Ramsey Theory and Repeated Communication Complexity (Review)

Vijay Bhattiprolu and David Wajc

December 8, 2014

Abstract

We summarize the 1995 paper of Alon and Orlitsky, "Repeated Communication and Ramsey Graphs" [1], in which the authors characterize the capacity of channels with adversarial noise which is unbounded in the amount of noise, but rather in the type of noise. A second problem they address is dual-source coding, in which the task is to minimize the amount of information needed to transmit one party's information, given that the receiving party has some "relevant" information. The authors relate these values to well-studied problems in graph theory and show, among others, that there exist channels in which the amount of information per use transmittable over multiple uses is exponentially larger than in one use alone, and there exists exist channels where one instance transmitted require arbitrarily many bits but any subsequent instances require only a little over one bit per instance. We make a few observations concerning improvements of their results.

1 The Model

1.1 Channel Coding

A *channel* consists of a finite input set X, a possibly infinite output set Y, and a non-empty subset of outputs per input, $S_x \subseteq Y$ for all $x \in X$. In each use, the *sender* transmits an input $x \in X$, and the *receiver* receives some $y \in S_x$. The goal of the sender and the receiver is to devise a communication protocol which allows them to transmit the maximum amount of information using such a channel, despite it being possibly noisy (i.e. some $x, x' \in X$ have overlapping output sets, $S_x \cap S_{x'} \neq \varnothing$).

Given such a noisy channel, we may concern ourselves with the possibility of zero-loss communication over it. Following Shannon, Alon and Orlitsky do just this. We now state the channel properties considered in their paper. As the channel C discussed will always be clear from context, we avoid any cumbersome notation which explicitly specifies C.

Definition 1.1. *The* single-use capacity *of C, denoted by $\gamma^{(1)}$, is the maximum number of bits the sender can communicate over C without error in a single use.*

Equivalently, $\gamma^{(1)}$ is the log of the largest number of inputs the sender can communicate an input over C without error in a single use. For example, for a completely noisy channel in which $S_x \cap S_{x'} \neq \varnothing$ for all $x \neq x'$, we have $\gamma^{(1)} = 0$. For a completely noise-free channel, in which $S_x \cap S_{x'} = \varnothing$ for all $x \neq x'$ we have $\gamma^{(1)} = \log |X|$.[1] Note that this value need not be integral.

Definition 1.2. *The* n-repeated use capacity *of C, denoted by $\gamma^{(n)}$, is the maximum number of bits the sender can communicate over C without error in n uses.*

[1] In this paper, all log's are taken to base 2.

Equivalently, $\gamma^{(1)}$ is the log of the largest number of inputs n-tuples the sender can communicate over C without error in n uses. Clearly $\gamma^{(n)} \geq (\gamma^{(1)})^n$, as the sender can repeat the single-use protocol n times. We shall see later that for some (random) channels much better can be done.

Definition 1.3. *For multiple uses, the* per-use *number of bits the users can convey over n uses of the channel C, as defined by Shannon [5], is*

$$C^{(n)} \triangleq \frac{\gamma^{(n)}}{n}.$$

As the sequence $\{\gamma^{(n)}\}$ is super-additive,[2] the sequence $C^{(n)}$ converges to a limit, $C^{(\infty)}$, known as Shannon's *zero-error capacity* of the channel C. Note that $C^{(\infty)} \geq C^{(n)}$ for all n.

Note that the above definitions are independent of the channel (input) size. This in some way does not capture the fact that we would hope a larger channel would be more useful. (Conversely, a larger channel which has the same capacity as a small channel will only prove more computationally expensive when trying to devise an efficient protocol for it.) This motivated Alon and Orlitsky to give the following definition.

Definition 1.4. *The* normalized n-use capacity *of a channel C is*

$$\tilde{C}^{(n)} \triangleq \frac{C^{(n)}}{\log |X|}.$$

1.2 Dual-Source Coding

A *dual source* consists of a finite set X, a possibly infinite set Y, and a *support set* $S \subseteq X \times Y$. In each instance, the *sender* wishes to transmit some $x \in X$, and the *receiver* wishes to learn x given that it has some $y \in Y$ such that $(x, y) \in S$. The goal of the sender and the receiver is to devise a communication protocol which allows the receiver to learn x while transmitting as few bits of information as necessary. Intuitively, the receiver knowing y and that $(x, y) \in S$, gives it a little bit of information on what x might be. We are interested in studying how much a support set can help the receiver and also if there are certain good support sets that are increasingly helpful as the number of instances increases.

Definition 1.5. *The* single-instance cost *of S, denoted by $\sigma^{(1)}$, is the minimum number of bits the sender must send the receiver for the receiver to learn a sender's single input x (given y).*

In case many instances of S are to be transmitted, the sender has inputs $x_1, \dots, x_n \in X$ and the receiver has inputs $y_1, \dots, y_n \in Y$ such that $(x_i, y_i) \in S$ for all $i \in [n]$. This setting calls for the following definition.

Definition 1.6. *The* n-instance cost *of S, denoted by $\sigma^{(n)}$, is the minimum number of bits the sender must send the receiver to convey the sender's n inputs $x_1, \dots, x_n$ (given $y_1, \dots, y_n$).*

Clearly $\sigma^{(n)} \leq (\sigma^{(1)})^n$, as the sender and receiver can repeat the single-instance protocol n times. We shall see later that for some dual sources much better can be done.

Definition 1.7. *For multiple uses, the* per-use *number of bits the users can convey over n uses of the channel S, as defined by Witsenhausen [6], is*

$$R^{(n)} \triangleq \frac{\sigma^{(n)}}{n}.$$

[2] In $n + m$ uses one can simulate one n-use and one m-use protocol, to transmit $\gamma^{(n+m)} \geq \gamma^{(n)} + \gamma^{(m)}$ bits.

