Fifty-Fifth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 3-6, 2017

A Solution to Cover's Problem for the Binary Symmetric Relay Channel: Geometry of Sets on the Hamming Sphere

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

***Abstract*—We solve a long-standing open problem posed by Cover and named "The Capacity of the Relay Channel," *Open Problems in Communication and Computation, Springer-Verlag, 1987*, in the case when the channels from the source to the relay and the destination are binary symmetric channels. Similar to our recent solution of this problem in the Gaussian case, we solve this problem by connecting it to high-dimensional geometry. However, our geometric approach in the binary case significantly deviates from the Gaussian case, since our treatment of the Gaussian case relied on an extension of the classical isoperimetric inequality on the Euclidean sphere, the counterpart of which does not exist on the Hamming sphere. Instead, we prove a Riesz rearrangement type inequality on the Hamming sphere, which allows us to develop a new upper bound on the capacity of the binary symmetric relay channel. Our argument (and consequently our upper bound) for the binary case is weaker than the one we obtained in the Gaussian case, but nevertheless strong enough to resolve Cover's problem.**

I. Introduction

Characterizing the capacity of the relay channel is a long-standing problem in network information theory, which remains open despite decades of research efforts [1]–[3]. A seemingly less demanding question that is of interest is the one posed by Cover in *Open Problems in Communication and Computation*, Springer-Verlag, 1987 [4]. This question was called "The Capacity of the Relay Channel" by Cover, however as we describe in the following it actually corresponds to a subquestion of the general capacity problem.

Consider a relay channel as depicted in Fig. 1, where the source X wants to send information to the destination Y with the help of the relay Z, and the relay Z can communicate to the destination Y via an isolated bit pipe of capacity C_0. Let $C(C_0)$ denote the capacity of this channel as a function of C_0. Then obviously $C(C_0)$ is non-decreasing in C_0, and we have $C(0) = \max_{p(x)} I(X;Y)$ and $C(\infty) = \max_{p(x)} I(X;Y,Z)$. Now, assuming that 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 as to achieve the maximum possible rate $C(\infty)$ on this channel, corresponding to full cooperation between the relay and the destination.

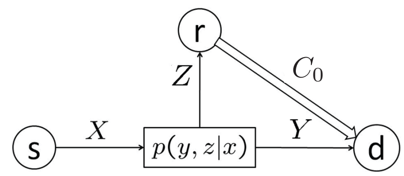

Fig. 1. A relay channel.

In our recent work [5]–[6], we answered this question in the Gaussian case and 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. Our proof was based on a geometric formulation of the problem and an application of Riesz' rearrangement inequality on the n-sphere due to Baernstein and Taylor [8] to obtain a strengthening of the classical isoperimetric inequality on the n-sphere.

In this paper, we further consider Cover's problem in the binary symmetric case, where both the X-Y and X-Z links are binary symmetric channels with some crossover probability p. We develop the analogous geometric formulation in this case, as well as a new discrete rearrangement inequality. Using these tools, we are able to resolve Cover's problem for the binary symmetric relay channel and show that C_0^* equals the Slepian-Wolf coding rate at which one can losslessly transfer the relay's observation to the destination.

II. Channel Model and Main Result

Consider the relay channel 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

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

This work was supported in part by NSF award CCF-1704624 and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370.

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 $C(C_0)$.

In the rest of the paper, we will focus on the binary symmetric case of the channel, where $p(y, z|x) = p(y|x)p(z|x)$ and

$$p(y|x) = \begin{cases} 1-p & , y = x \\ p & , y \neq x \end{cases}$$

$$p(z|x) = \begin{cases} 1-p & , z = x \\ p & , z \neq x \end{cases}$$

with X, Y and Z all taking values from $\mathcal{H} = \{0, 1\}$ and $0 < p < \frac{1}{2}$. It is easy to check that in this case

$$C(\infty) = \max_{p(x)} I(X; Y, Z) = 1 + H(p * p) - 2H(p),$$

where the maximum is attained with the uniform input distribution on $\mathcal{H}$,

$$H(r) = -r \log r - (1-r) \log(1-r), \forall r \in [0, 1]$$

is the binary entropy function, and

$$p_1 * p_2 := p_1(1-p_2) + p_2(1-p_1) \ .$$

Note that in order to achieve $C(\infty)$, a natural strategy is to use a simple Compress-and-Forward scheme with only Slepian-Wolf binning, i.e. the so-called Hash-and-Forward (H-F) [7], to faithfully transfer the relay's observation Z^n to the destination by treating Y^n as correlated side information so that the destination can decode the source message based on both Z^n and Y^n. Implementing this strategy requires a rate $C_0 = H(p * p)$, where $H(p * p)$ is the conditional entropy $H(Z|Y)$ when X is uniformly distributed on $\mathcal{H}$, and this leads to an upper bound on C_0^*, namely

$$C_0^* \leq H(p * p).$$

Note that this H-F upper bound converges to 0 as $p \to 0$; but interestingly it converges to 1 as $p \to 1/2$, even though $C(\infty)$ itself is diminishing in this regime.

On the converse side, we can apply the cut-set bound to obtain the following lower bound on C_0^*:

$$C_0^* \geq H(p * p) - H(p),$$

which significantly deviates from the H-F upper bound on C_0^*. See Fig. 2. Note that especially as $p \to 1/2$ while achievability requires a full bit of C_0 to support the diminishing $C(\infty)$ rate, the cut-set bound potentially allows to achieve the diminishing $C(\infty)$ rate at diminishing C_0. Fig. 2 also plots a previous lower bound we obtain for C_0^* in [9], which shows that C_0^* is strictly larger than 0 when $p \to 1/2$.

