aligned}$$

the relation (15) can be used to argue that the set of $\mathbf{z}$'s jointly typical with $\mathbf{x}$ that are mapped to the given $\mathbf{i}$, denoted by

$$A_{\mathbf{x}}(\mathbf{i}) = \left\{\mathbf{z} \in \text{Shell}\left(\mathbf{x}, \sqrt{nB(N_1 \pm \epsilon)}\right) : f(\mathbf{z}) = \mathbf{i}\right\},$$

will occupy a volume

$$|A_{\mathbf{x}}(\mathbf{i})| \doteq 2^{nB\left(\frac{1}{2}\log 2\pi e N_1 \sin^2\theta_n\right)}, \tag{16}$$

on this thin shell.

Assume now that the set $A_{\mathbf{x}}(\mathbf{i})$ were a spherical cap as illustrated in Fig. 4. In general, a spherical cap on $\text{Shell}\left(\mathbf{x}, \sqrt{nB(N_1 \pm \epsilon)}\right)$ can be defined as a ball in terms of the angle, i.e.,

$$\text{Cap}(\mathbf{z}_0, \phi) = \left\{\mathbf{z} \in \text{Shell}\left(\mathbf{x}, \sqrt{nB(N_1 \pm \epsilon)}\right) : \angle(\mathbf{z}_0, \mathbf{z}) \leq \phi\right\}$$

where we will refer to $\mathbf{z}_0$ as the pole and ϕ as the angle of the cap. Using the volume formula for the hyperspherical cap and characterizing the exponent of such a volume (c.f. [6, Appx. A]), it can be shown that the volume in (16) would correspond to an angle of θ_n for the spherical cap. Now, a straightforward computation would yield the following result: Let $V_n = |\text{Cap}(\mathbf{z}_0, \theta_n) \cap \text{Cap}(\mathbf{z}_1, \omega_n)|$ where $\angle(\mathbf{z}_0, \mathbf{z}_1) = \pi/2$ and $\theta_n + \omega_n > \pi/2$. Then,

$$\Pr\left(|A_{\mathbf{x}}(\mathbf{i}) \cap \text{Cap}(\mathbf{Z}, \omega_n)| \dot{\geq} V_n \middle| \mathbf{x}\right) \to 1 \text{ as } B \to \infty. \tag{17}$$

In words, if we take a $\mathbf{z}$ uniformly at random on the shell and draw a spherical cap centered at $\mathbf{z}$ with angle $\omega_n > \pi/2 - \theta_n$, then with high probability the intersection volume of this cap with the cap $A_{\mathbf{x}}(\mathbf{i})$ will be approximately lower bounded by V_n. This statement follows from the fact that in high dimensions most of the volume of the shell is concentrated around any equator, and in particular the equator at angle $\pi/2$ from the pole of $A_{\mathbf{x}}(\mathbf{i})$. Therefore, as the dimension nB gets large, for almost all $\mathbf{z}$'s, the intersection volume of the two spherical caps will be approximately given by V_n (see Fig. 5), which can be shown to be

$$V_n \doteq 2^{nB\left(\frac{1}{2}\log 2\pi e N_1(\sin^2\theta_n - \cos^2\omega_n)\right)},$$

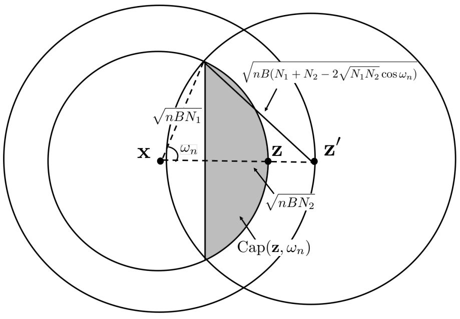
by using the volume formula for the intersection of two hyperspherical caps and characterizing the exponent of this volume (c.f. [6, Appx. B]). One of the main technical steps we develop in [5] (see also [6]) is to show that the statement (17) holds for any arbitrary set $A_{\mathbf{x}}(\mathbf{i})$ with volume given in (16), not only when $A_{\mathbf{x}}(\mathbf{i})$ is a spherical cap as we assumed above. Note that this is proved by using the Riesz rearrangement inequality to strengthen the isoperimetric inequality on the sphere and show that the spherical cap is the extremal set, not only for minimizing the volume of its neighborhood as done by the classical isoperimetric inequality, but also the extremal set when one is interested in minimizing the total intersection volume at given distance.