Again, as the sequence $\{\sigma^{(n)}\}$ is sub-additive,[3] the sequence $R^{(n)}$ converges to a limit, $R^{(\infty)}$, known as Witsenhausen's *zero-error capacity* of S. Note that $R^{(\infty)} \geq R^{(n)}$ for all n.

Finally, again mirroring the structure of the channel coding part of the paper, Alon and Orlitsky consider the normalized n-instance rate of S, defined below.

Definition 1.8. *The* normalized n-use rate *of a dual source S is*

$$\tilde{R}^{(n)} \triangleq \frac{R^{(n)}}{\log|X|}.$$

1.3 Relation to Graph Theory

A channel C as defined above has a *characteristic graph* G, defined as follows: every input symbol $x \in X$ is a vertex, and vertices x, x' neighbor if their output sets overlap; that is, $S_x \cap S_{x'} \neq \emptyset$. Note that every graph $G = (V, E)$ is the characteristic graph of a code, where the input set is $X = V$, output set is $Y = E$ and the individual output sets are the edges coinciding with the vertex $S_x = (u, v) \in E$.

Similarly, a dual-code S as defined above has a *characteristic graph* G, defined as follows: every symbol $x \in X$ is a vertex, and vertices x, x' neighbor if their fan-out sets overlap; that is, $S_x \cap S_{x'} \neq \emptyset$. Note that every graph $G = (V, E)$ is the characteristic graph of a dual-code, where $X = V$, $Y = E$ and the support set is $S = \{(v, e) \mid v \in E\}$.

Given this relation between our problems and graphs, we may well wonder whether natural properties of such problem instances can be quantified using properties of their characteristic graph. Indeed, Alon and Orlitsky show that this is the case. Before stating their results, we formally define some graph-theoretic quantities and operations used throughout the paper.

1.3.1 Useful Graph-Theoretic Definitions

Given an undirected graph $G = (V, E) = (V[G], E[G])$, we say a subset of vertices $U \subseteq V$ forms a *clique* (resp. *independent set*) if all (resp., none) of its endpoints are interconnected by edges in E. The *clique number* of G, denoted by $\omega(G)$, is the size of its largest clique. The *independence number* of G, denoted by $\alpha(G)$, is the size of its largest independent set. The *complement* of a graph G, denoted by $\bar{G}$, is a graph on the same vertex set as G with an edge $(u, v) \in E[\bar{G}]$ iff $(u, v) \notin E[G]$. Clearly $\alpha(G) = \omega(\bar{G})$. A *coloring* of a graph is an assignment of colors to its vertices such that neighboring vertices have distinct colors. The *chromatic number* of a graph G, denoted $\chi(G)$, is the minimum number of colors of a coloring of G.

The *AND product* of n graphs $G_1 = (V_1, E_1), \ldots, G_n = (V_n, E_n)$ is the graph $G_1 \wedge \cdots \wedge G_n$ whose vertex set is the Cartesian product $V_1 \times \cdots \times V_n$ with an edge between distinct n-tuples $(v_1, \ldots, v_n)$ and $(v'_1, \ldots, v'_n)$ iff $v_i = v'_i$ or $(v_i, v'_i) \in E_i$ for all $i \in [n]$. The *OR product* of n graphs, $G_1 \vee \cdots \vee G_n$, is defined similarly, with an edge between its vertices $(v_1, \ldots, v_n)$ and $(v'_1, \ldots, v'_n)$ iff $(v_i, v'_i) \in E_i$ for *some* $i \in [n]$. As we would expect from the notation and De Morgan's Laws, we have

$$\overline{G_1} \wedge \cdots \wedge \overline{G_n} = \overline{G_1 \vee \cdots \vee G_n}.$$

For the special case where G_i are all equal to the same graph G, we denote by $G^{\wedge n}$ and $G^{\vee n}$ the n-fold AND product and OR product of G with itself. We will concern ourselves with the maximum independence number of an AND product graph with all "multiplicands" having independence number at most l_i, denoted by

$$\rho(l_1, \ldots, l_n) \triangleq \max\{\alpha(G_1 \wedge \cdots \wedge G_n) \mid \alpha(G_i) \leq l_i \;\; \forall i \in [n]\}.$$

For the case where all graphs G_i are equal, we define

$$\rho_n(l) \triangleq \max\{\alpha(G^{\wedge n}) \mid \alpha(G) \leq l\}.$$

[3] In order to transmit $n + m$ uses one can simulate one n-instance and one m-instance protocol, requiring a total of $\sigma^{(n)} + \sigma^{(m)}$ bits. Consequently, $\sigma^{(n)} + \sigma^{(m)} \geq \sigma^{(n+m)}$.

Graph Definitions Specific to Channel Coding. The *Ramsey Number* $r(l_1,\dots,l_n)$ for some integer $n \ge 1$ is the largest integer r such that there exists an n-coloring (i.e. an assignment of values in $[n]$) of the edges of a complete graph on r vertices, K_r, where every monochromatic clique of color i has size at most l_i. Formally,

$$r(l_1,\dots,l_n) \triangleq \max\{r \mid \exists\chi : E[K_r] \to [n], \text{ s.t. } \forall i \in [n], \forall U \subseteq V \text{ satisfying } \chi_{|U\times U} = i, |U| \le l_i\}.$$

For a single color we have $r(l) = l$ and if some l_i is zero we have $r(l_1,\dots,l_n) = 0$. For the balanced case where all l_i are equal to some l, we abbreviate

$$r_n(l) \triangleq r(\overbrace{l,\dots,l}^{n}).$$

We will later rely on the following upper and lower bounds on 2-color Ramsey Numbers, $r_n(l)$, often taught in combinatorics classes (See [2]).

$$\sqrt{2}^l \le r_2(l) < 4^l \tag{1}$$

Graph Definitions Specific to Dual-Source Coding. The Kneser graph $K = K(u,t)$ consists of all $\binom{u}{t}$ t-element subsets of $[u]$. Two vertices are connected iff they are disjoint. Every vertex can be colored with one of its elements, hence $\mathcal{X}(K) \le u$. Lovasz [4] showed that $\mathcal{X}(K) = u - 2t + 2$.