The main result of this paper is an exact characterization of C_0^*, which resolves the above dichotomy and shows that to achieve $C(\infty)$, it is necessary to losslessly transfer the relay's

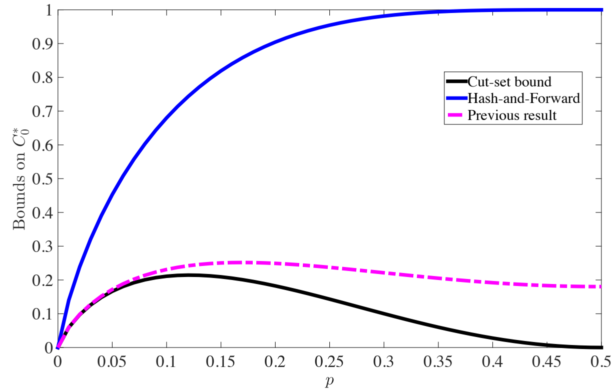

Fig. 2. Bounds on C_0^*.

observation to the destination at the Slepian-Wolf coding rate $H(p * p)$.

Theorem 2.1: $C_0^* = H(p * p)$

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 2.2: The capacity $C(C_0)$ satisfies

$$\begin{cases} C(C_0) \leq 1 - H(p) + C_0 - a & (2) \\ C(C_0) \leq 1 - H(p) + \min_{l \in [p*p-2rp, p*p]} h_a(l) - a & (3) \end{cases}$$

for some $a \in [0, H(p)]$, where $h_a(l)$ is given by

$$\begin{aligned} h_a(l) = &\ H(l) - 2rpH\left(\frac{l - p*p + 2rp}{4rp}\right) \\ & - 2r(1-p)H\left(\frac{l - p*p - 2rp + 2r}{4r(1-p)}\right) \\ & - H(p) + a + 2r \end{aligned} \quad (4)$$

and r is defined such that

$$H(p) - a = pH\left(\frac{r}{p}\right) + (1-p)H\left(\frac{r}{1-p}\right). \quad (5)$$

In Fig. 3 we plot this new bound under $p = 0.1$ together with the celebrated cut-set bound and the Compress-and-Forward rate. We also plot a previous upper bound we obtain in [9], which is actually tighter than the new bound in this paper for small C_0 but becomes loose when C_0 gets large. Note that from the figure one can visually observe that our new bound reaches the value $C(\infty)$ only when $C_0 \geq H(p*p)$, which leads to the conclusion in Theorem 1.1. This is formally proved in the next section.

III. Proofs

The proof of Theorem 2.2 follows from the below lemma, which is the main technical focus of this paper and whose proof is outlined in Section III-C. We now state this lemma and show how it leads to the bound in Theorem 2.2, which is then used to establish Theorem 2.1.

Lemma 3.1: Let I_n be a discrete random variable and X^n, Y^n and Z^n be n-length binary random vectors which form the Markov chain $I_n - Z^n - X^n - Y^n$, where Z^n and Y^n are independent Bernoulli vectors given X^n such that $X^n \oplus$

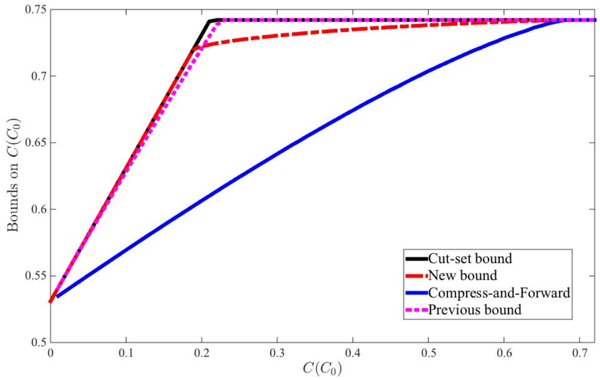

Fig. 3. Bounds on $C(C_0)$.

$Z^n, X^n \oplus Y^n \overset{\text{i.i.d.}}{\sim}$ Bern(p) and $I_n = f_n(Z^n)$ is a deterministic function of Z^n. Let $H(I_n|X^n)$ be denoted by na_n. Then we have

$$H(I_n|Y^n) \leq n \min_{l_n \in [p*p-2r_n p, p*p],} h_{a_n}(l_n),$$

where $h_{a_n}(l_n)$ and r_n are similarly defined as in (4) and (5) respectively.

A. Proof of Theorem 2.2

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

$$\begin{aligned} nR &\leq I(X^n; Y^n, I_n) + n\mu && (6)\\ &= I(X^n; Y^n) + I(X^n; I_n|Y^n) + n\mu \\ &= I(X^n; Y^n) + H(I_n|Y^n) - H(I_n|X^n) + n\mu && (7)\\ &\leq nI(X_Q; Y_Q) + H(I_n|Y^n) - H(I_n|X^n) + n\mu && (8)\\ &\leq n(1 - H(p)) + H(I_n|Y^n) - H(I_n|X^n) + n\mu, && (9) \end{aligned}$$

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 $I(X_Q; Y_Q)$ is upper bounded by $1 - H(p)$ with the maximum being attained with the uniform distribution of X_Q on $\mathcal{H}$.

Now we use Lemma 3.1 to upper bound the difference $H(I_n|Y^n) - H(I_n|X^n)$ in (9). It is trivial to observe 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, and therefore we can bound the difference of the two entropy terms in (9) and conclude that for any achievable rate R,

$$R \leq 1 - H(p) + \min_{l_n \in [p*p-2r_n p, p*p],} h_{a_n}(l_n) - a_n + \mu, \quad (10)$$

where $a_n = n^{-1} H(I_n|X^n)$.

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

$$R \leq 1 - H(p) + C_0 - a_n + \mu, \quad (11)$$

which simply follows from (9) by upper bounding $H(I_n|Y^n)$ with nC_0 and plugging in the definition of a_n. Combining (10) and (11) concludes the proof of the theorem.