Next we translate the probabilistic statement (17) with respect to $\mathbf{Z}$ to that with respect to $\mathbf{Y}$. Particularly, for any $\mathbf{Z}$ let $\mathbf{Z}' = \sqrt{N_2/N_1}\mathbf{Z}$. Since given $\mathbf{x}$ the typical $\mathbf{z}$'s lie on a thin shell centered at $\mathbf{x}$ and of radius $\sqrt{nBN_1}$, the scaled version $\mathbf{z}'$ then typically lie on the thin shell with the same center and of radius $\sqrt{nBN_2}$. Furthermore, using the cosine formula it can be shown that for a typical $\mathbf{z}$, the spherical $\text{Cap}(\mathbf{z}, \omega_n)$ is contained by the Euclidean ball centered at $\mathbf{z}'$ and of radius $l_n = \sqrt{nB(N_1 + N_2 - 2\sqrt{N_1N_2}\cos\omega_n)}$. See Fig. 6. This combined with (17) immediately yields that

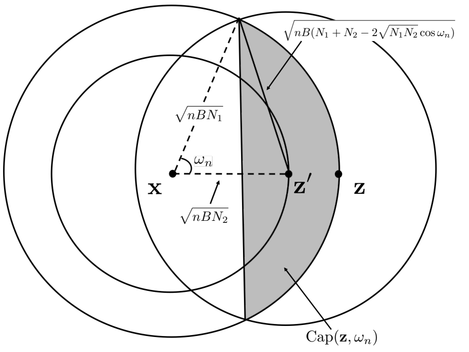
$$\Pr\left(|A_{\mathbf{x}}(\mathbf{i}) \cap \text{Ball}(\mathbf{Z}', l_n)| \dot{\geq} V_n \middle| \mathbf{x}\right) \to 1 \text{ as } B \to \infty.$$

Since $\mathbf{Z}'$ and $\mathbf{Y}$ have the same distribution given $\mathbf{x}$, the above relation still holds with $\mathbf{Z}'$ replaced by $\mathbf{Y}$.

Now noting that the above statement holds for any typical $(\mathbf{x}, \mathbf{i})$ pair, we can eliminate the conditioning with respect to $\mathbf{x}$ and reach the following conclusion regarding the random vectors $(\mathbf{I}, \mathbf{Y})$ with high probability: if we take $\mathbf{Y}$ and draw a Euclidean ball of radius l_n around it, the volume of the

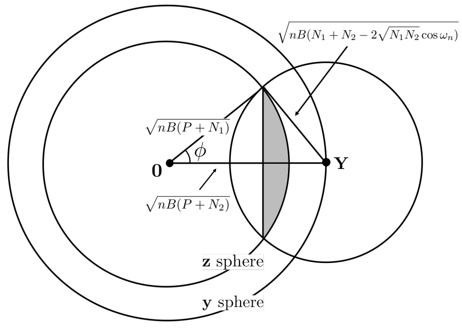


(a) $N_1 \leq N_2$

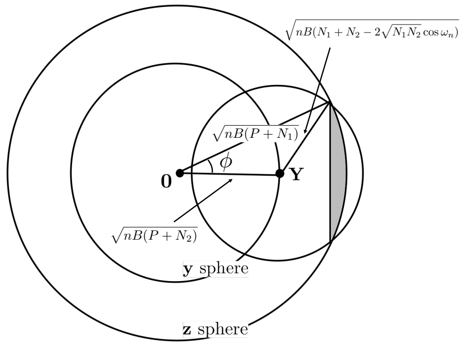


(b) $N_1 > N_2$

Fig. 6. Euclidean ball contains the cap.