Let A be a finite Abelian group. A set $K \subseteq A$ is symmetric if $-K = K$. The *Cayley graph* over A with respect to a symmetric set K has A as its vertex set and distinct vertices $a, b \in A$ are connected iff $a - b$ (hence also $b - a$) is in K. All operations involving vertices are performed in A.

2 Results

2.1 Channel Coding

As noted above, every channel C can be represented by a graph G, and every graph G represents some channel. This relation allows one to fully characterize the power of noisy channels for communication by relating the seemingly disparate topics of repeated communication and graph theory, yielding some rather surprising results.

Recall that the single-use capacity of C, $\gamma^{(1)}$, denotes the number of bits of information transmittable in one use of the channel, and generally $\gamma^{(n)}$ is the n-use capacity of C. While it is clear that $\gamma^{(n)} \ge n \cdot \gamma^{(1)}$, and in particular $\gamma^{(2)} \ge 2\gamma^{(1)}$, Alon and Orlitsky showed that there exist an exponential improvement in number of bits transmittable in two uses compared to one. They also showed that this is asymptotically tight.

Theorem 2.1. *There exist channels C such that $\gamma^{(2)} \ge 2^{\gamma^{(1)}-1}$.*
On the other hand, for all channels C it holds $\gamma^{(2)} < 2^{\gamma^{(1)}+1}$.

They prove these claims by relating the single-use capacity to the independence number of the channel's characteristic graph, and its multi-use capacity independence number of the n-fold AND product of the graph with itself. They then relate the largest discrepancy between the single- and multi-use capacity to Ramsey Numbers. In particular, they show that $\rho_n(l) \triangleq \max\{\gamma^{(n)}|\gamma(1) \le \log l\} = \log r_n(\log l)$. Known bounds on 2-color Ramsey Numbers $r_2(k)$ yield the bounds of the theorem. For the n-use per-use capacity of C, $C^{(n)} = \gamma^{(n)}/n$, as $C^{(\infty)} \ge C^{(2)}$, and $C^{(2)} = \gamma^{(2)}/2$, the above yields the following corollary.

Corollary 2.2. *There exist channels such that their per-use capacity is exponentially larger than their single-use capacity. That is,*

$$C^{(\infty)} \geq 2^{C^{(1)-2}} = 2^{\gamma^{(1)}-2}$$

Alon and Orlitsky ask whether this is the largest possible improvement for repeated communication versus single-use communication. However, the correspondence between Ramsey Theory and this problem implies that resolving this question would resolve a long-standing open problem of Erdős. (More on this when we discuss the proof of the above theorem.)

Finally, for the normalized n-use capacity of the channel, $\tilde{C}^{(n)} = C^{(n)} / \log|X| = \gamma^{(n)}/n \log|X|$, they prove the existence of a channel of size X for any $|X| \equiv 0 \mod 4$ with single-use capacity $\gamma^{(1)} \leq \log\lceil 2\log|X|\rceil$ and double-use capacity $\gamma^{(2)} \geq \log|X|$, implying the following theorem.

Theorem 2.3. *For every $\epsilon > 0$ there exists a channel such that $\tilde{C}^{(1)} \leq \epsilon$ but $\tilde{C}^{(\infty)} \geq \tilde{C}^{(2)} \geq \frac{1}{2}$.*

They also show that the above is essentially the best possible, as the next theorem asserts.

Theorem 2.4. *For every channel C and integer $n \geq 2$, $\tilde{C}^{(n)} - \tilde{C}^{(1)} \leq 1 - \frac{1}{n}$. In particular, $\tilde{C}^{(2)} - \tilde{C}^{(1)} \leq \frac{1}{2}$*

2.2 Dual-Source Coding

As noted above, every dual source S can be represented by a graph G, and every graph G represents some dual source. As with channel coding, this relation allows one to fully characterize the power of dual-source coding in graph theoretic terminology, with surprising results.

Recall that the single-use rate of S, $\sigma^{(1)}$, denotes the number of bits of the sender needs to send the receiver to convey $x \in X$, given the receiver has some $y \in Y$ satisfying $(x, y) \in S$, and generally $\sigma^{(n)}$ is the n-use rate of S, denoting the equivalent number of bits necessary to transmit n instances $x_1, \ldots, x_n \in X$. While it is clear that $\sigma^{(n)} \leq n \cdot \sigma^{(1)}$, and in particular $\sigma^{(2)} \geq 2\sigma^{(1)}$, the authors showed that for arbitrarily-large values of $\sigma^{(1)}$ there exist dual sources where a second instance requires only a few more bits, as asserted by the following theorem.

Theorem 2.5. *There exist dual codes S such that $\sigma^{(2)} \leq \sigma^{(1)} + 6$.*

They prove this theorem by relating the single-use rate to the chromatic number of the channel's characteristic graph, and its multi-use capacity chromatic number of the n-fold AND product of the graph with itself. Using this characterization they then show the above discrepancy in single- and double-instance rates can be achieved by dual sources with Kneser Graphs as characteristic graphs. Using the same graphs (with different parameters) they then show that for arbitrarily-large values of $\sigma^{(1)}$ there exist dual sources where a single instance requires to transmit $\sigma^{(1)}$ bits, whereas the per-instance number of bits necessary over a long sequence is effectively one per instance, as the following theorem states.