B. Proof of Theorem 2.1

To see Theorem 2.2 implies Theorem 2.1, we will use bounds (2)–(3) to prove that $C(C_0) < C(\infty)$ if C_0 is strictly less than $H(p*p)$. First consider bound (2). If here $a = H(p)$, then clearly $C(C_0) < C(\infty)$ for $C_0 < H(p*p)$. Otherwise, if $a < H(p)$, then (3) will bound $C(C_0)$ away from $C(\infty)$. In particular, note that for any $a < H(p)$, the difference $h_a(l) - a$ in (3) satisfies

$$h_a(p*p) - a = H(p*p) - H(p) \quad (12)$$

and $h_a(l)$ is increasing at $p*p$, or more precisely,

$$h_a'(p*p) > 0.$$

Therefore, as long as $a < H(p)$, the minimization of $h_a(l) - a$ with respect to l in (3) yields a value strictly smaller than the R.H.S. of (12), and thus $C(C_0)$ is also strictly less than $C(\infty)$ if $a < H(p)$. This proves Theorem 2.1.

C. Proof Outline for Lemma 3.1

In this section we provide a proof outline for Lemma 3.1. At a high-level, this lemma tries to mimic the corresponding lemma we derived for the Gaussian case [6, Lemma 2.1]. However, our key technical argument in this paper is different from the Gaussian case. The key step in proving Lemma 2.1 in [6] is to prove a generalization of the classical isoperimetric inequality on the high-dimensional Euclidean sphere by using Riesz' rearrangement inequality. The classical isoperimetric inequality on the Euclidean sphere states that among all sets on the sphere with a given area the spherical cap has the smallest boundary or more generally the smallest neighborhood. Our generalization in [6] shows that the spherical cap is the extremal set not only in terms of minimizing the area of its neighborhood, but roughly speaking also in terms of minimizing its total intersection volume with the neighborhood of a randomly chosen point on the sphere (where the ω-neighborhood of a point on the sphere is defined as the set of all points of geodesic distance at most ω to this point). When trying to extend this argument to the binary case, we face a major difficulty; there is no counterpart for the classical isoperimetric inequality on the Hamming sphere. In particular, it is known that the set with smallest boundary on the Hamming sphere may not be a spherical cap [10], [12] and no general characterizations are known despite this being a natural question in discrete geometry of significant interest since [10], [12]. We circumvent this problem by proving a discrete version of the Riesz rearrangement inequality on the Hamming sphere. We detail the contribution of our work from a discrete geometric perspective after we outline the proof of the lemma in the following.

Recall our goal in Lemma 3.1 is to bound $H(I_n|Y^n)$ 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 i.i.d. Bernoulli vectors given X^n, and $I_n = f_n(Z^n)$ is a deterministic mapping of Z^n to a set of integers. Similarly to the Gaussian case, one can verify that if $H(I_n|X^n) = 0$ then $H(I_n|Y^n) = 0$ as follows: $H(I_n|X^n) =$

0 implies that given the transmitted codeword X^n, there is no ambiguity about I_n, or equivalently all Z^n sequences jointly typical with X^n are mapped to the same I_n. However, since Y^n and Z^n are statistically equivalent given X^n, they share the same typical set given X^n and since this typical set is mapped to a single I_n value, this implies that I_n can be also determined based on Y^n and therefore $H(I_n|Y^n) = 0$. Based on this intuition, the goal therefore is to prove an upper bound on $H(I_n|Y^n)$ for a given value of $H(I_n|X^n)$.

Following a similar line of thought, if $H(I_n|X^n)$ is fixed to a certain non-zero value, say $H(I_n|X^n) = na_n$ with $a_n \in (0, H(p)]$, this roughly speaking implies that the typical Z^n's surrounding an X^n are now mapped to multiple I_n values. This argument can be made precise as follows: Consider the following B-length i.i.d. sequence

$$\{(X^n(b), Y^n(b), Z^n(b), I_n(b))\}_{b=1}^{B}, \tag{13}$$

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, 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

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

Now we can apply a standard typicality argument to say that for any typical $(\mathbf{x}, \mathbf{i})$ pair,[1]

$$p(\mathbf{i}|\mathbf{x}) = P(f(\mathbf{Z}) = \mathbf{i}|\mathbf{X} = \mathbf{x}) \doteq 2^{-nBa_n}. \tag{14}$$

This probabilistic statement can be translated into the following geometric picture: Given $\mathbf{x}$, typical $\mathbf{y}$ and $\mathbf{z}$ sequences will be approximately uniformly distributed on an ϵ-thin Hamming spherical shell centered at $\mathbf{x}$ and of radius np, defined as

$$\begin{aligned}&\text{Shell}\,(\mathbf{x}, nB(p-\epsilon), nB(p+\epsilon))\\ &:= \left\{\mathbf{a} \in \mathcal{H}^{nB} : d(\mathbf{a}, \mathbf{x}) \in [nB(p-\epsilon), nB(p+\epsilon)]\right\}\end{aligned}$$

where $\epsilon \to 0$ as $B \to \infty$. The relation (14) can then be used to argue that the set of $\mathbf{z}$'s jointly typical with $\mathbf{x}$ that are mapped to a given $\mathbf{i}$, denoted by

$$A_{\mathbf{x}}(\mathbf{i}) = \{\mathbf{z} \in \text{Shell}\,(\mathbf{x}, nB(p-\epsilon), nB(p+\epsilon)) : f(\mathbf{z}) = \mathbf{i}\},$$

has cardinality

$$|A_{\mathbf{x}}(\mathbf{i})| \doteq 2^{nB[H(p)-a_n]} \tag{15}$$

on this thin shell. This translation between probabilities and cardinalities of sets is immediate since $\mathbf{y}$ and $\mathbf{z}$ are distributed approximately uniformly on the shell.