(a) $N_1 \leq N_2$



(b) $N_1 > N_2$

Fig. 7. The cap on $\mathbf{z}$ sphere.

intersection of the set $A(\mathbf{I})$ with this ball is lower bounded by $|A(\mathbf{I}) \cap \mathrm{Ball}\,(\mathbf{Y}, l_n)| \dot{\geq} V_n$, where $A(\mathbf{I})$ is defined as $A(\mathbf{I}) = \{\mathbf{z} \in \mathbb{R}^{nB} : f(\mathbf{z}) = \mathbf{I}\}$. This puts an upper limit on the number of possible values of $\mathbf{I}$ given $\mathbf{Y}$. To get a tighter bound, we can incorporate the fact that typical $\mathbf{x}$'s lie on a thin shell centered at $\mathbf{0}$ of radius $\sqrt{nBP}$, and typical $\mathbf{z}$'s and $\mathbf{y}$'s lie on shells of radii $\sqrt{nB(P+N_1)}$ and $\sqrt{nB(P+N_2)}$, respectively. Therefore the number of possible values for $\mathbf{I}$ given $\mathbf{Y}$ can be bounded by the ratio of the spherical cap volume

$$\left|\mathrm{Shell}\left(\mathbf{0}, \sqrt{nB(P+N_1 \pm \epsilon)}\right) \cap \mathrm{Ball}\left(\sqrt{nB(P+N_2)}\mathbf{e}, l_n\right)\right|,$$

where $\mathbf{e}$ is any arbitrary unit vector, to the volume each possible $\mathbf{i}$ occupies from this cap, i.e. V_n. See Figure 7. To calculate the cap volume, i.e. the volume of the shaded area in Fig. 7, we use the cosine formula to conclude that

$$\sin^2\phi = \frac{P(N_1+N_2) + N_1N_2\sin^2\omega_n - 2P\sqrt{N_1N_2}\cos\omega_n}{(P+N_1)(P+N_2)}$$

where ϕ is the angle of the cap, and therefore the cap volume can be shown to be

$$\begin{aligned} &\dot{\leq} 2^{nB\left(\frac{1}{2}\log 2\pi e(P+N_1)\sin^2\phi\right)} \\ &= 2^{nB\left(\frac{1}{2}\log 2\pi e\frac{P(N_1+N_2)+N_1N_2\sin^2\omega_n - 2P\sqrt{N_1N_2}\cos\omega_n}{P+N_2}\right)}, \end{aligned}$$

and its ratio to V_n is

$$\dot{\leq} 2^{nB\left(\frac{1}{2}\log\frac{P(N_1+N_2)+N_1N_2\sin^2\omega_n - 2P\sqrt{N_1N_2}\cos\omega_n}{N_1(P+N_2)(\sin^2\theta_n - \cos^2\omega_n)}\right)}.$$

This in turn imposes the following bound on $H(I_n|Y^n)$:

$$n\left(\frac{1}{2}\log\frac{P(N_1+N_2)+N_1N_2\sin^2\omega_n - 2P\sqrt{N_1N_2}\cos\omega_n}{N_1(P+N_2)(\sin^2\theta_n - \cos^2\omega_n)}\right).$$

The upper bound (5) in Lemma 2.1 then follows by noting that the above argument holds for any $\omega_n > \pi/2 - \theta_n$.

References

[1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[2] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, 1979.

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

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

[5] X. Wu, L. Barnes, A. Ozgur, "Cover's open problem: "The capacity of the relay channel"," *Proc. of 54th Annu. Allerton Conf. Commun., Control, Comput.*, Illinois, 2016.

[6] X. Wu, L. Barnes, A. Ozgur, "Cover's open problem: "The capacity of the relay channel"," submitted to *IEEE Trans. Inform. Theory*. Available: https://arxiv.org/abs/1701.02043

[7] X. Wu and A. Ozgur, "Improving on the cut-set bound via geometric analysis of typical sets," in *Proc. of 2016 International Zurich Seminar on Communications*.

[8] X. Wu, A. Ozgur, L.-L. Xie, "Improving on the cut-set bound via geometric analysis of typical sets," accepted to *IEEE Trans. Inform. Theory*. Available: http://arxiv.org/abs/1602.08540

[9] X. Wu and A. Ozgur, "Cut-set bound is loose for Gaussian relay networks," in *Proc. of 53rd Annu. Allerton Conf. Commun., Control, Comput.*, Illinois, Sept. 29–Oct. 1, 2015.

[10] X. Wu and A. Ozgur, "Cut-set bound is loose for Gaussian relay networks," submitted to *IEEE Trans. Inform. Theory*. Available: http://arxiv.org/abs/1606.01374

[11] X. Wu and A. Ozgur, "Improving on the cut-set bound for general primitive relay channels," *Proc. of IEEE Int. Symp. Inf. Theory*, 2016.