Theorem 2.6. *For every $\epsilon, \sigma^{(1)} > 0$, there exist dual sources such that their single-instance rate is $\sigma^{(1)}$ but their per-instance multiple-instance is ϵ-close to one. That is,*

$$R^{(1)} = \sigma^{(1)} \quad \textit{but} \quad R^{(\infty)} \leq 1 + \epsilon$$

Finally, for the normalized 2-instance rate of the source, $\tilde{R}^{(n)} = R^{(n)} / \log|X| = \sigma^{(n)}/n \log|X|$, they prove the existence of a channel of size X for any $|X| \equiv 1 \mod 4$ a prime power with single-instance cost $\sigma^{(1)} = \log|X| - O(\log\log|X|)$ and double-instance cost $\sigma^{(2)} \leq \log|X|$, and consequently implying the following theorem.

Theorem 2.7. *For every $\epsilon > 0$ there exists a source such that $\tilde{R}^{(1)} \geq 1 - \epsilon$ but $\tilde{R}^{(\infty)} \leq \tilde{R}^{(2)} \leq \frac{1}{2}$.*

We also see that the above is essentially the best possible, as the next lemma asserts.

Lemma 2.8. *For every dual-source and integer n, $\tilde{R}^{(1)} - \tilde{R}^{(n)} \geq \left(1 - \frac{1}{n}\right) \tilde{R}^{(1)}$.*

3 Proofs

3.1 Channel Coding

As alluded to in Section 2 above, the capacity of a channel and the independence number of its characteristic graph are related. In particular, any set of inputs the sender can send without any possible collision imply that no two symbols' fan-out sets intersect, and in particular the subgraph induced by these symbols forms an independent set in the graph G. Any (zero-error) single-use protocol would therefore require the sender and receiver to agree on an independent set of G. This implies the following easy observation.

Observation 3.1. *The single-use capacity of a code C with characteristic graph G is equal to*

$$\gamma^{(1)} = \log \alpha(G).$$

A repeated, n-use protocol has the sender send some n-tuple of X (equivalently, of $V[G]$), $\bar{x} = (x_1, \dots, x_n)$, following which the receiver receives some n-tuple of Y, $(y_1, \dots, y_n) \in S_{x_1} \times \dots \times S_{x_n}$. This can be thought of as using some larger channel C_n, with input set X^n and output set Y^n, and $S_{\bar{x}} = S_{x_1} \times \dots \times S_{x_n}$ for all $\bar{x} = (x_1, \dots, x_n)$. For two input symbols $\bar{x} = (x_1, \dots, x_n)$ and $\bar{x'} = (x'_1, \dots, x'_n)$ to have their fan-out sets intersect, we must have $S_{x_i} = S_{x'_i}$ for all $i \in [n]$. In other words, we find that the characteristic graph of the n-use channel C_n is simply the n-fold AND product of G with itself, $G^{\wedge n}$. As a consequence of this and the above observation we obtain the following characterization of the n-use capacity of the channel.

Corollary 3.2. *The n-use capacity of a code C with characteristic graph G is equal to*

$$\gamma^{(n)} = \log \alpha(G^{\wedge n}).$$

In order to obtain a characterization of the maximum $\gamma^{(n)}$ as a function of $\gamma^{(1)}$, Alon and Orlitsky define the following:

$$\rho(l_1, \dots, l_n) \triangleq \max\{\omega(G_1 \vee \dots \vee G_n) \mid \omega(G_i) \le l_i \text{ for all } i \in [n]\}.$$

Recall that $\overline{G_1} \wedge \dots \wedge \overline{G_n} = \overline{G_1 \vee \dots \vee G_n}$ and $\alpha(G) = \omega(\bar{G})$, as noted above. Consequently, we find that $\rho(l_1, \dots, l_n)$ can be re-written in a form more clearly-relevant to the problem at hand, namely $\rho(l_1, \dots, l_n) = \max\{\alpha(G_1 \wedge \dots \wedge G_n) \mid \alpha(G_i) \le l_i \text{ for all } i \in [n]\}$.[4] In similar fashion Alon and Orlitsky define $\rho_n(l) \triangleq \max\{\alpha(G^{\wedge n}) \mid \alpha(G) \le l\}$. The following two theorems yield a characterization of $\rho_n(l)$.

Theorem 3.3. [5] *For all integers $n \ge 1$ and $l \ge 0$,*

$$\rho(l, \dots, l) = r(l, \dots, l) = r_n(l).$$

Theorem 3.4. *For every integer $n \ge 1$,*

$$\rho_n(l) = \rho(l, \dots, l).$$

Before outlining the proofs of these theorems, we show why this implies bounds on $\gamma^{(2)}$ as a function of $\gamma^{(1)}$, given by the following theorem.

Theorem 3.5. *There exist channels C such that $\gamma^{(2)} \ge 2^{\gamma^{(1)}-1}$.*

On the other hand, for all channels C it holds $\gamma^{(2)} < 2^{\gamma^{(1)}+1}$.

[4] The reason to consider this "negation" of the value of interest is to simplify the proofs to follow.

[5] This theorem is in fact stated for possibly different l_i, but as we will not make use of this general formulation, we only state this restricted version of the theorem.