To upper bound $H(I_n|Y^n)$, we will apply a simple packing argument similar to the one we developed in the Gaussian case [5]–[6]. In particular, consider a typical $\mathbf{y}$ and draw a Hamming ball around it with certain radius nBl_n, denoted by

$$\text{Ball}(\mathbf{y}, nBl_n) := \left\{\mathbf{a} \in \mathcal{H}^{nB} : d(\mathbf{a}, \mathbf{y}) \le nBl_n\right\}$$

where the normalized radius $l_n \in [0, 1]$. Obviously, each sequence in this ball is mapped to an $\mathbf{i}$ value by f. Let $A(\mathbf{i})$ denote the set of binary sequences that are mapped to a given value $\mathbf{i}$, i.e. $A(\mathbf{i}) = \left\{\mathbf{z} \in \mathcal{H}^{nB} : f(\mathbf{z}) = \mathbf{i}\right\}$. Then $|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})|$ denotes the number of sequences in $\text{Ball}(\mathbf{y}, nBl_n)$ that are mapped to the value $\mathbf{i}$. If we can characterize $|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})|$ for each typical $\mathbf{i}$, or less ambitiously find a lower bound on $|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})|$ that holds for any typical $\mathbf{i}$, then the total number of possible $\mathbf{i}$'s given $\mathbf{y}$ can be upper bounded by the ratio

$$\frac{|\text{Ball}(\mathbf{y}, nBl_n)|}{|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})|}.$$

This implies the following bound on the conditional entropy of $\mathbf{I}$ given $\mathbf{Y}$,

$$H(\mathbf{I}|\mathbf{Y}) \le \log \frac{|\text{Ball}(\mathbf{y}, nBl_n)|}{|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})|}, \tag{16}$$

which, divided by B, gives a bound on $H(I_n|Y^n)$ since $H(\mathbf{I}|\mathbf{Y}) = B\,H(I_n|Y^n)$ due to the fact that $(\mathbf{I}, \mathbf{Y})$ are generated i.i.d. from the distribution of (I_n, Y^n).

The remaining task is then to lower bound $|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})|$ for a typical $(\mathbf{y}, \mathbf{i})$ pair. To this end, note that

$$|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})| \ge |\text{Ball}(\mathbf{y}, nBl_n) \cap A_{\mathbf{x}}(\mathbf{i})|$$

for any $\mathbf{x}$ that is jointly typical with $(\mathbf{y}, \mathbf{i})$, since by definition $A_{\mathbf{x}}(\mathbf{i}) \subseteq A(\mathbf{i})$. As we have previously argued in (15), $|A_{\mathbf{x}}(\mathbf{i})| \doteq 2^{nB[H(p)-a_n]}$ for any typical $(\mathbf{x}, \mathbf{i})$ pair. Moreover, any $\mathbf{y}$ that is typical with this $(\mathbf{x}, \mathbf{i})$ pair will be uniformly distributed on $\text{Shell}\,(\mathbf{x}, nB(p-\epsilon), nB(p+\epsilon))$. Therefore the problem of characterizing $|\text{Ball}(\mathbf{y}, nBl_n) \cap A_{\mathbf{x}}(\mathbf{i})|$ for set of jointly typical $\mathbf{x}, \mathbf{y}, \mathbf{i}$ triple reduces to the following geometric problem: characterize the intersection of the set $A_{\mathbf{x}}(\mathbf{i}) \subseteq \text{Shell}\,(\mathbf{x}, nB(p-\epsilon), nB(p+\epsilon))$ of cardinality $|A_{\mathbf{x}}(\mathbf{i})| \doteq 2^{nB[H(p)-a_n]}$ with the nBl_n-neighborhood of a randomly and uniformly chosen point on the shell. Note that this intersection in general depends on the shape of $A_{\mathbf{x}}(\mathbf{i})$, which is dictated by the relay mapping and can be arbitrary. We therefore need to find a lower bound on $|\text{Ball}(\mathbf{y}, nBl_n) \cap A_{\mathbf{x}}(\mathbf{i})|$ that holds for any subset $A_{\mathbf{x}}(\mathbf{i})$ of the shell with given cardinality. In [5]–[6], we achieve this by showing that the intersection is minimal when $A_{\mathbf{x}}(\mathbf{i})$ is a spherical cap (by developing a generalization of the classical isoperimetrical inequality on the Euclidean sphere) and computing the intersection when $A_{\mathbf{x}}(\mathbf{i})$ is a spherical cap. As mentioned earlier, there is no counterpart of the classical isoperimetric inequality on the Euclidean sphere in Hammming space. Instead, we lower bound the cardinality of the intersection by showing that it would be minimal if $A_{\mathbf{x}}(\mathbf{i})$ were a *stable* set. These stable sets take the place of the spherically symmetrized subsets that appear in the Euclidean setting, i.e. spherical caps, but are more general than a spherical cap on the

[1]For the purpose of this outline, the term "typical" sequence or a pair of "typical" sequences can be understood to mean "almost all" sequences or pair of sequences that one can observe from the prescribed distribution. Once a certain geometric property is established for a set of sequences with probability approaching one, it is easy to formally argue that we can restrict our attention to this high probability set and ignore remaining "atypical" sequences.

Hamming sphere. (For example, stable sets on the Hamming sphere include lower dimensional spheres and spherical caps as the two intuitively extremal special cases.) To show this we develop a certain symmetrization argument which gives rise to a discrete version of the Riesz rearrangement inequality on the Hamming sphere. However, since a *stable* set is a general concept, much less specific than a spherical cap or a lower dimensional sphere, we cannot simply compute the intersection $|\text{Ball}(\mathbf{y}, nBl_n) \cap A_{\mathbf{x}}(\mathbf{i})|$ by assuming $A_{\mathbf{x}}(\mathbf{i})$ is stable. Instead we develop a (possibly loose) lower bound on this intersection by arguing that each stable set of size $2^{nB[H(p)-a_n]}$ needs to contain a lower dimensional subsphere of a certain size and the intersection can be lower bounded by stacking the points in $A_{\mathbf{x}}(\mathbf{i})$ on this lower dimensional sphere. This allows to obtain the following lower bounded on this intersection

$$|\text{Ball}(\mathbf{y}, nBl_n) \cap A_{\mathbf{x}}(\mathbf{i})| \dot{\geq} 2^{nBV_n},$$

where

$$\begin{aligned} V_n := & \, 2r_n pH\left(\frac{l_n - p*p + 2r_np}{4r_np}\right) \\ & + 2r_n(1-p)H\left(\frac{l_n - p*p - 2r_np + 2r_n}{4r_n(1-p)}\right) \\ & + H(p) - a_n - 2r_n \end{aligned} \tag{17}$$

for the following range of l_n:

$$l_n \in [p*p - 2r_np, p*p], \tag{18}$$

where r_n in both (17) and (18) is defined such that

$$H(p) - a_n = pH\left(\frac{r_n}{p}\right) + (1-p)H\left(\frac{r_n}{1-p}\right). \tag{19}$$

This result is stated and proved in the next section. Because of this, one can conclude that for a typical $(\mathbf{y}, \mathbf{i})$ pair and l_n satisfying (18),

$$|\text{Ball}(\mathbf{y}, nBl_n) \cap A(\mathbf{i})| \dot{\geq} 2^{nBV_n},$$

which, when plugged into (16), immediately yields the desired bound on $H(I_n|Y^n)$.

IV. Discrete Geometry

Consider the shell

$$\text{Shell}\,(\mathbf{x}, nB(p-\epsilon), nB(p+\epsilon))$$

where $\epsilon \to 0$ as $nB \to \infty$, and some arbitrary subset A of the shell with $|A| = 2^{nB[H(p)-a_n]}$. Our goal in this section is to show that for any $\delta > 0$ and nB sufficiently large,

$$\Pr\left(|A \cap \text{Ball}(\mathbf{Y}, nB(l_n+\delta))| \geq 2^{nB(V_n-\delta)}|\mathbf{x}\right) \geq 1-\delta\,.$$

Interestingly, the above can be shown by restricting our attention to a single sphere that is a subset of $\text{Shell}\,(\mathbf{x}, nB(p-\epsilon), nB(p+\epsilon))$. In particular, note that there are only polynomially many spheres that make up this shell, and therefore there must exist some p_{nB} with $p_{nB} \to p$ and

$$|A \cap \text{Sphere}(\mathbf{x}, nBp_{nB})| \geq 2^{nB[H(p)-a_n-\epsilon_1]}$$

for some $\epsilon_1 \to 0$ as $nB \to \infty$. Define

$$\Gamma_\delta(V) = \left\{\mathbf{a} \in \mathcal{H}^{nB} : \min_{\mathbf{b} \in V} d(\mathbf{a}, \mathbf{b}) \leq \delta\right\}$$

as the δ-neighborhood of a set $V \subseteq \mathcal{H}^{nB}$ and let $A_s = A \cap \text{Sphere}(\mathbf{x}, nBp_{nB})$. By the triangle inequality,

$$\Gamma_{nB\delta}\left(\left\{\mathbf{y} : |A_s \cap \text{Ball}(\mathbf{y}, nBl_n)| \geq 2^{nB(V_n-\delta)}\right\}\right) \tag{20}$$

$$\subseteq \left\{\mathbf{y} : |A_s \cap \text{Ball}\,(\mathbf{y}, nB(l_n+\delta))\,| \geq 2^{nB(V_n-\delta)}\right\}\,. \tag{21}$$

If the intersection size in (20) is greater than or equal the threshold $2^{nB(V_n-\delta)}$ for most $\mathbf{y} \in \text{Sphere}(\mathbf{x}, nBp_{nB})$, then using the blowing-up lemma [13], (20) can be made to have probability approaching one conditionally on $\mathbf{x}$, where $\delta \to 0$ as $nB \to \infty$. Thus (21) will also have conditional probability approaching one.

With this in mind we will make the following notational simplifications. First, we will state the main lemma of this section for general $n \to \infty$ with the understanding that there is no difficulty in applying it to dimension $nB \to \infty$. Second, we will restrict our attention to the Hamming sphere $S \subset \mathcal{H}^n$ of radius np, or

$$S = \{y \in \mathcal{H}^n \mid d(0, y) = np\}\,.$$

Formally, we are actually considering a sequence of Hamming spheres with radius np_n (which must be an integer) and $p_n \to p$ as $n \to \infty$. Third, we define $B_l(y)$ to be the spherical cap centered at $y \in S$ with normalized radius l i.e.

$$B_l(y) = \{y' \in S \mid d(y, y') \leq nl\}\,.$$

The main lemma of this section is as follows:

Lemma 4.1: Suppose y is a random variable with uniform distribution over S, $A \subseteq S$ with $|A| = 2^{n[H(p)-a_n]}$, and $2p(1-p) - 2r_np \leq l_n \leq 2p(1-p)$. Then

$$\Pr\left(|A \cap B_{l_n}(y)| \geq 2^{n(V_n-\delta)}\right) \to 1$$

as $n \to \infty$ where V_n is defined in (17), r_n is defined by (19), and $\delta \to 0$ as $n \to \infty$.

We will first introduce a notion of rearrangement in this discrete setting. To accomplish this we use a generalization of the two-point symmetrization process seen in [8]. This discrete version is also presented in [10] and [12] while investigating isoperimetric inequalities. Let $\sigma_{ij} : S \to S$ be the transposition that switches elements i and j so that for $y = (\mathrm{y}_1, \ldots, \mathrm{y}_n)$,

$$\sigma_{ij}(\ldots, \mathrm{y}_i, \ldots, \mathrm{y}_j, \ldots) = (\ldots, \mathrm{y}_j, \ldots, \mathrm{y}_i, \ldots)\,.$$

For an arbitrary subset A of S, define the positive section, negative section, and zero section of A associated with i, j as follows:

$$A_{ij}^+ = \{y \in A \mid \mathrm{y}_i = 1, \mathrm{y}_j = 0\}$$

$$A_{ij}^- = \{y \in A \mid \mathrm{y}_i = 0, \mathrm{y}_j = 1\}$$

$$A_{ij}^0 = \{y \in A \mid \mathrm{y}_i = \mathrm{y}_j\}$$

The i,j symmetrization of A is then

$$A_{ij}^* = A_{ij}^0 \cup (A_{ij}^+ \cup \sigma_{ij}(A_{ij}^-)) \cup (A_{ij}^- \cap \sigma_{ij}(A_{ij}^+)) \ . \quad (22)$$

The first thing to note about A_{ij}^* is that $|A| = |A_{ij}^*|$. This follows because any element y of A which is excluded from the negative section because of the intersection in the final term of (22) will have its counterpart $\sigma_{ij}(y)$ included in positive section because of the union in the middle term.