Proof. Let C be a code with single-use capacity $\gamma^{(1)} = \log l$. Then its characteristic graph G satisfies $\alpha(G) = l$. On the other hand, combining Theorems 3.3 and 3.4 we have $\rho_n(l) = \rho(l) = r_n(l)$. Plugging in the upper bound on Ramsey Numbers stated in 1, $r_2(l) < 4^l$, we find that

$$\gamma^{(2)} = \log \alpha(G^{\wedge 2}) \leq \log r_2(l) < 2l = 2^{\gamma^{(1)}+1}.$$

On the other hand, as all graphs define a channel, the graph G which achieves $\rho_2(l) = r_2(l)$ induces a channel with $\gamma^{(1)} = \log \alpha(G) \leq \log l$ and, this time relying on the lower bound in 1,

$$\gamma^{(2)} = \log \alpha(G^{(2)}) \geq r_2(l) \geq \log(\sqrt{2})^l = \frac{1}{2} \cdot l \geq 2^{\gamma^{(1)}-1}.$$

□

We now return to Theorems 3.3 and 3.4 and sketch their proofs. We begin with the proof of Theorem 3.4, which will allow us to slightly simplify the notation in the proof of Theorem 3.3.

Proof of Theorem 3.4. Clearly $\rho_n(l) \leq \rho(l, \dots, l)$, so we proceed to show $\rho_n(l) \geq \rho(l, \dots, l)$. Let $\rho = \rho(l_1, \dots, l_n)$. Let $G_1, \dots, G_n$ be graphs with $\omega(G_i) \leq i$ for all $i \in [n]$ such that $G = G_1 \vee \cdots \vee G_n$ satisfies $\omega(G) = \rho$. Let $S = \{(x_1^1, \dots, x_n^1), \dots, (x_1^\rho, \dots, x_n^\rho)\}$ be a clique of size ρ in G. Let H be a graph with vertex set $[\rho] \times [n]$, with an edge between vertices $(a, i), (b, i)$ iff $(x_i^a, x_i^b) \in E[G_i]$. By construction, $\omega(H) \leq \max\{\omega(G_i)\} \leq l$. On the other hand, the set $\{((a,1), \dots, (a,n)) \mid a \in [\rho]\}$ is a clique, as every $a, b \in [\rho]$ satisfy $((x_1^a, \dots, x_n^a), (x_1^b, \dots, x_n^b)) \in E[G] = E[G_1 \vee \cdots \vee G_n]$, and in particular there exists some i for which $(x_i^a, x_i^b) \in E[G_i]$. □

Proof of Theorem 3.3. $\rho_n(l) \geq r_n(l)$. Let $r = r_n(l)$. Fix an n-coloring χ of the complete graph with vertex set $[r]$ such that all monochromatic cliques have size at most l. Let G_i be a copy of the subgraph induced by the i-colored edges under χ. Then for all i the largest clique in G_i has size $\omega(G_i) \leq l$ for all $i \in [n]$. On the other hand, $\{(1, \dots, 1), \dots, (r, \dots, r)\}$ is a clique in $G_1 \vee \cdots \vee G_n$, as for every edge (u, v) is colored by some color $i \in [n]$ and therefore $(u, v) \in E[G_i]$

$r_n(l) \geq \rho_n(l)$. Let $\rho = \rho_n(l)$. Let $G_1, \dots, G_n$ be graphs with $\omega(G_i) \leq l$ such that $\omega(G_1 \vee \cdots \vee G_n) = \rho$. Let $S = \{(x_1^1, \dots, x_n^1), \dots, (x_1^\rho, \dots, x_n^\rho)\}$ be a clique of size ρ in $G^{\vee n}$. We n-color K_ρ by assigning each edge (u, v) with $u, v \in [\rho]$ some arbitrary color i such that $(x_i^u, x_i^v) \in E[G_i]$, as guaranteed by the fact that $((x_1^u, \dots, x_n^u), (x_1^v, \dots, x_n^v)) \in E[G_1 \vee \cdots \vee G_n]$. Every monochromatic set has size at most l, as an i-monochromatic set implies a clique of same size in G. □

3.1.1 ($\star$) Constant-Factor Improvement

Note that Erdős's familiar probabilistic proof method by which the lower bound on $r_2(c)$ (and therefore $C^{(2)}$) is achieved can be extended to show that 3-use per-use capacity is slightly better than 2-use per-use capacity, as the following theorem asserts.

Theorem 3.6. *There exit channels for which the 3-use per-use capacity satisfies*

$$C^{(3)} \geq \frac{\log(3)}{3} \cdot 2^{C^{(1)}-1}.$$

Proof. Given the above proofs, it suffices to prove a lower bound on $r_3(s)$. Indeed, by randomly 3-coloring the edges of a complete graph on $N = \lceil 3^{(s-1)/2}) \rceil$ vertices, the expected number of monochromatic cliques of size s is bounded from above by

$$\binom{N}{s} \cdot 3^{1-\binom{s}{2}} \leq \frac{N^s}{s!} \cdot 3^{-(s(s-1))/2} < 3^{(s(s-1))/2} \cdot 3^{-(s(s-1)/2)} = 1$$

In particular, by integrality of the number of monochromatic s-cliques, there exists a 3-coloring of k_N with no s-cliques. Consequently $\rho_3(s-1) = r_3(s-1) \geq N = \lceil 3^{(s-1)/2} \rceil \Rightarrow \rho_3(s) \geq 3^{s/2}$. As with the above proof, there exists a graph G which achieves $\rho_3(s)$ and therefore has $\gamma^{(1)} = \log s$ and $\gamma^{(3)} = \log \rho_3(s) = \log r_3(s) \geq \log 3^{s/2} = \log 3 \cdot s/2 = \log 3 \cdot 2^{\gamma^{(1)}-1}$. The theorem follows. □

Contrast the above per-use capacity of $C^{(3)} = \frac{\log 3}{6} \cdot 2^{\gamma^{(1)}} \approx 0.264 \cdot 2^{\gamma^{(1)}}$ to that given in the original paper, of $C^{(2)} = 0.25 \cdot 2^{\gamma^{(1)}}$. Generally, the same proof shows that for all $n \geq 2$ there exists a channel with n-use capacity satisfying $\gamma^{(n)} \geq \log n \cdot 2^{\gamma^{(1)}-1}$. However, as $\log n/n$ decreases for all $n \geq 3$ (verifiable, for example, by confirming that $(\log_2 x/x)' < 0$ for all $x \geq 3$), we find that the 3-use per-use capacity is the most for which "delving into the internals of the proofs" give any improvement over the original paper, and that this improvement is only a constant multiplicative improvement at that.[6]

3.1.2 Normalized Per-Use Capacity

This subsection concerns the normalized per-use capacity of a code C with v-vertex characteristic graph, $\tilde{C}^{(n)} = \frac{\gamma^{(n)}}{n \cdot \log v}$. We restate and outline a proof of the relevant theorems, 2.3 and 2.4, below.

Theorem 3.7. *For every $\epsilon > 0$ there exists a channel such that $\tilde{C}^{(1)} \leq \epsilon$ but $\tilde{C}^{(\infty)} \geq \tilde{C}^{(2)} \geq \frac{1}{2}$. Conversely, for every channel and integer n, $\tilde{C}^{(n)} - \tilde{C}^{(1)} \leq 1 - \frac{1}{n}$, and in particular $\tilde{C}^{(2)} - \tilde{C}^{(1)} \leq \frac{1}{2}$.*

Proof (Sketch). The upper bound on the discrepancy between $\tilde{C}^{(n)}$ and $\tilde{C}^{(1)}$ follows from the simple observation that for every v-vertex graph G and integer n, $\alpha(G^{\wedge n}) \leq v^{n-1} \cdot \alpha(G)$; therefore

$$\tilde{C}^{(n)} - \tilde{C}^{(1)} = \frac{\log \alpha(G^{\wedge n})}{n \cdot \log v} - \frac{\alpha(G)}{\log v} \leq \frac{\log v \cdot (n-1) + \alpha(G)}{n \cdot \log v} - \frac{\alpha(G)}{\log v} \leq 1 - \frac{1}{n}.$$

To prove the existence of channels (nearly) meeting this bound, Alon and Orlitsky focus on *self-complementary* graphs. These are graphs that are isomorphic to their complement. That is, there exists some permutation π such that $(u, v) \in E \iff (\pi(u), \pi(v)) \notin E$. They start by observing that any self-complementary v-vertex graph G satisfies $\alpha(G^{\wedge 2}) \geq v$.[7] They then proceed to give a probabilistic argument showing the existence of such a self-complementary v-node graph with $\alpha(G) = O(\log v)$ for all $v \in 4\mathbb{Z}$. For such a graph G the channel has $\tilde{C}^{(2)} = \frac{O(\log \log v)}{\log v}$, and $\tilde{C}^{(2)} = \frac{\log v}{2 \log v} = \frac{1}{2}$. Taking v large enough yields the claimed bound. We outline Alon and Orlitsky's existential proof of such a self-complementary graph below.

Let $v = 4a$ be some integer multiple of 4 and let $\mathbb{Z}_v$, the integers modulo v be the set of vertices (all addition here is modulo v). Define an equivalence relation where two edges $\{x, y\}$ and $\{x', y'\}$ by adding $0, a, 2a$ or $3a$ to both elements of the edge. That is, $\{x, y\} \sim \{x', y'\} \Rightarrow \{x', y'\} \in \{\{x + ia, y + ia\} \mid i \in 0, 1, 2, 3\}$. Choose the edge set randomly by picking half the edges of each equivalence class randomly. For two-edge classes (in case $x = y + 2a$) pick one of the two edges. For four-edge equivalence class pick a pair of edges that can be obtained from each other by adding $2a$ to both elements. That is, $\{x, y\}$ and $\{x + 2a, y + 2a\}$ for some x and y. By construction, the resulting graph is self-complementary under the permutation $\pi(v) \to v + a$. On the other hand, any set of vertices which contains elements $x, x + a$ and $y, y + a$ is not independent (again, by construction), but for any set of vertices of size $2\lceil \log v \rceil$ without such vertices the probability of the set being independent is such that the expected number of such independent sets is less than 1, and therefore there exists such a graph G with no independent set of the above size, or $\alpha(G) = O(\log v)$, as claimed above. □

3.1.3 ($\star$) A Generalization

We note that in general, for every self-complementary v-node graph G (i.e., a graph for which there exists some permutation π such that for all $u, v \in V$, exactly one of $(u, v), (\pi(u), \pi(v))$ are in E), G's n-fold AND product with itself, $G^{\wedge n}$, effectively has as large an independent

[6] Given that Alon literally wrote the book on the probabilistic method [2], we suppose this was not stated for the sake of brevity and/or simplicity.

[7] See the next subsection for a proof of a more general claim.

number as one could hope for. That is, $\alpha(G^{\wedge n}) \geq v^{n-1}$. To see this, consider the set of vertices $S = \{(v_1, \pi(v_1), v_2, v_3 \ldots, v_n) \mid v_i \in V \ \forall i \in [n]\}$. As all $u_1, v_1 \in E$ have either $(u_1, v_1) \notin E$ or $(\pi(u_1), \pi(v_1)) \notin E$, we find that S is an independent set of size v^{n-1}. Consequently, the construction of Alon and Orlitsky gives us the following, more general result.