For $i < j$ consider the repeated application of this symmetrization operation to the subset A. After finitely many applications for each $i < j$, eventually we will arrive at a subset A^* such that any further i,j symmetrization will not change the subset. Call such a subset A^* a *stable* subset. These stable subsets will take the place of the spherically symmetrized subsets that appear in the Euclidean setting.

Lemma 4.2: For any normalized radius l and subsets $A, C \subset S$,

$$\sum_{x,y \in S} 1_A(x) 1_C(y) 1_{B_l(y)}(x) \leq \sum_{x,y \in S} 1_{A^*}(x) 1_{C^*}(y) 1_{B_l(y)}(x) \ , \quad (23)$$

where A^*, C^* are stable sets constructed via the symmetrization process above from A, C respectively.

Proof: Define the functional

$$I(A,C) = \sum_{x,y \in S} 1_A(x) 1_C(y) 1_{B_l(y)}(x) \ .$$

The claim is that

$$I(A,C) \leq I(A_{ij}^*, C_{ij}^*)$$

for any $i < j$ and therefore by repeated application of symmetrization we get the desired result.

Now to check this claim. Any point $y \in C_{ij}^+ \cup C_{ij}^0$ will also appear in C_{ij}^* and cannot have fewer elements of A_{ij}^* within distance nl than it had elements of A. This is because any elements that were removed during symmetrization were replaced by reflected versions that are no further away from y.

Now consider any point $y \in C_{ij}^-$. Suppose $y \in C_{ij}^*$ which implies $\sigma_{ij}(y)$ was already in C. Any points x which were moved by symmetrization to outside of $B_l(y)$ will be moved to inside $B_l(\sigma_{ij}(y))$ and the sum will remain unchanged. The last case is if $y \notin C_{ij}^*$ which implies $\sigma_{ij}(y) \in C_{ij}^*$. Any element x which was in $B_l(y)$ but is not in $B_l(\sigma_{ij}(y))$ must have a counterpart $\sigma_{ij}(x) \in B_l(\sigma_{ij}(y))$ which will therefore be incorporated in the sum. ■

For any function $f : S \to \mathbb{R}_{\geq 0}$ let y_i be an ordering of all $y \in S$ such that $f(y_1) \leq f(y_2) \leq \ldots \leq f(y_{|S|})$. Write

$$f(y) = f(y_1) + \sum_{i=2}^{|S|} \left(f(y_i) - f(y_{i-1})\right) 1_{\{f(y) \geq f(y_i)\}}(y) \ .$$

Let $g : S \to \mathbb{R}_{\geq 0}$ be similarly written as

$$g(y) = g(y_1') + \sum_{i=2}^{|S|} \left(g(y_i') - g(y_{i-1}')\right) 1_{\{g(y) \geq g(y_i')\}}(y)$$

where $g(y_1') \leq \ldots \leq g(y_{|S|}')$. Next, let

$$f^*(y) = f(y_1) + \sum_{i=2}^{|S|} \left(f(y_i) - f(y_{i-1})\right) 1_{\{f(y) \geq f(y_i)\}^*}(y)$$

with the corresponding definition for g^*.

Corollary 4.1: For any functions $f, g : S \to \mathbb{R}_{\geq 0}$ and nonincreasing function $K : \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$,

$$\sum_{x,y \in S} f(x) g(y) K(d(x,y)) \leq \sum_{x,y \in S} f^*(x) g^*(y) K(d(x,y)) \ . \quad (24)$$

Proof: Writing

$$K(n) = \sum_{i=1}^{\infty} (K(i-1) - K(i)) 1_{\{n < i\}}(n) \ ,$$

each term in the summations in (24) can be written as a product of three summations. Using Lemma 2 the inequality holds for each term in that product. ■

Lemma 4.2 is a discrete version of the Riesz rearrangement inequality seen in [8]. This is even more explicit the way it is written in Corollary 4.1. In fact, an alternative way to define f^* and g^* is via a discrete analogue of the polarization process seen in [8] and [11]. Using this next lemma, and with some further work, we will be able to use this rearrangement inequality in a similar way as the continuous version was used in [5]–[6].

Define

$$\psi(y) = \sum_{x \in S} 1_A(x) 1_{B_l(y)}(x) = |A \cap B_l(y)|$$

and

$$\psi'(y) = \sum_{x \in S} 1_{A^*}(x) 1_{B_l(y)}(x) = |A^* \cap B_l(y)| \ .$$

We can rewrite (23) by choosing C to correspond to the k elements of S that give the k largest values of $\psi(y)$. In this case we get

$$\begin{aligned} \max_{y_1,\ldots,y_k} \sum_{j=1}^{k} \psi(y_j) &\leq \sum_{y \in S} 1_{C^*}(y) \psi'(y) \\ &\leq \max_{y_1,\ldots,y_k} \sum_{j=1}^{k} \psi'(y_j) \end{aligned} \quad (25)$$

where the maxima also require $y_i \neq y_m$ for $i \neq m$.

Lemma 4.3: Consider y to be a random variable with uniform distribution over S. Given arbitrary $A \subset S$, suppose

$$\Pr(|A^* \cap B_l(y)| \geq v_n) \to 1$$

as $n \to \infty$. Then

$$\Pr\left(|A \cap B_l(y)| \geq \frac{v_n}{2}\right) \to 1 \ .$$

Proof: Let y_i be an ordering across all $y \in S$ such that $\psi_i = \psi(y_i)$, $\psi_1 \geq \psi_2 \geq \ldots \geq \psi_{|S|}$ and let ψ_j' be ordered in the corresponding way.

Let ξ_n be the largest integer such that $\psi_{\xi_n} \geq \frac{v_n}{2}$ and let ξ'_n be the largest integer such that $\psi'_{\xi'_n} \geq v_n$. Any value n such that $\xi_n \geq \xi'_n$ does not pose any difficulty. Let n_k be the subsequence of n values such that $\xi_{n_k} < \xi'_{n_k}$. If there does not exist such a subsequence then there is nothing to prove.