Theorem 3.8. *For every $\epsilon > 0$, $n \geq 2$, there exists a channel such that $\tilde{C}^{(1)} \leq \epsilon$ but $\tilde{C}^{(\infty)} \geq \tilde{C}^{(n)} \geq 1 - \frac{1}{n}$. Conversely, for every integer $n \geq 2$ and every channel, $\tilde{C}^{(n)} - \tilde{C}^{(1)} \leq 1 - \frac{1}{n}$.*

3.2 Dual-Source Coding

The smallest number of possible messages the sender must transmit for a single instance of S is $\mathcal{X}(G)$, the chromatic number of the dual-source's characteristic graph. Intuitively, the sender and receiver agree in advance on a coloring of G. Given x, the sender transmits its color. The receiver, having y, can determine x because there is exactly one element of X with this color that is jointly possible with y. Conversely, it is easy to see that if two connected vertices are assigned the same message, an error can result. Thus,

$$\sigma^{(1)} = \log \mathcal{X}(G)$$

It is not hard to see, that multiple instances of a dual-source with characteristic graph G, can be viewed as a single instance of a dual-source with characteristic graph $G^{\wedge n}$. Thus we are interested in finding graphs that have arbitrarily large $\mathcal{X}(G)$, yet $\mathcal{X}(G^{\wedge n})$ is not too large compared to $\mathcal{X}(G)$.

But first, we see that the best one can do is given by,

Lemma 3.9. *For every dual-source and integer n, $\tilde{R}^{(1)} - \tilde{R}^{(n)} \geq \left(1 - \frac{1}{n}\right) \tilde{R}^{(1)}$.*

Proof. For every graph G and integer n, $\mathcal{X}(G^{\wedge n}) \geq \mathcal{X}(G)$. Hence $\tilde{R}^{(n)} \geq \tilde{R}^{(1)}/n$ and the claim follows. □

On to the first main theorem,

Theorem 3.10. *Let K be the Kneser graph $K(u, t)$. Then,*

$$\mathcal{X}(K^{\wedge n}) \leq \mathcal{X}(K^{\vee n}) \leq \left\lceil \left(\frac{u}{t}\right)^n n \ln \binom{u}{t} \right\rceil$$

Proof. First, observe that if every vertex $(v_1, \ldots, v_n)$ in $K^{\vee n}$ is colored with $z \in [u]^n$ such that $\forall i, z_i \in v_i$ (we refer to this as a representative coloring), then we obtain a valid coloring of $K^{\vee n}$. Indeed, consider any 2 distinct vertices $(v_1, \ldots, v_n)$ and $(u_1, \ldots, u_n)$ that have the same color and are adjacent. Since they're adjacent, there is some i, such that v_i is adjacent to u_i in K, implying $v_i \cap u_i = \phi$. Then the colors assigned must be different (by definition of representative coloring) - a contradiction.

Now pick $m = \left\lceil \left(\frac{u}{t}\right)^n n \ln \binom{u}{t} \right\rceil$ colors $z \in [u]^n$ independently and uniformly at random. The probability of a particular color being a representative color for a particular vertex is $\left(\frac{t}{u}\right)^n$. Thus the probability of no color being representative for the vertex is at most $\left(1 - \left(\frac{t}{u}\right)^n\right)^m$. The claim then follows by a union bound over the $\binom{u}{t}^n$ vertices. □

Finally, by simple algebraic manipulation we get,

Corollary 3.11. *For every $\epsilon, \sigma^{(1)} > 0$, there exist dual sources such that their single-instance rate is $\sigma^{(1)}$ but their per-instance multiple-instance is ϵ-close to one. That is,*

$$R^{(1)} = \sigma^{(1)} \quad \textit{but} \quad R^{(\infty)} \leq 1 + \epsilon$$

We next show a slightly weaker counterpart theorem to that presented by Alon and Orlitzky, which still has the same qualitative implications on normalized rate.

Theorem 3.12. *For every prime power $v \equiv 1 \mod 4$, there is a v-vertex graph G, such that,*

$$\mathcal{X}(G) \geq v/O(\log^3 v) \qquad \text{and} \qquad \mathcal{X}\left(G^{\wedge 2}\right) \leq v$$

This immediately yields,

Corollary 3.13. *For every $\epsilon > 0$ there exists a source such that $\tilde{R}^{(1)} \geq 1 - \epsilon$ but $\tilde{R}^{(\infty)} \leq \tilde{R}^{(2)} \leq \frac{1}{2}$.*

Proof of 3.12 The essence of the proof is in finding self-complementary Cayley graphs with low independence number, as testified by the following lemma:

Lemma 3.14. *Every self-complementary Cayley graph over A, has $\mathcal{X}\left(G^{\wedge 2}\right) \leq |A|$.*

Proof. Let π be a bijection mapping G onto its complement. The mapping

$$c(x, y) = x - \pi(y)$$

has range $|A|$. We'll see that it colors $G^{\wedge 2}$. Suppose that $c(x, y) = c(x', y')$. We claim that either $(x, y) = (x', y')$ or $\{(x, y), (x', y')\} \notin G^{\wedge 2}$. By the definition of c,

$$x - x' = \pi(y) - \pi(y').$$

Now either $x = x'$, implying $\pi(y) = \pi(y)$, and the vertices (x, y) and (x', y') coincide, or, $x \neq x'$, implying that $y \neq y'$, and by the earlier equation, the definition of G, and self-complementarity, $\{x, x'\} \in G$ iff $\{\pi(y), \pi(y')\} \in G$ iff $\{y, y'\} \in G$, showing that (x, y) and (x', y') are not connected $G^{\wedge 2}$. □

Now for the existence of Cayley graphs with such properties,

Theorem 3.15. *For every prime power $v \equiv 1 \mod 4$, there is a self-complementary cayley graph over F_v, with independence number $O(\log^3 v)$.*

The proof follows directly from the following two lemmas:

A set $T \subseteq A$ is said to have heavy differences if $T - T$ contains $|T|^2/2$ distinct non-zero elements.

Lemma 3.16. *For every prime power $v \equiv 1 \mod 4$, there is a self-complementary cayley graph over F_v, containing no $\Omega(\log v)$ size independent set with heavy-differences.*

Proof. Consider any order 4 subgroup $\{1, a, -1, -a\}$ of F_v^*. For each coset $\{x, xa, -x, -xa\}$, randomly and independently choose either x and $-x$ to go into K, or xa and $-xa$ to go into K. This determines a Cayley graph G over F_v.

Clearly, $\pi(x) = ax$ is a bijection under which self-complementarity follows.

Fix a t-element independent set with heavy-differences. If $T - T$ contains nonzero elements t_1, t_2 with $t_1 = at_2$ then (exactly) one of them is in K and T cannot be independent. Otherwise, the probability that none of the $\geq t^2/4$ distinct pairs of non-zero elements $x, -x$ in $T - T$ belongs to K, is $\leq 2^{-t^2/4}$. Now the claim follows by a union bound over all possible t-element sets. □

Lemma 3.17. *Every s-element set in an Abelian group of odd order contains a $\Omega(s)^{1/3}$-element subset whose non-zero differences are all distinct.*

Proof. Let S be an s-element abelian group of odd order. Start with $A_0 = \{x_0\}$ for any $x_0 \neq 0$ in S. For all $i < (2s)^{1/3}$, we add a vertex x_i to A_{i-1} to get $A_i = A_{i-1} \cup \{x_i\}$ such that for any i, $x_i \in S \smallsetminus A_{i-1}$ satisfies, $(x - A_{i-1}) \cap ((A_{i-1} - T) \cup (A_{i-1} - x)) = \phi$. This clearly implies the claim, and we need only show that such x_i exists for all $i < (2s)^{1/3}$.

We thus need to show that for every $T \subseteq S$ of size $t \leq (2s)^{1/3}$, there is an element $x \in S \smallsetminus T$ satisfying the above property. For any T, Let

$$T' = \{a + b - c : a, b, c \in T\} \cup \{(a + b)/2 : a, b \in T\}$$

For any $g \in S$, $g/2$ exists because S,has odd order. If $x - T$ intersects $(T - T) \cup (T - x)$, then $x \in T'$. But,

$$|T'| \leq t + \binom{t}{2} + t(t-1) + \binom{t}{2}(t-2) < s$$

Hence $S \smallsetminus T' \neq \phi$, and the claim follows. □

4 Discussion

The constructions in this paper mostly relied (either implicitly or explicitly) on the probabilistic method, and are as such existential rather than constructions of explicit examples. It might be interesting to "derandomize" these proofs. Note that the relations to graph theoretic problems would in that case imply similar results for the graph theoretic problems. In general a question this paper brings up is whether these characterizations of communication problems with graph theoretic terms could prove useful in obtaining better bounds for the relevant graph theoretic questions. We state these explicitly below.

4.1 Channel Coding

The gap in capacity between single- and double-use implies the existence of channels with Shannon capacity exponentially larger than their single-use capacity, $C^{(\infty)} \geq C^{(2)} \geq 2^{C^{(1)}-2}$. However, Alon and Orlitsky achieve this only by considering two-use protocols. A natural question posed by the authors can be restated thus: "Is this the largest possible gap, or can more be achieved by considering $n \geq 2$-use protocols?" Formally, they ask for a lower bound, given some fixed c, on

$$C^{(n)} = \frac{\gamma^{(n)}}{n} = \max\{\log \alpha(G^{\wedge n}) \mid \alpha(G) \leq 2^c\}n = \frac{\log \rho_n(2^c)}{n} = \frac{\log r_n(2^c)}{n}$$

The authors asked in particular whether $C^{(\infty)}$ grows arbitrarily compared to $C^{(1)}$, which in turn implies that $C^{(n)} = \log r_n(2^c)/n$ grows as a function of n, and therefore $r_n(2^c) \geq 2^{\omega(n)}$. This question was posed by Erdős earlier, and so the resolution of their question would earn its solver a check for $X\$$ and eternal fame (at least within some circles).

4.2 Dual-sources

Alon and Orlitsky showed that for every $\epsilon > 0$ there is a dual-source such that $\tilde{R}^{(1)} \geq 1 - \epsilon$ but $\tilde{R}^{(\infty)} \leq \tilde{R}^{(2)} \leq \frac{1}{2}$. **Open Question:** Is it true that for every $\epsilon > 0$ there is a dual-source such that $\tilde{R}^{(1)} \geq 1 - \epsilon$ but $\tilde{R}^{(\infty)} \leq \epsilon$? A positive answer would have interesting applications to communication complexity. Karchmer, Raz, and Wigderson [3] related certain open problems in computational complexity to the number of bits required to communicate multiple instances of problems with high communication complexity. Improved dual-source results as above, thus have direct applications in this realm.

Alon and Orlitsky proved that for every prime power $v \equiv 1 \mod 4$, there is a self-complementary cayley graph over F_v, with independence number $O(\log^2 v)$. **Open Question:** It would also be interesting to see whether this can be reduced to $\log v$.

References

[1] N. Alon and A. Orlitsky, *Repeated communication and ramsey graphs*, Information Theory, 1994. Proceedings., 1994 IEEE International Symposium on, Jun 1994, pp. 315–.

[2] Noga Alon and Joel H Spencer, *The probabilistic method*, John Wiley & Sons, 2004.

[3] Mauricio Karchmer, Ran Raz, and Avi Wigderson, *Super-logarithmic depth lower bounds via the direct sum in communication complexity*, computational complexity **5** (1995), no. 3-4, 191–204 (English).

[4] L Lovász, *Kneser's conjecture, chromatic number, and homotopy*, Journal of Combinatorial Theory, Series A **25** (1978), no. 3, 319 – 324.

[5] C.E. Shannon, *The zero error capacity of a noisy channel*, Information Theory, IRE Transactions on **2** (1956), no. 3, 8–19.

[6] H. Witsenhausen, *The zero-error side information problem and chromatic numbers (corresp.)*, Information Theory, IEEE Transactions on **22** (1976), no. 5, 592–593.