For $n = n_k$, first note that

$$\sum_{j=1}^{|S|} \psi_j = \sum_{j=1}^{|S|} \psi'_j = |A||B_l(y)| \ .$$

We will split these sums up into three parts: from $j = 1$ to ξ_{n_k}, from $j = \xi_{n_k} + 1$ to ξ'_{n_k}, and from $j = \xi'_{n_k} + 1$ to $|S|$. In the first part we have

$$\sum_{j=1}^{\xi_{n_k}} \psi_j \leq \sum_{j=1}^{\xi_{n_k}} \psi'_j$$

where the inequality is due to (25). In the last part we have

$$\sum_{j=\xi'_{n_k}+1}^{|S|} \psi_j \leq \frac{v_n}{2}(|S| - \xi'_{n_k}) + \sum_{j=\xi_{n_k}+1}^{|S|} \psi'_j$$

because the rightmost term is greater than or equal to zero and ψ_j is less than $\frac{v_n}{2}$ for the range of the sum. Putting all of these together we get

$$\sum_{j=\xi_{n_k}+1}^{\xi'_{n_k}} \psi'_j \leq \frac{v_n}{2}(|S| - \xi'_{n_k}) + \sum_{j=\xi_{n_k}+1}^{\xi'_{n_k}} \psi_j$$

and because $\psi'_j \geq v_n$ and $\psi_j \leq \frac{v_n}{2}$ over the range of this sum,

$$\frac{v_n}{2}(\xi'_{n_k} - \xi_{n_k}) \leq \frac{v_n}{2}(|S| - \xi'_{n_k}) \ .$$

■

Lemma 4.3 reduces our problem to showing that Lemma 4.1 is true for stable sets A. From now on we will assume that A is stable. The defining property of a stable set that will be useful is that if $y = (\mathrm{y}_1, \ldots, \mathrm{y}_n) \in A$, and if $\mathrm{y}_j = 1$ and $\mathrm{y}_i = 0$ for $i < j$, then $\sigma_{ij}(y)$ must also be in A. If this were not the case then A^*_{ij} would not equal A. Using this property we can see that the following subset U must be contained in any stable set A with the prescribed cardinality:

$$U = \left\{ y \in S \ \middle| \ \begin{array}{l} y_i = 1, \ i = 1, \ldots, n(p - r_n) \\ y_i = 0, \ i = n(p + r_n), \ldots, n \end{array} \right\} \ .$$

Let $y_0 \in S$ be the string with np ones in the first np components and then zeros in the remaining $n(1-p)$ components. The reason U must be contained in any stable A is that a stable set with cardinality $2^{n[H(p)-a_n]}$ must contain at least one point with distance greater than or equal to $2r_n n$ from y_0. The value r_n is defined by (19) to be such that a spherical cap the same size as A would have radius $2r_n n$, so this set is too large to have only points with distance less than $2r_n n$ from y_0. By stability, if a point with distance $2r_n n$ or greater from y_0 is in A then all of U must be in A. Note that there is actually a vanishing ϵ_n term that we have ignored in this analysis. The value of r_n only represents the radius of a spherical cap up to the addition of some $\epsilon_n \to 0$. We can ignore this extra term by letting U be defined by some slightly smaller r_n with the difference approaching zero as $n \to \infty$. This will not affect the resulting exponents.

Lemma 4.4: Consider y to be a random variable with uniform distribution over S, $A \subseteq S$ with $|A| = 2^{n[H(p)-a_n]}$, and $2p(1-p) - 2r_n p \leq l_n \leq 2p(1-p)$. Define

$$\Delta_n = l_n - (2p(1-p) - 2r_n p) \ .$$

Then

$$\Pr(|U \cap B_{l_n}(y)| \geq 2^{n(V_n - \delta)}) \to 1$$

as $n \to \infty$, where

$$V_n \geq 2r_n p H\left(\frac{\Delta_n}{4r_n p}\right) + 2r_n(1-p) H\left(\frac{r_n(1-2p) + \Delta_n/2}{2r_n(1-p)}\right) \ ,$$

r_n is defined by (19), and $\delta \to 0$.

Proof: A typical element $y \in S$ will have p proportion of ones in each of the following three blocks of components:

(i) $i = 1, \ldots, n(p - r_n)$
(ii) $i = n(p - r_n) + 1, \ldots, n(p + r_n)$
(iii) $i = n(p + r_n) + 1, \ldots, n$.

Note these typical elements are distance $2np(1-p) - 2nr_n p$ away from the set U. For some $l_n \geq 2p(1-p) - 2r_n p$, we would like to count how many elements of U are within distance nl_n of a typical $y \in S$.

Fix any typical $y \in S$. To get from y to one of the closest elements of U, there are are $2nr_n\left(\frac{1}{2} - p\right)$ ones from block (iii) that need to be moved to block (ii). We can also pick up to $n\Delta_n/2$ ones from within block (ii) to move to different locations within block (ii). This gives at least

$$\binom{2nr_n p}{n\Delta_n/2}\binom{2nr_n(1-p)}{2nr_n\left(\frac{1}{2} - p\right) + n\Delta_n/2}$$

elements. The first term is for the $n\Delta_n/2$ ones that were chosen from the $2nr_n p$ ones in block (ii). The second term is for choosing the $2nr_n\left(\frac{1}{2} - p\right) + n\Delta/2$ locations out of $2nr_n(1-p)$ spots within block (ii) that did not start with ones. These are the locations that will be filled with ones from both blocks (ii) and (iii). Characterizing the exponent of this expression gives the bound on V_n in the lemma. ■

We can now proceed with the proof of Lemma 4.1. The proof will build on the all of the previous lemmas, as well as another symmetrization-like procedure that will further simplify stable sets.

Proof (Lemma 4.1): For stable A, define A' to be a subset

$$A' = \{a \in A \mid d(y_0, a) \leq 2nR_n\}$$

where R_n is such that

$$|\{a \in A \mid d(y_0, a) = 2nR_n\}| \geq \frac{2^{n[H(p)-a_n]}}{2np+1} \ .$$

The reason such an R_n must exist is that there are at most $2np + 1$ different distances that a point in S can be from y_0. The following two properties of A' can be easily verified:

(i) $\liminf_{n\to\infty} R_n - r_n \geq 0$
(ii) A' is stable.

Instead of thinking about A as a set, consider a more general setting where A is replaced by a function $f : S \to \mathbb{R}_{\geq 0}$ and we are interested in the "generalized intersection value"

$$\psi(y) = \sum_{x\in S} f(x) 1_{B_l(y)}(x)$$

under the constraint that

$$\sum_{y\in S} f(y) = |A| \ .$$

Starting from $f(y) = 1_A(y)$, consider the following procedure:

for $i = 1, \dots, |\{a \in A \mid d(y_0, a) = 2nR_n\}|$,

1: pick one $y \in A'\backslash U$ such that $A'\backslash\{y\}$ is still stable and $d(y_0, y) = 2nR_n$
2: $A' \leftarrow A'\backslash\{y\}$
3: let $U_i = \{u \in U \mid f(u) = \min_{y\in U} f(y)\}$
4: pick one $y' \in U_i$ such that the set $U_i' = (U\backslash U_i) \cup \{y'\}$ is stable
5: $f(y) \leftarrow f(y) + 1_{\{y'\}}(y) - 1_{\{y\}}(y)$

Loosely speaking, this process takes each point $y \in A'\backslash U$ and stacks it on top of U – filling the resulting function in layer by layer. The loop will terminate with some

$$f(y) \geq \left\lfloor \binom{2^{n[H(p)-a_n]}}{2np+1} \Big/ \binom{2^{2nr_n}}{2} \right\rfloor 1_U(y) \ . \qquad (26)$$

The y, y' that are chosen in steps 1 and 4 can always be found such that the intended sets are stable. Explicitly, one could pick y that is the smallest candidate element in the lexicographic order. Any element that was relying on such a y for stability would need to be even smaller in the lexicographic order. Similarly, y' could be picked to be the largest candidate element in the lexicographic order.

The claim is that after each iteration of this loop, $f = f^*$ and the functional

$$\max_{y_1,\dots,y_k} \sum_{j=1}^{k} \psi(y_j) = \max_{y_1,\dots,y_k} \sum_{j=1}^{k} \sum_{x\in S} f(x) 1_{B_{l_n}(y_j)}(x) \qquad (27)$$

does not decrease for any fixed k. The property $f = f^*$ follows immediately from the stability of $A'\backslash\{y\}$ and U_i' and the fact that A' starts as a stable set.

We would like to show that (27) does not decrease on each iteration of the loop. We can assume the $y_1, \dots, y_k$ that give the maximum in (27) form a stable set. If they did not, then by applying Corollary 1 with $g = 1_{\{y_1,\dots,y_k\}}$ we would achieve at least the same value with $g^* = 1_{\{y_1,\dots,y_k\}^*}$. For these $y_1, \dots, y_k$ the following inequality holds:

$$\sum_{j=1}^{k} \sum_{x\in S} 1_{\{y'\}}(x) 1_{B_{l_n}(y_j)}(x) \geq \sum_{j=1}^{k} \sum_{x\in S} 1_{\{y\}}(x) 1_{B_{l_n}(y_j)}(x) \ . \qquad (28)$$

Due to property (i) of A', we can assume $R_n \geq r_n$ by defining U with a slightly smaller r_n with the difference approaching zero as $n \to \infty$. This means y' can be reached from y by shifting ones from the right to the left. Each time a one is shifted to the left, the number of elements in $\{y_1, \dots, y_k\}$ within distance nl_n of that point cannot decrease. This proves (28) which proves the claim.

Finally, using the the fact that (27) does not decrease during the construction of (26), we can exactly reproduce the proof of Lemma 4.3 to show that if

$$\Pr\left(\sum_{x\in S} f(x) 1_{B_{l_n}(y)}(x) \geq v_n\right) \to 1 \ ,$$

then

$$\Pr\left(\sum_{x\in S} 1_A(x) 1_{B_{l_n}(y)}(x) \geq \frac{v_n}{2}\right) \to 1 \ .$$

Furthermore,

$$\sum_{x\in S} f(x) 1_{B_{l_n}(y)}(x) \geq$$
$$\left\lfloor \binom{2^{n[H(p)-a_n]}}{2np+1} \Big/ \binom{2^{2nr_n}}{2} \right\rfloor \sum_{x\in S} 1_U(x) 1_{B_{l_n}(y)}(x)$$

so using Lemma 4.4 we arrive at the final result.

■

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] Y.-H. Kim, "Coding techniques for primitive relay channels," in *Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput.*, Monticello, IL, Sep. 2007.
[8] Albert Baernstein II and B.A. Taylor, "Spherical rearrangements, subharmonic functions, and $*$-functions in n-space," *Duke Mathematical Journal*, vol. 43, no. 2, pp. 245–268, 1976.
[9] X. Wu, A. Ozgur, L.-L. Xie, "Improving on the cut-set bound via geometric analysis of typical sets," *IEEE Trans. Inform. Theory*, vol. 63, pp. 2254–2277, April 2017.
[10] L. H. Harper, "On a problem of Kleitman and West," *Discrete Mathematics*, vol. 93, no. 2-2, pp. 169-182, 1991.
[11] A. Burchard, "A Short Course on Rearrangement Inequalities," `http://www.math.toronto.edu/almut/rearrange.pdf`, June 2009.
[12] B. List, "A Vertex Variant of the Kleitman-West Problem," in *Operations Research and Discrete Analysis*, Springer Netherlands, 1997.
[13] K. Marton, A simple proof of the blowing-up lemma, IEEE Trans. on Information Theory, vol. 32, no. 3, pp. 445446, May 1986.